

The Role of Tides and Differential Rotation in the Angular Momentum Evolution of Low Mass Stars

Jessica Birky

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Rory Barnes, Chair

James Davenport

Eric Agol

Program Authorized to Offer Degree:
Astronomy

©Copyright 2026
Jessica Birky

University of Washington

Abstract

The Role of Tides and Differential Rotation in the Angular Momentum Evolution of Low Mass Stars

Jessica Birky

Chair of the Supervisory Committee:
Rory Barnes
Astronomy

Stellar rotation is a fundamental property that governs magnetic activity, angular momentum evolution, and the dynamical interactions of stars with their companions. Despite the transformative datasets provided by *Kepler* and *TESS*, extracting physical information about stellar surfaces and tidal interactions from photometric time series remains challenging due to model degeneracies, computational expense, and the lack of physically interpretable statistical frameworks. This thesis develops novel statistical methods to address these challenges and applies them to advance our understanding of stellar rotation, tidal dissipation, and starspot variability in low-mass stars and binary systems.

In the second chapter, I present `alabi`, an open-source framework for accelerating Bayesian inference with computationally expensive forward models. By training Gaussian process surrogates via active learning, `alabi` reduces the number of expensive model evaluations by orders of magnitude while preserving posterior accuracy, achieving speedups of 10–1000× for problems with up to 64 dimensions.

In the third chapter, I apply `alabi` to investigate the prospects for constraining equilibrium tidal dissipation in low-mass binary stars. Using global sensitivity analysis and simulated Bayesian inference with coupled tidal, stellar evolution, and magnetic braking models, I show that constraining the tidal quality factor $\mathcal{Q}$ for individual systems is fundamentally limited by degeneracies with unknown initial conditions, even with ideal observational precision. Moreover, I identify population-level approaches, particularly overdensities in the observed

distribution of orbital periods, rotation periods, and eccentricities, as a more promising path forward.

In the fourth chapter, I derive a physically motivated Gaussian process kernel for stellar photometric variability from first principles. Starting from an analytic starspot model, I obtain a closed-form autocovariance function with hyperparameters (the equatorial rotation period, differential rotation shear, stellar inclination, spot lifetime, and spot emergence/decay timescale) that map directly onto measurable physical properties. I validate the kernel against numerical simulations, demonstrate that it captures harmonic structure absent in standard phenomenological kernels, and show that a banded Cholesky solver reduces the computational scaling from $\mathcal{O}(N^3)$ to $\mathcal{O}(Nb^2)$, where b is the bandwidth of the covariance matrix. Applied to Kepler-63, the kernel recovers rotation, differential rotation, and inclination consistent with independent transit and spectroscopic measurements. This result constitutes the first demonstration that differential rotation can be inferred from one-dimensional photometric lightcurves using a physically interpretable Gaussian process.

ACKNOWLEDGEMENTS

To my mom, Linda Petersen, for always supporting my interests and curiosities, no matter how arcane. Thank you for all the opportunities you have given me to try new things, for always prioritizing giving me the best education possible, and for letting me have any pet that I could ever want.

I want to thank my many mentors throughout my career, who have helped shape my view of the universe and how I approach problems. Thank you to Desire Whitmore, who first introduced me to research and took the time to mentor me and teach me programming and physics, despite not having a whole lot to offer as a highschool student. Thank you to my undergraduate advisor Adam Burgasser and his group at UC San Diego (particularly Chris Theissen, Dino Hsu, and Christian Aganze) for teaching me many of the fundamental skills of how to do research, write papers, and present at conferences. Perhaps most importantly, Adam has taught me the importance of metacognitive reflection, and to approach every problem with a growth mindset. Adam has always been extremely dedicated to mentoring undergrads and advocating for better equity in the department and in astronomy more broadly, and I am grateful to have had him as a mentor during my formative years as a scientist. I have often looked back on my experiences with Adam and returned to his example as I mentored others during my PhD. Thank you to my other undergrad advisor David Hogg, for his infectious enthusiasm for science and for making me realize that anything can be an interesting problem when you're working with people you enjoy. David taught me the importance of asking well-posed questions and has shaped much of my intuition for how I approach problems and think about data and statistics.

Thank you to my PhD advisor, Rory Barnes, who has consistently encouraged me to seek out new ideas that I am curious about. Rory has taught me a lot about how to write proposals, and how to maintain a positive outlook despite many setbacks and failures during my PhD. He has always been positive about my work and believed in my abilities, even when I had significant doubts about myself. I admire and respect his receptiveness to criticism and his desire to always improve, which are qualities I try to emulate in my own work. Thank you to my other main PhD mentor, Jim Davenport, for always being someone I can bounce ideas off of, and for getting me to explore new areas, including encouraging me to use APO. Thank you to my committee member Eric Agol, for being there to answer my questions about math, and for being one of the kindest and most humble scientists I know.

Finally, I wouldn't be here without the support of everyone in the UW grad community, who have helped me get through the rough times in grad school and have been some of my greatest friends. I am especially proud of my cohort Sam Gilbert-Janizek and Héctor Delgado Diaz, who have made it with me through this PhD starting from the pandemic and ending in a historical bad job market. I particularly want to thank Tom Wagg, who gave me a lot of advice on postdoc apps and provided me with the funding that made it possible to finish this last few quarters of my PhD; my officemate David Wang, for keeping

all of the office plants alive and for planning so many fun social events and ski trips; Gabby Suissa, for always inviting me over to hang out with her cats and for making me come on a spontaneous trip to NYC during the week that I was stressed about finishing my thesis and making me have fun; Oge Okoronkwo, for always being down for anything and for hosting the party office; V Hurtado, for their sense of humor and for reminding me to take care of myself; Andy Tzanidakis and Bonni Choi for being there to commiserate with me when I was stressed over job applications; Bethlee Lindor for always keeping it real and reminding me to not get caught up in bullshit; and Dino Bektesevic, Rudy Garcia, and Tyler Gordon, for being the wise old grads of my generation.

CONTENTS

Acknowledgements	i
Contents	iii
1 Introduction	1
1 Stellar Rotation	2
2 Statistical Methods	7
3 The Dynamical Evolution of Binary Systems	8
4 Measuring Differential Rotation	12
5 Thesis Work	22
2 alabi: Active Learning for Accelerated Bayesian Inference	25
1 Introduction	26
2 Methods	27
3 Implementation	37
4 Results and Performance	40
5 Discussion	57
6 Conclusion	64
3 Prospects of Constraining Equilibrium Tides in Low-Mass Binary Stars	65
1 Introduction	66
2 Model	69
3 Methods	76
4 Results	81
5 Discussion	103
6 Conclusion	109
1 Equilibrium Tide Equations	111
4 A Physically Motivated Gaussian Process Kernel for Stellar Rotation	116
1 Introduction	118
2 Lightcurve Model	121
3 Gaussian Process	128
4 Deriving the Mean Function	132
5 Deriving the Analytical Kernel for Starspot Variability	134
6 Physical Interpretation of Kernel Hyperparameters	149
7 GP Solver Implementation	161
8 MCMC Sampling the Kernel Hyperparameters	164
9 Numerical Results	168
10 Case Study: Application to Kepler-63	188
11 Discussion	202
12 Conclusion	204

1	Fourier Coefficients of the Visibility Function	205
2	Fourier Transform of the Trapezoidal Envelope	207
3	Autocorrelation of the Squared Envelope	209
5	Conclusion	212
1	Summary	213
2	Synthesis	214
3	Future Directions	215
	Bibliography	217

CHAPTER 1

INTRODUCTION

1	Stellar Rotation	2
1.1	Angular momentum evolution and gyrochronology	2
1.2	Starspots on convective envelopes	3
1.3	Differential Rotation and Stellar Dynamo Processes	3
1.4	Rossby Regimes	5
1.5	Tidal Interactions in Binary and Planetary Systems	6
1.6	Open Problems in the Field	7
2	Statistical Methods	7
2.1	Sensitivity Analysis	7
2.2	Bayesian Inference and MCMC	8
2.3	Surrogate Models and Active Learning	8
3	The Dynamical Evolution of Binary Systems	8
3.1	Tidal Circularization and Synchronization	9
3.2	The Mysterious Population of Subsynchronous Binaries	10
4	Measuring Differential Rotation	12
4.1	Parameterization of Differential Rotation	12
4.2	Challenges of Inferring Differential Rotation from Photometry	13
4.3	Differential Rotation from Spectral Line Profiles	15
4.4	Doppler Imaging	16
4.5	Photometric Period Variations	17
4.6	Individual Spot Modeling	18
4.7	Transit Spot-Crossing Events	19
4.8	Asteroseismology	19
4.9	Summary and Comparison of Methods	20
4.10	Findings from the Literature	21
5	Thesis Work	22
5.1	Paper I: <code>alabi</code>	22
5.2	Paper II: Binary Tides	23
5.3	Paper III: GP models for Stellar Variability	23

1 Stellar Rotation

Stellar rotation is a fundamental property of stars that has profound implications for the structure, evolution, activity, magnetic fields, mass loss, angular momentum distribution within stars and stellar systems. Understanding stellar rotation is crucial for interpreting a wide range of astrophysical phenomena, from the formation of planetary systems to the end stages of stellar evolution.

Recent advances in large-scale surveys such as *Kepler*, *K2*, *TESS*, and *Gaia* have delivered high-precision photometric timeseries and astrometric data, providing an unprecedented opportunity to study stellar rotation across a wide range of stellar types, evolutionary stages, and Galactic environments for nearly a million stars (McQuillan et al., 2014; Boyle et al., 2026). Concurrently with the influx of data, there has been significant progress in theoretical modeling and computational techniques for studying stellar rotation. These advances have enabled the development of more sophisticated models of stellar interiors, magnetic fields, and angular momentum transport processes.

My thesis seeks to advance our understanding of two important processes related to stellar rotation: (1) to understand how stellar rotation and tidal dissipation shape the long-term spin-orbit dynamical evolution of binary stars, and (2) to understand rotational dynamics and starspot properties on the surfaces of stars by developing novel methodology. Furthermore, I develop and apply novel statistical methods for bridging the gap between complex, computationally expensive models and the large datasets that are now available.

In this section I discuss the scientific motivation for studying stellar rotation and differential rotation. I review the current state of the field, including the open problems that motivate the work in this thesis. I also provide an overview of the statistical methods that I develop and apply in this thesis, and I summarize the main results of each chapter.

1.1 Angular momentum evolution and gyrochronology

Over time stars lose angular momentum through magnetized stellar winds, causing them to spin down as they age (Skumanich, 1972; Soderblom, 1983; Barnes, 2003a). The rate of this spin-down depends on the star’s mass, age, metallicity, and magnetic field strength, and is a key factor in the long-term evolution of stellar rotation. The change in angular momentum is driven by the torque due to stellar wind, in which ions flow out from the star along open magnetic field lines and cause the star to eventually lose angular momentum from mass loss (known as magnetic braking). However, the rate at which stars lose mass is poorly constrained and is not directly detectable for stars aside from our Sun.

Early work found that the rotation period scales according to age as $P_{\text{rot}} \propto t^{-1/2}$ (Skumanich, 1972), opening up the field of gyrochronology for inferring stellar ages based on their rotation period (Barnes, 2003b, 2007). More recent work (Matt et al., 2015; Breimann et al., 2021; Gordon et al., 2021; Lu et al., 2024) has made use of modern surveys (*e.g.*, *Kepler*) and age-dated stars from open clusters to calibrate more flexible, precise gyrochronology relations. Furthermore, recent studies such as van Saders et al. (2016) have discovered evidence for weakened magnetic braking based on rapidly rotating old field stars, suggesting that the magnetic braking process may be more complex than originally thought. Despite

these developments with new observations, our understanding of stellar spin-down is still empirical and lacks a solid theoretical understanding.

1.2 Starspots on convective envelopes

Much of our constraints on stellar rotation come from photometric variability driven by starspots. In particular, low-mass stars ($M \lesssim 1.5 M_{\odot}$) possess outer convective envelopes in which turbulent motions interact with rotation and magnetic fields to drive a stellar dynamo (Brun & Browning, 2017). Where magnetic flux tubes emerge through the stellar surface, intense local fields suppress convective energy transport, creating regions that are cooler and darker than the surrounding photosphere, which are observed as starspots (Strassmeier, 2009a). As direct, observable tracers of the subsurface magnetic field, starspots provide a window into the internal dynamo process and the geometry of magnetic flux emergence.

On the Sun, spots are confined to low latitudes ($\lesssim 35^{\circ}$), emerge as bipolar pairs, and follow the butterfly pattern of equatorward migration over the 11-year activity cycle (Hathaway, 2015). This activity cycle of the Sun has been well-established with observations dating back centuries (Schwabe & Schwabe, 1844). Sunspots are relatively small, with sizes ranging from $\sim 10 - 1000$ MSH (millionths of a solar hemisphere) and lifetimes ranging from days to months (Forgács-Dajka et al., 2021).

The spot properties for stars beyond our Sun can differ significantly. Doppler imaging of other stars has revealed that starspot properties can differ dramatically from the solar case. Starspots on active stars can be orders of magnitude larger than sunspots (covering up to tens of percent of the stellar surface) and can appear at all latitudes, including the rotational poles (Strassmeier, 2009a; Brun & Browning, 2017). These differences suggest that activity cycles could operate in fundamentally different regimes as a function of rotation rate, convective structure, and stellar mass, and that starspot distributions on other stars cannot simply be understood as scaled-up versions of the solar pattern.

1.3 Differential Rotation and Stellar Dynamo Processes

Differential rotation (the latitudinal variation in angular velocity) plays a fundamental role in the angular momentum evolution and activity cycles of a star. Figure ?? shows the solar differential rotation profile as revealed by helioseismology (Schou et al., 1998) and reproduced by recent numerical simulations (Käpylä et al., 2023). The Sun exhibits a solar-like differential rotation pattern, with the equator rotating faster than the poles.

Numerical simulations of rotating convection in stellar envelopes have shown that the magnitude and geometry of differential rotation depend on the stellar mass, rotation rate, and convective structure (Ballot et al., 2007; Augustson et al., 2012; Guerrero et al., 2013a; Brun et al., 2017a). These simulations predict a transition from solar-like differential rotation (fast equator) to anti-solar differential rotation (fast poles) as the stars spin down with age (Guerrero et al., 2013a). Furthermore, stronger differential rotation can reduce the angular momentum loss of a star and may be a plausible explanation for weakened magnetic braking, but further work is needed to observationally validate this hypothesis (Ireland et al., 2022; Tokuno et al., 2023).

Differential rotation also plays a fundamental role in the activity cycles of a star. In an effect known as a dynamo process, the spot coverage and magnetic activity on the Sun varies between maximum and minimum on a timescale of ~ 11 years as the polarity of the dipole magnetic fields periodically reverses (Käpylä et al., 2023). This activity cycle on the Sun arises due to the coupled effects of rotation, surface convection, and magnetic fields.

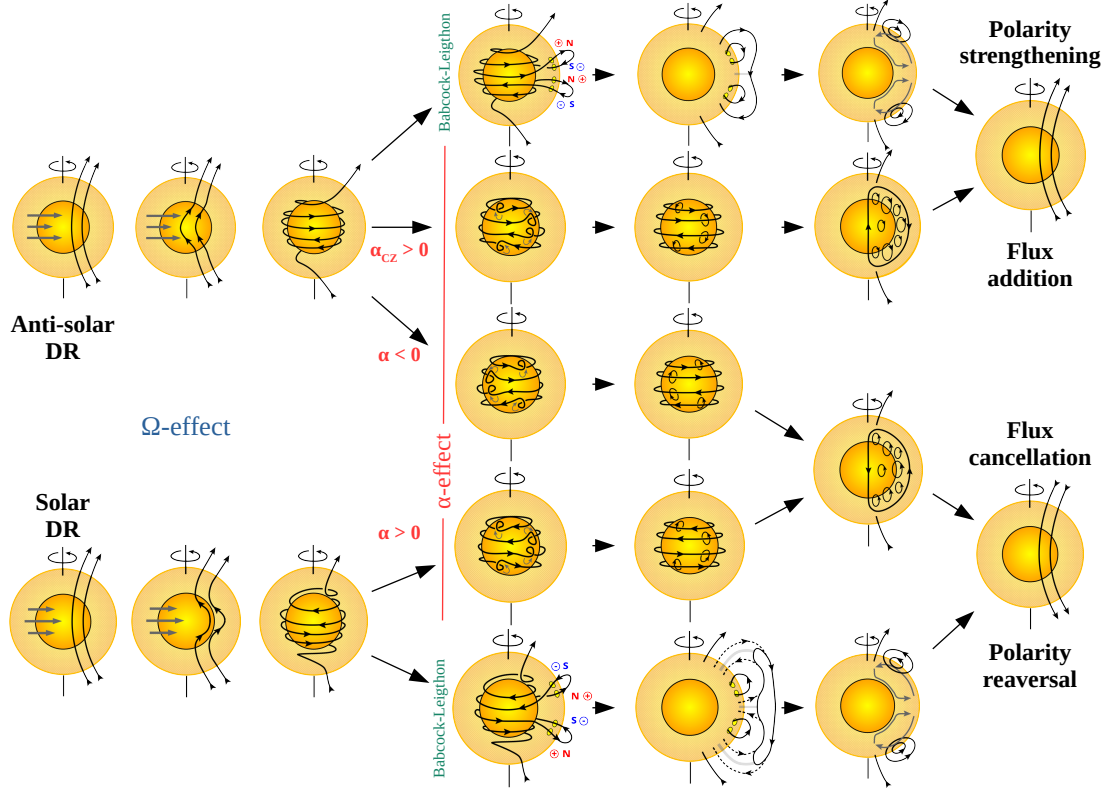


Figure 1.1: A diagram from Noraz et al. (2022b) (Figure 9 of Noraz et al. (2022b), reproduced under the creative commons license) shows how three mechanisms can drive a stellar dynamo. **Ω-effect:** starting from an initial poloidal (latitudinally-oriented) magnetic field, differential rotation twists field lines into a toroidal (longitudinally-oriented) configuration. From there, polarity reversal can occur through either the **α-effect**, where turbulence along the toroidal field creates small-scale poloidal fields, or the **BL-effect**, where bipolar magnetic regions (starspots) generate poloidal fields. These α or BL poloidal fields in combination with the sign of differential rotation (whether a star spins faster at the equator or poles) cause magnetic flux to either cancel (resulting in a dynamo), or strengthen (no dynamo).

Figure 1.1 illustrates the key mechanisms driving a stellar dynamo, highlighting the roles of differential rotation, the α -effect, and the BL-effect. The two rows starting from the left side show two different initial configurations of the magnetic field (poloidal and toroidal), and the three columns show how these fields can be transformed into each other through different physical processes. The key mechanism is the Ω -effect, in which differential rotation shears poloidal field lines into a toroidal configuration. Polarity reversal can then occur through either the α -effect, where turbulence along the toroidal field creates small-scale poloidal fields, or the BL-effect, where bipolar magnetic regions (starspots) generate poloidal fields.

The combination of these processes determines whether a dynamo can be sustained or not.

In the Sun, helioseismology has revealed two regions of strong rotational shear: the tachocline at the base of the convection zone and the near-surface shear layer (NSSL) in the outer $\sim 5\text{--}10\%$ of the solar radius. For decades, the tachocline was considered the primary seat of the solar dynamo. However, [Vasil et al. \(2024\)](#) recently demonstrated that the magneto-rotational instability (MRI), driven by the strongly inwardly increasing differential rotation in the NSSL, can explain several key observational constraints that tachocline-based models struggle to reproduce. In particular, the near-surface MRI naturally accounts for the spatiotemporal structure of torsional oscillations, the observed hemispherical helicity sign rule, and inferred subsurface magnetic field amplitudes.

These results from [Vasil et al. \(2024\)](#) suggest that the solar dynamo originates much closer to the surface than traditionally assumed, with implications for understanding magnetic cycles across stellar types – including fully convective stars that lack a tachocline entirely. Observational constraints on surface differential rotation are therefore critical for testing dynamo theories.

1.4 Rossby Regimes

The properties of starspot distributions encode information about several key physical processes, and many of these properties are governed by the Rossby number $Ro = P_{\text{rot}}/\tau_{\text{conv}}$, the ratio of the rotation period to the convective turnover timescale. The Rossby number determines the relative importance of rotational (Coriolis) forces to buoyancy-driven convection, and is the single most important dimensionless parameter for characterizing the stellar dynamo regime.

Spot lifetimes are set by the turbulent magnetic diffusivity η_t at the stellar surface, with $\tau_{\text{decay}} \sim r_{\text{spot}}^2/\eta_t$. Since η_t is determined by the same convective motions that define τ_{conv} , spot decay timescales are intrinsically linked to the Rossby number: stars with longer τ_{conv} (lower Ro at fixed rotation) have lower surface diffusivities and correspondingly longer-lived spots ([Giles et al., 2017](#)). The latitude distribution of spot emergence also depends on Ro , as it reflects the geometry of magnetic flux generation and the trajectories of rising flux tubes through the convection zone – processes that change qualitatively between the rapidly rotating ($Ro \ll 1$) and slowly rotating ($Ro \gg 1$) dynamo regimes ([Brun & Browning, 2017](#)).

Perhaps most importantly, differential rotation itself is predicted to undergo a sharp transition as a function of Rossby number (Figure ??). Numerical simulations of rotating convection ([Gastine et al., 2013](#); [Guerrero et al., 2013a](#); [Käpylä et al., 2023](#)) consistently find that rapidly rotating stars ($Ro \ll 1$) exhibit solar-like differential rotation (fast equator, slow poles), driven by Reynolds stresses from rotationally organized convective columns. As the Rossby number increases past $Ro \sim 1$, the Coriolis force weakens, convection becomes less organized, and turbulent angular momentum mixing drives a transition to anti-solar differential rotation (fast poles, slow equator).

However, this transition remains largely untested observationally, because measuring differential rotation for large samples of stars has been extremely challenging. In particular, the latitude distribution of starspots on stars other than the Sun remains poorly constrained from disk-integrated photometry, which cannot uniquely determine the surface location of individual spots ([Strassmeier, 2009a](#); [Basri & Shah, 2020](#)). Furthermore, the relationship be-

tween spot properties and fundamental stellar parameters (mass, age, rotation rate) is largely empirical, with theoretical predictions from flux-transport dynamo models only beginning to be tested against large photometric surveys (Brun & Browning, 2017).

Finally, fully convective M dwarfs, which lack a tachocline and were long predicted to be unable to sustain large-scale magnetic fields, are in fact among the most magnetically active stars, exhibiting strong fields and prominent starspot modulations, challenging traditional dynamo paradigms (Brun & Browning, 2017). Developing physically motivated statistical models that can extract spot and rotation properties from the vast photometric datasets provided by *Kepler* and *TESS* is a central goal of this thesis.

1.5 Tidal Interactions in Binary and Planetary Systems

Stellar rotation can also become coupled with the orbital dynamics of binary and planetary systems through tidal interactions. In any gravitationally bound system, the differential gravitational field of a companion raises tidal deformations on the star, and the dissipation of energy associated with these deformations drives a secular exchange of angular momentum between the stellar spin and the orbit (Hut, 1981; Leconte et al., 2010; Ferraz-Mello et al., 2008). The long-term outcome of this exchange depends critically on the star’s rotation state: tides drive the system toward a minimum-energy configuration in which the orbit is circular, the spin axis is aligned with the orbital axis, and the rotation period is synchronized with the orbital period (Counselman, 1973; Hut, 1980). The timescale on which this equilibrium is reached depends on the efficiency of tidal dissipation, parameterized by the tidal quality factor Q .

These tidal processes produce a rich variety of observable dynamical outcomes. Tidal circularization drives eccentric orbits toward $e = 0$, a trend clearly observed in binary populations where short-period systems are preferentially circular (Zahn & Bouchet, 1989; Meibom et al., 2006). Tidal synchronization locks the stellar rotation period to the orbital period, as confirmed for the majority of close binaries with $P_{\text{orb}} \lesssim 10$ days (Lurie et al., 2017; Hobson-Ritz et al., 2025).

Beyond simple synchronization, tidal torques can capture bodies into higher-order spin-orbit resonances – as exemplified by Mercury’s 3:2 resonance (Goldreich & Peale, 1966; Correia & Laskar, 2004) – and can drive the spin axis into Cassini states, equilibrium configurations in which the spin axis, orbit normal, and invariable plane normal remain coplanar while precessing together (Colombo & Shapiro, 1966; Peale, 1969). The obliquity evolution toward Cassini states has implications for the climate stability of terrestrial exoplanets (Heller et al., 2011) and for interpreting the observed distribution of spin-orbit misalignments in binary and exoplanetary systems (Albrecht et al., 2022).

Understanding stellar rotation is therefore essential for interpreting the dynamical states of binary and planetary systems across a wide range of scales. The rotation period, differential rotation profile, and obliquity of a star all influence how efficiently tidal energy is dissipated and how angular momentum is redistributed.

1.6 Open Problems in the Field

In summary, stellar rotation is a fundamental property that influences a wide range of astrophysical phenomena. Despite significant advances in observational capabilities and theoretical modeling, many open questions remain regarding the physical processes that govern stellar rotation and its coupling to orbital dynamics. This thesis investigates two important aspects of stellar rotation: the role of tidal interactions in binary systems and the development of statistical methods for analyzing stellar variability.

In particular, I focus on the following open problems:

How do we perform inference with complex and computationally expensive models?

What do learn about tidal synchronization from eclipsing binaries in large scale surveys and how can we constrain tidal Q ?

How can we infer differential rotation and starspot properties from stellar lightcurves?

In the next sections I review the motivation for the methods I use to answer these questions. I also discuss where the results of this thesis fit into the broader context of the field.

2 Statistical Methods

A recurring theme throughout this thesis is the challenge of performing rigorous statistical inference with complex and computationally expensive astrophysical models. The physical models described in the preceding sections involve solving coupled differential equations or evaluating complex forward models that can require seconds to minutes per evaluation. When combined with the need to explore high-dimensional parameter spaces, traditional inference methods become prohibitively expensive. In this section, I provide an overview of the statistical methods that I develop and apply in this thesis. These include global sensitivity analysis for understanding model behavior, Bayesian inference and Markov chain Monte Carlo (MCMC) for parameter estimation, and surrogate modeling with Gaussian processes to accelerate inference.

2.1 Sensitivity Analysis

Before attempting full Bayesian inference on a complex model, it is often valuable to first understand *which* input parameters most strongly influence the model outputs. Global sensitivity analysis (Sobol, 2001; Saltelli et al., 2010) provides a systematic framework for decomposing the variance of model outputs into contributions from individual input parameters and their interactions. Unlike local sensitivity methods that probe the response around a single point, global sensitivity analysis explores the full range of the input parameter space and captures nonlinear effects.

In the context of this thesis, sensitivity analysis is applied in Paper II (Chapter 3) to characterize how observable properties of binary stars (orbital period, rotation period, and

eccentricity) depend on the tidal quality factor Q , initial conditions, and other model inputs. This analysis reveals which parameters are constrainable in principle, which are degenerate, and where observational efforts should be focused – questions that must be answered before investing computational resources in full posterior inference.

2.2 Bayesian Inference and MCMC

Bayesian inference provides the natural framework for constraining physical parameters from observational data, combining prior knowledge with the likelihood of the data given a model to obtain the posterior distribution (Gelman et al., 2013). Markov chain Monte Carlo (MCMC) methods (Hogg et al., 2010; Speagle, 2020a) are the workhorse for sampling posterior distributions, but they require evaluating the likelihood function at each proposed step in the chain. Achieving convergence typically demands 10^4 – 10^6 likelihood evaluations, which becomes impractical when each evaluation involves running an expensive forward model.

This computational bottleneck is particularly severe for the problems addressed in this thesis. In Paper II, for example, each likelihood evaluation requires integrating the coupled tidal, stellar evolution, and magnetic braking equations over Gyr timescales – a calculation that takes ~ 1 – 10 s per evaluation. Naively running MCMC with 10^5 samples would require days to weeks of wall-clock time for a single system, and the problem scales further when considering populations of binaries or exploring the sensitivity of results to model assumptions.

2.3 Surrogate Models and Active Learning

To bridge the gap between expensive forward models and the demands of MCMC sampling, a powerful approach is to replace the likelihood function with a fast *surrogate model* trained to emulate the expensive model’s predictions (Cranmer et al., 2020). Gaussian process (GP) regression is a natural choice for building such surrogates as it provides not only a predictive mean but also a principled estimate of predictive uncertainty that can be exploited through *active learning*: iteratively selecting new training points where the surrogate is most uncertain and concentrating expensive model evaluations in the high-likelihood regions that matter most for posterior inference.

In Paper I (Chapter 2), I develop *alabi*, an open-source framework that implements this surrogate-accelerated inference pipeline. Once trained, the GP surrogate can be evaluated in microseconds, enabling millions of MCMC samples at negligible cost. This approach reduces the number of expensive model evaluations by factors of thousands while maintaining the accuracy of the posterior, and is directly applied in Paper II to constrain tidal dissipation in binary stars. The methodology is general and applicable to any problem where the likelihood is expensive to evaluate but smooth in parameter space.

3 The Dynamical Evolution of Binary Systems

In the third chapter of my thesis, I investigate the long-term dynamical evolution of low-mass binary stars under the influence of tidal interactions, magnetic braking, and stellar

evolution. The goal is to understand how these processes shape the observed distribution of orbital periods, rotation periods, and eccentricities in binary populations, and to constrain the efficiency of tidal dissipation through comparison with observations.

In this section, I introduce the key physical processes that govern binary evolution, review the current state of the field regarding tidal circularization and synchronization, and discuss the open problem of subsynchronous rotation in short-period binaries.

3.1 Tidal Circularization and Synchronization

The long-term evolution of a binary system is governed by the interplay of several physical processes (Figure 1.2). Tidal interactions between the two stars lead to an exchange of angular momentum between the stellar spins and the orbit, driving the system toward a state of minimum energy in which the orbit is circular, the spin axes are aligned with the orbital axis, and the rotation periods are synchronized with the orbital period (Counselman, 1973; Hut, 1980). The timescale on which this equilibrium is reached depends on the efficiency of tidal dissipation, parameterized by the tidal quality factor Q .

The dynamical evolution of low-mass binary stars is governed by the interplay of several physical processes, including tidal interactions, magnetic braking, and stellar evolution.

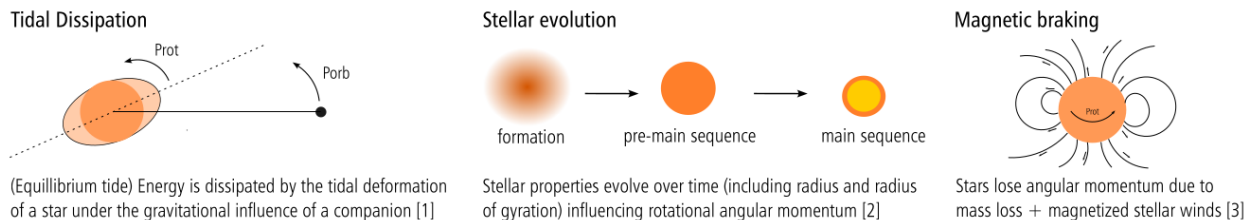


Figure 1.2: A schematic diagram of the various physical processes that govern the dynamical evolution of a binary system, including stellar evolution (Baraffe et al., 2015), magnetic braking (?), and equilibrium tides (Hut, 1981).

The gravitational interaction between components of a binary system raises tidal deformations on each star, and the dissipation of tidal energy drives the system towards a state of minimum energy (Counselman, 1973; Hut, 1980). Over long timescales (Myr–Gyr), this process causes binary orbits to circularize ($e \rightarrow 0$) and stellar rotation periods to synchronize with the orbital period ($P_{\text{rot}} \rightarrow P_{\text{orb}}$), trends that are broadly confirmed by observations (Meibom et al., 2006, 2015). Theories of tidal deformation fall into two broad categories: *equilibrium tide* models (Hut, 1981), which assume a hydrostatic tidal bulge with a phase lag parameterized by the tidal quality factor Q , and *dynamical tide* models (Zahn, 1975), which consider the excitation of internal waves driven by the gravitational influence of the companion. For low-mass stars with radiative cores and convective envelopes, the relevant dissipation mechanisms include turbulent viscosity acting on the equilibrium tide in convective layers (Zahn, 2008) and the damping of internal gravity waves in radiative zones (Zahn, 1975; Ogilvie, 2013; Barker, 2020).

Observationally, the circularization of binary orbits as a function of orbital period and age has long served as the primary diagnostic for constraining tidal efficiency. Early work introduced the *circularization cut-off period* P_{cut} (the orbital period below which most binaries

in a coeval population have circular orbits) as a metric for tidal dissipation strength (Mayor & Mermilliod, 1984; Mathieu & Mazeh, 1988; Mathieu et al., 2004). Meibom & Mathieu (2005) refined this approach with a functional fit to the period–eccentricity distribution. However, recent studies have challenged the utility of circularization as a tidal constraint: Bashi et al. (2023) found from *Gaia* spectroscopic binaries that the circularization period depends more strongly on stellar effective temperature than on cluster age, and Mirouh et al. (2023) used population synthesis to show that constraining tidal efficiency from circularization alone is difficult due to the inefficiency of main-sequence tides and strong dependence on initial orbital configurations. Importantly, Mirouh et al. (2023) demonstrated that tidal *synchronization* (the fraction of binaries with $P_{\text{rot}} \approx P_{\text{orb}}$ at a given age) better captures the age-dependent signature of tidal dissipation than circularization, motivating the use of rotation period measurements from photometric surveys.

Despite decades of effort, published estimates of the tidal quality factor Q for solar-type stars span orders of magnitude, from $\sim 10^5$ to 10^8 (Meibom et al., 2006; Jackson et al., 2009; Hansen, 2010; Patel & Penev, 2022; Penev & Schussler, 2022). This discrepancy reflects both the inherent difficulty of the inference problem (including degeneracies between Q and unknown initial conditions) and ongoing debate over whether individual system-by-system constraints (Barker, 2020) or population-level hierarchical approaches (Cameron & Jardine, 2018) are more appropriate. Motivated by these challenges, in Paper II (Chapter 3) I apply statistical methods described in Section 2 to investigate the prospects for constraining tidal dissipation parameters from binary populations, and to understand the degeneracies and limitations of such inferences.

3.2 The Mysterious Population of Subsynchronous Binaries

While equilibrium tidal models can explain the general trends of circularization and synchronization, they do not predict the existence of a population of short-period binaries with rotation periods systematically longer than their orbital periods, or in other words, a state of *subsynchronous rotation*. Lurie et al. (2017) conducted the first large-scale study of tidal synchronization in *Kepler* eclipsing binaries by visually inspecting the lightcurves of 2,278 systems and measuring rotation periods for 816 with detectable starspot modulations. They found that 79% of eclipsing binaries with orbital periods shorter than 10 days are tidally synchronized, with a clear transition near $P_{\text{orb}} \approx 10$ days from predominantly circular, synchronized systems to eccentric, pseudosynchronized systems.

However, the most striking discovery was a population of short-period eclipsing binaries with rotation periods $\sim 13\%$ slower than their orbital periods. Lurie et al. (2017) attributed this subsynchronous population to differential rotation, arguing that if the stellar equator is tidally locked but starspots preferentially emerge at higher latitudes where the rotation rate is slower (for solar-like differential rotation), the photometrically measured rotation period would appear systematically longer than the orbital period.

Following from the Lurie et al. (2017) discovery of subsynchronous binaries in the *Kepler* survey, I mentored teams of undergraduates during my PhD to investigate whether these systems are also found in the *TESS* survey. Hobson-Ritz et al. (2025) classified 4,584 lightcurves from the *TESS* eclipsing binary catalog according to out-of-eclipse variability type, identifying 584 high-confidence eclipsing binaries with starspot modulations. As shown in Figure

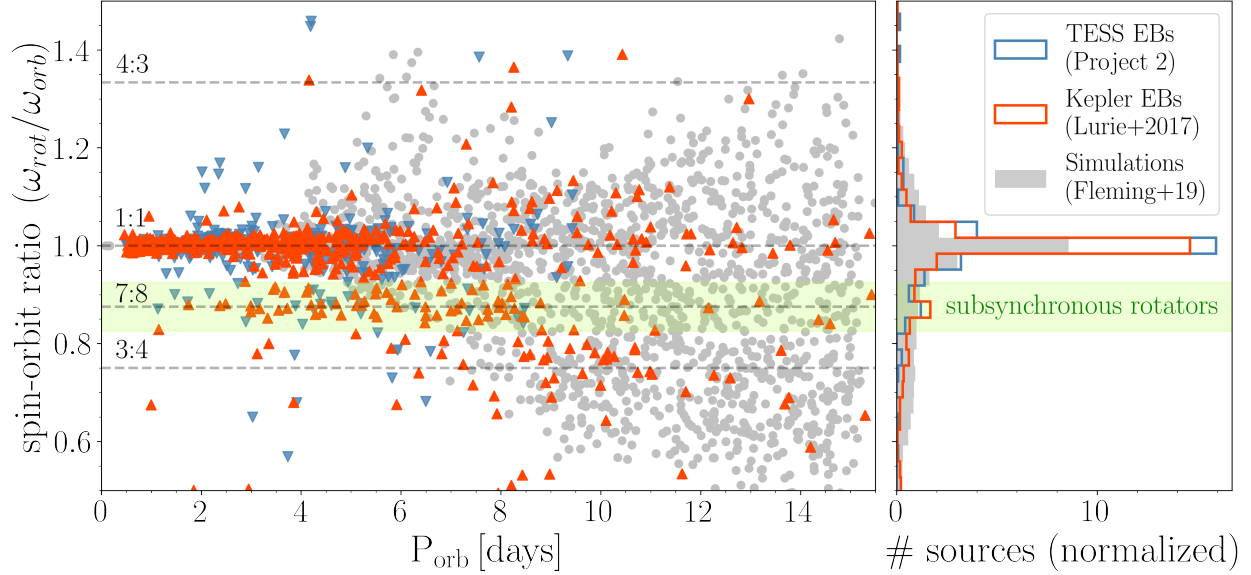


Figure 1.3: The distributions of orbital period (x-axis) and spin-orbit ratio (y-axis) for the *Kepler* (red) and *TESS* (blue) eclipsing binary samples, showing the distinct synchronous and subsynchronous populations. Equilibrium tides (gray) can explain the synchronized population with 1:1 spin-orbit ratio, but not the existence of the subsynchronous population with $\sim 7:8$ spin-orbit ratio (Fleming et al., 2019).

1.3, our analysis confirmed the existence of two distinct populations in the spin-orbit period distribution: a synchronous population ($P_{orb} \approx P_{rot}$) and a subsynchronous population with a characteristic ratio of $8 P_{orb} \approx 7 P_{rot}$, consistent with the $\sim 13\%$ period excess found by Lurie et al. (2017).

Using Bayesian model comparison, Hobson-Ritz et al. (2025) demonstrated that a bimodal distribution is a significantly better fit than a unimodal distribution for both the *Kepler* and *TESS* samples individually, as well as the combined dataset. The subsynchronous systems were found to have low eccentricities and showed no strong dependence on primary mass for FGK-type stars, suggesting the mechanism responsible is common across a broad range of low-mass binaries.

In summary, my paper Birky et al. (2025) shows that reliable constraints on tidal dissipation parameters cannot reliably be obtained for individual binary systems, and that a population approach is likely to be a more viable path forward. Furthermore, my work in Hobson-Ritz et al. (2025) confirms the existence of a population of subsynchronous binaries in the *TESS* survey, consistent with the earlier *Kepler* results, and demonstrates that this population is statistically significant and not an artifact of data quality or selection effects. The physical mechanism responsible for the subsynchronous population remains an open question, but the leading hypothesis is that it arises from differential rotation combined with preferential spot emergence at higher latitudes. Therefore, developing robust methods for measuring differential rotation from photometric time series is a critical step toward understanding the origin of the subsynchronous population and its implications for tidal interactions in binary systems. This motivates the development of the Gaussian Process

kernel method for inferring differential rotation, which I present in Chapter 4.

An additional line of evidence connecting tidal interactions and stellar activity comes from Wang et al. (2024), who analyzed starspot lifetimes on 311 detached eclipsing binaries from *Kepler*. They found that the median starspot lifetime on tidally interacting binaries is longer than on single stars at the same rotation period, and that the relative lifetime correlates with orbital parameters. This analysis was performed using a Gaussian process with a quasi-periodic kernel, which has limited interpretability, and does not include the effects of differential rotation (Nicholson & Aigrain, 2022, ; Paper III). Despite these limitations, this study suggests that tidal forces could influence the convective dynamics on the surface of stars in close binaries, further motivating the need to develop physically motivated models for starspot evolution and differential rotation.

4 Measuring Differential Rotation

While differential rotation is highly important to many fundamental processes in stellar evolution, it is also notoriously difficult to measure. The main challenge is that the signatures of differential rotation are often subtle and can be easily masked by other phenomena such as spot evolution, activity cycles, or instrumental systematics. Additionally, many methods for measuring differential rotation rely on indirect proxies (e.g., period variations, line profile shapes) that can be influenced by multiple factors, leading to degeneracies in the interpretation.

In this section, I review the main methods that have been developed to measure stellar differential rotation, including their underlying principles, advantages, and limitations. I also discuss the key challenges that arise in interpreting these measurements and how they motivate the development of new methodologies, such as the Gaussian Process kernel approach presented in this thesis.

4.1 Parameterization of Differential Rotation

The most common parameterization of surface differential rotation is through the angular rotation rate Ω as a function of latitude Φ :

$$\Omega(\Phi) = \Omega_{\text{eq}}(1 - \kappa \sin^2 \Phi), \quad (1.1)$$

where Ω_{eq} is the angular rotation rate at the equator and κ is a dimensionless parameter that quantifies the degree of differential rotation. For $\kappa > 0$, the star exhibits solar-like differential rotation, with the equator rotating faster than the poles. For $\kappa < 0$, the star exhibits anti-solar differential rotation, with the poles rotating faster than the equator. The magnitude of κ indicates how strongly the rotation rate varies with latitude, with larger $|\kappa|$ corresponding to stronger differential rotation.

This parameterization captures the essential latitudinal dependence of rotation and is widely used in both observational analyses and theoretical models of stellar rotation. Figure 18.16 of Gray (1976) illustrates this parameterization, showing how the angular velocity varies with latitude for different values of κ for the Sun.

4.2 Challenges of Inferring Differential Rotation from Photometry

Extracting differential rotation from disk-integrated photometry has long been a challenging problem, with recent studies concluding that it may be a hopeless endeavor. As [Strassmeier \(2009a\)](#) summarizes:

Measurements of the rotational period of stars with spots also allow the detection of differential surface rotation... Unfortunately, such measurements are affected by complex spot distributions, active region growth and decay, local migration due to, e.g. diffusion, spot cycles, etc., that all can either mask out the differential-rotation signal or even introduce a spurious signal. *Moreover, any one-dimensional analysis, i.e. an analysis based on disk-integrated data like time-series photometry, cannot resolve the ambiguity of whether the polar or the equatorial regions rotate faster or slower.*

The fundamental difficulty is that disk-integrated lightcurves collapse the two-dimensional surface brightness distribution into a one-dimensional time series, creating severe degeneracies between spot properties. Furthermore, the signal of differential rotation (a small difference in rotation periods at different latitudes) must be disentangled from the much larger effects of spot emergence, evolution, and decay, which operate on similar timescales and produce similar signatures in periodogram analyses.

An alternative to period-tracking methods is to fit analytic spot models directly to the lightcurve, attempting to recover individual spot parameters (latitude, longitude, size, and contrast) and thereby infer differential rotation from the latitude-dependent rotation rates of resolved spots. However, two systematic studies have demonstrated that this approach is severely limited by intrinsic degeneracies. [Walkowicz et al. \(2013\)](#) used synthetic lightcurves generated with the **Cheetah** starspot modeling code to investigate the information content of broadband precision photometry. They showed that in the absence of additional constraints, the stellar inclination and spot latitude are highly degenerate: different combinations of inclination and latitude can produce nearly identical lightcurves, particularly for inclinations within 10° – 20° of one another. When differential rotation is included, the changing longitude separation between spots breaks the symmetry of the hemispheres, but the dominant period recovered via periodogram analysis is typically that of the largest spot—even when multiple spots with periods representative of the star’s differential rotation exist. They concluded that while spot modeling can provide valuable information on spot distributions, “given the inherent degeneracies and uncertainty present in starspot models, caution should be exercised in their interpretation.”

[Basri & Shah \(2020\)](#) extended this analysis to realistic multi-spot configurations with varying spot numbers, lifetimes, differential rotation, and inclinations. Using an extensive grid of 3,000 simulated lightcurves per parameter combination, they found that the single-double dip ratio (SDR) – a metric characterizing whether the lightcurve exhibits one or two dips per rotation – contains some statistical information about differential rotation when measured over many rotations, but is much less meaningful within a few rotations. The SDR is strongly correlated with the rotation period of the star but only weakly sensitive to the differential rotation shear, and this sensitivity is further confounded by spot lifetime: for short-lived spots, the SDR distributions with and without differential rotation are nearly

indistinguishable. Only for long-lived spots ($\tau \gtrsim 20$ rotation periods) at sufficiently different latitudes does the differential shear produce a detectable drift in the lightcurve morphology, and even then, independent knowledge of the spot lifetime is required to break the degeneracy. They concluded that “differential rotation may be nearly impossible to accurately infer from light curves alone unless spots live for many rotations,” and that the Sun and solar-type stars (whose spots typically survive only one to a few rotations) represent the most difficult case.

Aigrain et al. (2015) conducted a comprehensive blind “hare-and-hounds” exercise to test the recoverability of stellar rotation and differential rotation from *Kepler*-like lightcurves. They generated 1,000 simulated lightcurves with realistic spot evolution, activity cycles, butterfly patterns, and differential rotation, injected them into actual quiescent *Kepler* lightcurves, and challenged five teams to recover the input parameters using a variety of methods (Lomb-Scargle periodograms, autocorrelation functions, wavelet analysis, and spot modeling). While the overall rotation period was well recovered by most teams (with values within 10% of the true value in 70% of cases) there was *little correlation between the reported and simulated values of the differential rotation shear*. This result demonstrates that differential rotation studies based on photometric lightcurves alone must be treated with extreme caution, at least for solar-type stars.

Morris et al. (2019b) reinforced this conclusion using the Sun itself as a benchmark. They reconstructed a solar lightcurve from archival sunspot records spanning 1917–1985 and attempted to recover the solar differential rotation signal using quasi-periodic Gaussian process regression. Despite having a noise-free lightcurve with known spot properties, they found that the apparent rotation period remained roughly constant at ~ 26.3 d throughout all phases of the activity cycle, with no detectable variation due to differential rotation. Even with “infinite” signal-to-noise, the stochastic nature of spot emergence at a broad range of latitudes at all phases of the cycle causes the measured period to reflect an average over the active latitudes rather than revealing the latitude-dependent rotation law. They concluded that measuring differential rotation from rotational modulation alone is “a fraught exercise.”

A separate but related challenge concerns the physical interpretability of Gaussian process models, which have become a common tool for modeling stellar variability (Angus et al., 2018a; Gordon et al., 2021; Morris et al., 2019b; Nicholson & Aigrain, 2022; Luger et al., 2021b; Foreman-Mackey et al., 2017, 2023). The most widely used GP covariance function for stellar rotation is the quasi-periodic (QP) kernel, which has four hyperparameters controlling the period, amplitude, evolution timescale, and harmonic complexity of the signal. Nicholson & Aigrain (2022) systematically tested the mapping between QP kernel hyperparameters and the physical parameters of a spot model by fitting GP models to simulated lightcurves with known spot properties. They found that while the GP rotation period and evolution timescale are reliably recovered, the harmonic complexity parameter Γ has no intuitive physical interpretation – it correlates only weakly with any single spot property and instead reflects a complex combination of spot number, size distribution, and lifetime. Furthermore, the GP hyperparameters derived from photometric lightcurves can differ significantly from those derived from radial velocity time series of the same star, indicating that the kernel parameters are not uniquely determined by the underlying stellar physics. Critically, the standard QP kernel contains no parameter that encodes differential rotation, meaning that any differential rotation signal is absorbed into the other hyperparameters in

an uncontrolled way.

These results collectively motivate the development of a physically grounded GP kernel – one whose hyperparameters map directly to measurable stellar properties including differential rotation – which is the subject of Paper III (Chapter 4) of this thesis.

4.3 Differential Rotation from Spectral Line Profiles

One approach to measuring stellar differential rotation comes from the analysis of spectral absorption line profiles rather than photometric time series. The Fourier Transform Method (FTM), developed and applied extensively by [Reiners & Schmitt \(2002a,b, 2003\)](#), exploits the fact that differential rotation leaves characteristic signatures in the Fourier transform of rotationally broadened spectral lines, as illustrated in Figure ?? from [Gray \(1976\)](#).

More specifically, the observed spectral line profile $D(\lambda)$ can be modeled as a convolution of an intrinsic line profile $F(\lambda)$ – which encodes thermal broadening, turbulence, damping, and instrumental effects – with a rotational broadening function $G(\lambda)$ that captures the Doppler shifts from the rotating stellar surface:

$$D(\lambda) = F(\lambda) * G(\lambda). \quad (1.2)$$

In Fourier space, convolution becomes multiplication:

$$d(\sigma) = f(\sigma) \cdot g(\sigma), \quad (1.3)$$

where σ is the Fourier frequency conjugate to velocity (in units of skm^{-1}), and d , f , g are the Fourier transforms of D , F , G , respectively.

For a rigidly rotating star with linear limb darkening $I(\theta)/I(0) = 1 - \epsilon + \epsilon \cos \theta$, the Fourier transform of the rotational broadening function has an analytic form ([Carroll, 1933](#); [Gray, 1977](#)):

$$g_r(\sigma) = \frac{2}{x(1 - \epsilon/3)} \left[(1 - \epsilon) J_1(x) + \epsilon \left(\frac{\sin x}{x^2} - \frac{\cos x}{x} \right) \right], \quad (1.4)$$

where $x = 2\pi\sigma\delta$, $\delta = \lambda_0 v \sin i / c$, and J_1 is the first-order Bessel function. This function consists of a central main lobe followed by a series of decaying sidelobes separated by zeros (nodes) at positions $\sigma_1, \sigma_2, \dots$ in Fourier frequency.

[Reiners & Schmitt \(2002b\)](#) shows that the zeros of the function $g_r(\sigma)$ can be used as a diagnostic for differential rotation for fast rotating stars. The zeros of $g(\sigma)$ are the roots of:

$$(1 - \epsilon) J_1(x) + \epsilon \left(\frac{\sin x}{x^2} - \frac{\cos x}{x} \right) = 0. \quad (1.5)$$

Since the intrinsic profile $f(\sigma)$ and turbulent broadening are smooth, non-zero functions in Fourier space, they multiply $g(\sigma)$ without moving its zeros. The zeros of the *observed* Fourier-transformed profile $d(\sigma)$ therefore coincide with the zeros of $g(\sigma)$, regardless of the (unknown) intrinsic line profile, which can be found without deconvolution or atmospheric modeling.

The zero positions are parameterized as dimensionless quantities q_n via:

$$\sigma_n = \frac{q_n(\epsilon)}{v \sin i}, \quad (1.6)$$

so the first zero σ_1 directly yields $v \sin i$ once ϵ is assumed. [Reiners & Schmitt \(2002b\)](#) showed that differential rotation systematically shifts the zero ratio q_2/q_1 , and that measuring the value of q_2/q_1 is a direct, model-independent indication that the star is *not* rotating rigidly. This effect is shown in Figure 18.16 of [Gray \(1976\)](#), which illustrates how the line profile of a rotating star is broadened by the Doppler effect, and how the presence of differential rotation can alter the shape of the line profile.

The FTM is a powerful method for detecting differential rotation, but it has specific requirements and limitations that must be considered when applying it to real data. In particular, it requires sufficiently high projected rotational velocity ($v \sin i \gtrsim 12 \text{ km s}^{-1}$) so that the Fourier zeros fall below the Nyquist frequency set by the spectral resolution R . It also requires high signal-to-noise ratio to reliably locate the zeros above the noise floor in Fourier space. However, this method is advantageous in that only a single epoch of high-resolution spectrum is needed. [Eiff & Reiners \(2012\)](#) applied this method to ~ 180 A–F stars and detected differential rotation in 56 stars, finding evidence for two distinct populations of differential rotators separated at the granulation boundary.

4.4 Doppler Imaging

Another spectroscopic approach is Doppler imaging, which reconstructs the two-dimensional surface temperature (or brightness) distribution of a rapidly rotating star from time-resolved high-resolution spectroscopy ([Vogt & Penrod, 1983](#); [Rice, 2002](#); [Piskunov & Rice, 1993](#)). As illustrated in Figure 18.31 from [Gray \(1976\)](#), cool spots cause reduced absorption at wavelengths corresponding to the Doppler velocity of their surface location, producing “bumps” that migrate through the line profile as the star rotates. By inverting a time series of line profiles covering at least one full rotation, the latitude and longitude of spot features can be recovered.

If spots are detected at different latitudes, their apparent migration rates over successive Doppler images directly reveal the latitude-dependent rotation law. This is the only method (besides direct solar observations) that can unambiguously determine both the magnitude and *sign* of differential rotation.

The principle underlying Doppler imaging is that a spot at latitude Φ and longitude Λ on a star with equatorial velocity v_* and inclination I produces a Doppler shift proportional to $\Delta v = v_* \cos \Phi \sin \Lambda \sin I$. As the star rotates, the spot’s longitude Λ changes linearly in time, causing the bump to migrate through the line profile at a rate set by the local rotation period. If the star rotates differentially, spots at different latitudes migrate at different rates, directly revealing $\Omega(\Phi)$.

Requirements and limitations: Doppler imaging is a viable method for systems with fast rotation ($v \sin i \gtrsim 15\text{--}20 \text{ km s}^{-1}$) so that the rotational broadening dominates the intrinsic line width and spot signatures are resolved in velocity space ([Vogt & Penrod, 1983](#); [Rice, 2002](#)). It also requires high spectral resolution ($R \gtrsim 40,000$), high signal-to-noise ($S/N \gtrsim$

200), and dense phase coverage (i.e., spectra must be obtained over at least one full rotation period, with sufficient cadence to resolve spot features) (Strassmeier, 2009a). For differential rotation measurements, at least two Doppler images at different epochs are needed to measure differential migration (Collier Cameron et al., 2002).

These demanding requirements limit the technique to bright, rapidly rotating stars and make it labor-intensive compared to photometric methods. Spot evolution between epochs can also complicate the interpretation, since features may grow, decay, or merge between successive maps (Strassmeier, 2009b). Despite these challenges, Doppler imaging remains the gold standard for measuring stellar surface differential rotation, with results available for ~ 70 individual late-type stars as of 2009 (Strassmeier, 2009b). More recent Doppler imaging studies have expanded this sample and revealed that single stars follow a clear trend of increasing relative shear with rotation period, while tidally locked binary members show significantly weaker differential rotation (Kóvári et al., 2017; Xiang et al., 2020). Klein et al. (2025) have recently demonstrated that Doppler imaging can be extended to slowly rotating Sun-like stars using new-generation high-stability spectrographs, successfully recovering the solar differential rotation from disk-integrated HARPS-N observations.

4.5 Photometric Period Variations

The most straightforward method for detecting stellar differential rotation from photometry is to measure changes in the apparent rotation period over time (Hall, 1972; Baliunas & Vaughan, 1985; Strassmeier & Bopp, 1992; Messina & Guinan, 1999; Walker et al., 2007). If spots exist at different latitudes on a differentially rotating star, each spot modulates the lightcurve at the local rotation period $P(\Phi)$. Over time, spots emerge and decay at different latitudes, causing the dominant photometric period to shift.

The differential rotation parameter is estimated from the range of observed periods:

$$\Delta P = P_{\max} - P_{\min}, \quad (1.7)$$

where $P_{\max}$ and $P_{\min}$ are the longest and shortest rotation periods observed during the monitoring baseline. The relative differential rotation is then:

$$\alpha = \frac{\Delta P}{P_{\max}} = \frac{\Delta \Omega}{\Omega_{\text{eq}}}, \quad (1.8)$$

where $\Delta \Omega = \Omega_{\text{eq}} - \Omega_{\text{pole}}$ is the difference in angular velocity between equator and pole.

Donahue et al. (1996) searched for period variations in chromospheric Ca II H&K emission and found the relation $\Delta P(P_{\text{rot}}) = 10^{-1.3 \pm 0.06} \cdot P_{\text{rot}}^{1.3 \pm 0.01}$, suggesting that slower rotators have larger absolute (but comparable relative) differential rotation. Hall (1991) found a similar correlation from broadband photometry, reporting $\alpha(P_{\text{rot}}) = 10^{-2.3 \pm 0.06} \cdot P_{\text{rot}}^{0.85 \pm 0.06}$ for their sample.

Limitations: This method has several important caveats: (1) it requires long monitoring baselines (years) to sample spots at a range of latitudes, making it impractical for short observing campaigns; (2) the latitude of each spot is unknown, so the measured ΔP provides only a *lower bound* on the true differential rotation; if spots are confined to a narrow latitude

band, ΔP underestimates α ; (3) active region growth and decay, spot migration, and activity cycles can all introduce spurious period changes that mimic differential rotation (Strassmeier, 2009b); (4) the method is inherently one-dimensional (disk-integrated photometry), so it cannot resolve the *sign* of differential rotation (whether equatorial or polar regions rotate faster).

4.6 Individual Spot Modeling

An alternative to periodogram-based approaches is to fit an analytic spot model directly to the lightcurve, treating each spot as a circular region on the stellar surface parameterized by its latitude, longitude, radius, and contrast relative to the photosphere. The analytic framework of Kipping (2012) provides a computationally efficient forward model in which the flux deficit from each spot depends on its projected area and limb-darkening-weighted visibility as the star rotates. When combined with prescriptions for spot emergence and decay (typically modeled as trapezoidal or exponential temporal profiles), this approach can generate realistic synthetic lightcurves from a specified number of spots with known properties, enabling direct comparison with observations.

The primary appeal of individual spot modeling is that the model parameters have clear physical meaning: one can in principle recover spot latitudes, sizes, lifetimes, and—if spots exist at different latitudes—the differential rotation rate. However, as discussed in the preceding section, this approach is severely limited by parameter degeneracies inherent to disk-integrated photometry (Walkowicz et al., 2013; Basri & Shah, 2020). Additionally, the number of spots on the stellar surface is itself unknown and must be determined through model selection, adding a discrete combinatorial dimension to the inference problem.

Ikuta et al. (2020) developed a Bayesian framework for starspot mapping that addresses several of these challenges. Their code uses the analytic spot model of Kipping (2012) with circular spots whose radii evolve via trapezoidal temporal profiles, and includes differential rotation as a free parameter through the standard solar-like rotation law $\Omega(\Phi) = \Omega_{\text{eq}}(1 - \kappa \sin^2 \Phi)$. Each spot is parameterized by its latitude, initial longitude, maximum radius, emergence duration, decay duration, stable duration, spot contrast, and reference time, yielding ~ 8 parameters per spot plus global stellar parameters (inclination, equatorial period, differential rotation, and spot contrast). To explore this high-dimensional, multi-modal posterior, they employ an adaptive parallel tempering algorithm that runs multiple MCMC chains at different “temperatures,” allowing chains to exchange states and thereby escape local minima. They additionally use an importance sampling algorithm to compute the Bayesian model evidence $\mathcal{Z}$ for models with different numbers of spots, enabling principled model selection.

Testing on synthetic *Kepler*-like lightcurves with two and three spots, Ikuta et al. (2020) demonstrated that the correct number of spots can be determined via model evidence comparison, with the true model preferred by $\Delta \log \mathcal{Z} > 900$ over alternatives. However, as shown in Ikuta et al. (2020) the problem is highly degenerate with respect to spot latitude, inclination, size, and contrast and that multiple combinations of these parameters can produce nearly identical lightcurves.

These results illustrate a fundamental tension in individual spot modeling: the approach is most informative when only a few large, long-lived spots dominate the lightcurve, but real

stars (particularly solar-like stars) are expected to host many smaller, shorter-lived spots whose individual contributions overlap and cannot be uniquely decomposed. This motivates statistical approaches that model the *ensemble* properties of spots rather than individual spot parameters, which is the philosophy underlying the GP kernel developed in Chapter 4 of this thesis.

4.7 Transit Spot-Crossing Events

For transiting exoplanet systems, the planet can serve as a spatial probe of the stellar surface. When a planet transits across a starspot, it temporarily occults the dark region, producing a brief brightening (bump) in the transit lightcurve. If the same spot is crossed on successive transits, the change in the spot’s apparent longitude between transits reveals the local rotation rate at the spot’s latitude.

This technique has been applied to a growing number of transiting exoplanet systems. [Sanchis-Ojeda et al. \(2013a\)](#) characterized the polar orbit of Kepler-63b and used transit spot-crossings to constrain the stellar obliquity and spot latitude distribution. [Morris et al. \(2017\)](#) used spot occultations by HAT-P-11b to show that the K4 dwarf hosts a low-latitude spot distribution strikingly similar to the Sun at activity maximum, providing evidence for a solar-like α - Ω dynamo. [Netto & Valio \(2020\)](#) constructed the first stellar butterfly diagram for Kepler-63 from 297 transit-mapped spots, revealing a bimodal latitude distribution and a short magnetic cycle. Most recently, [Araújo & Valio \(2023\)](#) used transit spot mapping of 13 stars spanning M to F spectral types to measure the dependence of differential rotation shear on effective temperature and rotation period, and [Sagynbayeva & Farr \(2026\)](#) applied a hierarchical Bayesian surface mapping framework to recover non-equatorial active latitude bands on Kepler-17 and Kepler-63.

The transit spot-crossing method is unique in that it provides *spatially resolved* information about the stellar surface without requiring rapid rotation. For favorably aligned systems, the transit chord can constrain the spot’s latitude (given the system geometry), breaking the latitude ambiguity inherent in disk-integrated photometry. It can also be applied to slowly rotating, faint stars observed by space missions (Kepler, TESS, PLATO) that are inaccessible to spectroscopic methods.

However, this method is limited in that it requires a transiting planet with a favorable orbital geometry. Furthermore, spot lifetimes must exceed the interval between transits for the same spot to be re-detected. The transit chord samples only a narrow strip of the stellar surface; different spot latitudes are only accessible if the orbital inclination and stellar obliquity provide a range of impact parameters, or if the spot latitude distribution extends across the chord.

4.8 Asteroseismology

Stellar oscillation modes are sensitive to the internal rotation profile of the star. In the presence of rotation, non-radial oscillation modes with azimuthal order m are split in frequency

by an amount proportional to the angular velocity:

$$\delta\nu_{n,l,m} = m \int_0^R \int_0^\pi K_{n,l,m}(r, \theta) \Omega(r, \theta) r dr d\theta, \quad (1.9)$$

where $K_{n,l,m}(r, \theta)$ are the rotational kernels that weight the contribution of each region to the frequency splitting. For solar-like oscillators, the rotational splitting of p -modes is dominated by conditions in the outer convective envelope, making the splittings sensitive to *surface* differential rotation.

Gizon & Solanki (2003) showed that differential rotation produces an asymmetry in the rotational multiplets: the $m = +1$ and $m = -1$ components are no longer equally spaced from the $m = 0$ central peak if the rotation rate varies with latitude. The magnitude of this asymmetry scales with the differential rotation parameter α .

However, in order to get good constraints on differential rotation, asteroseismology requires extremely high-precision, long-duration photometry to resolve individual oscillation modes (Kepler-quality data over months to years). The rotational splitting is typically very small ($\sim 0.1\text{--}1\ \mu\text{Hz}$ for solar-like stars), requiring low-noise data and careful peak-fitting. The sensitivity to *latitudinal* differential rotation (as opposed to radial rotation gradients) is limited for low-degree modes ($l \leq 3$), which are the only modes observable in disk-integrated light.

Despite these challenges, asteroseismology offers the unique advantage of probing *internal* differential rotation, complementing the surface-only constraints from photometric and spectroscopic methods. Lund et al. (2014) demonstrated through forward modeling that asteroseismology can in principle distinguish between solar-like and anti-solar rotation profiles, with optimal sensitivity at stellar inclinations of $\sim 60^\circ\text{--}70^\circ$. Nielsen et al. (2015) compared asteroseismic and starspot rotation periods for five Sun-like stars, finding agreement within $\sim 4\%$, consistent with weak differential rotation. Benomar et al. (2018) detected latitudinal differential rotation in 13 of 40 *Kepler* Sun-like stars, all showing solar-like profiles but with shear approximately five times greater than the solar value – a result that significantly exceeds predictions from numerical simulations. Schunker et al. (2016) proposed ensemble asteroseismic fitting as a path to detecting systematic variations in radial differential rotation across stellar populations with future missions such as PLATO.

4.9 Summary and Comparison of Methods

Table 1.1 summarizes the key characteristics of each method.

Each method probes differential rotation through a different physical channel and carries distinct systematic biases. Photometric period variations and the GP approach developed in this work are the most widely applicable, requiring only time-series photometry of spotted stars, but they differ fundamentally in how they extract the differential rotation signal: period tracking measures ΔP over long baselines as spots appear at different latitudes, while the GP kernel encodes the differential rotation parameter κ directly in the shape of the autocovariance function (through the latitude-dependent broadening of the PSD harmonic peaks).

Table 1.1: Comparison of observational methods for measuring stellar differential rotation.

Method	Data Required	Applicable To
Spectral Line FT	Single spectrum	Fast rotators ($v \sin i \gtrsim 12 \text{ km s}^{-1}$)
Doppler imaging	Multi-epoch spectroscopy	Fast rotators
Transit spot-crossing	Transit lightcurves	Transiting systems
Asteroseismology	Photometric time series	Solar-like oscillators
GP kernel fitting (this work)	Photometric time series	All spotted stars

4.10 Findings from the Literature

Several decades of observational work have established broad empirical trends in how differential rotation varies with stellar properties, though significant tension remains between different measurement techniques and between observations and theory.

Dependence on effective temperature: The most robust observational result is a strong positive correlation between the absolute differential rotation shear $\Delta\Omega$ and stellar effective temperature. From Doppler imaging of 10 late-type stars spanning G2 to M2, [Barnes et al. \(2005\)](#) found a steep power-law dependence $\Delta\Omega \propto T_{\text{eff}}^{8.9 \pm 0.3}$, with M dwarfs showing surface shear an order of magnitude weaker than early G stars. This temperature dependence is broadly consistent with mean-field models of differential rotation ([Küker & Rüdiger, 2011](#)), in which the vigor of convective angular momentum transport (which scales with the depth and velocity of the convective envelope) controls the resulting latitudinal shear. [Eiff & Reiners \(2012\)](#) extended the Fourier transform method to ~ 180 A–F stars and found evidence for two populations of differential rotators: rapidly rotating A stars near the granulation boundary and moderately rotating mid- to late-F stars, with the frequency of detections decreasing for hotter and faster-rotating stars. The Doppler imaging compilation of [Kővári et al. \(2017\)](#) confirmed the temperature trend across a broader sample, finding $\Delta\Omega \propto T_{\text{eff}}^{5.8 \pm 1.0}$ from an unweighted fit to all available measurements.

Dependence on rotation period: The dependence of differential rotation on the stellar rotation rate is more subtle and method-dependent. [Barnes et al. \(2005\)](#) found that $\Delta\Omega$ is nearly independent of rotation rate ($\Delta\Omega \propto \Omega^{0.15 \pm 0.10}$), implying that the relative shear $\alpha = \Delta\Omega/\Omega_{\text{eq}}$ increases strongly with rotation period (i.e., slowly rotating stars have larger relative shear). This trend is confirmed by [Kővári et al. \(2017\)](#), who found $\alpha \approx P_{\text{rot}}[\text{d}]/200$ for single late-type stars. However, members of close binary systems show significantly weaker differential rotation at the same rotation period, with a shallower slope $\alpha \approx (0.0049 \pm 0.0001)P_{\text{rot}}[\text{d}]$, suggesting that tidal forces constrain the surface shear ([Kővári et al., 2017](#)).

From photometric period variations, [Reinhold & Gizon \(2015\)](#) found $\alpha \propto P_{\text{min}}^{0.71}$ for $\sim 12,300$ active *Kepler* stars cooler than 6700 K, broadly consistent with the Doppler imaging results. However, as [Reinhold & Gizon \(2015\)](#) acknowledge, the detected period differences can arise from both differential rotation and spot evolution, and no method currently exists to reliably separate these effects in disk-integrated photometry.

Transit spot mapping: Using starspot transit mapping of 13 slowly rotating stars ($P_{\text{rot}} \geq 4.5$ days) spanning M to F spectral types, [Araújo & Valio \(2023\)](#) found a significant anti-

correlation between $\Delta\Omega$ and the average rotation period (Pearson $r = -0.77$), but only a weak correlation between $\Delta\Omega$ and effective temperature ($r = 0.21$). The transit mapping technique probes differential rotation at the specific latitudes sampled by the transit chord, which may explain some of the quantitative differences with Doppler imaging and photometric results that sample different latitude ranges.

Asteroseismology: Asteroseismic measurements have produced some of the most surprising results. [Benomar et al. \(2018\)](#) detected latitudinal differential rotation in 13 of 40 *Kepler* Sun-like stars, all showing solar-like profiles (equator faster than poles), but with a median shear from equator to 45° latitude approximately five times greater than the solar value. This unexpectedly large shear significantly exceeds predictions from 3D convection simulations ([Brun et al., 2017a](#)) and poses a major challenge to theoretical models. In contrast, [Nielsen et al. \(2015\)](#) found no statistically significant difference between asteroseismic (envelope-averaged) and starspot (surface) rotation periods for five Sun-like stars, consistent with weak radial differential rotation.

Simulations: Three-dimensional magnetohydrodynamic simulations of rotating convection predict that the transition between solar-like and anti-solar differential rotation depends on the fluid Rossby number ([Guerrero et al., 2013a](#); [Brun et al., 2017a](#)). [Brun et al. \(2017a\)](#) found a scaling $|\Delta\Omega| \propto (M_*/M_\odot)^{1.93} R_{\text{of}}^{-0.76}$ calibrated from 15 ASH simulations spanning 0.5 – $1.1 M_\odot$, and [Augustson et al. \(2012\)](#) reported $\Delta\Omega \propto M^{5.7} \Omega^{0.8}$ for F-type stars. These simulations generally predict weaker latitudinal shear than is observed asteroseismically ([Benomar et al., 2018](#)), suggesting that important physics (potentially related to turbulent anisotropy, magnetic feedback, or near-surface dynamics ([Vasil et al., 2024](#))) is missing from current models.

Summary of tensions: The observational picture is broadly consistent across methods: differential rotation shear increases with effective temperature and decreases (in absolute terms) for faster rotators. However, significant quantitative disagreements persist between techniques, and the asteroseismic shear values substantially exceed both photometric/Doppler estimates and theoretical predictions. These tensions motivate the development of new, physically grounded methods – such as the GP kernel approach presented in this thesis – that can provide independent constraints on differential rotation from the large photometric datasets now available.

5 Thesis Work

5.1 Paper I: *alabi*

In Paper I (Chapter 2), I present *alabi* (Active Learning for Accelerated Bayesian Inference), an open-source Python package for performing Bayesian inference with computationally expensive models. *alabi* trains a Gaussian process surrogate model to emulate the likelihood function and employs active learning to iteratively improve predictive performance in high-likelihood regions, enabling accurate posterior estimation while reducing the number of expensive model evaluations by factors of thousands. The package provides a uniform interface to multiple MCMC and nested sampling backends, and I demonstrate speedups of 10 – $1000\times$

on problems with up to 64 dimensions when the likelihood evaluation time exceeds $\gtrsim 1$ s.

This work lays the methodological foundation for the subsequent chapters: the surrogate-accelerated inference framework developed here is directly applied in Paper II to constrain tidal dissipation in binary stars, where individual forward model evaluations require integrating coupled differential equations over Gyr timescales.

`alabi` has wide-ranging applications, including Birky et al. (2025), which used `alabi` to perform simulated inference of binary evolution, Birky et al. (2021a) to infer the XUV luminosity history of exoplanet host, Trappist-1. I have also developed the package `vplanet_inference` for performing statistical inference with `vplanet` and `alabi`, which has been used in Gilbert-Janizek et al. (2026) to optimize a large parameter space and identify planetary evolution models that can reproduce Earth-like habitable conditions, as well as Barnes et al. (2026) to constrain the range of cumulative XUV flux on GJ 1132 b.

5.2 Paper II: Binary Tides

In Paper II (Chapter 3), I assess the prospects and limitations of constraining equilibrium tidal dissipation in low-mass binary stars. Using global sensitivity analysis and simulated Bayesian inference (accelerated with `alabi`), I systematically characterize how sensitive observable properties (orbital period, rotation period, and eccentricity) are to the tidal quality factor Q and other model inputs, including initial conditions, stellar evolution, and magnetic braking. My analysis reveals that constraining Q from individual systems is fundamentally challenging due to strong degeneracies between Q and the unknown initial orbital and rotational states, helping to explain the orders-of-magnitude scatter ($Q \sim 10^5$ – 10^8) reported in the literature for solar-type stars.

I further examine the intriguing population of subsynchronous binaries discovered in the *Kepler* eclipsing binary catalog, in which $\sim 15\%$ of systems exhibit rotation periods $\sim 13\%$ slower than their orbital periods. This population has been hypothesized to arise from differential rotation, in which starspots at high latitudes rotate more slowly than the tidally synchronized equator, motivating the development of methods to measure differential rotation from photometric lightcurves – the subject of Paper III.

5.3 Paper III: GP models for Stellar Variability

In Paper III (Chapter 4), I derive a physically motivated Gaussian process kernel for modeling starspot variability from photometric lightcurves. Starting from an analytic starspot flux model, I obtain a closed-form kernel function whose hyperparameters directly correspond to physical quantities: equatorial rotation period, differential rotation, stellar inclination, spot lifetimes, and spot emergence/decay timescales. I validate the analytic kernel against numerical Monte Carlo simulations and show that it captures harmonic structure and non-sinusoidal periodicity absent in standard quasi-periodic kernels, while providing significant computational speedups through a sparse banded matrix solver that reduces GP inference complexity from $\mathcal{O}(N^3)$ to $\mathcal{O}(Nb^2)$.

I apply this GP model to *Kepler*-63 and demonstrate that I can recover an inclination consistent with spectroscopic $v_* \sin I$ measurements and spot properties consistent with transit

spot-crossing analysis, including differential rotation and spot lifetimes. This work represents the first demonstration that differential rotation can be inferred from one-dimensional photometric lightcurves using a physically interpretable Gaussian process, opening the door to measuring surface differential rotation across the large stellar samples observed by *Kepler* and *TESS*.

CHAPTER 2

ALABI: ACTIVE LEARNING FOR ACCELERATED BAYESIAN INFERENCE

1	Introduction	26
2	Methods	27
	2.1 Bayesian Inference	27
	2.2 Gaussian Process	28
	2.3 Training with Active Learning	31
	2.4 Hyperparameter Optimization	33
	2.5 Scaling the Training Data	35
	2.6 Markov Chain Monte Carlo	35
	2.7 Performance Metrics	35
3	Implementation	37
	3.1 Basic Example	37
	3.2 Advanced Usage	39
4	Results and Performance	40
	4.1 Bayesian Optimization	41
	4.2 Benchmark Problems	42
	4.3 MCMC Results on the True Posteriors	44
	4.4 GP performance with degeneracies and multimodalities	45
	4.5 GP performance scaling with number of dimensions	47
	4.6 Training data scaling	52
	4.7 Hyperparameter optimization choices	54
5	Discussion	57
	5.1 How many training points should I use to train the surrogate model?	57
	5.2 How do I tell if the surrogate model has converged?	57
	5.3 Which posterior sampler should I use for my problem?	58
	5.4 When is it worth training a surrogate model for my likelihood?	58
	5.5 What hyperparameter settings should I use for my problem?	59
	5.6 Parallelization	60
	5.7 Limitations & Future Work	61
6	Conclusion	64

1 Introduction

The exponential growth in data availability and model complexity across computational sciences is transforming the landscape of scientific inference. Although this data abundance offers unprecedented opportunities for discovery, it also presents computational challenges that strain traditional analytical approaches. Simulations, or forward models, that predict the outcomes of different conditions are highly useful for gaining insight from datasets. Bayesian inference provides a natural framework for incorporating prior knowledge, quantifying uncertainty, and making robust decisions under uncertainty (Gelman et al., 2013; Murphy, 2012). However, the computational demands of Bayesian methods often render them impractical for complex, high-dimensional problems that characterize modern scientific applications.

The computational bottleneck in Bayesian inference stems primarily from the need to evaluate complex likelihood functions to properly explore the parameter space. Traditional approaches, particularly Markov chain Monte Carlo (MCMC) methods (Hogg et al., 2010; Speagle, 2020a), require extensive sampling to achieve convergence, making them prohibitively expensive for models with costly likelihood evaluations. This limitation is especially pronounced in applications that involve large-scale simulations, partial differential equations, or models with intricate hierarchical structures, where each likelihood evaluation may require substantial computational resources.

Recent advances in machine learning offer promising avenues for addressing these computational challenges (e.g. Cranmer et al., 2020). This paper introduces `alabi`¹: a comprehensive framework for accelerated Bayesian inference that uses machine learning techniques to overcome the computational limitations of traditional methods. Our approach addresses the critical need for scalable inference tools in computationally expensive scientific models by providing efficient posterior approximations while maintaining statistical rigor. We present both the theoretical foundations and the practical implementation of `alabi`, demonstrating its effectiveness in various benchmarks and providing guidance to practitioners facing similar computational challenges.

Other similar packages for performing active learning with Gaussian processes include APPROXPOSTERIOR (Fleming & VanderPlas, 2018), `botorch` (Balandat et al., 2020), Bayesian Optimization (Nogueira, 2014), `scikit-optimize` (Head et al., 2020), GPry El Gammal et al. (2023), BOBE Cohen et al. (2026), and MCMC-BO Yi et al. (2024). One common challenge between these types of framework is that they require many steps to train accurately. These steps often involve choosing different algorithms and tuning many free parameters, which can be a steep learning curve for new users and require many trials and experiments to understand. Furthermore, there are a number of common failure modes in which the steps of the process can easily become numerically unstable and fail (such as poorly conditioned matrices, poor hyperparameter optimization, poor GP convergence, or poor MCMC convergence). In practice, these challenges make it difficult to train surrogate models for Bayesian problems, as it is often difficult to diagnose where an underlying problem is occurring and identify the setting that will fix the issue.

If not conditioned well, GP models also suffer from the “curse of dimensionality” problem,

¹<https://github.com/jbirky/alabi>

in which the number of training points needed to achieve convergence scales exponentially with the number of dimensions. This effect is well documented in the GP literature (e.g., see [Binois & Wycoff 2022](#) for a detailed review). In this paper, we discuss how `alabi` improves its training performance for high dimensions over a naïvely implemented GP.

The goal of this work is not to present novel algorithms or replacements for existing codes, but rather to develop a “best practice” guide on how a user might choose different configurations for their problem. `alabi` is designed to create a flexible, yet user-friendly framework for easily testing between different GP surrogate models and/or MCMC samplers, and includes comprehensive documentation on a variety of benchmark problems.

The remainder of this paper is structured as follows. Section 2 presents the theoretical framework underlying `alabi`, including the machine learning techniques employed for posterior approximation. Section 3 details the software implementation, user interface, and provides comprehensive examples in Python. Section 4 evaluates `alabi`’s performance across multiple benchmark problems, comparing accuracy and computational efficiency against traditional methods. Finally, Section 5 discusses the practical implications of our results, identifies current limitations, and provides guidance for selecting appropriate inference strategies for different problem cases.

2 Methods

The `alabi` Python package provides a user-friendly and efficient platform for performing Bayesian inference with computationally expensive models. Given a likelihood function defined by the user (usually constructed from a forward model and observational data), `alabi` trains a Gaussian process (GP) surrogate model to emulate the likelihood function. As described in this section, the GP is trained in an iterative process designed to minimize the overall computational expense. This training process consists of two parts: the Gaussian process and an active learning algorithm.

2.1 Bayesian Inference

In typical Bayesian inference problems, the goal is usually to sample the posterior distribution:

$$P(\Theta|\mathcal{D}) = \frac{P(\mathcal{D}|\Theta) \cdot P(\Theta)}{P(\mathcal{D})}, \quad (2.1)$$

where $P(\mathcal{D}|\Theta)$ is the likelihood probability of a model with parameters Θ given data $\mathcal{D}$. $P(\Theta)$ is known as the prior which incorporates previously known information, $P(\mathcal{D})$ is known as the evidence, and $P(\Theta|\mathcal{D})$ is the posterior.

The fundamental challenge in computing the posterior distribution lies in evaluating the evidence term $P(\mathcal{D})$, which requires integrating over the entire parameter space:

$$P(\mathcal{D}) = \int P(\mathcal{D}|\Theta) \cdot P(\Theta) d\Theta. \quad (2.2)$$

This integral is analytically intractable for most realistic models, particularly for high-dimensional parameter spaces. Direct numerical integration methods become computationally prohibitive because the volume of the parameter space grows exponentially with each additional dimension. MCMC methods circumvent this problem by generating samples from the posterior distribution without explicitly computing the evidence term, since the normalization constant cancels out in the acceptance ratios used by MCMC algorithms. However, MCMC sampling remains computationally intensive because it requires evaluating the likelihood function $P(\mathcal{D}|\Theta)$ at each proposed step in the chain, and achieving adequate sampling of the posterior often requires millions of likelihood evaluations to ensure proper convergence, particularly when dealing with complex, multi-modal, or highly correlated parameter spaces.

For likelihood functions that are slow to evaluate, `alabi` is designed to speed up MCMC computations by replacing the likelihood function with a *surrogate model*, which once trained, can quickly be evaluated for millions of samples for robust MCMC sampling. As a conceptual overview, Figure 2.1 shows the workflow for traditional MCMC inference, while Figure 2.2 shows the workflow for MCMC inference using `alabi`. In both cases the user input and output are the same: the user defines the likelihood function and prior function for their problem, and the framework returns the computed posterior samples. While under the hood `alabi` requires more steps, `alabi` is modularized such that it remains easy to use on any black-box likelihood function. This modularity allows the user a high degree of flexibility to define their likelihood and prior functions as is appropriate for their problem, while abstracting away the details of surrogate model training.

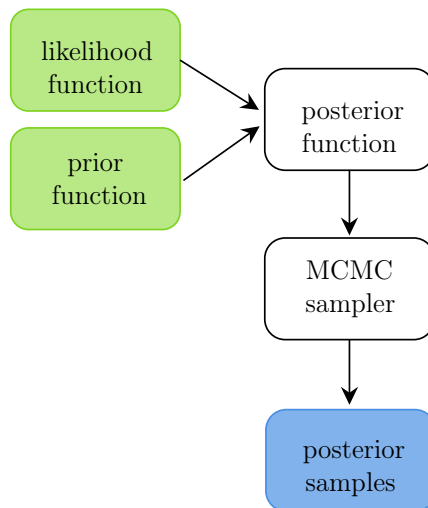


Figure 2.1: Schematic showing the steps involved in using a traditional MCMC sampler.

2.2 Gaussian Process

The key difference between `alabi` (Fig. 2.2) and standard MCMC (Fig. 2.1) is the use of GP regression to emulate the likelihood function. GPs provide a flexible and powerful framework for modeling and inference in various domains, including machine learning, statistics, and computational science. A GP defines a distribution over functions, such that any finite

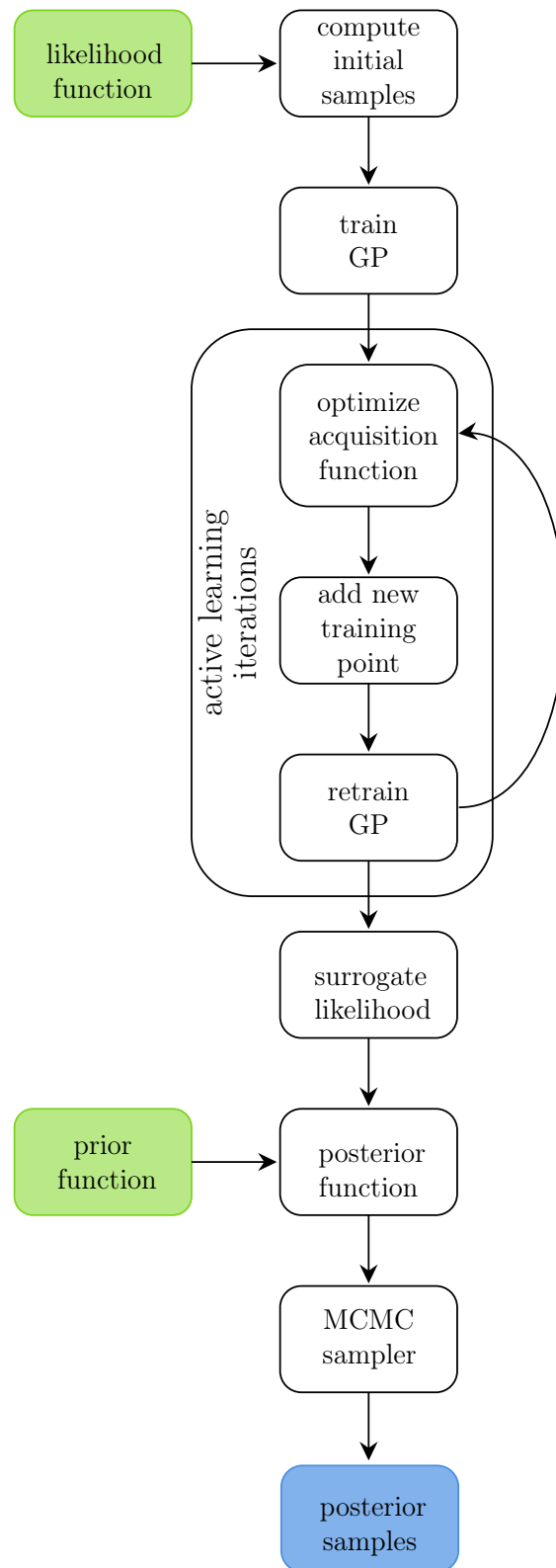


Figure 2.2: Schematic of the `alabi` framework. The user defines their likelihood and prior functions as inputs to `alabi`, and `alabi` returns samples of the posterior.

set of function values follows a multivariate Gaussian distribution. This property makes GPs particularly suitable for modeling complex and nonlinear relationships in data. In this section we give a brief overview of GPs; for an in-depth review see [Rasmussen & Williams \(2006\)](#).

A GP is defined by a mean function, denoted as $m(\mathbf{x})$, and covariance function (or kernel), denoted as $k(\mathbf{x}, \mathbf{x}')$. The mean function captures the overall trend of the underlying function, while the covariance function characterizes the dependency structure and smoothness properties. Mathematically, a GP is defined by a multivariate Gaussian:

$$\mathcal{GP}(\mathbf{x}) \sim \mathcal{N}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (2.3)$$

where $\mathbf{x} \in \mathbb{R}^m$ is an input vector with m dimensions, $\mu(\mathbf{x})$ is the mean function representing the expected value of the function at each input point $\mathbf{x}$, and $k(\mathbf{x}, \mathbf{x}')$ is the covariance function modeling the correlation between function values at different input points $\mathbf{x}$ and $\mathbf{x}'$. This kernel function, chosen by the user, effectively quantifies the smoothness and spatial properties of the GP.

An example of a commonly used kernel function that performs well on a wide range of functions is the squared exponential kernel:

$$k_{\text{SE}}(r) = A \cdot \exp\left(-\frac{1}{2}r^2\right), \quad (2.4)$$

where r is the scaled separation ($r = |\mathbf{x} - \mathbf{x}'|/\ell$) between two given input points $\mathbf{x}$ and $\mathbf{x}'$, ℓ is a kernel hyperparameter that defines the length scale of the correlation in each dimension, and A is the hyperparameter that defines the amplitude.

Other options for kernels compatible with `alabi` include the Matern-3/2 kernel:

$$k_{\text{M3/2}}(r) = A \cdot \left(1 + \sqrt{3}r\right) \exp\left(-\sqrt{3}r\right), \quad (2.5)$$

or the Matern-5/2 kernel:

$$k_{\text{M5/2}}(r) = A \cdot \left(1 + \sqrt{5}r + \frac{5r^2}{3}\right) \exp\left(-\sqrt{5}r\right). \quad (2.6)$$

Figure 2.3 compares the exponential squared, Matern-3/2, and Matern-5/2 kernels. While the shapes of each function are qualitatively similar, the Matern-3/2 kernel has a steeper peak at small separations, while the exponential squared kernel is flatter at small separations. In Section 4 we test each of these kernels on different examples and discuss their performance.

The kernel function is used to construct the covariance matrix K of the GP. Following [Ambikasaran et al. \(2015\)](#); [Rasmussen & Williams \(2006\)](#), we add a diagonal white noise term to the kernel

$$K_{ij} = k(x_i, x_j) + \sigma_n^2 \mathbb{I}, \quad (2.7)$$

where K_{ij} are the elements of covariance matrix $K \in \mathbb{R}^{n \times n}$, σ_n^2 is the variance of the Gaussian white noise distribution, and $\mathbb{I}$ is the identity matrix.

The free parameters of the covariance matrix are known as *hyperparameters* of the GP.

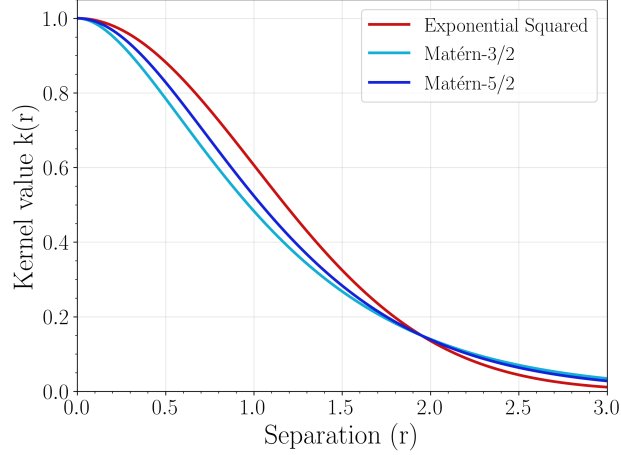


Figure 2.3: Plot comparing the shape of different kernel functions, with length scale $\ell = 1$ and amplitude $A = 1$.

These hyperparameters, which we denote as θ include parameters such as the length scales of each dimension, amplitude, and white noise variance, *e.g.*, $\theta = \{l_1, \dots, l_m, A, \sigma_n, \dots\}$. In Section 2.4 we discuss different methods for tuning these free parameters.

2.3 Training with Active Learning

Training the GP surrogate model consists of two steps: first, an initial sampling phase in which the GP is trained on evenly sampled grid points, and second, an active learning sampling phase in which new training points are iteratively selected to optimally improve GP convergence.

New points are chosen by first evaluating the predicted output of the current iteration of the GP. For any point in the input domain, x^* , the GP outputs a distribution of outcomes $y^* = \mathcal{GP}(x^*)$. Taking the expectation values over the distribution, one can compute the predictive mean and variance at point x^* as:

$$\mathbb{E}[y^*] = k(x_*, \mathbf{x})^T K^{-1} \mathbf{y} \quad (2.8)$$

$$\text{Var}[y^*] = k(x_*, x_*) - k(x_*, \mathbf{x})^T K^{-1} k(x_*, \mathbf{x})^T. \quad (2.9)$$

To avoid confusion, $k(x_*, x_*) \in \mathbb{R}$ is a scalar with the kernel function evaluated at x_* , while $k(x_*, \mathbf{x}) \in \mathbb{R}^n$ is a vector between test point x_* and training data $\mathbf{x}$, and $K \in \mathbb{R}^{n \times n}$ is the matrix computed with the training data $\mathbf{x}$ and white noise (Eqn. 2.7).

`alabi` implements several different algorithms known as *acquisition functions* to perform active learning. One of the most common acquisition functions is called the Expected Improvement (EI; Jones et al., 1998), which is

$$U_{\text{jones}}(\mathbf{x}) = (\mu_t(\mathbf{x}) - f_{\text{best}} - \zeta)\Phi(z) + \sigma_t(\mathbf{x})\Psi(z), \quad (2.10)$$

where $z = (\mu - f_{\text{best}} - \zeta)/\sigma$, μ_t and σ_t^2 are the predictive mean and variance of the GP, calculated from the kernel using Equations (2.8) and (2.9) at the test point $\mathbf{x}$, f_{best} is the op-

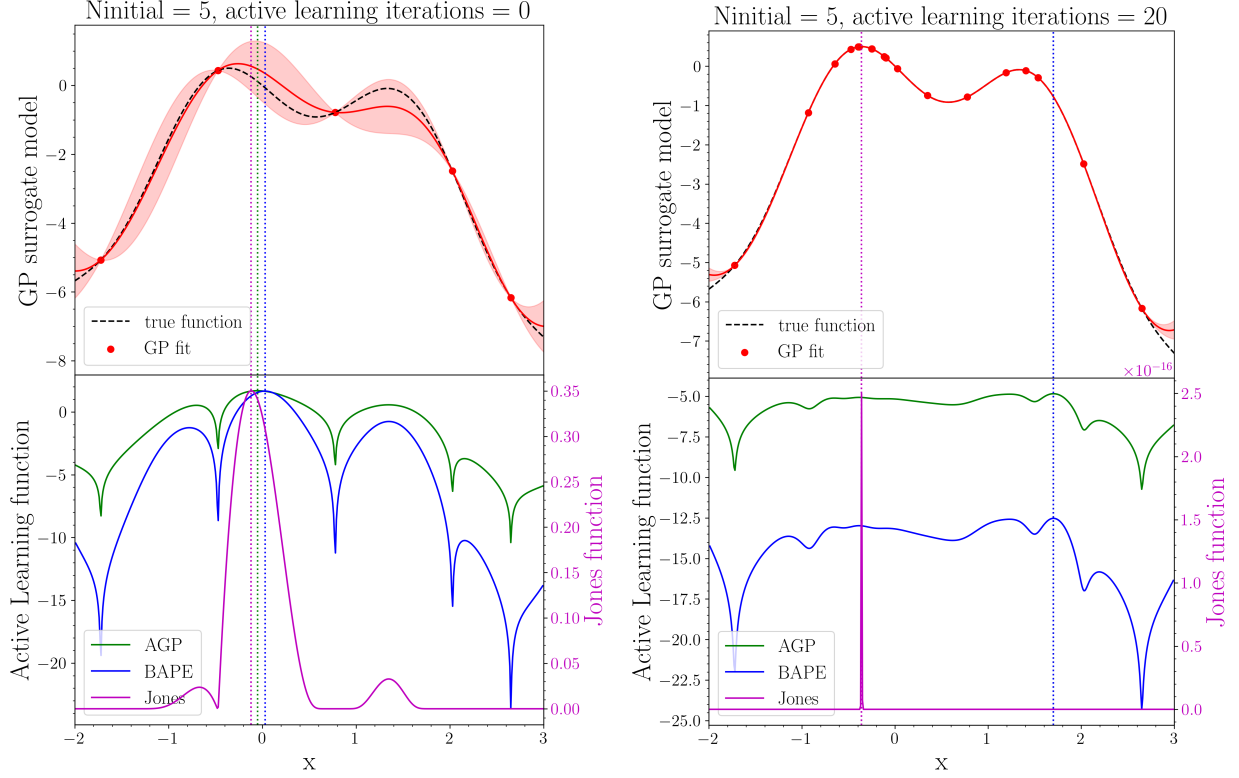


Figure 2.4: Demonstration of a GP trained on a 1D probability distribution function. The left panels show the GP conditioned on $N_{\text{initial}} = 5$ initial training points, while the right panel shows the GP conditioned on $N_{\text{initial}} = 5$ initial training points plus $N_{\text{active}} = 20$ active learning points. The top panels show the true function in black compared to the GP prediction conditioned on the training points in red, where the red line shows the GP mean prediction and the shaded red region show the $1\text{-}\sigma$ standard deviation of the prediction. The bottom panel shows the Jones, BAPE and AGP acquisition functions Equations (2.10–2.12) as a function of input parameter x , with the left y-axis (in black) showing the magnitude of the BAPE and AGP functions, and the right y-axis (in purple) showing the magnitude of the Jones function. The dashed vertical lines show the maximum value of each acquisition function, indicating where the next training point will be chosen according to that algorithm.

timal value in the current training sample, Φ is the standard normal cumulative distribution function, and Ψ is the standard normal probability distribution function. There is also a free parameter ζ , which tunes the algorithm’s priority for exploration vs. exploitation. A higher $\zeta (\gg 0)$ value focuses on exploring a wider range of parameter space, while a lower $\zeta (\approx 0)$ value focuses on converging to the local optimum. As such, a higher ζ is less susceptible to converging to a local minimum and better for finding the global maximum (important for multimodal functions), while a lower ζ is better for quicker convergence. The Jones function is ideal for problems in which the user aims to find the local/global minimum or maximum of a computationally expensive function.

For Bayesian inference problems, `alabi` implements acquisition functions known as Bayesian

Active Learning for Posterior Estimation (BAPE; [Kandasamy et al., 2017](#)), and Adaptive Gaussian Process (AGP; [Wang & Li, 2018](#)). The BAPE and AGP acquisition functions prioritize selecting new training points where the GP has high uncertainty and high probability. The acquisition functions for the BAPE and AGP are defined as:

$$U_{\text{bape}}(\mathbf{x}) = \exp[2\mu_t(\mathbf{x}) + \sigma_t^2(\mathbf{x})] \cdot (\exp[\sigma_t^2(\mathbf{x})] - 1) \quad (2.11)$$

$$U_{\text{agp}}(\mathbf{x}) = \mu_t(\mathbf{x}) + \frac{1}{2} \log |2\pi \exp[\sigma_t^2(\mathbf{x})]|, \quad (2.12)$$

where μ and σ^2 are the predictive mean and variance, Equations (2.8–2.9).

Figure 2.4 shows an example of each acquisition function applied to a GP surrogate model. Top panels show the mean and variance prediction of the GP. The bottom panels show active learning acquisition functions, where blue is the BAPE function ([Kandasamy et al., 2017](#)), green is the AGP function ([Wang & Li, 2018](#)), and purple is the Jones function ([Jones et al., 1998](#)) plotted on a separate y scale. The vertical dashed lines indicate the optimal next point to evaluate the model to improve the GP fit. While the Jones function is concentrated on the maximal function value, the BAPE and AGP functions optimize for selecting new points with high probability and high GP prediction uncertainty.

2.4 Hyperparameter Optimization

`alabi` implements two different options for GP hyperparameter optimization: (1) marginal likelihood optimization, and (2) cross-validation.

2.4.1 Marginal Likelihood Approach

In the marginal likelihood optimization approach, hyperparameters are chosen by solving for the set of $\theta = \{l_1, \dots, l_m, A, \sigma_n, \dots\}$ that maximizes the log marginal likelihood:

$$\log P(\mathbf{y}|\mathbf{x}, \theta) = -\frac{1}{2}\mathbf{y}^T K^{-1}\mathbf{y} - \frac{1}{2} \log |K| - \frac{n}{2} \log 2\pi, \quad (2.13)$$

where K is the covariance matrix with elements given by Eqn. 2.7. The term $-\frac{1}{2}\mathbf{y}^T K^{-1}\mathbf{y}$ accounts for the data fit (which decreases with length scale), the term $-\frac{1}{2} \log |K|$ serves as a complexity penalty (which increases with length scale), and the term $-\frac{n}{2} \log 2\pi$ acts as a normalization constant (see Section 5.4 of [Rasmussen & Williams, 2006](#)).

In principle, the complexity penalty should steer the solution away from over-fitting to the data. However, performance can be poor for high dimensions, as the length scale needed to properly condition a GP increases with the number of dimensions ([Hvarfner et al., 2024](#)). This difference may not be immediately intuitive, but if we consider randomly sampled points from a unit cube, the average pairwise distance between any given points increases as the number of dimensions of the space increases. For stationary kernels, the kernel value decreases to 0 as the distance between two points increases.

Thus, to maintain correlations in high dimensions, larger length scales are needed. To address this issue, we adopt the length scale regularization function from [Hvarfner et al.](#)

(2024):

$$P(\ell_i) = \log \mathcal{N} \left(\mu_\ell + \frac{\log(N_{\text{dim}})}{2}, \sigma_\ell \right) \quad (2.14)$$

where N_{dim} is the number of dimensions of the parameter space, $\mu_\ell = 2$ is the mean offset, and $\sigma_\ell = 1$ is the standard deviation for the length scale distribution. Hvarfner et al. (2024) show that the addition of this regularization function can improve the performance of simple stationary kernels in high dimensions without resorting to more complex kernel functions.

Optionally (for higher dimensional problems), the user can apply the Hvarfner et al. (2024) length scale regularization term by setting the regularization weight $w_r > 0$. Thus, the optimization problem becomes:

$$\hat{\theta}_{ml} = \arg \max_{\theta} \left[\underbrace{\log P(\mathbf{y}|\mathbf{x}, \theta)}_{\text{marginal likelihood}} + \underbrace{w_r \cdot P(\ell_i)}_{\text{regularization}} \right] \quad (2.15)$$

using the marginal likelihood (2.13) and the length scale regularization term (2.14). In practice, we use the `george` (Ambikasaran et al., 2015) Python implementation for GP training, and optimize hyperparameters numerically using `scikit-optimize` (Head et al., 2020). `alabi` automatically makes use of the gradient function, enabling the user to use gradient-based optimization methods such as the Newton-CG algorithm (Gould et al., 1999), which can speed up the optimization by a factor $\sim 100\times$ compared to non-gradient methods such as the Nelder-Mead algorithm (Nelder & Mead, 1965).

2.4.2 Cross-validation Approach

Another method for optimizing GP hyperparameters is the k-fold cross-validation approach. In the k-fold cross validation approach (Stone, 1974; Kohavi, 1995), the training sample is divided into k evenly distributed subsamples, indexed by $j \in \{1, \dots, k\}$. To compute the best set of hyperparameters θ , we select a subset $\mathbf{x}_j$ and compute the GP prediction while training on all samples not including set j (which we denote as $\mathbf{x}_{/j}$). We then compute the mean squared error between the GP prediction for $\mathbf{x}_j$, and the true values $\mathbf{y}(\mathbf{x}_j)$. This calculation is repeated for all subsets $j \in \{1, \dots, k\}$, and the total cross-validation score for a given set of hyperparameters is the sum of the mean squared error over all j . Thus, the optimal cross-validation hyperparameters are given by:

$$\hat{\theta}_{cv} = \arg \min_{\theta} \sum_{j=1}^k [\mathcal{GP}(\mathbf{x}_j; \mathbf{x}_{/j}, \theta) - \mathbf{y}(\mathbf{x}_j)]^2, \quad (2.16)$$

where $\mathcal{GP}(\mathbf{x}_j; \mathbf{x}_{/j}, \theta)$ is the GP prediction for the sample $\mathbf{x}_j$ trained on dataset $\mathbf{x}_{/j}$ with hyperparameters θ , and $\mathbf{y}(\mathbf{x}_j)$ is the true value. This method is often more robust at preventing over-fitting than the marginal likelihood approach (2.4.1), however, it is usually more computationally expensive to perform as it requires retraining the GP k times per set of hyperparameters considered.

2.5 Scaling the Training Data

In many multi-dimensional problems, the ranges of interest for each parameter may have very different length scales, which can sometimes be orders of magnitude different. To better condition the data set for the GP fitting, `alabi` implements different scaling options. By default `alabi` uses a minmax scaler to scale training inputs to the range $[0,1]$ when training the GP, and implements transform functions so that the user can easily call the GP model in real coordinates.

Another documented issue for training GPs in high dimensions is the boundary sampling issue, in which the acquisition function optimization (2.3) tends to over sample the edges of the prior boundaries (Swersky, 2017). To account for cases in which boundary clustering becomes an issue with active learning, `alabi` implements the option for applying the beta warping function from Swersky (2017):

$$w_i(x_i) = \int_0^{x_i} \frac{u^{\alpha-1}(1-u)^{\beta-1}}{\mathcal{B}(\alpha, \beta)} du, \quad (2.17)$$

where $\mathcal{B}$ is the beta distribution with scale parameters α and β .

2.6 Markov Chain Monte Carlo

`alabi` is modularized to interface with different MCMC packages for sampling posterior distributions, given any black-box likelihood function. Currently `alabi` provides integrated support for the `emcee` affine-invariant sampling package (Foreman-Mackey et al., 2013), as well as the nested sampling (Skilling, 2004) packages `dynesty` (Speagle, 2020b) and `pymultinest` (Feroz et al., 2009; Buchner et al., 2014) and `ultranest` (Buchner, 2021). For a detailed description of each algorithm, see their respective papers. `alabi` provides a unified interface for conveniently testing different samplers on their problem.

2.7 Performance Metrics

The performance of `alabi` depends on how well the GP surrogate model converges to the true model after training the model with active learning. To quantify how well the surrogate model is performing, we compute the Kullback-Leibler Divergence and test error metrics to assess the convergence between the surrogate and true model as a function of active learning iteration.

2.7.1 Kullback-Leibler Divergence

The dissimilarity between two probability distributions P and Q can be quantified by the Kullback-Leibler (KL) divergence metric:

$$D_{\text{KL}}(P\|Q) = \int_{\mathcal{D}} p(x) \log \left| \frac{p(x)}{q(x)} \right| dx \geq 0, \quad (2.18)$$

where $P(x)$ and $Q(x)$ represent two probability density functions, and the integral is taken over the prior domain $\mathcal{D}$. The KL divergence measures the average difference in the logarithm of the ratio of probabilities between the two distributions. It indicates how much information is lost when using Q to approximate P . If $D_{\text{KL}}(P\|Q) = 0$, it implies that the two distributions are identical, whereas a larger KL divergence value indicates greater dissimilarity.

When the distributions $P(x) \sim \mathcal{N}(\mu_1, K_1)$ and $Q(x) \sim \mathcal{N}(\mu_2, K_2)$ are both k -dimensional Gaussian distributions, KL divergence between two $P(x)$ and $Q(x)$ can be written as:

$$D_{\text{KL}}(P\|Q) = \frac{1}{2} \left[\log \frac{|K_2|}{|K_1|} + \mu^T K_2^{-1} \mu + \text{tr}(K_2^{-1} K_1) - k \right], \quad (2.19)$$

where K_1 is the covariance matrix of P , K_2 is the covariance matrix of Q , and $\mu = \mu_1 - \mu_2$ is the difference in the mean of P and Q . We apply this KL divergence metric in Section 4 to benchmark how well sampling with `alabi` compares to the ground truth.

For the general case, when the distributions P and/or Q are not Gaussian, there are a variety of different methods for estimating the KL divergence, which typically involve computing $D_{\text{KL}}(P\|Q)$ using density estimates of P and Q (see [Perez-Cruz 2008](#) and the citations therein). Following this approach, we numerically estimate the KL divergence by first sampling the distributions using MCMC (2.6) to get samples $x_i \sim p(x)$. We then use kernel density estimation ([Rosenblatt, 1956](#); [Parzen, 1962](#)) to get the estimated density functions $\tilde{p}(x)$ and $\tilde{q}(x)$, and estimate the integral in Equation (2.18) as the sum over all samples x_i :

$$\begin{aligned} \hat{D}_{\text{KL}}(P\|Q) &= \sum_{x_i \in \mathcal{D}} \tilde{p}(x_i) \log \left| \frac{\tilde{p}(x_i)}{\tilde{q}(x_i)} \right| \\ &= \mathbb{E}_{x_i} [\log \tilde{p}(x_i) - \log \tilde{q}(x_i)], \end{aligned} \quad (2.20)$$

which is equivalent to taking the mean over the difference in log distributions.

2.7.2 Test Sample Error

In most Bayesian inference problems that one would want to use `alabi` for, it would be impractical to compute the KL divergence between the true distribution and the surrogate model, as it would either require having an analytic function for the likelihood, or require sampling the true function. As a more practical measure of convergence, `alabi` also includes the option of computing the *test error*. These test points are computed from the true model, but are held out from the training sample so that the accuracy of the surrogate model can be quantified while not being biased by overfitting the GP. For example, given a set of N_{test} test samples, we can compute the mean squared error as:

$$\hat{\Delta}_{\text{mse}} = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left(\frac{y_{\text{gp}}(x_i) - y_{\text{true}}(x_i)}{y_{\text{true}}(x_i)} \right)^2, \quad (2.21)$$

or we can compute the average percent error as:

$$\hat{\Delta}_{\text{prc}} = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left| \frac{y_{\text{gp}}(x_i) - y_{\text{true}}(x_i)}{y_{\text{true}}(x_i)} \right| \times 100, \quad (2.22)$$

where $y_{\text{true}}(x_i)$ is the prediction from the true model and $y_{\text{gp}}(x_i)$ is the prediction from the surrogate model at each point x_i for $i = 1, \dots, N_{\text{test}}$ in the test sample.

3 Implementation

3.1 Basic Example

The `alabi` package builds off of previous work known as APPROXPOSTERIOR (Fleming & VanderPlas, 2018), with improved modularity, interface with additional sampling packages, as well as user-friendly features for saving and benchmarking results. In this section we provide a minimum use case example to familiarize users with the `alabi` workflow.

First the user defines a function to train the GP surrogate model. For the case of a Bayesian inference problem this function would represent the posterior, $\ln \text{Posterior} \propto \ln \text{Likelihood} + \ln \text{Prior}$. As a simple example, we define a 1-dimensional multimodal function to train (Figure 2.4):

```
1 import numpy as np
2
3 def lnp(x):
4     d1 = 0.6 * np.exp(-(x + 2)**2/2)
5     d2 = 0.4 * np.exp(-(x - 2)**2/2)
6     return np.log((d1 + d2)/np.sqrt(2*np.pi))
```

The next step is to train the GP surrogate model using `alabi`. In this example, we start with an initial training sample of `ntrain=10` points, evenly sampled from the prior space using the Sobol sampler (`sampler="sobol"`).

```
1 import alabi
2 from alabi import SurrogateModel
3
4 kernel = "ExpSquaredKernel"
5 savedir = f"results/{kernel}"
6
7 sm = SurrogateModel(lnlike_fn=lnp,
8                     bounds=[(-5,5)],
9                     savedir=savedir)
10
11 # Randomly sample 10 initial training points
12 sm.init_samples(ntrain=10, sampler="sobol")
```

We then initialize the GP model by fitting it to the initial training sample points. At this stage, we can specify the kernel function (e.g., the exponential squared kernel, Eqn. 2.4) as well as the additional hyperparameters (in this case an amplitude and mean parameter).

`alabi` also has the option to set a white noise parameter, where the variable `white_noise` sets the logarithm of the noise scale. The white noise option adds uncorrelated noise along the diagonal of the covariance matrix (Eqn. 2.7) that improves the numerical stability of computing the GP.

```
1 # Fit the GP to the initial training sample
2 sm.init_gp(kernel=kernel,
3           fit_amp=True,
4           fit_mean=True,
5           white_noise=-12)
```

After the GP surrogate model is trained on an initial sample of training points, the next step is to iteratively sample new training points using active learning. At this stage, the user can specify the number of iterations, `niter` and active learning algorithm, `algorithm` (with options including "bape", "agp", or "jones"). The `save_progress` option (set to `True` or `False`) allows the user to automatically cache results to a pickle file each time the hyperparameters are optimized, which is useful if your computation may be interrupted or restarted.

```
1 # Run the BAPE active learning algorithm for 20 iterations
2 sm.active_train(niter=20, algorithm="bape", gp_opt_freq=5, save_progress=True)
3 sm.plot(plots=["gp_all"])
```

As the GP trains on more data points, `alabi` has the option to improve the fit by re-optimizing the hyperparameters after a specified number of iterations set by `gp_opt_freq`. In this example we re-optimize the hyperparameters every 5 active learning iterations.

After training the GP, the MCMC sampler can be run to sample the surrogate model, e.g., using the `emcee` and/or `dynesty` package:

```
1 # Run MCMC sampler with emcee
2 sm.run_emcee(nwalkers=10, nsteps=int(5e4))
3 sm.plot(plots=["emcee_all"])
4
5 # Run MCMC sampler with dynesty
6 sm.run_dynesty()
7 sm.plot(plots=["dynesty_all"])
```

At this stage, the user can define the prior function that they want to use to sample their posterior. This modularity allows for the user to test different prior assumptions without retraining the surrogate model on the likelihood function. If the prior function is left unspecified, `alabi` will default to using a uniform prior.

Figure 2.5 shows the probability density of the true distribution compared to the sampling with the GP surrogate model with `emcee` and `dynesty`. Results are automatically saved to a pickle file that can be loaded and rerun:

```
1 sm = alabi.cache_utils.load_model_cache(savedir)
```

`alabi` also saves a summary of each training run in a human-readable text file that summarizes the configuration parameters used.

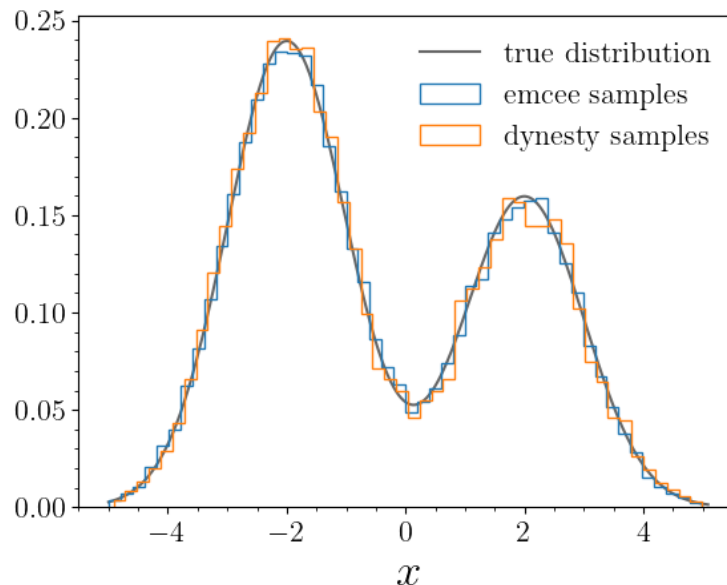


Figure 2.5: Probability density of the true distribution (black) compared to the MCMC samples drawn from the GP surrogate model after 10 initial training samples + 20 active learning training samples with `emcee` (blue) and `dynesty` (orange).

3.2 Advanced Usage

For more advanced usage, there are a number of different settings and hyperparameters of the training process that can be adjusted to improve performance according to the user’s problem, for example, the choices of: kernel function, the active learning algorithm, the number of initial or active learning training points, or the MCMC sampler. `alabi` also includes functionality for running the GP training and MCMC sampling in parallel to make use of multi-core hardware. To see examples and discussion of advanced use cases (such as parallelization, defining custom priors, or creating diagnostic plots), see the most recent documentation².

²<https://jbirky.github.io/alabi/build/index.html>

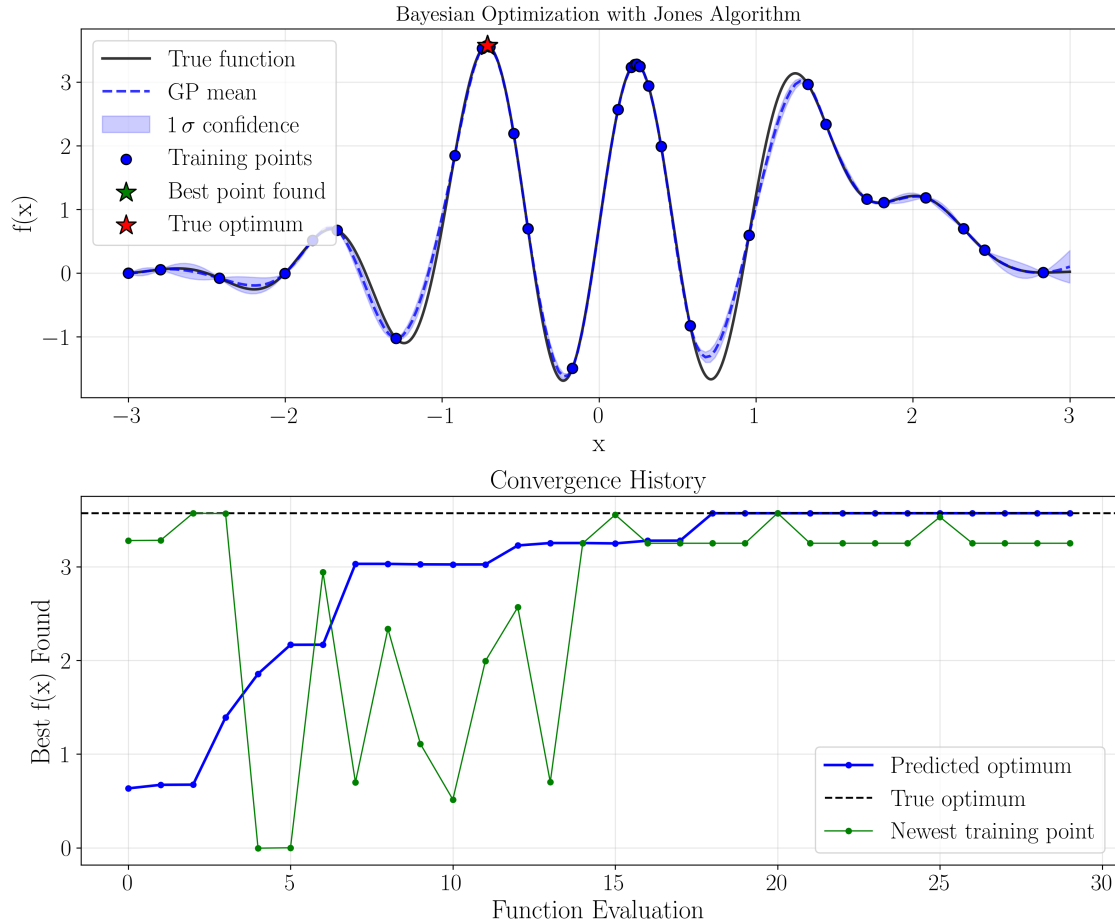


Figure 2.6: An example using `alabi` for Bayesian optimization on a 1-dimensional multi-modal function. The top panel shows the GP prediction and training points compared to the true function. The bottom panel shows the convergence history of the GP as a function of iteration.

4 Results and Performance

In this section, we provide some illustrative examples of how different `alabi` configuration settings perform on a variety of test cases, including distributions with degeneracies, multi-modalities, and high dimensionality. It is not always easy to determine which settings are most appropriate for a given problem. However, in the following section, we present several illustrative examples for a variety of test cases to provide an intuition for how to select different settings in practice. We furthermore demonstrate the computational efficiency of these methods in terms of computation time, number forward model evaluations. We also compare the performance of the `emcee`, `dynesty`, `ultranest`, and `pymultinest` MCMC sampling algorithms for each of the test cases.

In order to assess `alabi`'s performance on different test cases, we show how `alabi` performs against several benchmarks where we know the ground truth. In most real-life use cases, the

user does not know the ground truth for their posterior function. This section is designed to provide a heuristic intuition for how `alabi` performs in different scenarios so that the user can choose the training settings (e.g., number of training iterations, kernel function, MCMC sampler) that are appropriate for their problem.

4.1 Bayesian Optimization

Bayesian optimization is a common application for this framework. Figure 2.6 shows how `alabi` can be used to maximize a multimodal peaks function. As seen in the top plot, points sampled by the Jones active learning algorithm are clustered around maximum function values to improve the precision of the optimum value. In this example, the GP optimum converges to the true optimum within 20 iterations.

4.2 Benchmark Problems

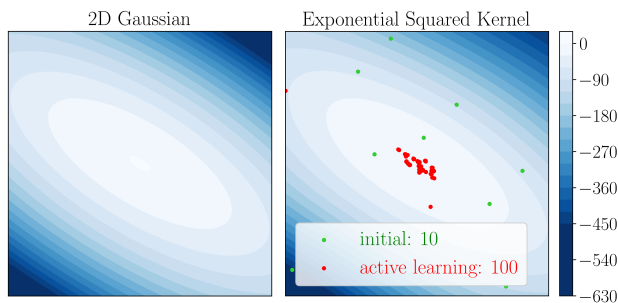


Figure 2.7: Two-dimensional Gaussian function (Example 4.2.1) with the true function (left panel) compared to the GP surrogate model (right panel). The green points highlight the quasi-uniform sampled initial training points, and the red points highlight training points sampled iteratively using active learning.

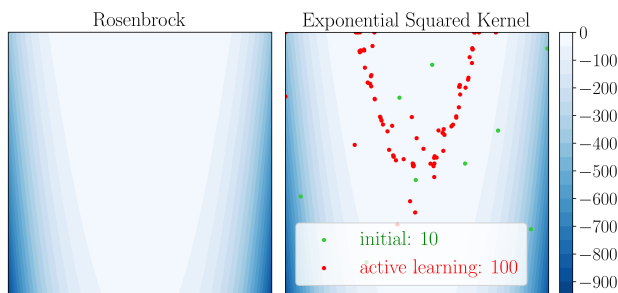


Figure 2.8: Rosenbrock function (Example 4.2.2).

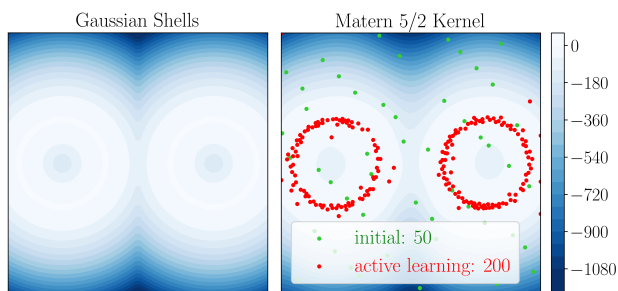


Figure 2.9: Gaussian shells function (Example 4.2.3).

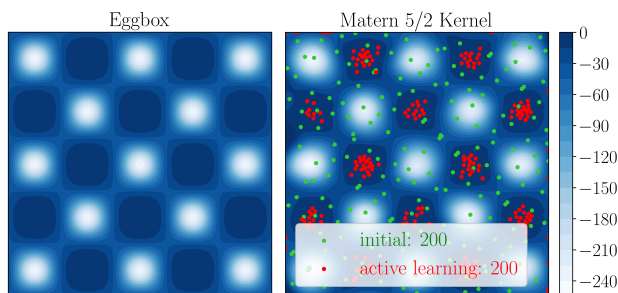


Figure 2.10: Eggbox function (Example 4.2.4).

4.2.1 Example 1: 2-dimensional Gaussian

As a first example, we test `alabi` on a simple unimodal Gaussian distribution, with mean and covariance parameters

$$f_1(\mathbf{x}) = \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.2 & -0.1 \\ -0.1 & 0.1 \end{bmatrix} \right). \quad (2.23)$$

Figure 2.7 shows the true function compared to the GP surrogate model trained with 10 initial training samples, and 100 active learning samples. For a simple unimodal distribution, not many training points are needed to replicate the true function, which we show quantitatively in Section 4.4.

4.2.2 Example 2: 2-dimensional Rosenbrock function

As a second example, we test the GP performance on a 2D function with a curved degeneracy feature given in Wang & Li (2018):

$$f_2(\mathbf{x}) = \exp \left[-\frac{1}{100}(x_1 - 1)^2 - (x_1^2 - x_2)^2 \right], \quad (2.24)$$

with the input defined over the prior domain $x_i \in [-5, 5]$. Figure 2.8 shows the true function compared to the GP surrogate model trained with 10 initial training samples, and 100 active learning samples.

4.2.3 Example 3: 2-dimensional Gaussian shells

As a third example, we test how the GP performs with highly correlated features using the Gaussian shells example from Speagle (2020b):

$$f_3(\mathbf{x}) = \mathcal{C}(\mathbf{x}|c_1, r_1, w_1) + \mathcal{C}(\mathbf{x}|c_2, r_2, w_2), \quad (2.25)$$

with the input defined over the domain $x_i \in [-6, 6]$ and each circle is given by the distribution:

$$\mathcal{C}(\mathbf{x}|c, r, w) = \frac{1}{\sqrt{2\pi}w^2} \exp \left[-\frac{1}{2} \frac{(|\mathbf{x} - c| - r)^2}{w^2} \right]. \quad (2.26)$$

Figure 2.9 shows the true function compared to the GP surrogate model trained with 50 initial training samples, and 200 active learning samples. Compared to the unimodal Gaussian and Rosenbrock functions, the shape of the Gaussian shells function is more complex and requires a greater number of training points to converge to the true model.

4.2.4 Example 4: 2-dimensional multimodal

As a fourth example, we test the GP performance on a highly multimodal 2D distribution known as the eggbox function (Feroz et al., 2019):

$$f_4(\mathbf{x}) = \exp \left\{ \left[2 + \cos \left(\frac{x_1}{2} \right) \cos \left(\frac{x_2}{2} \right) \right]^5 \right\} \quad (2.27)$$

with the input defined over the domain $x_i \in [0, 1]$. Figure 2.10 shows the true function compared to the GP surrogate model trained with 100 initial training samples, and 250 active learning samples. Out of all of the 2D examples we show, the eggbox function is the most difficult to model and takes the most number of training points to converge.

4.3 MCMC Results on the True Posteriors

alabi provides a flexible, and easy to use interface to switch between different popular MCMC samplers. We first see how each MCMC sampler performs when sampling the true functions. For this test, we use the configuration settings from Table 2.1 as the key word options for each sampler, in the following example:

```

1 from alabi import SurrogateModel
2 from alabi import utility as ut
3 from funtools import partial
4 from scipy.stats import multivariate_normal
5
6 gaussian_2d_fn = partial(
7     multivariate_normal.logpdf,
8     mean=np.zeros(2), cov=np.eye(2)*0.1)
9 bounds = [(-3, 3) for _ in range(2)]
10
11 # Set up uniform prior function (emcee format)
12 prior_fn = partial(ut.lnprior_uniform, bounds=bounds)
13
14 # Set up uniform prior transform (dynesty/ultranest format)
15 prior_transform = partial(
16     ut.prior_transform_uniform, bounds=bounds)
17
18 # Use alabi as a wrapper for different MCMC samplers
19 sm = SurrogateModel(gaussian_2d_fn, bounds=bounds)
20
21 sm.run_emcee(like_fn=sm.true_log_likelihood,
22             prior_fn=prior_fn, **emcee_kwargs)
23
24 sm.run_dynesty(like_fn=sm.true_log_likelihood,
25               prior_transform=prior_transform,
26               sampler_kwargs=dynesty_sampler_kwargs,
27               run_kwargs=dynesty_run_kwargs)
28
29 sm.run_ultranest(like_fn=sm.true_log_likelihood,
30                 prior_transform=prior_transform,
31                 sampler_kwargs=ultranest_sampler_kwargs,
32                 run_kwargs=ultranest_run_kwargs)
33
34 sm.run_pymultinest(
35     like_fn=sm.true_log_likelihood,
36     prior_transform=prior_transform,
37     sampler_kwargs=pymultinest_kwargs)

```

Table 2.2 summarizes how each MCMC sampler performs in sampling different functions in terms of computational runtime, evidence error, number of effective posterior samples,

and sampling efficiency. Each of the nested samplers was run with 1000 live points and run until an evidence error of $\text{dlogz} < 0.5$ was achieved.

Each of the samplers perform well at sampling unimodal distributions (2D Gaussian and rosenbrock), however only the nested samplers (`dynesty`, `ultranest`, and `pymultinest`) converged for the multimodal examples (Gaussian shells and eggbox), while `emcee` failed to converge, even running for longer chains of 1×10^5 steps. In terms of total computation time, `pymultinest` generally performed the quickest of all samplers, but performed similarly to the other samplers in terms of the efficiency of computing the number of effective samples per second.

Table 2.1: MCMC sampler configuration parameters

Description	Parameter	Value
emcee		
Number of walkers	<code>nwalkers</code>	10
Steps per walker	<code>nsteps</code>	1×10^4
Burn-in steps	<code>burn</code>	1×10^3
dynesty		
Number of live points	<code>nlive</code>	1,000
Bounding method	<code>bound</code>	multi
Sampling method	<code>sample</code>	auto
Posterior weight fraction	<code>pfrac</code>	1.0
Maximum iterations	<code>maxiter</code>	2×10^4
Initial evidence tolerance	<code>dlogz_init</code>	0.5
UltraNest		
Number of live points	<code>min_num_live_points</code>	1,000
Target effective samples	<code>min_ess</code>	2,000
Evidence tolerance	<code>dlogz</code>	0.5
KL divergence tolerance	<code>dKL</code>	0.5
PyMultiNest		
Number of live points	<code>n_live_points</code>	1,000
Evidence tolerance	<code>evidence_tolerance</code>	0.5
Sampling efficiency	<code>sampling_efficiency</code>	0.8
Multimodal detection	<code>multimodal</code>	True
Maximum modes	<code>max_modes</code>	20

4.4 GP performance with degeneracies and multimodalities

Next we use the true posteriors (computed in the previous section 4.3) as a baseline for comparing the performance of GP surrogate models. We explore how well the GP converges for the different 2-dimensional examples (Examples 4.2.1–4.2.4), by computing the KL divergence (2.18) between the true distribution and GP surrogate distribution. Figure 2.11 shows how well the surrogate performs on the benchmark examples in Section 4.2. Each GP surrogate model was trained using 10 initial training samples, and run for 250 active learning

Table 2.2: MCMC Sampler Performance Summary

Sampler	gaussian_2d	rosenbrock	gaussian_shells	eggbox
Runtime (seconds)				
emcee	13.7	6.0	5.6	4.2
dynesty	25.3	12.6	61.0	13.5
ultranest	15.3	13.8	18.5	15.1
pymultinest	3.9	4.2	3.7	3.3
Log Evidence ($\log Z \pm \text{error}$)				
emcee	—	—	—	—
dynesty	-3.65 ± 0.04	-2.55 ± 0.03	-1.65 ± 0.03	-4.19 ± 0.03
ultranest	-3.62 ± 0.12	-2.50 ± 0.07	-1.75 ± 0.08	-4.18 ± 0.08
pymultinest	-3.58 ± 0.01	-2.53 ± 0.01	-1.71 ± 0.01	-3.81 ± 0.03
Effective Samples				
emcee	5,290	4,090	550	590
dynesty	15,860	14,820	15,270	14,631
ultranest	9,686	8,100	7,465	8,777
pymultinest	3,474	2,790	2,820	3,153
Sampling Efficiency (samples/sec)				
emcee	386	684	98.9	142
dynesty	627	1.2k	251	1.1k
ultranest	634	588	403	583
pymultinest	891	667	755	960

iterations. The colors denote three different kernels: Exponential Squared (blue; Eq. 2.4), Matern 3/2 (red; Eq. 2.5), and Matern 5/2 (green; Eq. 2.6). Each line shows the median KL divergence value over 30 random trials with the same training configuration, and the upper and lower shaded regions show the 25th and 75th percentiles.

For the 2D Gaussian example, all kernels perform similarly well. The GP converges quickly with just the 10 initial samples for the exponential squared and Matern-5/2 kernels, and within ~ 10 active learning iterations for the Matern-3/2 kernel. For the Rosenbrock example, the exponential squared kernel performs the best (converging within ~ 20 active learning samples), while the Matern-3/2 kernel performs the worst (taking ~ 120 active learning iterations to converge).

For the last two examples, we find that the kernel performance is the opposite: the Matern-3/2 kernel performs the best, while the exponential squared kernel plateaus and does not achieve the same performance even when adding more active learning samples. This difference is due to the profile of the kernel function. Referring back to Figure 2.3, the Matern-3/2 kernel has a steeper peak at small separations, whereas the exponential squared kernel is flatter at small separations, and the Matern-5/2 kernel falls somewhere between the two. This steeper peak of the Matern-3/2 kernel performs better at reproducing functions with tightly correlated modes or features (which we see in the performance of the Gaussian shells and eggbox example), while the flatter exponential squared kernel performs better on functions with broader features (which we see in the 2D gaussian and Rosenbrock examples).

4.5 GP performance scaling with number of dimensions

We next test how well a GP surrogate model can replicate functions of higher dimensions by training the surrogate on an N-dimensional Gaussian distribution:

$$f_5(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{N_{\text{dim}}} \det \mathbf{\Sigma}}} \exp \left[-\frac{1}{2}(\mathbf{x} - \mu)^T \mathbf{\Sigma}^{-1}(\mathbf{x} - \mu) \right], \quad (2.28)$$

where N_{dim} is the number of dimensions, μ is the mean vector, and $\mathbf{\Sigma}$ is the covariance matrix. For the true parameters of the N-dimensional Gaussian, we choose a mean vector μ centered at zero in each dimension, and generate a random covariance matrix with elements Σ_{ij} such that $\mathbf{\Sigma}$ is positive definite.

Next, we estimate the number of training samples required for the surrogate model to converge when trained on Gaussian distributions with 16, 32, 48, and 64 dimensions. Figure 2.12 shows the test error measured as the average percent error. Here N_{init} refers to the number of initial training samples, N_{active} refers to the number of active learning samples, and N_{dim} refers to the number of dimensions of the problem. For each dimension shown with different colored lines, we train the GP using $10 \times N_{\text{dim}}$ initial samples and run for an additional 1000 active learning iterations. We train the GP using an exponential squared kernel, perform active learning with the BAPE algorithm, and optimize the hyperparameters every 50 iterations using a 16-fold cross validation.

Figure 2.13 shows the true log-likelihood vs. the predicted log-likelihood of the surrogate model. The left column shows the initial GP fit compared to the final GP fit after the 1000 active learning iterations. For ≤ 32 dimensions, the GP performs converges to less

than 1% error, given only the initial training samples ($N_{\text{init}} = 10 \times N_{\text{dim}}$). For 48 and 64 dimensions, the fit is initially poor, but converges to $< 1\%$ error within 400 and 1000 iterations respectively. In all cases shown in Figure 2.13, the GP fits well to all of the training sample points, but the panels with poor fits to the test sample (e.g. the initial fit for 48D and 64D). These cases with low training error but poor test error mean that the GP is not generalizing well to points that it wasn't trained on, showing that the GP can be prone to overfitting without enough training data.

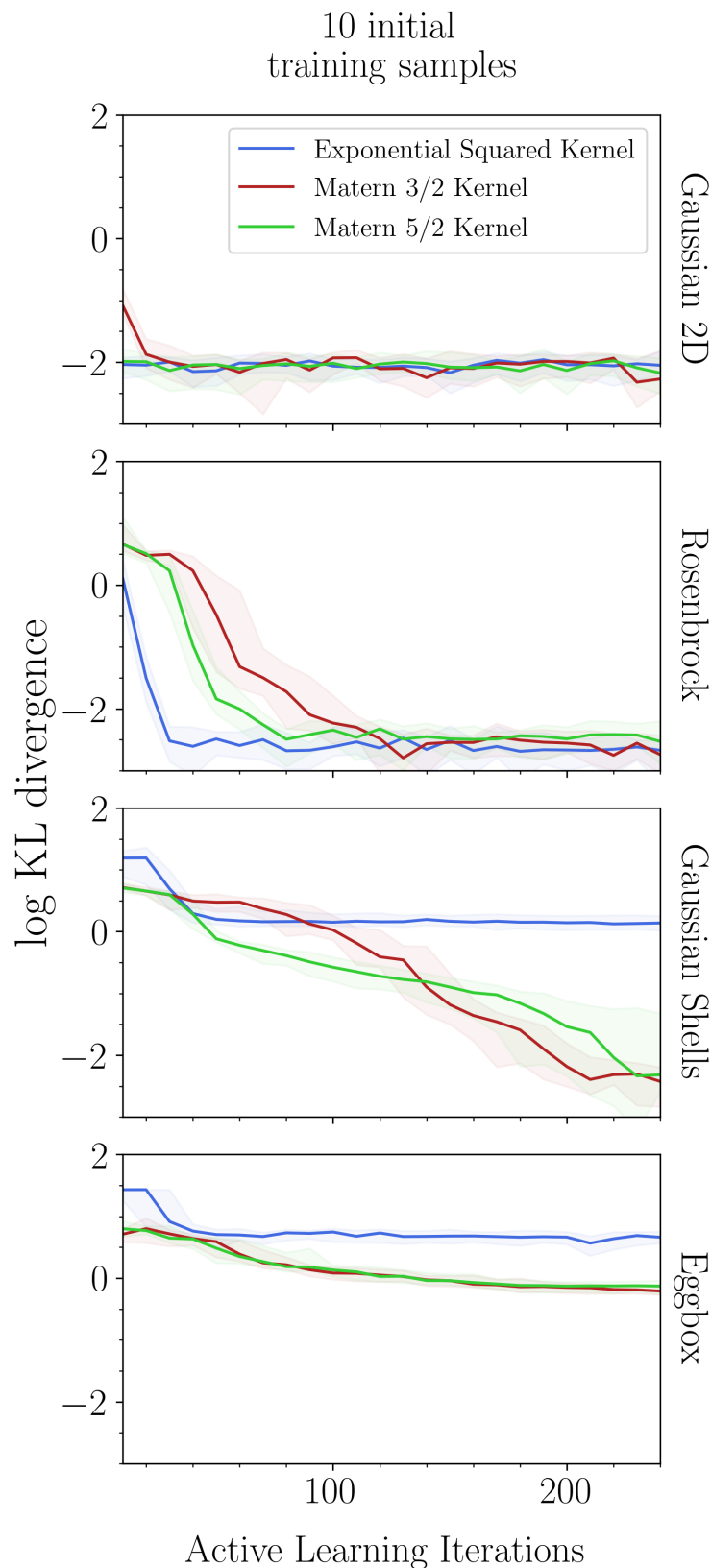


Figure 2.11: Performance of different GP kernel functions (exponential squared, Matern-3/2, and Matern-5/2) perform for each of the four different 2-dimensional examples (the Gaussian 2D, rosenbrock, gaussian shells, and eggbox functions), using `dynesty` to sample each function. The y-axis shows the KL divergence metric computed between the GP surrogate model and the true model, where values closer to zero mean better convergence.

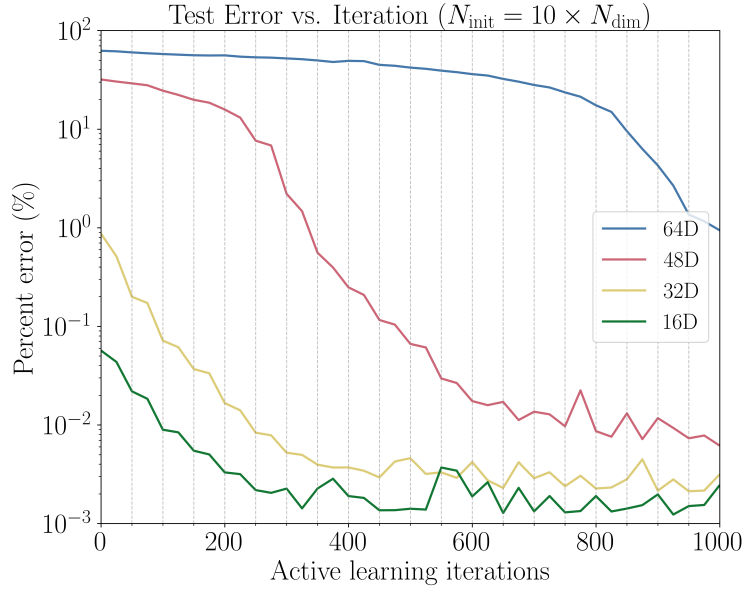


Figure 2.12: Surrogate model convergence as a function of active learning iterations (x-axis) and number of dimensions (colored lines). For each dimension, we train a surrogate model with $10 \times N_{\text{dim}}$ initial training samples and compute how well the model converges in terms of the percent error (y-axis) computed using test samples held out from the training sample. The vertical dashed lines show the iterations in which we re-optimized the hyperparameters (every 50 iterations).

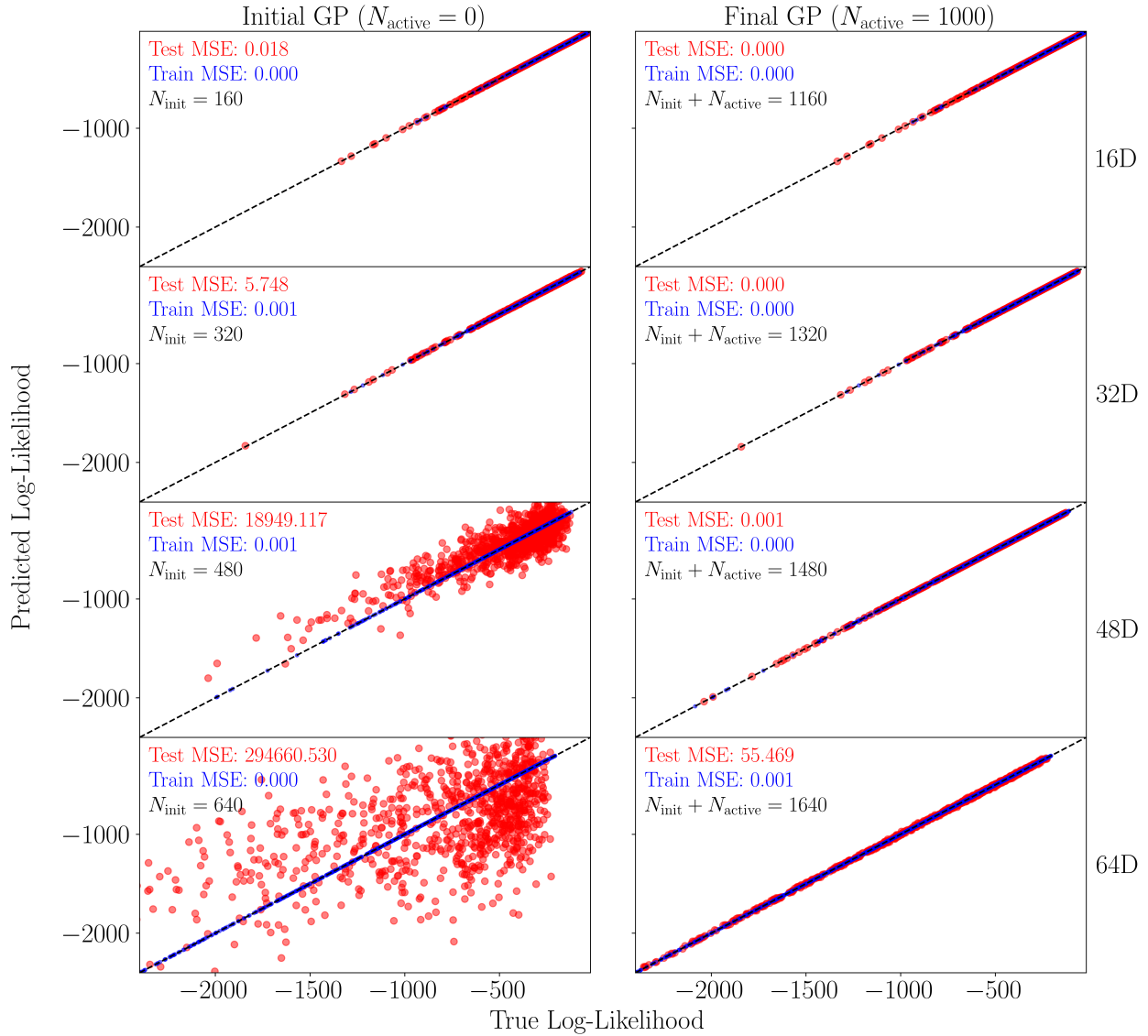


Figure 2.13: True log-likelihood (x-axis) vs. the predicted log-likelihood (y-axis), demonstrating the accuracy of the surrogate model for the initial GP fit (left column) and the final GP fit (right column). The initial GP is trained on $N_{\text{init}} = 10 \times N_{\text{dim}}$ samples, and the final GP is trained on $N_{\text{init}} + N_{\text{active}} = 10 \times N_{\text{dim}} + 1000$ samples. The rows show the initial and final GP fits for 8, 16, 32, 48, and 64 dimensional Gaussians, taken from the same runs as Figure 2.12. The blue points show the predictions and mean squared error training sample, and the red points show that of the test sample.

4.6 Training data scaling

As discussed in Section 2.5, scaling the input training data is an important preprocessing step that affects how the GP hyperparameters—particularly the kernel length scales—are optimized. When inputs are scaled to a common range (e.g., $[0, 1]$ via a min-max scaler), the length scales become easier to interpret and constrain: values of order ~ 0.1 – 1 correspond to correlations spanning a meaningful fraction of the parameter domain. Without scaling, the length scale bounds must be matched to the raw coordinate ranges of each parameter, which vary between problems.

To demonstrate the practical effect of length scale bounds on surrogate model quality, we train a GP on a 4-dimensional correlated Gaussian likelihood:

$$f_6(\mathbf{x}) = \mathcal{N} \left(\mathbf{0}, \begin{bmatrix} 1.00 & 0.60 & 0.30 & -0.20 \\ 0.60 & 1.00 & 0.10 & 0.40 \\ 0.30 & 0.10 & 1.00 & 0.50 \\ -0.20 & 0.40 & 0.50 & 1.00 \end{bmatrix} \right), \quad (2.29)$$

with each parameter defined over the domain $\theta_i \in [-4, 4]$. Both configurations use $N_{\text{init}} = 40$ initial Latin hypercube samples, $N_{\text{active}} = 60$ BAPE active learning iterations, the exponential squared kernel, min-max input scaling to $[0, 1]$, L-BFGS-B hyperparameter optimization, and cross-validation re-optimized every 20 iterations. The configurations differ only in the allowed range of the kernel length scale hyperparameters, as summarized in Table 2.3.

Table 2.3: Hyperparameter configurations for the 4D Gaussian length scale comparison. Both configurations share the same kernel, optimizer, input scaling, and training procedure. The only difference is the allowed range of the GP length scale hyperparameters ℓ .

Setting	Config A	Config B
Kernel	ExpSquared	ExpSquared
θ scaling	MinMax $[0, 1]$	MinMax $[0, 1]$
HP opt. method	L-BFGS-B	L-BFGS-B
$\log_{10}(\ell)$ range	$[-1, 1]$	$[-4, -2]$
$\log_{10}(A)$ range	$[3, 5]$	$[3, 5]$
White noise ($\log \sigma_n$)	-12 (fixed)	-12 (fixed)
Fit amplitude	True	True
Fit mean	True	True

`alabi` includes a `plot_gp_predictions_1D` diagnostic that visualizes the GP surrogate by sweeping each input parameter individually while holding the remaining parameters fixed at a reference point (typically the surrogate maximum found via `find_max_surrogate`). For each dimension, the plot shows the GP predictive mean, the $\pm 2\sigma$ uncertainty band, individual function draws sampled from the GP posterior, and the training data. This diagnostic is particularly useful in problems with more than two dimensions, where 2D contour plots can only show pairwise projections.

Figure 2.14 shows the 1D diagnostic for Configuration A, where the length scale bounds are set to $\log_{10}(\ell) \in [-1, 1]$. Because the inputs have been scaled to $[0, 1]$, length scales of

order ~ 0.1 – 1 are appropriate for capturing the broad, smooth structure of the Gaussian likelihood. The optimized length scales are $\log_{10}(\ell) \approx 0.2$ – 0.5 across the four dimensions. The GP mean closely tracks the true log-likelihood (dashed red line), the uncertainty bands are narrow near training data, and the individual GP draws are smooth and well-behaved.

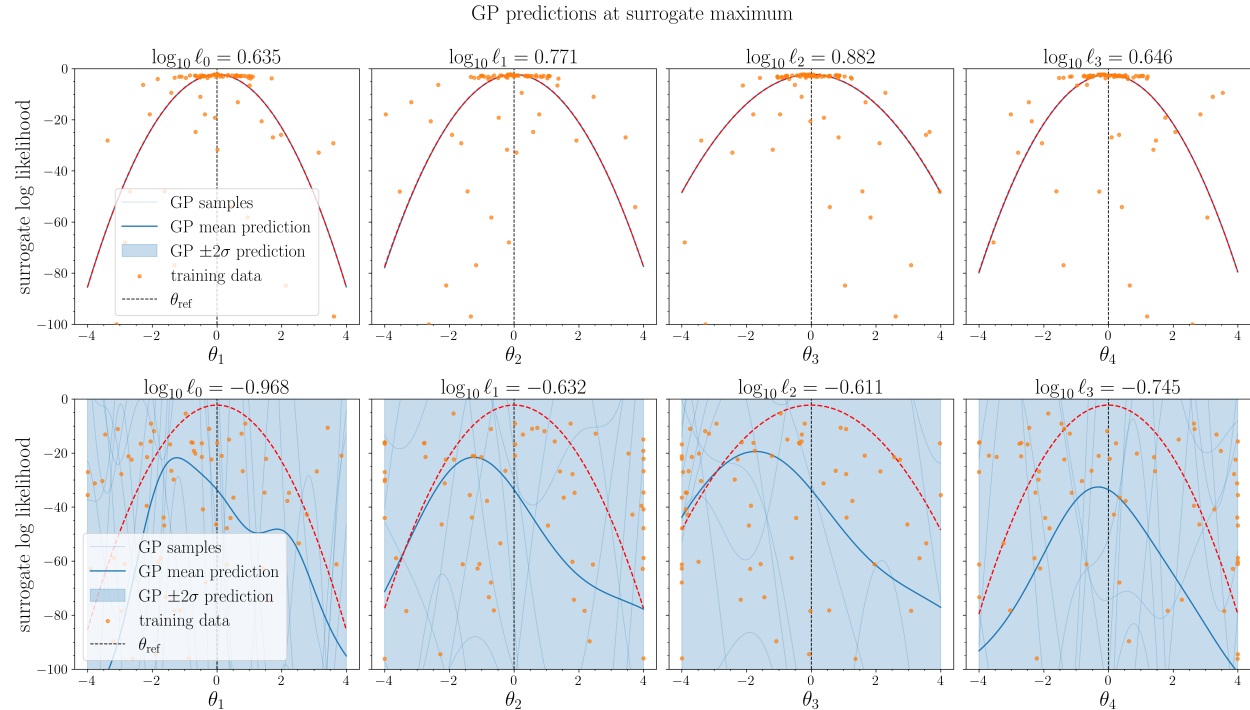


Figure 2.14: 1D GP prediction diagnostic for Configuration A (top panel), with well-chosen length scale bounds ($\log_{10}(\ell) \in [-1, 1]$) v.s. Configuration B (bottom panel), with poorly chosen length scale bounds ($\log_{10}(\ell) \in [-4, -2]$). Each panel sweeps one parameter through its full range while holding the other three fixed at the surrogate maximum. The blue line shows the GP predictive mean, the shaded blue region shows the $\pm 2\sigma$ uncertainty, the light blue lines show individual function draws from the GP posterior, the orange points show the training data, and the dashed red line shows the true log-likelihood. For Configuration A, the optimized length scales ($\log_{10} \ell \approx 0.2$ – 0.5) produce smooth, well-constrained predictions. In Configuration B, the GP posterior draws are highly oscillatory, the uncertainty bands are much wider, and the mean prediction deviates from the true log-likelihood (dashed red line). This overfitting behavior results from length scales that are far too short for the smoothness of the underlying function.

Figure 2.14 shows the same diagnostic for Configuration B, where the length scale bounds are restricted to $\log_{10}(\ell) \in [-4, -2]$, corresponding to length scales of 10^{-4} – 10^{-2} . These very short length scales force the GP to treat each training point as nearly independent, producing a highly oscillatory interpolation. The individual GP posterior draws vary wildly between training points, the uncertainty bands are substantially wider, and the GP mean deviates significantly from the true model. This is a clear signature of overfitting: the GP memorizes the training data but fails to learn the underlying smooth structure of the likelihood surface.

In practice, a surrogate model exhibiting this behavior would produce unreliable posterior estimates when sampled with MCMC.

This comparison illustrates several practical guidelines for diagnosing and avoiding overfitting due to length scale misconfiguration. Firstly, it is usually good practice to scale inputs before setting length scale bounds. When inputs are scaled to $[0, 1]$, length scales of order ~ 0.1 – 1 (i.e., $\log_{10}(\ell) \in [-1, 1]$) are a reasonable starting range for most smooth functions. Secondly, a good visual diagnostic for checking length scale overfitting is to plot the 1D fits around a given point (e.g. the maximum likelihood estimate; Figure 2.14). If the GP posterior draws appear highly oscillatory (“wiggly”) or the uncertainty bands are unexpectedly wide, the length scales are likely too short. Finally, it is good to check the optimized length scale values and check if any of the values are converging towards the lower end of the bounds. In this case, the user may need to adjust the length scale regularization penalty (Equation 2.14) to a higher mean and sigma value.

4.7 Hyperparameter optimization choices

The quality of the GP surrogate model depends critically on the choice of hyperparameters used during training. Key choices include the kernel function, data scaling strategy, and optimization settings. Rather than manually tuning these, `alabi` supports systematic testing of multiple configurations to identify the best settings for a given problem. In this section, we demonstrate this approach using the eggbox benchmark function (Example 4.2.4).

We define a base GP configuration and then vary three settings: the kernel function (exponential squared, Matern-3/2, and Matern-5/2), the input scaler (no scaling, min-max, and standard scaler), and the output scaler (no scaling, min-max, and standard scaler). This produces $3 \times 3 \times 3 = 27$ combinations. Table 2.4 summarizes the base configuration shared by all combinations, and Table 2.5 shows the variable settings tested.

Table 2.4: Base GP configuration shared across all hyperparameter (HP) combinations. These settings are held fixed while the kernel, input scaler, and output scaler are varied.

Setting	Value
Fit amplitude	True
Fit mean	True
Fit white noise	False
White noise ($\log \sigma_n$)	-12
HP opt. method	L-BFGS-B
HP selection	8-fold cross-validation
$\log_{10}(A)$ range	$[-1, 1]$
$\log_{10}(\ell)$ range	$[-2, 2]$
N_{init}	50
Initial sampler	Latin hypercube

For each of the 27 configurations, we fit the GP to the initial training sample of $N_{\text{init}} = 50$ points and evaluate performance using the mean squared error (MSE) on a held-out test set of 1000 points. Table 2.6 shows the five best-performing configurations ranked by test MSE.

Table 2.5: Variable hyperparameter settings tested in the grid search. Each combination of kernel, input scaler (θ), and output scaler (y) is evaluated by fitting a GP to the initial training sample and computing the test set MSE on 1000 held-out points.

Setting	Options tested
Kernel	ExpSquared, Matern-3/2, Matern-5/2
θ scaler	None, MinMax, Standard
y scaler	None, MinMax, Standard

Using the best configuration (Config 7), we then run 200 active learning iterations with the BAPE algorithm, re-optimizing the GP hyperparameters every 20 iterations. Figure 2.15 shows the test MSE as a function of active learning iteration. To assess the effect of random variance in the training process, we repeat the same configuration 10 times with different random seeds. Figure 2.15 shows the test MSE trajectories for all 10 trials. The test MSE generally decreases steadily from ~ 2000 to ~ 50 over 200 iterations with some random variations between the 10 trials generated by different random seeds. While there is some spread in the convergence paths—particularly during the first ~ 100 iterations where different active learning samples explore different regions of parameter space—all trials converge to similarly low test MSE values by iteration 200.

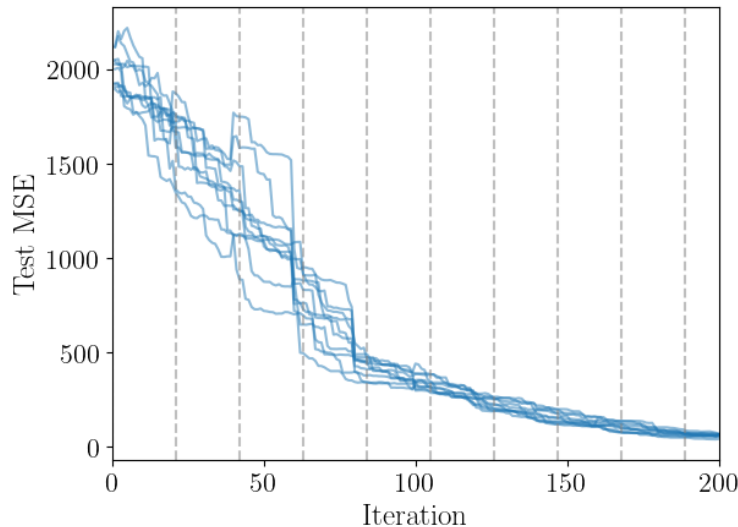


Figure 2.15: Test set MSE over 10 independent trials using the same hyperparameter configuration (Config 7 in Table 2.6). Each blue line represents a different random seed. The spread illustrates the stochastic variance in the active learning process, while the overall convergence trend is consistent.

Finally, we compare the active learning performance of the top 5 initial configurations from Table 2.6. Figure 2.16 shows the test MSE trajectories for each. The configuration with the lowest initial test MSE (Config 7, exponential squared kernel with standard input scaling and min-max output scaling) also achieves the lowest final MSE after 200 active learning iterations. However, the ranking among configurations can change during active

learning: Config 11 (Matern-3/2 with no input scaling) starts with a higher initial MSE but converges faster than Configs 24 and 6. Notably, Config 24 (Matern-5/2 with standard input scaling) shows minimal improvement during active learning, indicating that the initial hyperparameter fit quality does not always predict active learning performance. In practice, if the test error plateaus or fails to converge, this may indicate that a different hyperparameter configuration should be tried.

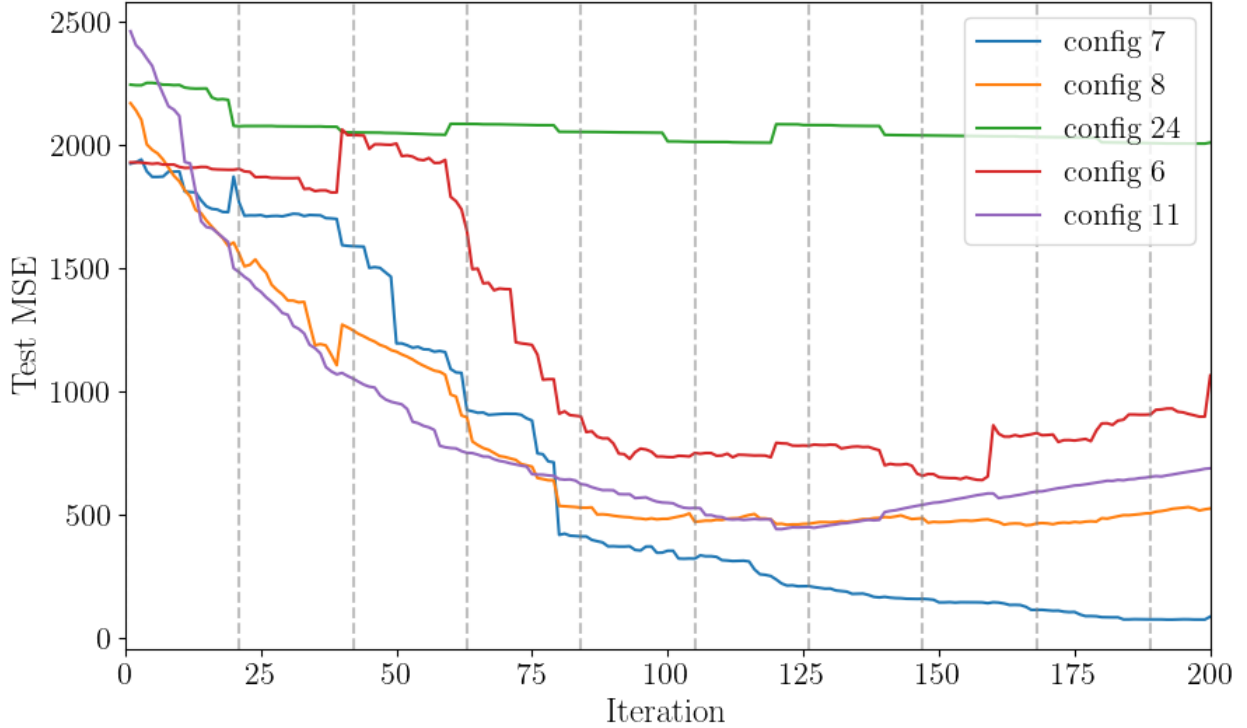


Figure 2.16: Comparison of active learning convergence for the top 5 hyperparameter configurations from Table 2.6 on the eggbox benchmark. Each colored line corresponds to a different configuration. The best initial fit (Config 7, blue) also achieves the best final performance, but the relative ranking of other configurations changes during training.

Table 2.6: Top 5 hyperparameter configurations ranked by initial test set MSE for the eggbox benchmark. All configurations use the base settings from Table 2.4. The best configuration uses the exponential squared kernel with standard input scaling and min-max output scaling.

Config	Kernel	θ scaler	y scaler	Test MSE
7	ExpSquared	Standard	MinMax	1993.4
8	ExpSquared	Standard	Standard	1998.5
24	Matern-5/2	Standard	None	2029.1
6	ExpSquared	Standard	None	2066.4
11	Matern-3/2	None	Standard	2067.5

In conclusion, there is no definitive choice of hyperparameter configurations that will work for every problem, and it often takes a bit of experimenting to find which configuration is suited for your problem. GP optimization and test error performance are inherently variable (Figure 2.15), and does not necessarily converge monotonically (sometimes the test error will increase for some iterations before improving). As we show in this section, testing a grid of different configurations is a principled and practical way of selecting training configuration parameters without adding a lot of computational overhead (i.e. it requires recomputing the GP multiple times, but uses the existing training sample and doesn't require computing any new forward model evaluations).

5 Discussion

In this section, we discuss `alabi`'s performance in terms of accuracy and computational speed, as well as known limitations. `alabi`'s accuracy at reproducing the true posterior depends on two key results: (5.1) how well the GP surrogate model has converged to the true function, and (5.3) how well the MCMC sampler is able to accurately sample the GP surrogate model.

5.1 How many training points should I use to train the surrogate model?

The amount of training points needed generally depends on the complexity and dimensionality of the problem. As a rule of thumb, a good starting place is to run $\sim 10 \times N_{\text{dim}}$ initial training samples to train the model. For well “behaved posteriors” (posteriors with unimodal structure contained within a reasonably constrained prior space), we find that initial samples of $\sim 10 \times N_{\text{dim}}$ are adequate, including for a relatively high number of dimensions. Depending on the error after the initial samples, additional active learning samples can be run as needed. `alabi`'s training can also be parallelized over multiple cores, allowing for the user to parallelize computing the initial training/test samples, as well as running multiple active learning chains at a time.

For more complex posteriors (those with multiple modes or degeneracies features), the user will likely need to train more/longer active learning chains to capture the high-likelihood regions. It is difficult to say a priori the exact number of samples that will be needed for a given problem, but a general strategy would be to run additional batches of $\sim 10 \times N_{\text{dim}}$ active learning samples (either sequentially or divided over multiple processes) and repeat until the surrogate model has converged to the desired accuracy based on the chosen test error metric (Section 2.7.2).

5.2 How do I tell if the surrogate model has converged?

Computing the prediction error (e.g. mean squared error, mean absolute error) is a straightforward and practical way to test how accurately the surrogate model is performing. Ideally, the user would compute at least $\sim 10 \times N_{\text{dim}}$ test samples, distributed across the prior space

of interest to compute the error. The “cutoff” test error value depends on the users problem and the desired level of accuracy. For Figures 2.12, we show the average percent error (2.22) so that the different examples can be compared relatively to one another.

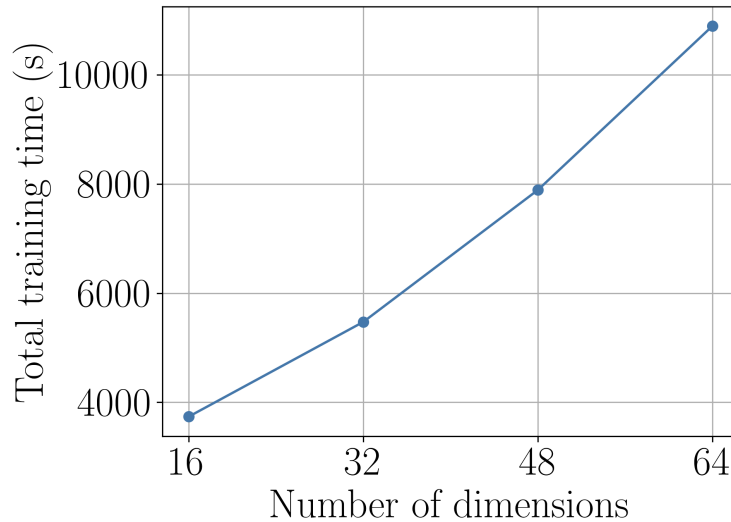


Figure 2.17

5.3 Which posterior sampler should I use for my problem?

Assuming that the surrogate model can converge to the true posterior, `alabi` can only perform as well as the underlying sampler that the user chooses. This sampling performance can depend on the posterior structure and number of dimensions.

For any sort of problem with “complex” features, such as multiple modes or degeneracies, nested sampling (through `dynesty`, `pymultinest`, or `ultranest`) tends to be much more reliable than affine-invariant sampling (such as `emcee`). For unimodal problems however, `emcee` offers more flexibility for easily defining custom prior functions, unlike nested sampling, which requires defining a prior transform function that maps points from a unit cube to the desired function. `emcee` is also simpler to run than nested sampling, requiring no free parameters to tune, whereas nested sampling has many.

Posterior sampling for high dimensional problems is also an open area of research (Jones & Hobert, 2001). Discussing the limitations of sampling methods is beyond the scope of this paper, however, we refer the reader to Dittmann (2024) for a discussion on hyperparameter tuning for nested sampling methods, as well as Rajaratnam & Sparks (2015) for a discussion on convergence for MCMC methods in high dimensions.

5.4 When is it worth training a surrogate model for my likelihood?

`alabi` can significantly reduce the number of forward model evaluations by orders of magnitude. However, the process of training `alabi` also introduces some computational overhead, and it isn’t fully advantageous to use `alabi` unless one’s forward model is computationally slow. Figure 2.18 breaks down how much computational overhead goes into each step

of `alabi` for the GP training time, objective function optimization, and hyperparameter optimization.

Figure 2.18 shows how `alabi`’s training time scales as a function of training iterations and number of dimensions. The top panel shows the scaling of the GP training time, which is the computation time required to compute the predictive mean and variance (Eqn. 2.8 and 2.9) by inverting the covariance matrix. The computational cost of inverting the covariance scales exponentially with the number of data points as $\mathcal{O}(n \log^2 n)$ using the `george` implementation of the GP (Ambikasaran et al., 2015).

For training the GP, the computation time scales exponentially with the number of training samples, but for ≤ 32 dimensions, it is not significantly more costly to run for a larger number of dimensions for a stationary kernel. The second panel of Figure 2.18 shows how long it takes to optimize the active learning function, used to select the next training point. In this example, we optimize the BAPE active learning function (2.11) during each iteration. The number of dimensions of the acquisition function is equal to the number of dimensions of the likelihood. Lastly, the bottom panel of Figure 2.18 shows the amount of time it takes to reoptimize the hyperparameters of the GP, which is performed every 50 active learning iterations. The number of hyperparameters is equal to the number of dimensions (one length scale parameter per dimension), plus two parameters (for the mean and amplitude parameters).

Figure 2.19 shows how the cumulative training time translates to the total MCMC sampling time (y-axis) as a function of forward model evaluation time (x-axis) for varying numbers of MCMC samples—i.e., the number of times the MCMC evaluates the model (teal showing for the true model, and gray showing for the surrogate model). For forward model evaluations $\gtrsim 1$ s, MCMC with the `alabi` performs a factor of 10 – 1000 $\times$ faster than MCMC with the true model, depending on the number of posterior samples computed. For `alabi`, the main computational overhead for training a GP surrogate model is mainly dependent on the number of training samples. Models with more dimensions take a larger number of training samples to converge, hence, the cumulative training time for higher dimensional problems is longer.

5.5 What hyperparameter settings should I use for my problem?

In practice, selecting GP hyperparameter settings (including the kernel function, data scaling, length scale bounds, and optimization method) is one of the most challenging aspects of training a surrogate model. A poor choice of settings can lead to overfitting, slow convergence, or numerically unstable fits, and the symptoms are not always obvious from the training loss alone. Sections 4.6 and 4.7 demonstrate two complementary strategies for addressing this challenge: visual diagnostics of GP fit quality, and systematic grid search over candidate configurations. Together, these provide a principled workflow for selecting hyperparameters.

We recommend the following approach: first, start with a grid search on the initial training sample. As demonstrated in Section 4.7, fit the GP using a grid of candidate configurations (varying the kernel, input scaler, and output scaler) and rank them by test set MSE on held-out points (Table 2.6). This step is inexpensive because it only requires fitting the GP to the initial training sample (no additional forward model evaluations), and it quickly narrows

the search space to a small number of promising settings. The grid search can be performed after the initial sampling phase (via `init_samples`) and before starting active learning.

For the top-ranked configurations, plot the 1D GP predictions around the maximum likelihood point (Section 4.6) to visually inspect the GP fit (Figure 2.14). In general, you want to check that: (a) the GP mean tracks the training data without excessive oscillation, (b) the uncertainty bands are reasonably narrow near training points, (c) the individual GP posterior draws are smooth and consistent, and (d) the optimized length scale values are not pinned at the boundary of their allowed range. If the GP draws appear overly “wiggly” or the uncertainty is unexpectedly large, the length scale bounds or penalty term (2.14) likely need to be adjusted.

From there, run active learning in batches and re-evaluate. Rather than committing to a full active learning run with a single configuration, run a moderate number of iterations (e.g., 50–100) and then re-check the test MSE and 1D diagnostic. As shown in Figure 2.16, the relative performance of configurations can change during active learning—a configuration that ranks well on the initial fit does not always converge the fastest. If the test error plateaus or increases, this is a signal to try a different configuration before investing further computational effort. This iterative approach prevents the user from wasting time on a setting that will not converge.

This workflow is designed to be practical: the grid search and diagnostic steps add minimal overhead relative to the cost of training the surrogate model, and they can save substantial time by identifying poor configurations early.

5.6 Parallelization

There are several processes that can be parallelized in `alabi`’s workflow:

1. **Initial training:** Computing the initial training and test samples can be done in parallel.
2. **Optimizing hyperparameters:** for the marginal likelihood approach (2.4.1) `alabi` will run multiple optimizations in parallel, or for the cross-validation approach (2.4.2) it will parallelize computing the cross-validation score over each candidate hyperparameter.
3. **Active learning chains:** `alabi` can run multiple active learning chains in parallel and combine all training samples at the end.
4. **MCMC sampling:** `alabi` can easily make use of the parallelized implementation for different samplers, allowing for parallelized sampling chains.

So while steps 1 – 4 must be run sequentially, the computations within each step can be distributed over multiple processors. For step 3, we find that parallelizing multiple chains while adding single points per iteration is more computationally efficient than running a single chain while adding multiple points per iteration. This efficiency arises for two reasons. First, because sequentially adding new points to the GP allows the GP to update the acquisition function each iteration, that each new point informs the model. On the other

hand, drawing multiple points at a time from the same acquisition function can cause multiple points to converge at the same location, producing redundant training data. Second, as noted in Section 5.4, the training time takes longer per iteration as the number of training samples increases. Therefore, by parallelizing over different chains, each chain remains relatively short (making computation time quicker) while adding unique information to the fit.

5.7 Limitations & Future Work

This paper compares the performance of commonly used kernels, active learning algorithms and MCMC samplers appropriate for the task of performing Bayesian inference with active learning, but it is not an exhaustive test of all possible combinations of algorithms. However, `alabi` is modularized so that new algorithms can be incorporated in the future, or custom defined by users. Further work would also implement more flexibility for different user-defined kernels.

Currently `alabi`'s use of gradient optimization is performed by taking analytic gradients of functions, but an alternative GP backend that is built off of an auto-differentiable framework could allow for improved optimization speed and flexibility.

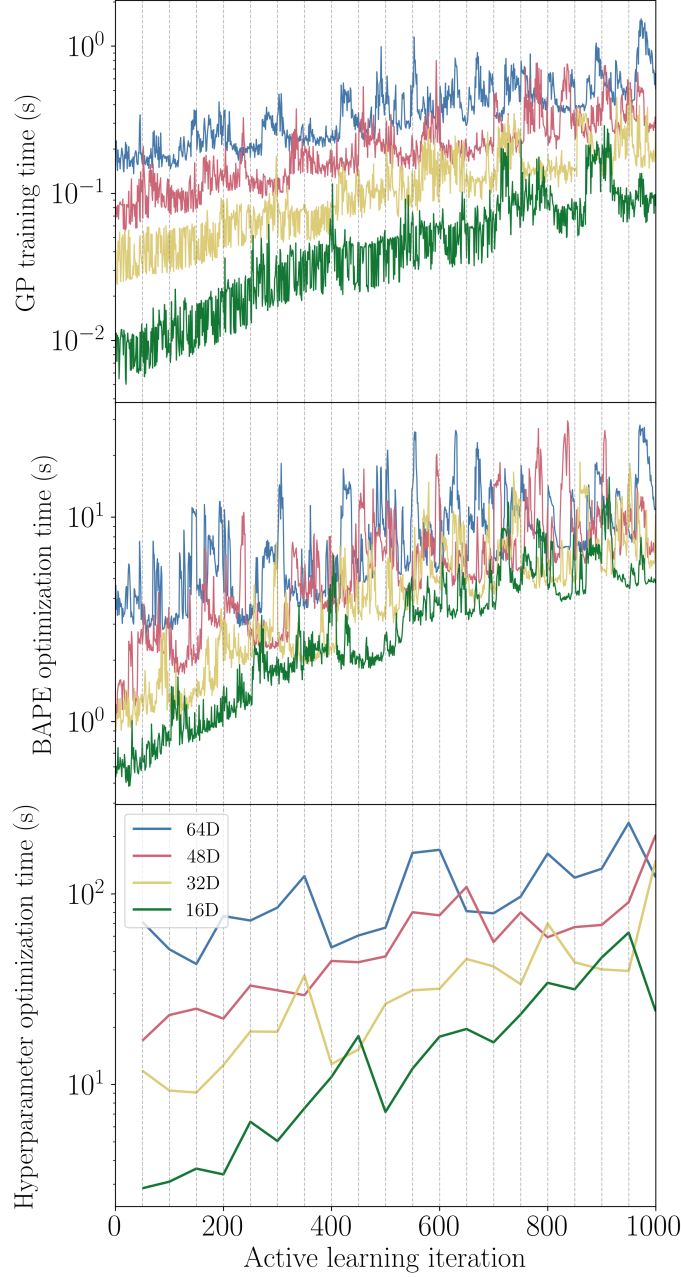


Figure 2.18: The computation time for training `alabi` as a function of the number of active learning iterations and number of dimensions. Each run (16, 32, 48 and 64D) is started with an initial training sample of $10 \times N_{\text{dim}}$ points, and run for 1000 active learning iterations, each iteration adding one additional training point. The top panel shows the time to train the GP on the new data points each iteration, the middle panel shows the time to optimize the BAPE active learning function each iteration, and the bottom panel shows the time = to reoptimize the GP hyperparameters every 50 iterations. The primary computational overhead for training the GP surrogate is the training points used, and the total computation time scales exponentially with the number of training points.

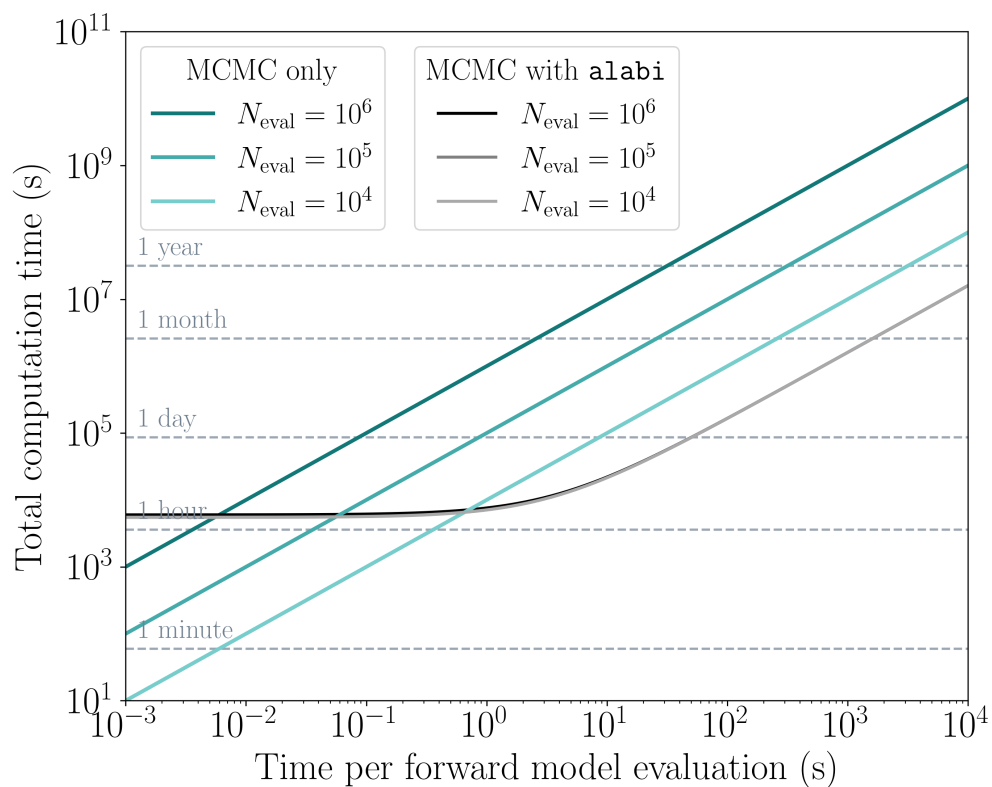


Figure 2.19: Total MCMC computation time as a function of forward model evaluation time. The teal lines show the total MCMC time when running the sampler with the true model for varying numbers of samples ($N_{\text{eval}} = 10^4, 10^5, 10^6$), while the gray lines show the total MCMC time when running the sampler with the `alabi` surrogate model. For forward models taking $\lesssim 1$ s to evaluate, the `alabi` training overhead exceeds the computational savings for computing MCMC samples, however for $\gtrsim 1$ s, MCMC with `alabi` performs a factor of $10 - 1000\times$ faster, depending on the number of posterior samples computed.

6 Conclusion

We have presented `alabi`: an open-source Python framework for accelerating Bayesian inference with computationally expensive forward models. This tool presents a way forward for performing simulation based inference with complex simulations and datasets. This framework has been applied to a variety of computational astrophysics problems (e.g., Birky et al., 2021b, 2025), and can furthermore be used for inference with any black-box forward model in any domain of computational science.

We evaluated `alabi` on a suite of benchmark problems spanning unimodal distributions, curved degeneracies, multimodal structure, and high dimensionality. Our main findings are: (1) Accuracy: `alabi` successfully trains surrogate models for problems with up to 64 dimensions (Section 4.5), and can achieve $< 1\%$ average percent $\sim 10 \times N_{\text{dim}}$ initial training points and up to 1000 active learning iterations. (2) Computational speedup: for forward models with evaluation times $\gtrsim 1$ s, `alabi` reduces the total number of model evaluations by orders of magnitude compared to direct MCMC sampling, yielding speedups of $10\text{--}1000\times$ (Section 5.4). (3) Hyperparameter selection: We demonstrated a principled workflow for selecting GP hyperparameters (Sections 4.6, 4.7, and 5.5) using a grid search and visual diagnostics test.

Each step of `alabi`’s workflow—initial sampling, hyperparameter optimization, active learning, and posterior sampling—can be parallelized across multiple processors, enabling the framework to scale to large computing clusters. `alabi` is designed to be modular: new kernels, acquisition functions, and sampling backends can be incorporated as they become available. The package, along with comprehensive documentation and tutorials for all benchmark problems presented in this paper, is publicly available at <https://github.com/jbirky/alabi>.

Acknowledgments

The authors would like to acknowledge David Fleming (formerly UW), James Davenport (UW), Eric Agol (UW), Tyler Gordon (UCSC), and Samantha Gilbert-Janizek (UW) for various constructive discussions in the process of this project. JB acknowledges funding support from NSF Graduate Research Fellowship grant number DGE-1762114 and a Scialog grant supported by the Heissing-Simmons Foundation.

Software:: APPROXPOSTERIOR (Fleming & VanderPlas, 2018), `dynesty` (Speagle, 2020b), `emcee` (Foreman-Mackey et al., 2013), `corner` (Foreman-Mackey, 2016), `george` (Ambikasaran et al., 2015), `matplotlib` (Hunter, 2007), `numpy` (Walt et al., 2011), `pymultinest` (Feroz et al., 2009; Buchner et al., 2014) `scikit-optimize` (Head et al., 2020), `scipy` (Jones et al., 2001), `ultranest` (Buchner, 2021).

The code for `alabi` is available on Github: <https://github.com/jbirky/alabi>.

CHAPTER 3

PROSPECTS OF CONSTRAINING EQUILIBRIUM TIDES IN LOW-MASS BINARY STARS

1	Introduction	66
2	Model	69
	2.1 Energy Dissipation and Torque due to Tides	70
	2.2 Stellar Evolution and Magnetic Braking	71
	2.3 Model State Space	72
	2.4 Numerical Simulations	72
3	Methods	76
	3.1 Sensitivity Analysis	76
	3.2 Simulated Likelihoods	77
	3.3 Posterior Sampling with Active Learning	78
4	Results	81
	4.1 Model Sensitivity Analysis	81
	4.2 Simulated Posteriors	91
	4.3 Degeneracies in the most simplified simulated model	98
5	Discussion	103
	5.1 Physical Origin of Short-period Binaries	104
	5.2 Outlook	105
	5.3 Limitations	109
6	Conclusion	109
1	Equilibrium Tide Equations	111
	1.1 Governing Equations	111
	1.2 Secular Evolution of Keplerian Elements	112

1 Introduction

Binary stars make up nearly half of the galactic population (Duquennoy & Mayor, 1991) and form a foundation for most of modern astrophysics. The dynamics of these systems are continuously evolving through a combination of nonlinear physical processes in which angular momentum can be exchanged between the orbit and rotations of the individual stars. As various studies have modeled so far (Witte & Savonije, 2002; Repetto & Nelemans, 2014; Penev et al., 2014; Bolmont & Mathis, 2016; Song et al., 2018; Fleming et al., 2019; Zanazzi & Wu, 2021), the dominant processes include tidal dissipation, stellar evolution, and magnetic braking. Tides cause orbital energy to be dissipated in the form of heat and cause angular momentum to exchange between the orbit and rotations, ultimately altering observable properties (orbital period, rotation period, eccentricity, and obliquity) over long timescales (Myr – Gyr). Additionally stellar evolution plays a role in dynamical evolution, as the radius and mass concentration alter the moment of inertia of the individual stars, thus influencing rotational angular momentum. Magnetic braking also plays an important role as stars lose angular momentum due to magnetized stellar winds and mass loss (Skumanich, 1972; Barnes, 2003a; Matt et al., 2015; Breimann et al., 2021). In other words, the influence of tides drives binary orbits towards a state of energy minimization (Counselman, 1973; Hut, 1980), causing systems to converge to long-term quasi-equilibrium states in which orbital periods become synchronized with rotation periods, and eccentricity decays towards circularization, which is broadly consistent with observations (Meibom et al., 2006, 2015).

The influence of stellar tides extends to many systems beyond just binary stars. Tidal interactions have been observed to cause profound dynamical effects across many scales, from exoplanet systems to Galactic populations. These effects can include tidal locking (Barnes, 2017), tidal heating (Barnes et al., 2013; Leconte et al., 2010; Levrard et al., 2007; Jackson et al., 2008), spin-orbit alignment (Heller et al., 2011; Albrecht et al., 2022), capture into spin-orbit resonances (Goldreich & Peale, 1966; Colombo & Shapiro, 1966; Correia & Laskar, 2004), as well as tidal destruction of exoplanets (Jackson et al., 2009; Hamer & Schlaufman, 2019). Additionally, on population and Galactic scales, tidal interactions play a role in the dynamics of triple and higher-order multistar systems (Ginat & Perets, 2021; Hamers et al., 2021), influencing the rate of capture, escape, or merging in multibody stellar encounters, as well as chaos in stellar clusters (e.g., Mardling & Aarseth, 2001).

Theories predicting stellar deformation can be broadly characterized as *equilibrium tide* (Hut, 1981), which assume a potential raises a hydrostatic tidal deformation on a star, or *dynamical tide* (Zahn, 1975), which assume the gravitational influence of a companion drives hydrodynamic or fluid motions. Most analytic models of tides apply a “lag-and-add” approach (Greenberg, 2009), in which the gravitational response of a star is considered to be the summation of elongated bulge components. Among the most simplified models of equilibrium tide which have been widely applied to stars and planets, the lag components are considered to be constant in phase or constant in time. (MacDonald, 1964; Kaula, 1966; Goldreich & Soter, 1966; Greenberg, 1974). More complex nonlinear mechanisms of tides have also been proposed for stars of different structures. For low-mass stars with radiative cores and convective envelopes, these models consider processes such as turbulent viscosity

that dissipates the equilibrium tide in convective layers (Zahn, 2008; Vidal & Barker, 2020), or dissipation due to internal gravity waves in radiative zones due to the dynamical tide (Zahn, 1975; Lai, 1997; Terquem et al., 1998; Witte & Savonije, 1999; Witte & Savonije, 2002; Ogilvie, 2013; Fuller & Lai, 2012; Barker & Ogilvie, 2009; Barker, 2020, 2022; Burkart et al., 2014; Zanazzi & Wu, 2021).

Current observational opportunities through large-scale stellar surveys offer significant promise for rigorously testing hypotheses of tides for different stellar types. From time domain surveys such as *Kepler* (Borucki et al., 2010), *K2* (Howell et al., 2014), and *TESS* (Transiting Exoplanet Survey Satellite; Ricker et al., 2015), eclipsing binaries are ideal observational laboratories for constraining stellar tides, as they enable direct measurements of fundamental stellar properties (masses, densities, temperatures, etc.), and undergo observable changes in orbital/rotational dynamics under the long-term influence of tides. Time-series photometry of eclipsing binaries provides constraints on orbital periods, rotation periods (from star-spot rotation), as well as approximate eccentricities (Matson et al., 2016) and mass estimates (Windemuth et al., 2019).

Already *Kepler* and *TESS* have revolutionized discoveries of tidal synchronization (Lurie et al., 2017; Hobson-Ritz et al., 2025) by discovering an intriguing sub-population of FGKM-type binaries in subsynchronous spin-orbit ratios in which the star has a slower rotation period than the orbital period. Fleming et al. (2019) demonstrated that equilibrium tide models can reproduce 1:1 synchronized binaries and generate a wide distribution of subsynchronously rotating binaries. However their model falls short of reproducing the tight overdensity of observed subsynchronous rotators discovered in Lurie et al. (2017) in which $\sim 15\%$ of the *Kepler* eclipsing binary population was found to have rotation periods $\sim 13\%$ slower than their orbital period. Lurie et al. (2017) suggested that this population of subsynchronous binaries could be the result of differential rotation in which a latitudinal shear along a star could produce starspots rotating slower than the equator at high latitudes. Furthermore, Lurie et al. (2017), Hobson-Ritz et al. (2025), and Jermyn et al. (2020) suggest that low-mass eclipsing binaries with measurable star-spot modulation are promising targets for investigating differential rotation on the surface of convective stars.

In addition to eclipsing binaries, spectroscopic follow-up has enabled precise masses and eccentricities of many binaries from radial velocity constraints. Multi-epoch surveys, such as the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al., 2015), have enabled the discovery and orbital constraints of thousands of spectroscopic binaries (Price-Whelan & Goodman, 2018; El-Badry et al., 2018; Price-Whelan et al., 2020; Kounkel et al., 2021). In some cases with detailed follow-up, it is additionally possible to constrain obliquities from individual binary systems through the Rossiter-McLaughlin effect (Mazeh, 2008; Albrecht et al., 2022), or infer the distribution of obliquities from a population of binaries hierarchically (Morton & Winn, 2014). Although ages are among the most challenging parameters to determine for stellar systems (Soderblom, 2010), *K2* and *TESS* have enabled Galactic coverage of many stellar populations, including binaries in open clusters, which are the gold standard when it comes to constraining orbital properties as a function of age (*e.g.*, Southworth & Clausen, 2006; David et al., 2015, 2016; Gillen et al., 2017; Torres et al., 2018).

Despite the drastic improvements in observational capabilities, it remains an open question as how to best approach constraining tidal models against data. Early attempts to constrain

tidal models considered a metric known as the circularization cut-off period (Mayor & Mermilliod, 1984; Mathieu & Mazeh, 1988; Mathieu et al., 2004), which determines the orbital period at which most binaries in a population have circularized (where $e \approx 0$ for systems with $P_{\text{orb}} < P_{\text{cut}}$). The idea behind this method relies on the assumption that binaries at close orbital separation will experience stronger tides, and should thus circularize faster — a trend which has been confirmed in observations of open clusters (Mathieu & Mazeh, 1988; Mathieu et al., 2004).

Later work by Meibom & Mathieu (2005) proposed a more robust metric known as the tidal circularization period. Instead of determining P_{cut} from the observations at which $e > 0$, they fit a functional form to observed P_{orb} and e , assuming that populations of binaries could originate from a distribution of initial orbital periods and eccentricities. This circularization period method did a better job at accounting for some uncertainty in initial conditions; however, it did not consider all prior uncertainties (e.g. rotation period, stellar evolution) or incorporate the constraints from stellar rotation periods. Recent studies have called these additional model factors into question, such as Bash et al. (2023), which finds evidence from *Gaia* spectroscopic binaries that the circularization period depends more on stellar effective temperature than open cluster age. Additionally a study by Mirouh et al. (2023) used population synthesis to examine the impacts of tides and initial orbital period distributions on the circularization and synchronization of binaries as a function of age (using data of systems from 8 open clusters). Their findings concluded that constraining tidal efficiency based on the circularization of binary populations is difficult or impossible due to the inefficiency of tides on the main-sequence, as well as strong dependencies on the initial orbital configuration. However, their work also showed that tidal synchronization better captures the age-dependent effects of tidal dissipation, and thus the fraction of synchronous binaries at a given age would be a more promising constraint on tides.

More generally, there are debates in the literature as to whether individual or population approaches to inferring tidal Q 's are more appropriate. Some approaches have focused on using forward modeling approaches to infer unknown parameters including the initial orbital states and the tidal dissipation rate (or equivalently the tidal Q) value that best reproduces observations of individual stars or planets (Goldreich & Soter, 1966). Studies such as Barker (2020) advocate for a system-by-system comparison of tidal Q values, given the potential dependence Q may have with stellar parameters or age. On the other hand, studies such as Cameron & Jardine (2018) advocate for a population based approach using a broad sample of binaries to constrain Q . The population approach has the advantage of accounting for uncertainties in the initial condition as a population parameter; however, it requires making an assumption about how Q depends on stellar parameters (in this case, Cameron & Jardine (2018) assume a functional dependence on stellar effective temperature). Resulting estimates of tidal Q for solar-like stars from the literature have spanned orders of magnitude with estimates of Q ranging from $\sim 10^5$ to 10^8 (Meibom et al., 2006; Jackson et al., 2009; Hansen, 2010; Patel & Penev, 2022; Penev & Schussler, 2022), bringing into question the underlying cause of the discrepancy. Given that constraints on stellar tides have far-reaching consequences on dynamics from exoplanetary to galactic scales, with decades of research investment, it is critical that we rigorously examine all uncertainties (both model and observational) to understand the prospects and limitations of constraining the tidal dissipation mechanisms.

From a theoretical perspective, there may be a variety of reasons why latent parameters may not be feasible to constrain: there may be infinite solutions, there may be finite solutions, or there may be no solutions. If solutions exist, it may be that the latent parameters do not (strongly) influence manifest parameters, or there may be significant degeneracies between latent parameters and initial conditions. As we argue in this paper, constraining tidal dissipation has been an underestimated challenge, even when considering the simplest formulations of tidal dissipation against simulated data.

In this paper, we assess the degree to which improved observations would resolve the discrepancy in the inferred values of tidal $\mathcal{Q}$ based on individual systems. For simplicity we will consider a commonly used and easily tractable model of the equilibrium tide, however the statistical methodology presented in Section 3 may also be applicable to more complex tidal dissipation models, including dynamical tide formulations. In this study, our aim is to answer the questions: *How do we constrain theories of tidal dissipation?* In particular, to what degree is model inference inhibited by unidentifiability or degeneracies inherent in the model formulation? To what degree do observational uncertainties limit model inference? Finally, what needs to be done next and how should we focus our efforts (*e.g.*, such as observational follow up)?

In this paper we characterize the prospects and limitations of constraining equilibrium tides in low-mass stars using global sensitivity analysis (Sobol, 2001; Saltelli et al., 2010), which has been widely used in applications from biological systems to finance (Saltelli et al., 2017). Although it is less widely applied to astrophysical simulations, as we present in this paper, sensitivity analysis is an effective way to systematically characterize how sensitive model outputs are to model inputs. Sensitivity analysis tells us which input parameters of the model are most influential to first order, however, it does not give us as much information about the correlation between model inputs. To further investigate the dominant effects of this non-linear model, we perform simulated Bayesian inference (using Active Learning for Accelerated Bayesian Inference, or *alabi*; Birky et al., in prep.) (see also Chapter 2) to test how observational uncertainties influence the quality of posterior constraints as well as examine the degeneracies between input parameters.

This paper is organized as follows. Section 2 defines the terminology used throughout the analysis, as well as describes the model assumptions and numerical implementations used for all of our simulations. Section 3 describes the challenges that arise when comparing tidal models to binary observations and describes the methodology for sensitivity analysis. Section 4 explains the results of the sensitivity analysis applied to the numerical simulations described in Section 2. Section 5 synthesizes the main findings and discusses some of the main limitations of this analysis. Finally, Section 6 discusses future observational and numerical work that is motivated from our results.

2 Model

The rotational and orbital evolution of low-mass stars ($M_* < 1.5M_\odot$) is influenced by a combination of physical processes, namely tidal dissipation (Section 2.1), stellar evolution (Section 2.2), and magnetic braking (Section 2.2). Here we provide an overview of the main

assumptions used in our model. In Section 2.3 we summarize the key model variables for the coupled system of ordinary differential equations and numerically solve the equations using the `VPLanet` implementation (Barnes et al. 2020; see also Fleming et al. 2018, 2019). The goal of this model set up is to analyze and establish the limitations of comparing equilibrium tidal models to datasets available in surveys (*e.g.* *Kepler*, *K2*, *TESS*, *APOGEE*) and the literature.

2.1 Energy Dissipation and Torque due to Tides

To consider the rate of energy dissipated over an orbit, it is common to introduce a quantity known as the “tidal quality factor”. This quantity is proportional to the maximum energy stored in tidal deformation divided by the average energy dissipated over an orbit. With E_0 defined as the maximum energy and E being the orbit energy as a function of time t , the tidal Q is defined as

$$Q^{-1} \equiv \frac{\text{energy dissipated over an orbit}}{\text{maximum energy stored in deformation}} \equiv \frac{1}{2\pi E_0} \oint \left(-\frac{dE}{dt} \right) dt. \quad (3.1)$$

A lower tidal Q value indicates that the star is more efficient at dissipating tidal energy. This parameterization of tidal Q can be translated into the secular evolution of the orbit by assuming that each Fourier component of the potential induces a tidal response on a perturbed star. In other words, each component of the potential drives the distortion of material at a particular frequency and phase lag, known as σ and ε respectively. This effective forcing frequency depends on both the mean motion n and the rotation frequency ω , where $\sigma = kn - m\omega$ and k and m are integers. The phase lag ε is equal to twice the geometric lag angle of the bulge orientation. The tidal Q is related to the phase lag (Equation 105, Efroimsky & Williams, 2009):

$$Q^{-1} = \frac{\tan \varepsilon}{1 - \left(\frac{\pi}{2} - \varepsilon \right) \tan \varepsilon} = \tan \varepsilon + \mathcal{O}(\varepsilon^2). \quad (3.2)$$

In general, the relationship between the phase lag and the effective forcing frequency $\varepsilon(\sigma)$ depends on the internal structure of the body. However, for small phase lags (assuming a weak friction approximation), $\tan \varepsilon(\sigma) \approx \varepsilon(\sigma)$, in which the response of a tidal bulge under the forcing frequency is analogous to a damped driven harmonic oscillator (Greenberg, 2009). The relation between ε and the tidal Q can be approximately linearized by assuming either a constant phase lag $\varepsilon(\sigma) \approx 1/Q$, where Q is held constant, or a constant time lag that assumes $\varepsilon(\sigma) \approx n\tau$, where the time lag, τ is held constant. These linear parameterizations of the equilibrium tide are hereafter referred to as the constant phase lag (CPL; Goldreich & Soter, 1966; Ferraz-Mello et al., 2008) and the constant time lag (CTL; Hut, 1981; Leconte et al., 2010) models. The derivations for the CTL and CPL models are well reviewed in literature, but we provide a summary of the assumptions and equations used in the Appendix Section 1.

2.2 Stellar Evolution and Magnetic Braking

In general, the relationship between external torques and the angular momentum evolution of a single star is written as

$$\mathcal{T}_{\text{rot}} = \frac{dJ_{\text{rot}}}{dt} = \mathcal{I} \frac{d\omega}{dt} + \omega \frac{d\mathcal{I}}{dt}, \quad (3.3)$$

where $J_{\text{rot}} = \mathcal{I}\omega$ is the rotational angular momentum, $\mathcal{I} = Mr_g^2 R^2$ is the moment of inertia (with r_g being the radius of gyration), and $\mathcal{T}_{\text{rot}}$ is the net rotational torque due to external influences (in the case of stellar evolution, the angular momentum lost due to stellar winds, $\mathcal{T}_{\text{rot}} = \mathcal{T}_{\text{mb}}$). **VPLanet** includes several implementations of magnetic braking models (Skumanich, 1972; Repetto & Nelemans, 2014; Reiners & Mohanty, 2012; Matt et al., 2015; Breimann et al., 2021) that simulate the rotational torque due to stellar winds. In this study we adopt the Matt et al. (2015, 2019) model for $\mathcal{T}_{\text{mb}}$ that considers a semi-analytic saturated magnetic braking law, and uses stellar evolution models from Baraffe et al. (2015) to determine the change in moment of inertia, $d\mathcal{I}/dt$.

Magnetohydrodynamic simulations of solar-like stellar winds (Matt et al., 2012) show that the rotational torque on a star can generally be written in semianalytic form. Making physically motivated assumptions, Matt et al. (2015) introduce a scaling relation for the surface magnetic field B_* and the stellar wind mass loss $\dot{M}_w$ as a function of stellar mass and Rossby number, in which the torque due to stellar winds can be written as:

$$\mathcal{T}_{\text{mb}} = \mathcal{T}_0 \begin{cases} \chi^2 \left(\frac{\omega}{\omega_\odot} \right) & \text{if saturated, } R_o \leq R_{o,\odot}/\chi \\ \left(\frac{\tau_{cz}}{\tau_{cz,\odot}} \right)^2 \left(\frac{\omega}{\omega_\odot} \right)^3 & \text{if unsaturated, } R_o > R_{o,\odot}/\chi \end{cases}. \quad (3.4)$$

The Rossby number is the ratio of the rotation period to the convective turnover timescale, $R_o = P_{\text{rot}}/\tau_{cz}$. The timescale τ_{cz} is computed using a relation from Cranmer & Saar (2011). The transition between saturated and unsaturated regimes is specified by $\chi \equiv R_{o,\odot}/R_{o,\text{sat}}$, where we adopt a value of $\chi = 10$ as used in Matt et al. (2015). The proportionality constant is

$$\mathcal{T}_0 = 6.3 \times 10^{30} \text{ erg} \left(\frac{R}{R_\odot} \right)^{3.1} \left(\frac{M}{M_\odot} \right)^{0.5}. \quad (3.5)$$

In the case of a binary star system, the rotational torque acting on an individual star is influenced by both the magnetic braking torque as well as the tidal torque. We make the assumption that the torque processes due to magnetic braking and tides are linearly independent, such that the net rotational torque is the sum of the two components $\mathcal{T}_{\text{rot}} = \mathcal{T}_{\text{mb}} + \mathcal{T}_{\text{tide}}$. We note that a non-linear coupling between $\mathcal{T}_{\text{mb}}$ and $\mathcal{T}_{\text{tide}}$ (e.g., considering the influence of tides on magnetic field structure of a star) would require significantly more theoretical work and magnetohydrodynamic simulations to accurately model, hence we consider only linear coupling following previous studies (Witte & Savonije, 2002; Repetto & Nelemans, 2014; Penev et al., 2014; Bolmont & Mathis, 2016; Song et al., 2018; Fleming et al., 2019; Zanazzi & Wu, 2021).

2.3 Model State Space

Adopting terminology from dynamical systems analysis (Willems, 2000), the characterization of each of our model free parameters is listed in Table 3.1. Model variables that govern the dynamical evolution of a system can be characterized as *state variables* (time-variable parameters of the coupled ordinary differential equations), *manifest variables* (time-fixed parameters that are observable), or *latent variables* (time-fixed parameters that are not directly observable).

Table 3.1: Definition of model parameters, where subscript 1 is the primary and subscript 2 is the secondary for the two stars in the binary system. Note that we use the variables (ω, a, e) in our model parameterization (Section 1) but the state of the system can be equivalently parameterized by $(P_{\text{rot}}, P_{\text{orb}}, e)$, with P_{rot} being the rotation period and P_{orb} being the orbital period, which are the values which we use to compare to observations.

<i>State Variables</i>	
a	semi-major axis of the orbit
e	eccentricity of the orbit
ω_1, ω_2	rotation frequency of each star
ψ_1, ψ_2	obliquity of each star (angle between spin axis and orbital angular momentum)
R_1, R_2	average radius of each star
r_{g1}, r_{g2}	radius of gyration of each star
<i>Manifest Variables</i>	
M_1, M_2	mass of each star
<i>Latent Variables</i>	
$\mathcal{Q}_1, \mathcal{Q}_2$	tidal quality factor
τ_1, τ_2	tidal time lag
k_2	Love number of second degree
τ_{cz}	convective turnover timescale

2.4 Numerical Simulations

We use the VPLanet package (Barnes et al., 2020) to evolve the system of differential equations using fourth order Runge-Kutta with adaptive time stepping. Our implementation follows Fleming et al. (2018) and Fleming et al. (2019) in which we numerically solve for each state variable (Table 3.1). A demonstration of evolution trajectories is shown in Figure 3.1 for the CTL model and Figure 3.2 for the CPL model. The state variables are coupled under the assumption of conservation of energy and angular momentum. Figures 3.3 and 3.4 illustrate how energy and angular momentum are exchanged between orbit and rotation over the course of 8 Gyr evolution for a $1M_{\odot} - 1M_{\odot}$ binary system according to both CTL and CPL implementations.

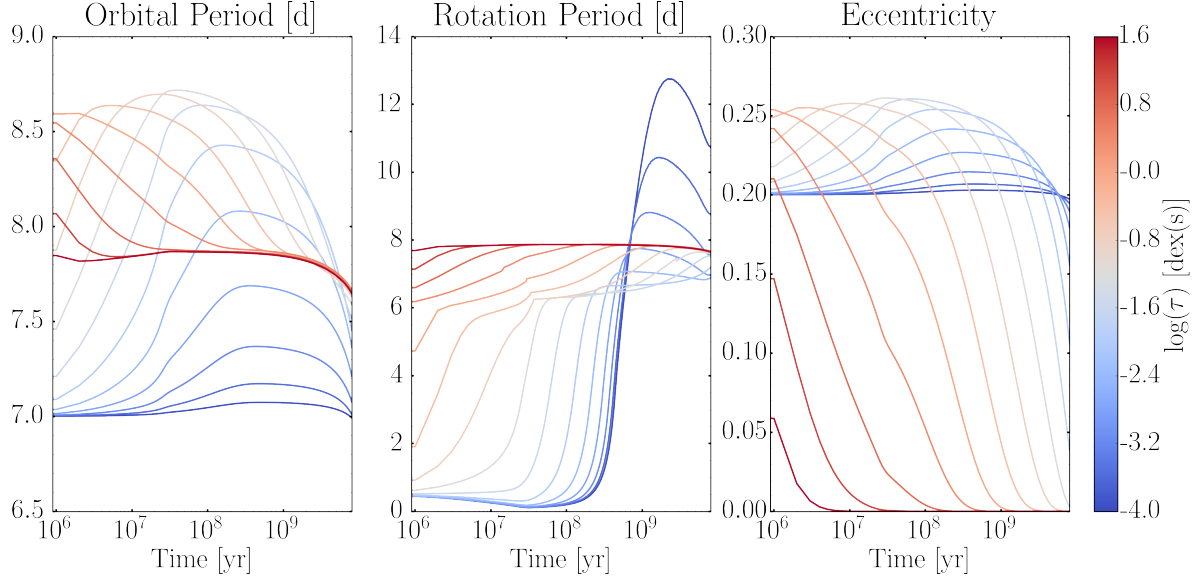


Figure 3.1: Simulations of equilibrium tide (CTL) coupled with stellar evolution. Panels show the evolution of orbital period (left), rotational period (center), and eccentricity (right) for tidal τ strengths in the range $-4.0 < \log(\tau) < 1.6$. Tidal τ influences the timescale of synchronization and circularization, where *higher* tidal τ results in more rapid evolution. Each track originates from the same initial conditions for a $1M_{\odot} - 1M_{\odot}$ mass binary ($P_{\text{orb},i} = 7.0$ d, $P_{\text{rot1},i} = P_{\text{rot2},i} = 0.5$ d, $e_i = 0.3$, $\psi_{1i} = \psi_{2i} = 10^{\circ}$; Table 3.2), and we vary only tidal τ to illustrate the effect of tides on the timescale of orbital evolution.

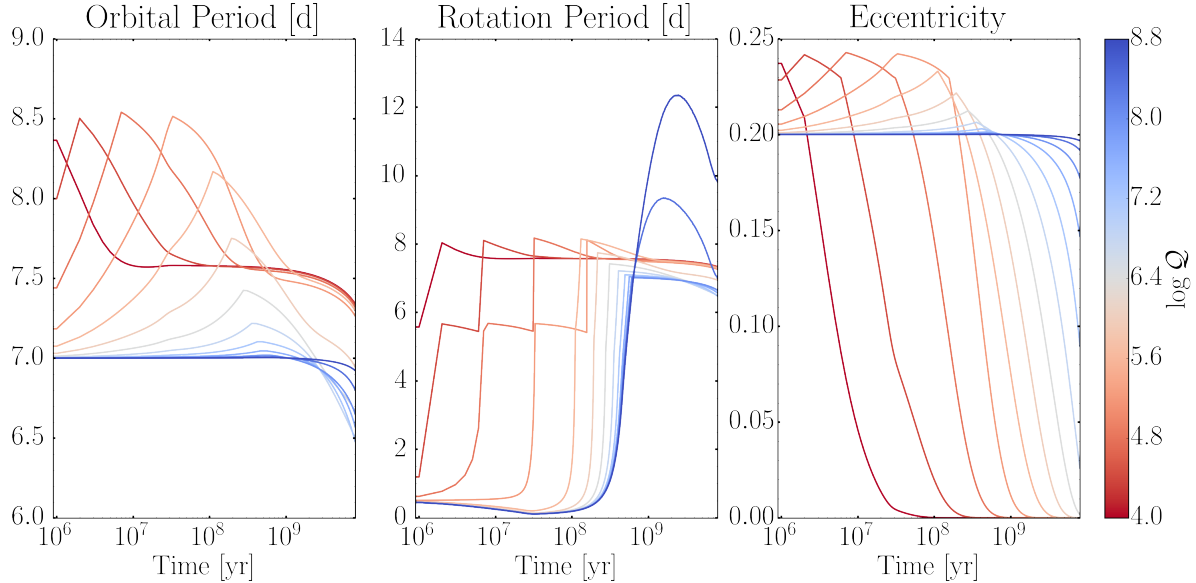


Figure 3.2: Same as Figure 3.1, but with the CPL model. Panels show tidal Q strengths in the range $4.0 < \log(Q) < 9.0$. Tidal Q influences the timescale of synchronization and circularization, where *lower* tidal Q results in more rapid evolution.

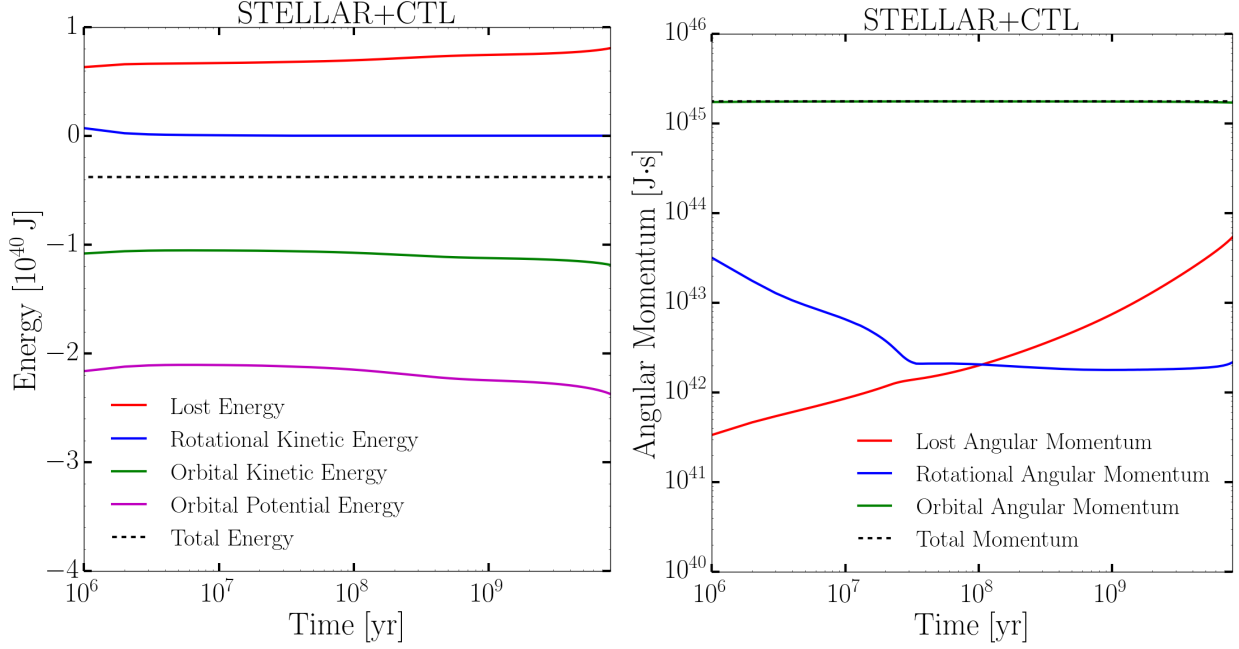


Figure 3.3: Exchange in energy (left panel) and angular momentum (right panel) between the rotation and orbit of stars in the binary system, as predicted by the CTL model (Hut, 1981) when coupled with stellar evolution (Baraffe et al., 2015) and magnetic braking (Matt et al., 2015), using fiducial initial conditions from Table 3.2 with $\log \tau = -1.0$. Lost energy (red line, left panel) is due to tidal heating, and lost angular momentum is due to stellar winds (red line, right panel). The black dashed line shows that energy and angular momentum remained conserved over time, showing that the coupling of our equations is behaving as expected.

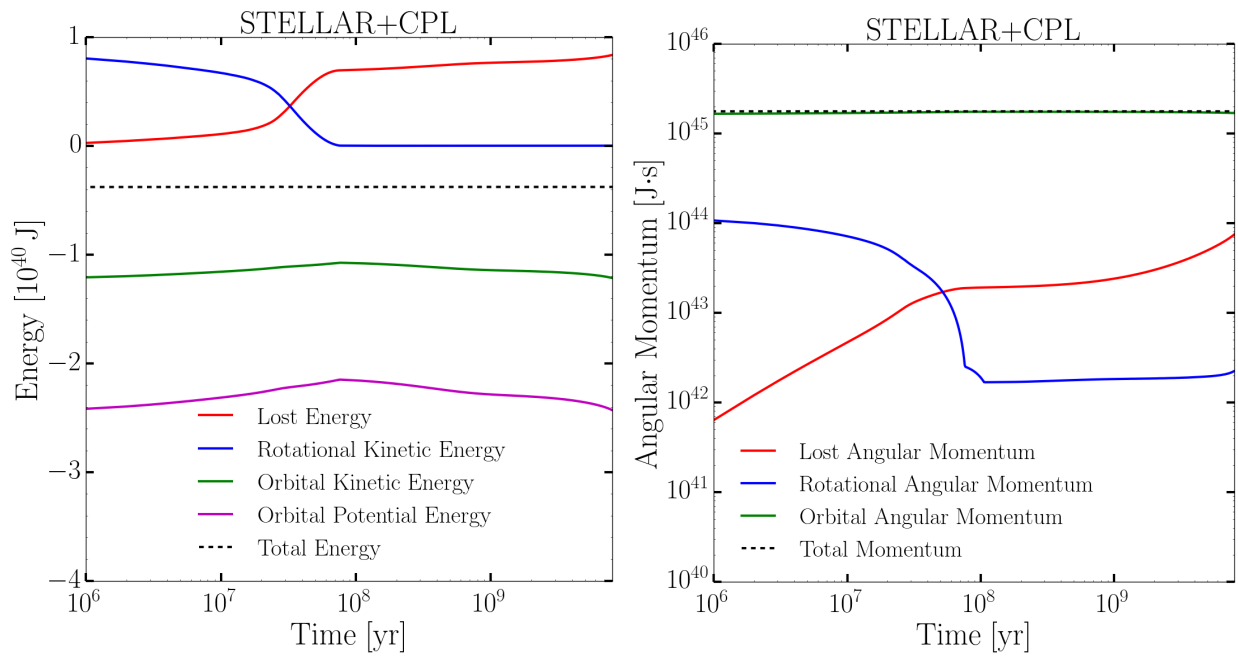


Figure 3.4: Same as Figure 3.3, but for the CPL model.

3 Methods

3.1 Sensitivity Analysis

Interpreting high-dimensional, nonlinear dynamical models is a non-trivial task. Understanding the potential for inference (e.g., identifying systems with the strongest constraints and inspecting degeneracies) requires a framework to assess how sensitive the model outputs are to each input. Overcoming these challenges requires developing a methodology for systematically analyzing high-dimensional parameter spaces to establish which input model parameters dominate the final state, and to identify which observable uncertainties are the most important to reduce (thus, informing observational priorities). To make this assessment, we perform sensitivity analyses. We also consider the best-case limits for inferring latent parameters of the models (Q or τ) by applying sensitivity analysis to a simulated likelihood with optimistic uncertainties based on observational studies of the most precisely constrained systems.

Variance-based global sensitivity analysis is a statistical technique used to identify and quantify the impact of input variables on the variance of a system's output. It is often used to assess the robustness and uncertainty of a model or simulation, as well as to identify the most influential inputs. Sensitivity analysis has been widely applied in many areas of quantitative model analysis, but has seen relatively few applications on astrophysical models (e.g., [Oleskiewicz & Baugh, 2019](#)).

The most well-known variance-based method is the Sobol sensitivity indices ([Sobol, 2001](#); [Saltelli et al., 2010](#)), which decompose the output variance into main effects and interaction effects. This method involves sampling the input space using a pseudo-random design and evaluating the output for each set of inputs. The resulting data are then used to calculate sensitivity indices for each input variable, which represent the contribution of that variable to the total variance of the output.

There are several types of Sobol sensitivity indices: first-order indices that measure the contribution of each individual input variable to the output variance, total-order indices that measure the contribution of each input variable, along with all possible interactions with other variables, to the output variance, and higher-order indices that measure the contribution of specific interactions between input variables to the output variance.

The procedure for computing Sobol sensitivity indices is as follows. Any model that can be described in the form $Y = f(\mathbf{X})$, where $\mathbf{X} \in \mathbb{R}^d$ is an input vector, and d is the number of input variables. If Y is a chosen univariate output of the model, Y can be decomposed in the form:

$$Y = f_0 + \sum_{i=1}^d f_i(X_i) + \sum_{i < j}^d f_{ij}(X_i, X_j) + \sum_{i < j < k}^d f_{ijk}(X_i, X_j, X_k) + \cdots + f_{1,\dots,d}(X_1, \dots, X_d), \quad (3.6)$$

where $f_0 = \mathbb{E}[Y]$ is the overall mean, $f_i(X_i) = \mathbb{E}[Y \mid X_i] - f_0$ captures the main effect of variable X_i , $f_{ij}(X_i, X_j) = \mathbb{E}[Y \mid X_i, X_j] - f_i - f_j - f_0$ captures the pairwise interaction between X_i and X_j , and higher-order terms are defined analogously. Taking the variance of

Y yields:

$$\text{Var}(Y) = \sum_{i=1}^d V_i + \sum_{i<j}^d V_{ij} + \cdots + V_{12\dots d} \quad (3.7)$$

where the variance components can be computed as:

$$V_i = \text{Var}_{X_{\sim i}}(E_{X_{\sim i}}[Y|X_i]), \quad (3.8)$$

$$V_{ij} = \text{Var}_{X_{ij}}(E_{X_{ij}}[Y|X_i, X_j]) - V_i - V_j. \quad (3.9)$$

Here the notation $X_{\sim i}$ is used to denote the set of all inputs except X_i . The first-order sensitivity indices, S_i , are defined as the fraction V_i attributed to model parameter i , divided by the total variance $\text{Var}(Y)$:

$$S_i = \frac{V_i}{\text{Var}(Y)}. \quad (3.10)$$

We follow the procedure of [Herman & Usher \(2017\)](#) for computing variances. First, a set of N random samples for $\mathbf{X}$ is drawn from a defined input domain. In practice, this sampling is performed using a pseudo-random approach, such as Sobol sampling ([Sobol, 2001](#)). The idea behind Sobol sampling is to create a low-discrepancy sequence of points that is more evenly spaced than a random sampling method and can efficiently span a large parameter space given a finite set of points. For each parameter i , we calculate the variance of the output due to that parameter, V_i , via computing the estimator:

$$\text{Var}_{X_i}(E_{X_{\sim i}}[Y|X_i]) \approx \frac{1}{N} \sum_{j=1}^N f(\mathbf{B})_j (f(\mathbf{A}_B^i)_j - f(\mathbf{A})_j), \quad (3.11)$$

where $\mathbf{A}$ and $\mathbf{B}$ are each $N \times d$ matrices of input parameters sampled using the Sobol (or other pseudorandom) method, and $\mathbf{A}_B^i$ denotes the matrix formed by taking all columns from $\mathbf{A}$ except the i -th, which is replaced by the i -th column of $\mathbf{B}$ ([Saltelli et al., 2010](#)). In our analysis, we use the **SALib** Python implementation ([Herman & Usher, 2017](#)) to compute the Sobol sensitivity indices.

3.2 Simulated Likelihoods

In addition to the sensitivity of each output parameter, we also consider how the relative uncertainties weight each output. To make this assessment, we consider a simulated likelihood function:

$$\ln \mathcal{L} = \sum_{j=1}^n \left(\frac{y_j - \hat{y}_j}{\hat{\sigma}_j} \right)^2, \quad (3.12)$$

where $y_j = \{P_{\text{rot}}, P_{\text{orb}}, e\}$ are the outputs of a given simulation compared to a fiducial simulation with mean values $\hat{y}_j$ and uncertainties $\hat{\sigma}_j$. In other words, y_j represents the model samples and $(\hat{y}_j, \hat{\sigma}_j)$ represent the simulated data.

We further simulate posterior probability estimates according to Bayes' Theorem,

$$\ln \mathcal{P} \propto \ln \mathcal{L} + \ln \Pi, \quad (3.13)$$

Input	Unit	Min range	Max range	Fiducial	Prior
M_1, M_2	$M_\odot$	0.1	1.1	1.0	1.0
ψ_{1i}, ψ_{2i}	deg	0	30	10	0
$P_{\text{rot}1,i}, P_{\text{rot}2,i}$	days	0.1	10.0	0.5	$\mathcal{U}(0.1, 10.0)$
$P_{\text{orb},i}$	days	0.1	12.0	7.0	$\mathcal{U}(0.1, 12.0)$
e_i		0.0	0.5	0.3	$\mathcal{U}(0.0, 0.5)$
$\log_{10}(\mathcal{Q}_1), \log_{10}(\mathcal{Q}_2)$		4.0	12.0	6.0	$\mathcal{U}(4.0, 12.0)$
$\log_{10}(\tau_1), \log_{10}(\tau_2)$	log (s)	-4.0	1.0	-1.0	$\mathcal{U}(-4.0, 1.0)$

Table 3.2: Ranges of input parameters considered in the sensitivity analysis simulations. Input parameters chosen for the fiducial simulation to compute the likelihood (Equation 3.12). Subscript 1 and 2 denote the primary and secondary of the system, while subscript i denotes an initial value of a time-varying parameter.

where $\ln \mathcal{P}$ is the sampled posterior distribution sampled according to an uninformative prior, $\ln \Pi$, and model likelihood fit is $\ln \mathcal{L}$ (Eqn. 3.12). Fiducial values for the simulated likelihood, as well as prior ranges, are given in Table 3.2.

We consider the limits of inference under optimistic uncertainties, adopting the values for $\hat{\sigma}_1$ and $\hat{\sigma}_2$ listed in Table 3.3. The uncertainties in Table 3.3 are chosen to represent the (optimistic) order of magnitude to which each parameter has been constrained in the literature using data from eclipsing binaries and radial velocities (David et al., 2015, 2016; Gillen et al., 2017; Torres et al., 2018). Particularly in time series photometry with long baselines such as *Kepler*, orbital periods can be highly precise (~ 1 sec). Rotation periods, on the other hand, are inherently more difficult to constrain, as they depend on the complex star-spot processes that make them observable (*e.g.* spot latitude, differential rotation, and emergence/decay timescale; Aigrain et al., 2015). However studies such as David et al. 2016 and Gillen et al. 2017 show that rotation periods for short-period binaries may be measured to uncertainties of $\sim 0.01 - 0.05$ d, by using a Gaussian process with a periodic kernel (see 5.4 of Rasmussen & Williams, 2006). While obliquities can in some cases be constrained by radial velocities (Triaud, 2018; Hatzes, 2019; Albrecht et al., 2007), given that the results of our sensitivity analysis (Section 4.1, Figures 3.5–3.6) suggest that obliquity is not a dominant factor in the final orbital state, for simplification in these posterior simulations we hold obliquity fixed at $\psi = 0$.

3.3 Posterior Sampling with Active Learning

We further examine the structure of the simulated posterior by sampling the points from (Eqn. 3.13) using the priors of Table 3.2. Systematically sampling points in a d -dimensional space at a density high enough to resolve high-posterior modes is a challenging problem especially as d becomes large and forward model evaluations are computationally expensive. In the case of our problem, we sample $d = 5$ input parameters (initial $P_{\text{rot},1}$, $P_{\text{rot},2}$, e , P_{orb} , and $\mathcal{Q}$). The computation time for our model can take on an order of a few seconds (integrated to ~ 10 s of Myrs) to ~ 30 seconds (integrated to ~ 10 Gyr).

Parameter	Description	Observation Method	Uncertainty
$P_{\text{rot1},f}, P_{\text{rot2},f}$	final rotation period [days]	LC autocorrelation function	0.01
$P_{\text{orb},f}$	final orbital period [days]	LC lomb scargle	10^{-5}
e_f	final eccentricity	LC eclipse + RVs	10^{-3}

Table 3.3: Model parameters that are observable from lightcurve (LC) photometry and/or spectroscopic radial velocity (RV) datasets. The subscript f denotes that these are the final values of each state variable, evaluated at a given age. The “Uncertainty” column reports typical uncertainties that are obtainable by certain types of observation, listed in the corresponding column of the “Observation Method” column. The uncertainties obtained from a method obviously depend on the quality of a particular dataset (instrument, observing conditions, quantity of observations, etc.); however, as argued in Section 4.1, the relative orders of magnitude for observable precision is important for considering any likelihood-based inference.

In order to efficiently sample and interpret the structure of our 5-dimensional posterior, we adopt the methodology presented in [Kandasamy et al. \(2017\)](#) in which we train a Gaussian process (GP) surrogate model to replicate the true posterior function and apply active learning to iteratively sample points concentrated at high probability. This methodology is also highly efficient for performing Markov chain Monte Carlo sampling with computationally expensive models ([Fleming & VanderPlas, 2018](#); [Fleming et al., 2020](#); [Birky et al., 2021b](#)), however in this analysis we just use it to visualize high-probability modes and degeneracies in the posterior. We perform this sampling using the package `alabi` (Active Learning for Accelerated Bayesian Inference; [Birky et al., in prep.](#)) (Chapter 2), an open-source Python implementation for training GP surrogate models to sample posteriors for computationally expensive forward models.

We train a Gaussian process defined by a mean and covariance function. The covariance function $k(x, x')$ measures the degree to which the value of the function at one point x is correlated with the value at another point x' . The covariance function is modeled according to a kernel function, a positive semidefinite function of input points x and x' , that determines the smoothness and complexity of the GP. For an in-depth review, see [Rasmussen & Williams \(2006\)](#).

To emulate the posterior, we use a square exponential kernel,

$$k(x, x') = \exp\left(-\frac{(x - x')^2}{2l^2}\right), \quad (3.14)$$

where x and x' are two given input points, and l is a length scale hyperparameter. We assume that each input dimension has its own scale value. The mean and length scale hyperparameters of the Gaussian process are numerically optimized by the GP’s likelihood of the training data using `scipy.optimize` ([Jones et al., 2001](#)).

In the first step, we draw a training sample of $N = 1000$ evenly sampled training points using the [Sobol \(2001\)](#) pseudorandom algorithm. Next, using the procedure of [Kandasamy](#)

[et al. \(2017\)](#) we apply the Bayesian Active Learning for Posterior Estimation (BAPE) algorithm to iteratively sample high posterior regions using the GP surrogate model for $N = 2000$ more samples. We demonstrate the resulting posterior samples in the form of a corner plot in the following section.

4 Results

4.1 Model Sensitivity Analysis

We apply the global sensitivity analysis method described in Section 3.1 to simulations that combine stellar evolution, magnetic braking, and equilibrium tide. We use a quasi-monte carlo sampling scheme (Saltelli et al., 2010) to evenly sample $N = 4096$ simulations spanning the input parameter space. The ranges of input parameters we consider are given in Table 3.2.

4.1.1 Sensitivity of the orbital period, rotation period, and eccentricity

Figures 3.5 and 3.6 visualize the results of the sensitivity analysis applied to the full set of model parameters. Figure 3.5 shows the sensitivity analysis applied to the CTL model, and the following Figure 3.6 shows the CPL model. Within each of the figures, there are three grids displaying the sensitivity for three different model output parameters: final primary rotation period (left grid), final orbital period (center grid), and final orbital eccentricity (right grid). The x-axis for each subpanel shows the model input parameters that were sampled over the ranges listed in Table 3.2, and the y-axis of each subpanel shows the age to which the set of N simulations evolved (ranging from 10 Myr to 10 Gyr). Each value in the grid (denoted by color and labeled with numerical value) is the first-order sensitivity index (see Section 3.1) corresponding to a given input parameter at a specific age. Thus, we can interpret which output parameters are most sensitive to which input parameters as a function of age.

In Figure 3.5 our sensitivity analysis shows that the final P_{orb} and e values are dominated by initial orbital conditions for systems across all ages (as shown by the rows with highest variance). Picking out the regions of age/initial condition parameter space with the highest variance due to $\mathcal{Q}$ (or τ), elucidates the observables (final P_{rot} , P_{orb} , e) that are most sensitive to the tidal dissipation. The analysis in Figures 3.5 and 3.6 suggest similar results for both the CTL and CPL models. In particular, the primary P_{rot} of young systems $\sim 50 - 100$ Myr, or the eccentricity or secondary P_{rot} of old systems $\sim 5 - 10$ Gyr hold the most promise for constraining tidal $\mathcal{Q}$ (or τ), in other words, whether or not tides were strong enough to circularize and synchronize most initial conditions by that age. The results shown in Figures 3.5 and 3.6 suggest that the final rotation periods tend to be more sensitive to tidal $\mathcal{Q}$ than the final orbital period or eccentricity. However, given that rotation periods are difficult to measure precisely from lightcurves ($\sigma \sim 0.01$ d optimistically, Section 3.2), rotation periods generally have less constraining power than orbital periods or eccentricities.

4.1.2 Effects of observational uncertainties

In addition to looking at the sensitivities of individual parameters, we can also look at the combined constraint of using measured P_{rot} , P_{orb} , e and their uncertainties. Figure 3.7 shows how the combined likelihood (computed at fiducial parameters; see Table 3.3) is sensitive to each input parameter. The likelihood sensitivity tests how sensitive the model goodness-of-fit would be when combining the fit from all of the model observables and their uncertainties.

Figure 3.7 shows that when uncertainties are taken into account, the initial P_{orb} completely dominates the model goodness of fit across all ages. This implies that inference of tidal τ or Q is very dependent (and may be systematically biased) by the prior chosen for the initial orbital period, which we explore further in Section 4.2.

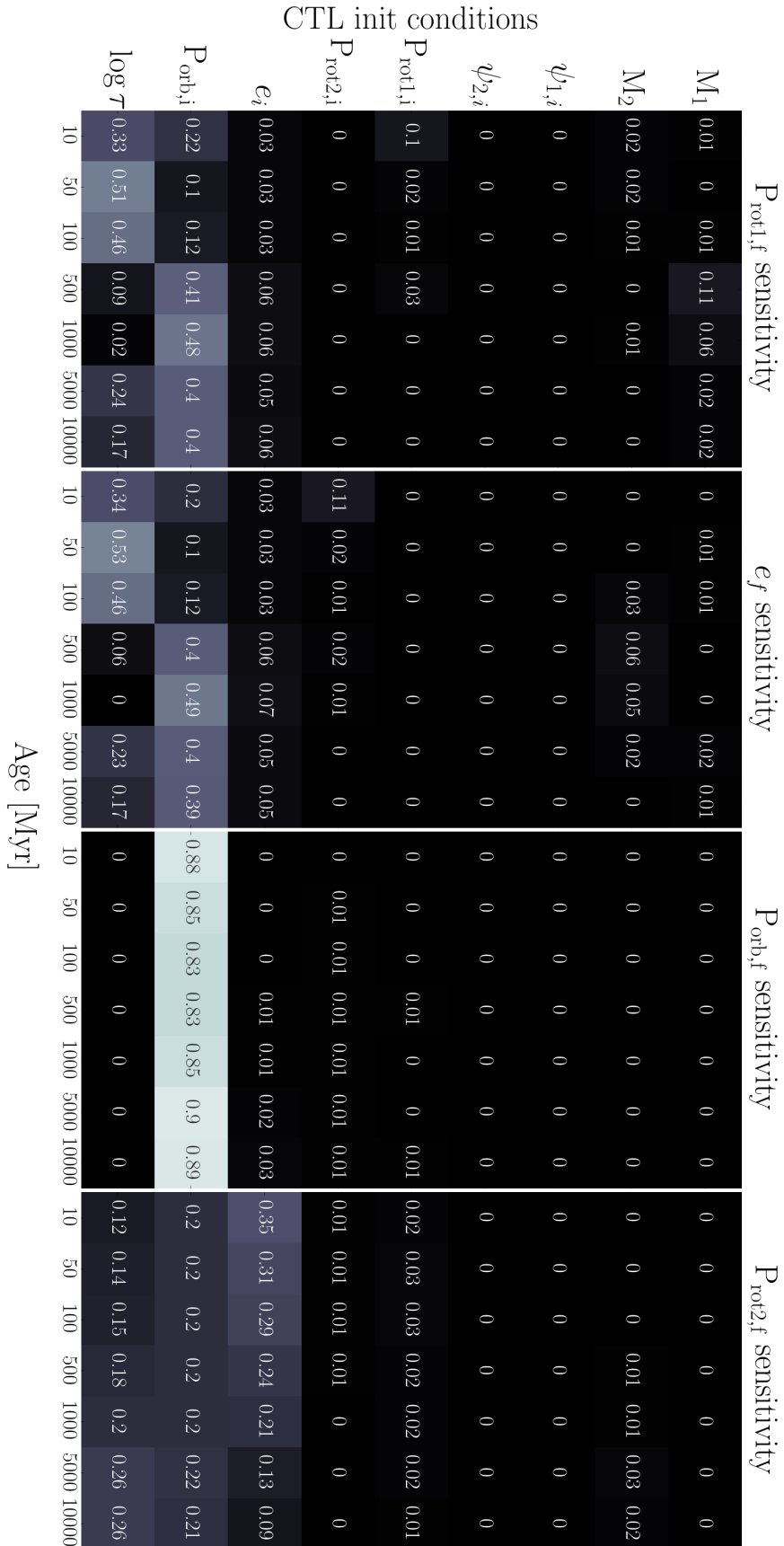


Figure 3.5: Global sensitivity analysis performed on a set of 4096 simulations with varying initial conditions (y-axis) and age (x-axis). Each panel quantifies the sensitivity of an output parameter (final P_{rot} , P_{orb} , e) in terms of the amount of variance each input parameter contributes to the distribution of final parameters.

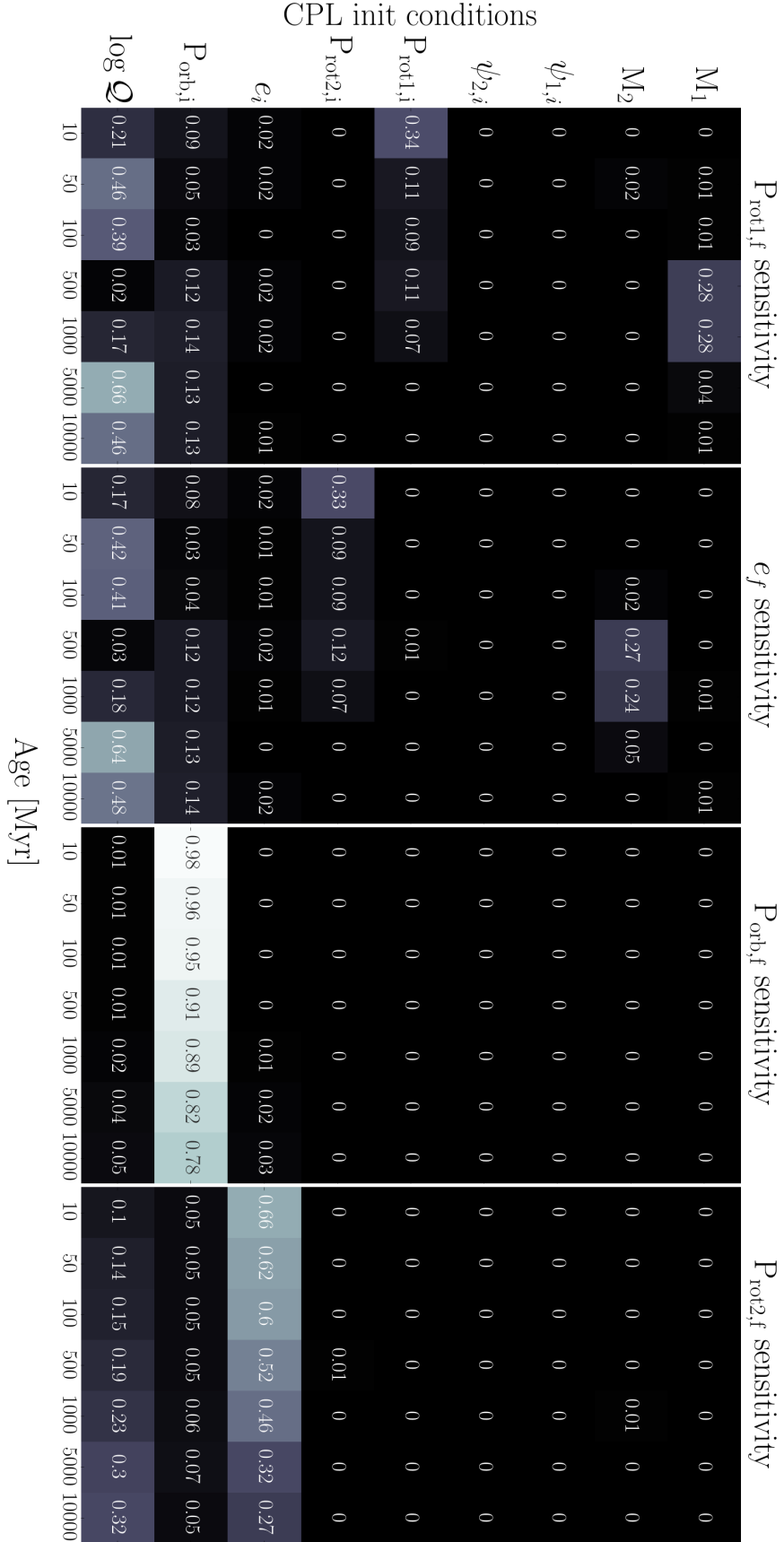


Figure 3.6: Global sensitivity analysis performed on a set of 4096 simulations with varying initial conditions (y-axis) and age of evolution (x-axis). Each panel quantifies the sensitivity an output parameter (final P_{rot} , P_{orb} , e) in terms of the amount of variance each input parameter contributes to the distribution of final parameters.

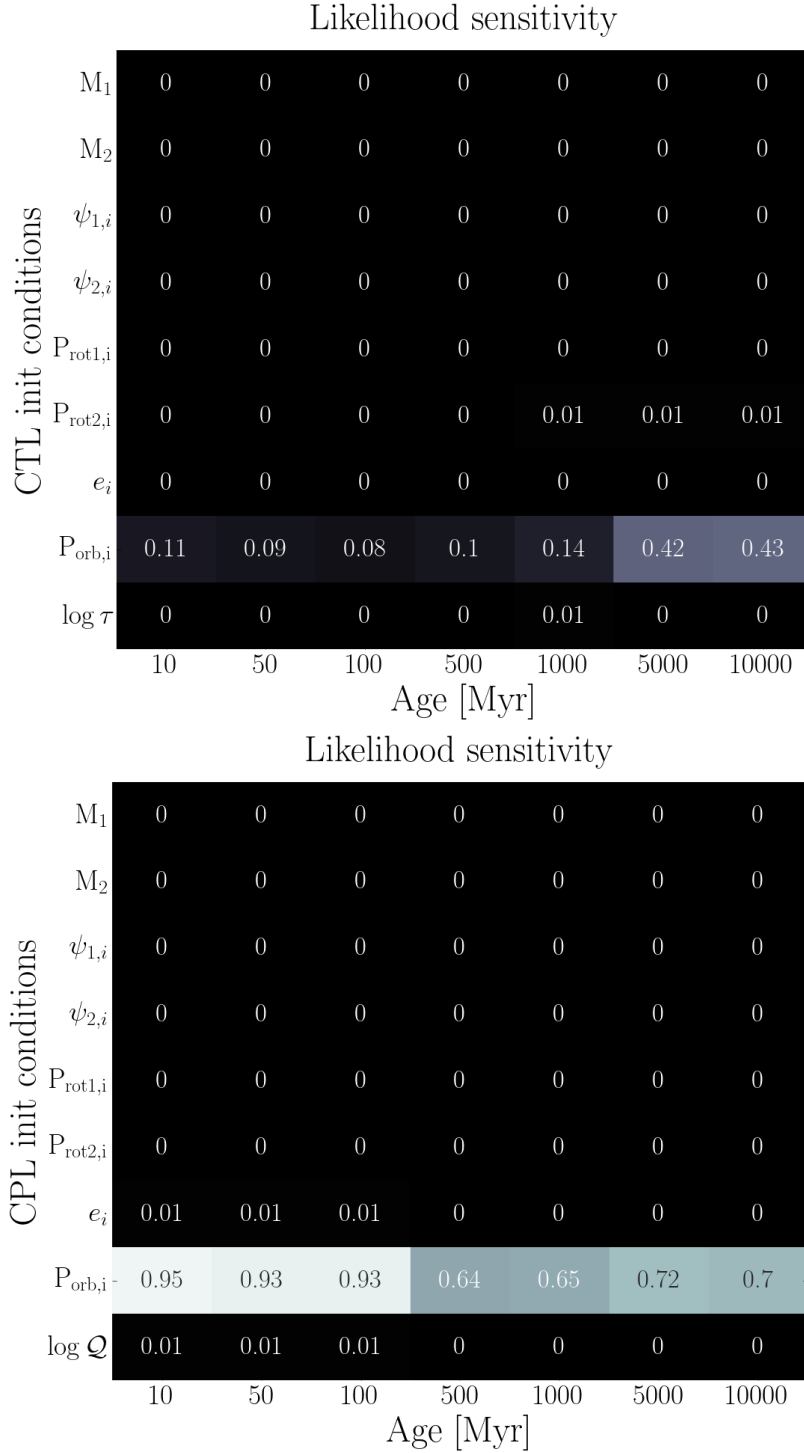


Figure 3.7: Likelihood sensitivity for the CTL (top) and CPL (bottom) model. The results for both models indicate that model goodness-of-fit (as quantified by a Gaussian likelihood, with uncertainties described in Section 3.2) is most sensitive to the initial orbital period.

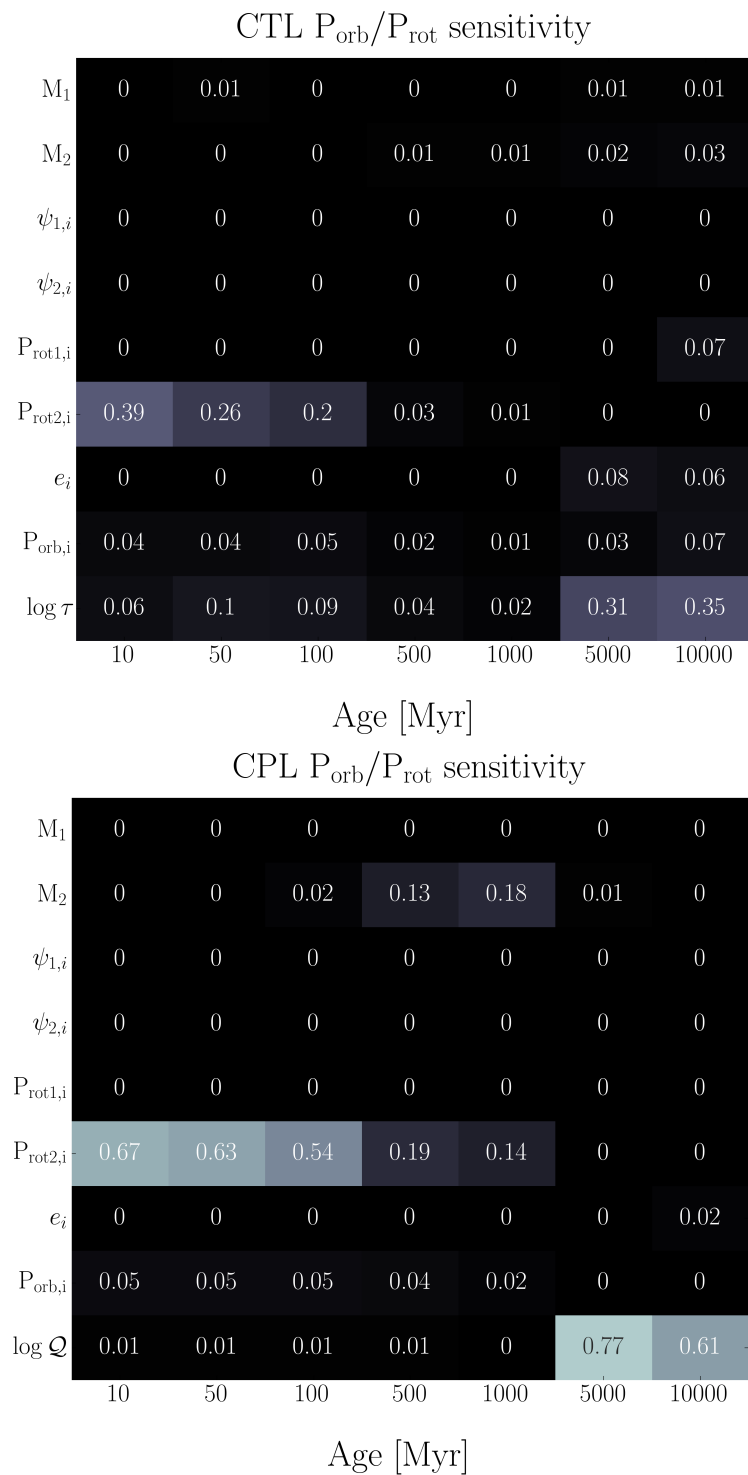


Figure 3.8: Sensitivity of the period ratio (orbital period / primary rotation period) for the CTL (top) and CPL (bottom) model. At ages of 5-10 Gyrs, τ and $\mathcal{Q}$ dominate the final period ratio $P_{\text{orb}}/P_{\text{rot1}}$.

4.1.3 Sensitivity of the period ratio

In summary, the final values of individual orbital parameters (Figures 3.5–3.6) and likelihood constructed from these parameters (Figure 3.7) are more dominated by the initial conditions than tidal τ or $\mathcal{Q}$. However, if we look at the sensitivity of the ratio $P_{\text{orb}}/P_{\text{rot}}$ as a function of evolution (Figure 3.8), we find that the sensitivity is less dependent on initial conditions at older ages. At ages of 5–10 Gyrs, the variance in the orbital period to rotation period ratio is dominated by τ and $\mathcal{Q}$ for the CTL and CPL models respectively.

Figure 3.9 examines the $P_{\text{orb}}/P_{\text{rot}}$ dependence on tidal parameters further. The blue scatterpoints show the initial distribution of parameters (with uniformly sampled initial conditions) and the black contours show the final distribution of parameters evolved to 5 Gyrs (where the period ratio is most dependent on tidal parameters, based on Figure 3.8). In Figure 3.9, the 5 Gyr simulations converge towards a bimodal distribution: one cluster with synchronized systems (highlighted in orange), and another cluster of subsynchronous, eccentric binaries. For the CTL model, there is a fairly clear separation where simulations with $\log \tau > 0$ are nearly all tidally locked. Similarly for the CPL model, most systems with $\log \mathcal{Q} < 6$ are tidally locked. For the range of weaker tides ($\log \tau < 0$, or $\log \mathcal{Q} > 6$) there is a higher variance in period ratio where some systems are synchronized, but a larger fraction of systems are subsynchronous.

Figures 3.10 and 3.11 investigate whether the bimodality between the subsynchronous and synchronous populations is distinguishable as a function of $\log \tau$ and $\log \mathcal{Q}$ respectively. Figure 3.10 shows the histograms of the sample categorized into three different populations: synchronized ($P_{\text{orb}}/P_{\text{rot1}} \approx 1$, orange), subsynchronous ($P_{\text{orb}}/P_{\text{rot1}} < 1$, blue) and supersynchronous ($P_{\text{orb}}/P_{\text{rot1}} > 1$, green), each at three different ages: 1 Gyr, 5 Gyrs, and 10 Gyrs. The supersynchronous population clusters near $P_{\text{orb}}/P_{\text{rot}} \approx 1.5$, consistent with the pseudo-synchronous equilibrium ratio $\omega/n = N(e)/\Omega(e)$ predicted by the CPL model for eccentric systems.

First we can note the importance of age on the synchronization fraction. The distributions evolved for 5 Gyrs closely overlap with distributions evolved for 10 Gyrs, but diverge with the distributions evolved for 1 Gyr, showing that most systems will have reached their final synchronization state within 5 Gyrs. This means that if we have a population of binaries that are roughly older than 5 Gyrs, we can place limits on $\log \tau$ based on the fraction of systems synchronized, without having to have precise age estimates for each of the individual systems, which is promising, as precise ages are observationally challenging for old systems Soderblom (2010).

Next we can look at the importance of $\log \tau$ on the synchronous fraction. Most importantly, we want to understand the limits where $\log \tau$ dominates over the effects of the initial conditions. As shown in Figure 3.10, after 5 Gyrs, nearly all orbits ($> 90\%$ marked by the purple dashed line) with $\log \tau \gtrsim 0$ are expected to become synchronized (shaded orange region), regardless of initial orbital configuration. On the opposite side (not shown on the bounds of the x-axis of the plot), it appears that $\log \tau \lesssim -4$ is roughly the lower limit where we expect the majority ($> 90\%$) of systems to be synchronized. Similarly for the CPL model in Figure 3.11, $\log \mathcal{Q} \lesssim 5.5$ would be the lower limit given a majority fraction of the population are subsynchronized, and $\log \mathcal{Q} \gtrsim 9$ would be the upper limit given a majority are synchronized. This observationally implies that we could measure the fraction

of synchronous and subsynchronous systems among populations of old (> 5 Gyr) systems: if a strong majority of systems are synchronized, then we could determine a lower bound on $\log \tau$ (or upper bound on $\log \mathcal{Q}$), and if a strong majority are subsynchronized, then we could determine an upper bound on $\log \tau$ (or lower bound on $\log \mathcal{Q}$).

We note that an upper or lower bound constraint would still leave several orders of magnitude of uncertainty in the tidal parameters, and would be roughly consistent with the range of constraints currently presented in literature (Meibom et al., 2006; Jackson et al., 2009; Hansen, 2010; Patel & Penev, 2022; Penev & Schussler, 2022). Gaining more precise constraints (within the middle regions $-4 \lesssim \log \tau \lesssim 0$, or $5.5 \lesssim \mathcal{Q} \lesssim 9$) becomes more difficult, however. Figures 3.10 and 3.11 suggest that the relative fractions of synchronous and subsynchronous systems vary consistently as a function of tidal parameter, and that these fractions could be compared to the fractions measured from observations to constrain the tidal parameters more precisely. These fractions are the result of using uniform priors for the initial conditions (Table 3.2). Further work would need to explore how different (possibly more realistic) distributions of initial conditions might change the synchronization fraction as a function of tidal parameter, and very careful observational work would need to be done obtain an unbiased sample with measured orbital periods and rotational periods.

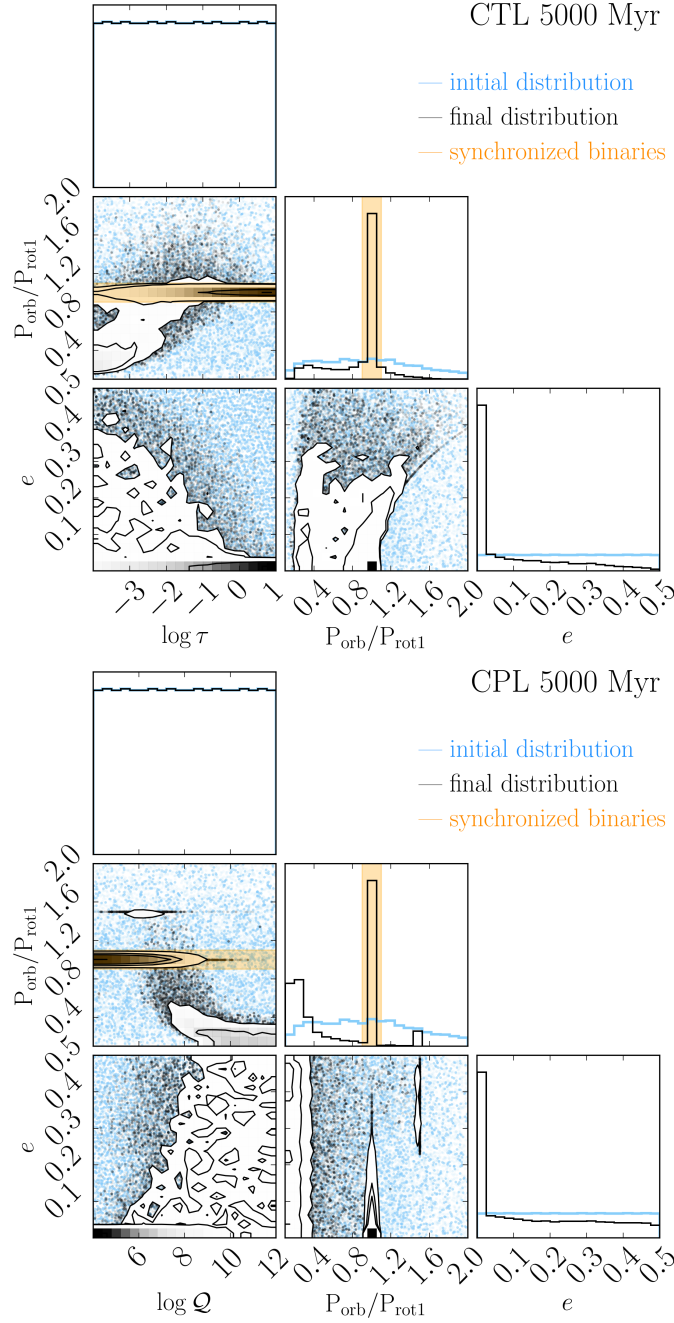


Figure 3.9: Distribution of orbital parameters for the CTL model (top) and CPL model (bottom) evolved to an age of 5 Gyr, taken from the same simulations as Figures 3.5–3.8. The blue scatterpoints show the initial distribution of parameters (sampled uniformly in eccentricity, orbital period, rotation period, and tidal τ or Q), the black contours show the final distribution of parameters after 5 Gyrs of evolution, and the orange range highlights synchronized binaries. At 5 Gyr, there is a clear bimodal separation between systems that are synchronized ($P_{\text{orb}}/P_{\text{rot1}} \approx 1$), and those that are subsynchronous ($P_{\text{orb}}/P_{\text{rot1}} < 1$).

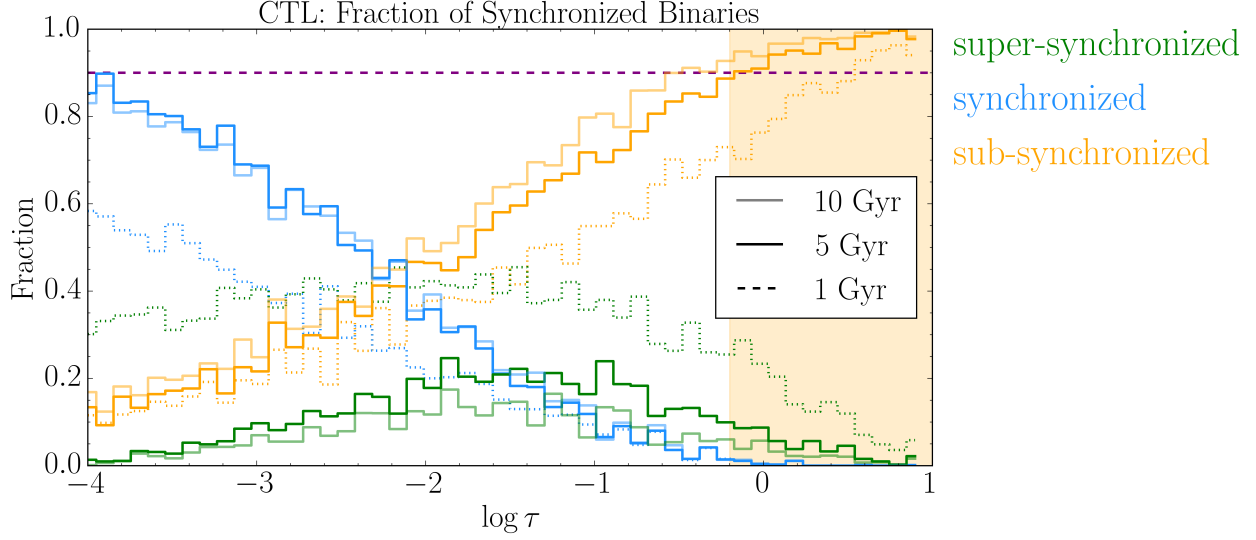


Figure 3.10: Fraction of binaries that are synchronized ($P_{\text{orb}}/P_{\text{rot1}} \approx 1$, orange), subsynchronous ($P_{\text{orb}}/P_{\text{rot1}} < 1$, blue) and supersynchronous ($P_{\text{orb}}/P_{\text{rot1}} > 1$, green) according to the CTL model. The dotted lines show the systems evolved to an age of 1 Gyr, the dark solid lines show 5 Gyr, and the light solid lines show 10 Gyr. For each bin of $\log \tau$, we compute relative fractions of synchronous, subsynchronous, and supersynchronous in each sample for a given age. The shaded orange region shows hypothetical the lower bound on $\log Q$ given that a majority of sources ($> 90\%$) are observed subsynchronous after 5 Gyr.

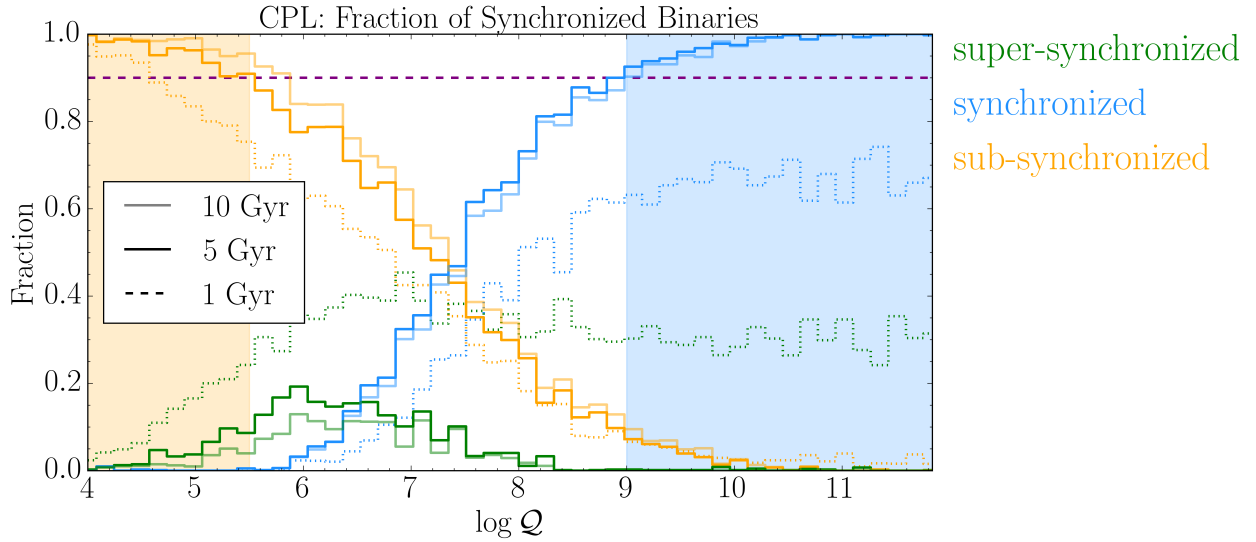


Figure 3.11: Same as Figure 3.10, but for the CPL model. The shaded orange region shows hypothetical the lower bound on $\log Q$ given that a majority of sources ($> 90\%$) are observed subsynchronous after 5 Gyr. The shaded blue region shows hypothetical the upper bound $\log Q$ given that a majority of sources ($> 90\%$) are subsynchronous after 5 Gyr.

4.2 Simulated Posteriors

Section 4.1 shows that, to first order, constraining tidal $\mathcal{Q}$ or τ is not very promising: the final orbital states are more sensitive to the initial orbital and rotational states than they are to the effects of the tidal dissipation efficiency. However, looking at Figures 3.5 and 3.6, we do see that tidal $\mathcal{Q}$ or τ accounts for a significant amount of variance in the final rotation periods and eccentricities (up to ~ 0.5 at some ages). For the sake of completeness, we further examine the covariance structure based on simulated posteriors and attempt to rigorously test whether informative constraints on $\mathcal{Q}$ or τ could be teased out from higher-order effects in the correlations between variables.

We generate 5-dimensional posteriors according to the prescription described in Section 3.2. In this test we allow for uninformed priors for five parameters (the initial states $P_{\text{rot1,i}}$, $P_{\text{rot2,i}}$, e_i , $P_{\text{orb,i}}$, and $\mathcal{Q}$) given that the initial states of a system are generally unknown. To isolate the effects of the other input parameters of the model, we fix the masses, initial obliquities, and ages to the true value. We performed a total of four simulated posterior tests encompassing both tidal models at young and old ages. These tests include: CTL at 50 Myr (Figure 3.12), CTL at 5 Gyr (Figure 3.13), CPL at 50 Myr (Figure 3.14), and CTL at 5 Gyr (Figure 3.15). We visualize the posteriors in the form of a corner plot, in which samples are colored according to their posterior probability, where darker blue represents higher probability values.

In simulations at young ages, the results are qualitatively similar for both CTL and CPL (Figures 3.12 and 3.14 respectively). Here, we see that the high posterior regions most strongly depend on the initial orbital period, which is consistent with the sensitivity analysis results. We also see that the high posterior regions are nearly flat in the previous range of $\mathcal{Q}$ (or τ). The reason is that at only a few Myrs of evolution, we expect hardly any evolution, even for stronger tides. Given that orbital periods are generally measured to the best level of uncertainty (Table 3.3), we find that sets of values that match around $P_{\text{orb,i}} \approx P_{\text{orb,f}}$ with weak values of $\mathcal{Q}$ generally tend to maximize the posterior. Our results suggest that we may expect young systems to provide a lower bound on $\mathcal{Q}$ (or an upper bound on τ). However even with optimistic uncertainties, this degeneracy between $P_{\text{orb,i}}$ and $\mathcal{Q}$ is poorly traced out.

The overfitting for initial orbital period is a trend that occurs not only for young ages, but persists for inference at old ages as well. The results for the CTL and CPL model for a system of 5 Gyr age are shown in Figures 3.13 and 3.15, respectively. At higher ages we see two degeneracy features emerge between the initial orbital period and initial eccentricity for both CTL and CPL models. To show this structure in more detail, Figure 3.16 shows a zoom-in for the marginal posterior of $(P_{\text{orb,i}}, e_i)$ for the CPL model. Here, we see that two main degeneracy features appear in the posterior points sampled. One of the features is similar to the young age posterior, which is flat across eccentricity with $P_{\text{orb,i}} \approx P_{\text{orb,f}}$, and corresponds to weak tidal dissipation (high $\mathcal{Q}$ or low τ). The other main feature shows a curved $P_{\text{orb,i}} - e_i$ degeneracy, which corresponds to stronger tidal dissipation (low $\mathcal{Q}$ or high τ). The curved degeneracy intersects with the true initial orbital period and eccentricity. However, the true values are statistically indistinguishable from other samples along the degeneracy.

CTL simulated posterior (50 Myr)

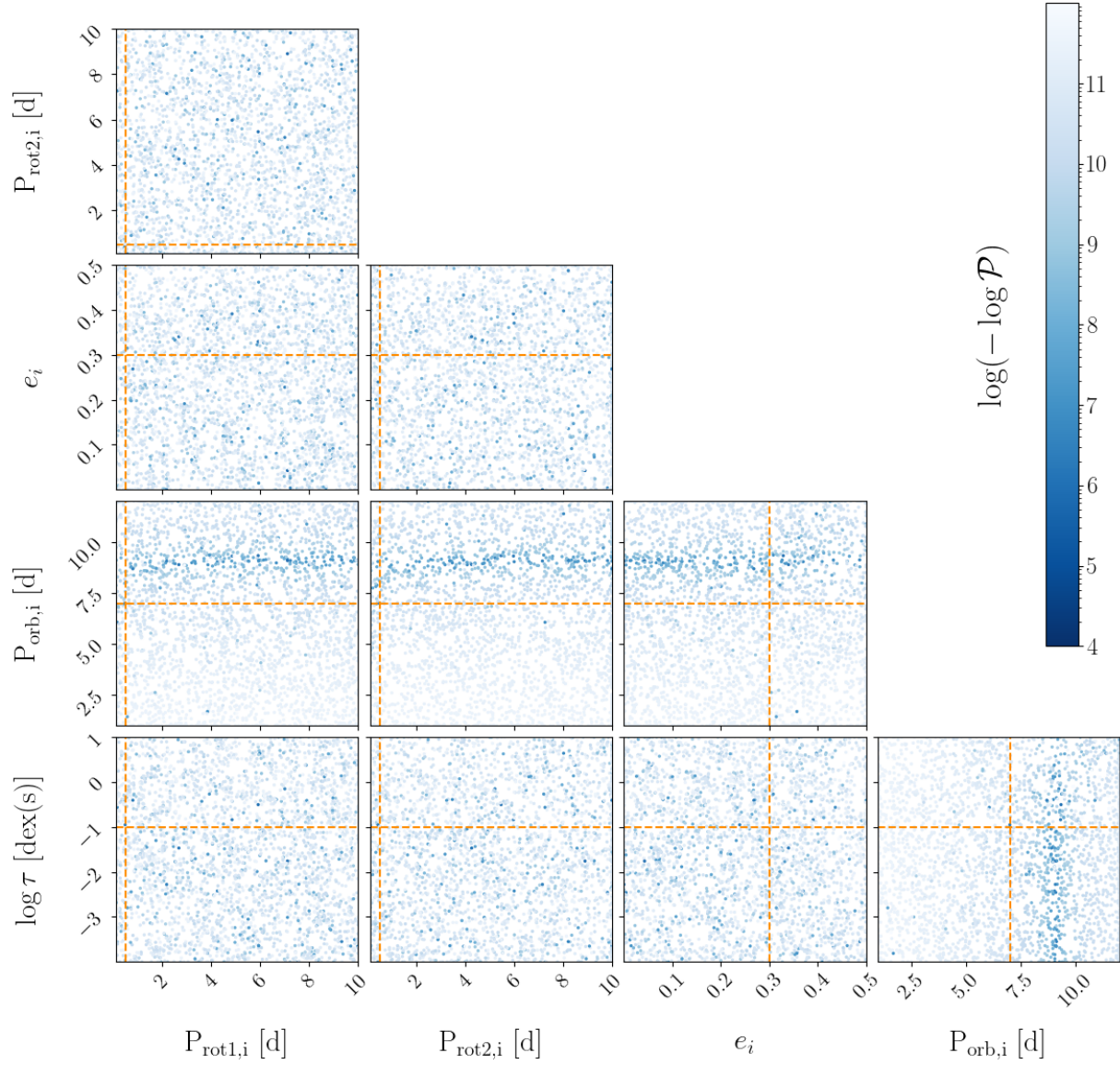


Figure 3.12: Samples from the simulated posterior for the CTL model at an early age of 50 Myr (blue points). The posterior function is sampled using a Gaussian process and active learning (Kandasamy et al., 2017) using the package `alabi` (Birky et al., in prep.) in order to visualize the structure of high posterior modes/degeneracies. Points are colored according to their posterior values. Since the posterior spans an order of magnitude range, we present these values as $\hat{P} = \log(-\log \mathcal{P})$, where lower values of $\hat{P}$ (darker blue) represent higher posterior probability. Orange dashed lines mark the true values of the initial parameters.

CTL simulated posterior (5 Gyr)

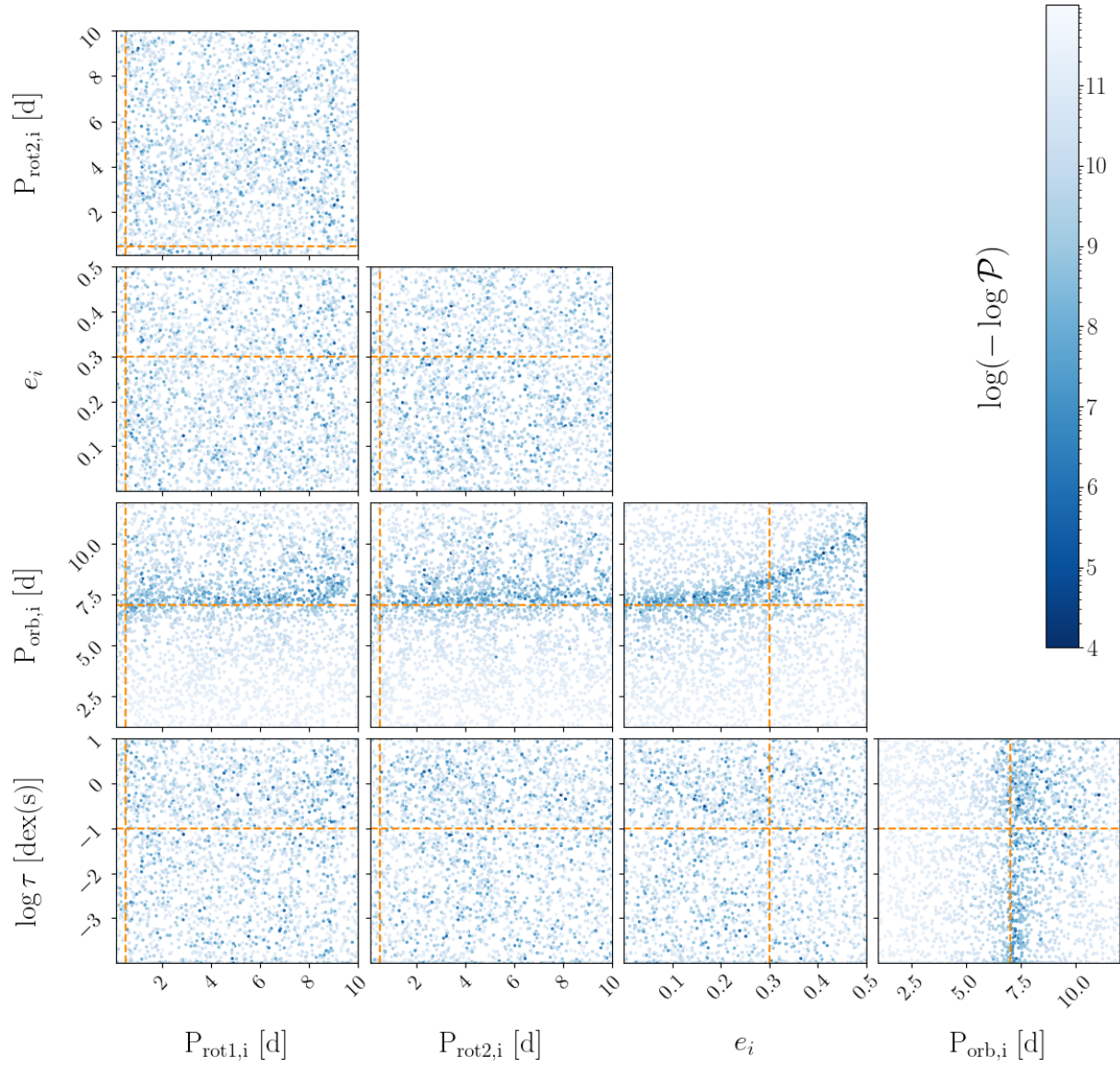


Figure 3.13: Same as Figure 3.12, but for 5 Gyr.

CPL simulated posterior (50 Myr)

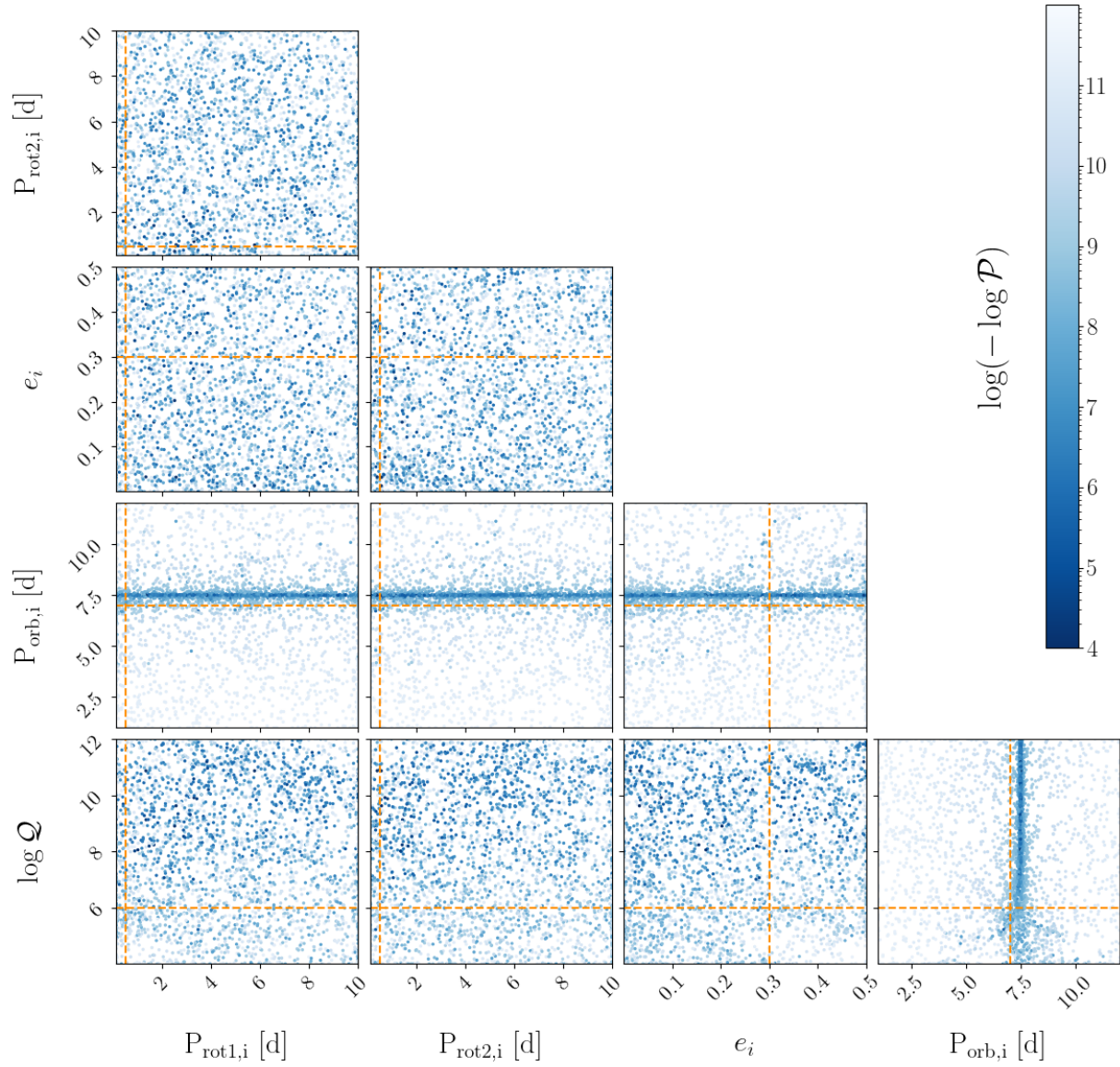


Figure 3.14: Same as Figure 3.12, but for the CPL model.

CPL simulated posterior (5 Gyr)

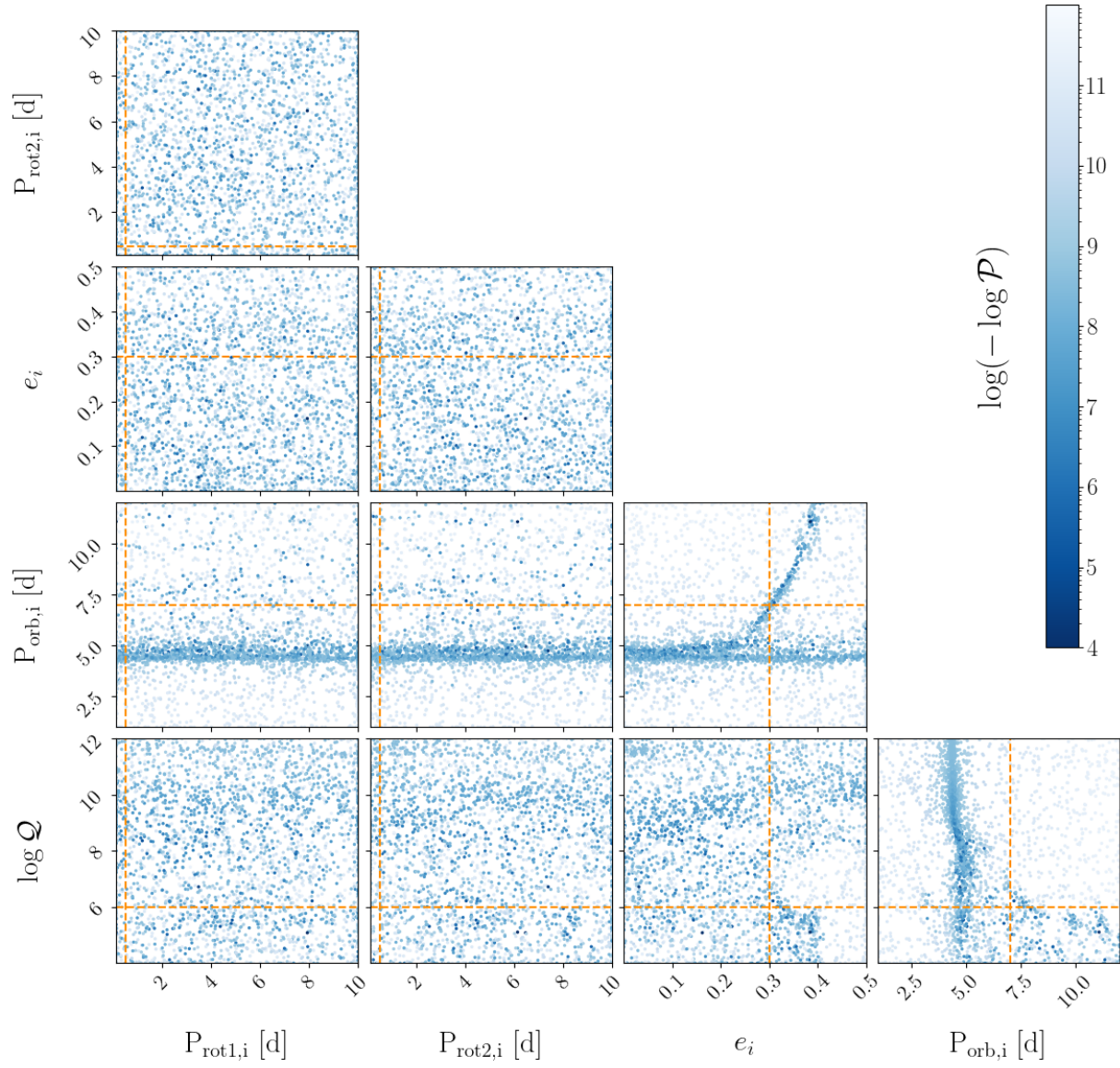


Figure 3.15: Same as Figure 3.14, but for 5 Gyr.

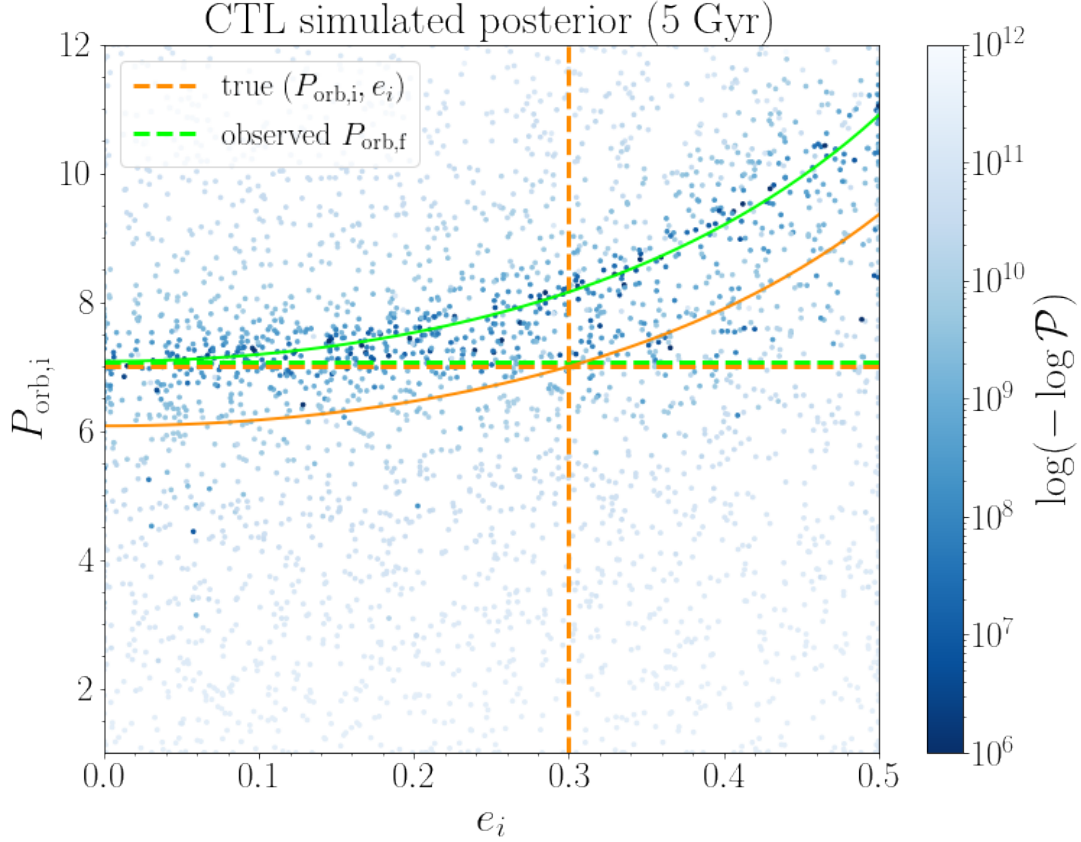


Figure 3.16: Samples from the simulated posterior for the CTL model at a late age of 5 Gyr (from Figure 3.13) zoomed in on the marginal distribution of initial orbital period and eccentricity. The dashed orange line shows the true initial orbital period that would be unknown, and the green dashed line shows the final orbital period that would be observed. The solid lines show the curve of $P_{\text{orb},i}$ and e_i for a constant value of orbital angular momentum. The orange solid curve shows the initial orbital angular momentum ($J_{\text{orb},i} = 1.80 \times 10^{45} \text{ kg} \cdot \text{m}^{-2}$), and the green solid line shows the final orbital angular momentum ($J_{\text{orb},f} = 1.61 \times 10^{45} \text{ kg} \cdot \text{m}^{-2}$). Instead of converging to the true initial orbital period, the posterior is biased towards picking solutions with $P_{\text{orb},i} \approx P_{\text{orb},f}$ (as seen by the overlapping horizontal dashed lines). Furthermore, we see that the high posterior samples lie along the curve of final orbital angular momentum, meaning that the posterior is more informative of the final conditions than it is of the initial conditions that we are trying to infer.

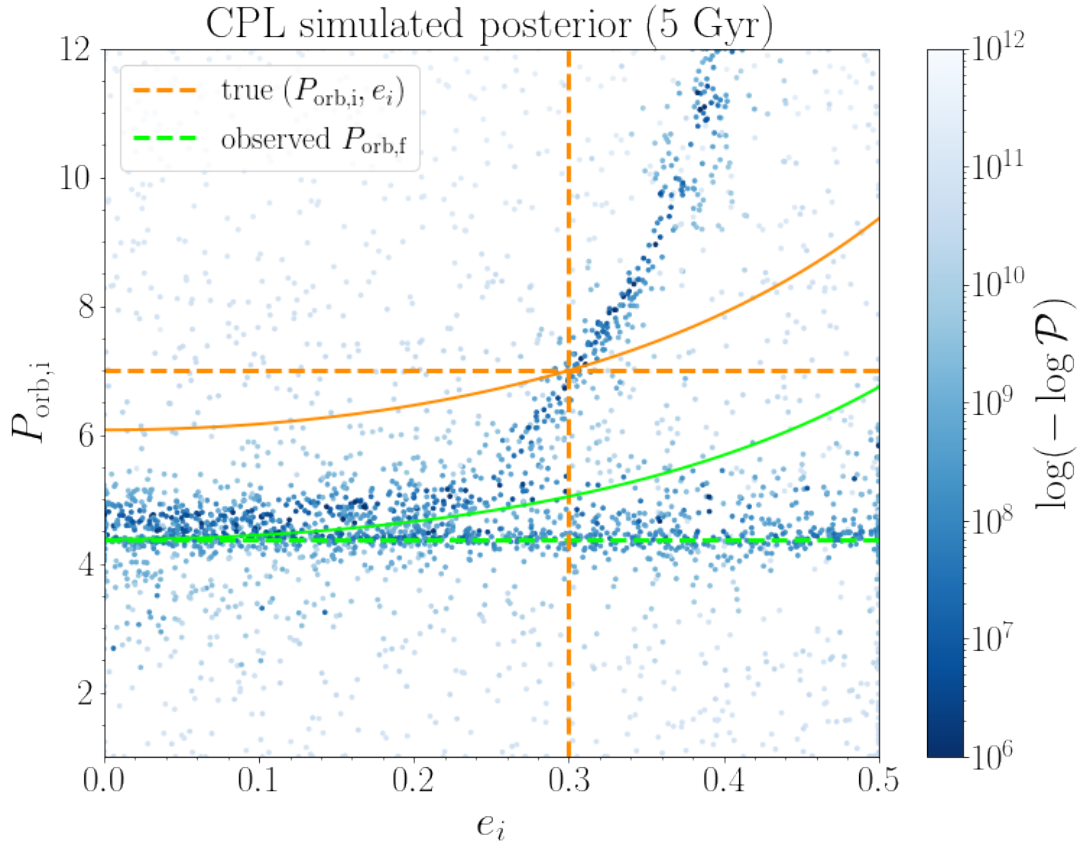


Figure 3.17: Same as Figure 3.16, but for the CPL model. Two degeneracy features emerge from the blue sampled posterior points: those flat across eccentricity with $P_{\text{orb},i} \approx P_{\text{orb},f}$ (with negligible values of tidal dissipation), and those with curved $P_{\text{orb},i} - e_i$ degeneracy (with stronger values of tidal dissipation).

4.3 Degeneracies in the most simplified simulated model

Finally to test the limitations of inference, we perform a likelihood recovery test by considering the most optimistic inference case possible: 1 free parameter ($\mathcal{Q}$ or τ), using simulated data with ideal uncertainties, see Table 3.3. This test is set up as follows: we run a simulation with fiducial initial values to compute a set of final values. This procedure is very similar to Section 4.2, but instead of five free parameters, we only vary one. The values of the fixed parameters are given in Table 3.4.

Parameter	M_1, M_2	ψ_{1i}, ψ_{2i}	$P_{\text{rot1},i}, P_{\text{rot2},i}$	$P_{\text{orb},i}$	e_i
Unit	$M_\odot$	deg	days	days	
Fiducial value	1.0	0	0.5	7.0	0.15

Table 3.4: Values of input parameters used in the 1-dimensional inference test. The subscript i an initial value of a time-varying parameter.

In Figures 3.18–3.20 we show the simulated 1-dimensional likelihood for the CTL and CPL models at three different ages: 50 Myr, 500 Myr, and 5 Gyr. The CTL model is shown on the left panels, and the CPL model is shown on the right panels. The true values of $\log \tau$ ($-3, -2, -1, 0, +1$) and $\log \mathcal{Q}$ ($4, 5, 6, 7, 8$) are indicated with dashed vertical lines. The likelihood values as a function of $\log \tau$ or $\log \mathcal{Q}$ are represented by distinct colored lines. For both CTL and CPL, the blue lines represent weaker tides and the red lines represent stronger tides (on the same color scale as Figures 3.1 and 3.2).

A single peak in a given curve represents a unimodal solution in which the true value of τ or $\mathcal{Q}$ could theoretically be recovered if τ or $\mathcal{Q}$ were the only free variable. However, degeneracies do occur for both the CTL and CPL models. The origin of these double peaks can be understood when we examine the evolution trajectories. In particular (as we noted in Section 4.1), the constraint on orbital period dominates the likelihood. In Figure 3.1 (showing the trajectories for varying τ) we see that there is a region in orbital period and eccentricity where solutions with different τ overlap. Similarly (but to a lesser extent), we also see in Figure 3.2 that there’s a region in the orbital period evolution of the CPL model where solutions with different $\mathcal{Q}$ overlap.

Figure 3.21 illustrates the degenerate solutions for $\log \tau = 0$: at 50 Myr the likelihood has peaks at $\log \tau = 0$ and -1.73 ; at age 500 Myr the likelihood has peaks at $\log \tau = 0$ and -2.17 ; and at age 5 Gyr the likelihood has peaks at $\log \tau = 0$ and 0.66 . The black lines in Figure 3.21 show the solution for $\log \tau = 0$, and the dashed lines show the solutions of the secondary peaks. Points where the solutions overlap are highlighted with dots. In the left panel of Figure 3.21 (orbital period), there are intersections with $\log \tau = 0$ at each of the ages, but in the other two panels (rotation period and eccentricity), there is only an intersection with $\log \tau = 0$ at 5 Gyr. Thus, we see that the orbital period dominates the likelihood constraint, and the double-peaked degeneracies seen in the CTL + STELLAR likelihood (Figures 3.18–3.20) occur when there are degeneracies in orbital period for different $\log \tau$ values.

We can gain further insight into the structure of the likelihood plots by examining the τ values where tidal locking, synchronization, and circularization happen. Tidal locking refers

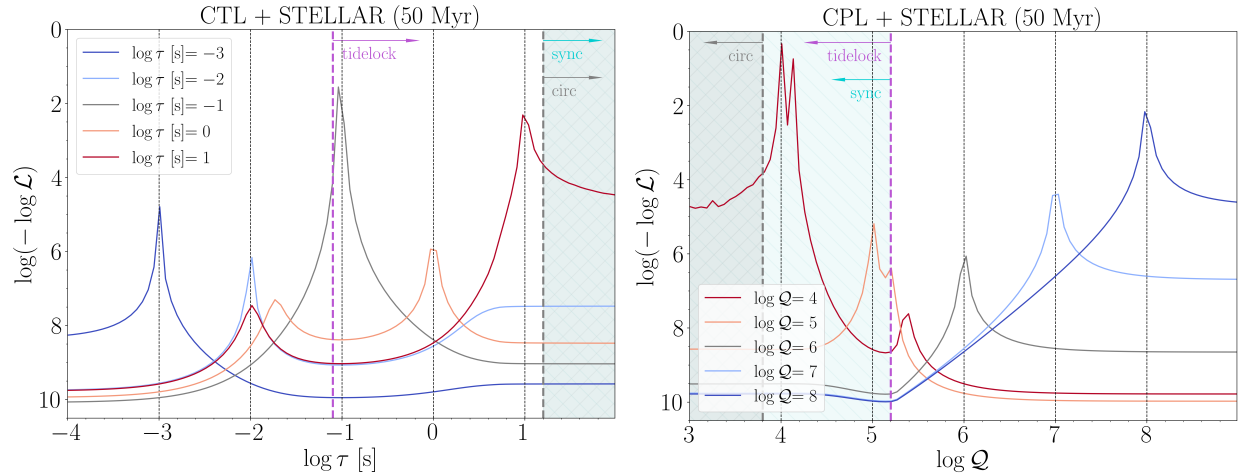


Figure 3.18: Simulated 1-dimensional likelihood for CTL model (left) and CPL model (right) evaluated at an age of 50 Myrs. Each of the colored lines shows the likelihood value as a function of varying τ or Q for different “true” values ($-3, -2, -1, 0, +1$) for τ , and ($4, 5, 6, 7, 8$) for Q . For both plots, blue lines represent weaker tides and red lines represent stronger tides. The y-axis shows the likelihood value when τ is allowed to vary and is compared to the fiducial simulation (with the values of Table 3.4, and uncertainties of Table 3.3). On the y-axis we compress the magnitude range for visualization by showing the likelihood as $\log(-\log \mathcal{L})$, where lower values of $\log(-\log \mathcal{L})$ (higher on the y-axis) represent higher likelihood values (similarly to Figures 3.12–3.15). The magenta dashed line and arrows indicate the range of tidal values that are strong enough to cause tidal locking by an age of 50 Myr. Similarly, the cyan arrows and shaded regions show the range of tidal values that are strong enough to cause synchronization, and the grey arrows and shaded regions show the range of tidal values that are strong enough to cause circularization.

to the point in the evolution at which there is no net tidal torque acting on the star. Synchronization means that the rotation period and orbital period are identical. Circularization occurs when the eccentricity reaches a value of 0. For the CTL model, tidal locking occurs first, followed by synchronization and circularization, which occur at the same time. For the CPL model, however, tidal locking and synchronization occur first at the same time, and circularization follows later.

Table 3.5 gives the values of τ or Q for tidal locking, synchronization, and circularization for each age in the likelihood plots (Figures 3.18–3.20). We also see in Figures 3.18–3.20 that there tends to be a symmetry to the degenerate peaks, where one solution occurs when the orbital period is rising, while the other solution occurs when the orbital period is decaying. This symmetry point is roughly the $\log \tau$ at which the stars become tidally locked. This “turn-over” in orbital period (as discussed in Fleming et al. 2018) happens when a system evolves from an initial state where the rotation period is faster than the orbital period. Assuming that the initial rotation periods of stars in a binary system is around the same distribution as the rotation periods of young single stars, it is reasonable to expect fast initial rotation periods of $P_{\text{rot}} \lesssim 1$ day (Stassun et al., 1999; Rebull et al., 2006; Marilli et al., 2007).

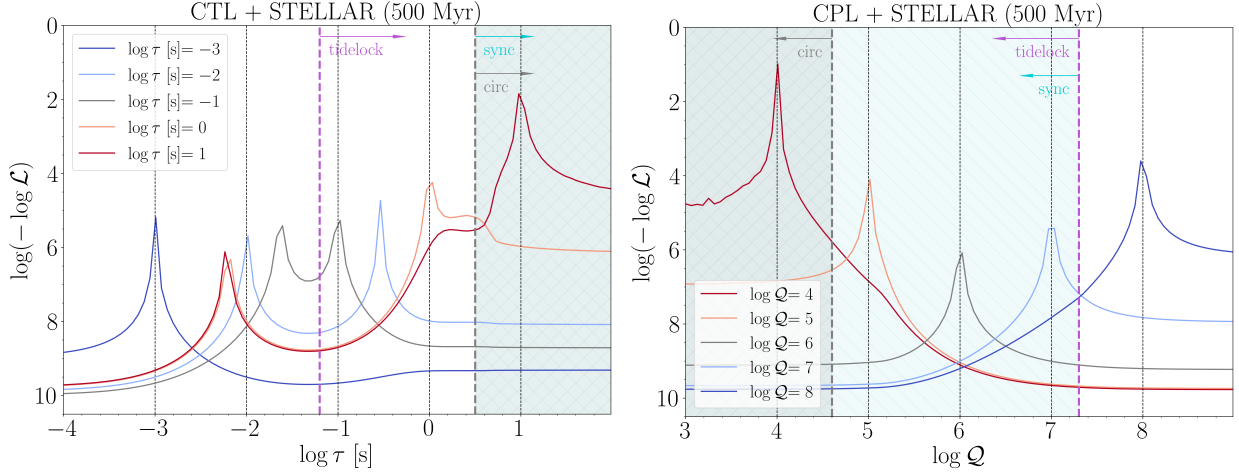


Figure 3.19: Same as Figure 3.18, but for ages 500 Myr.

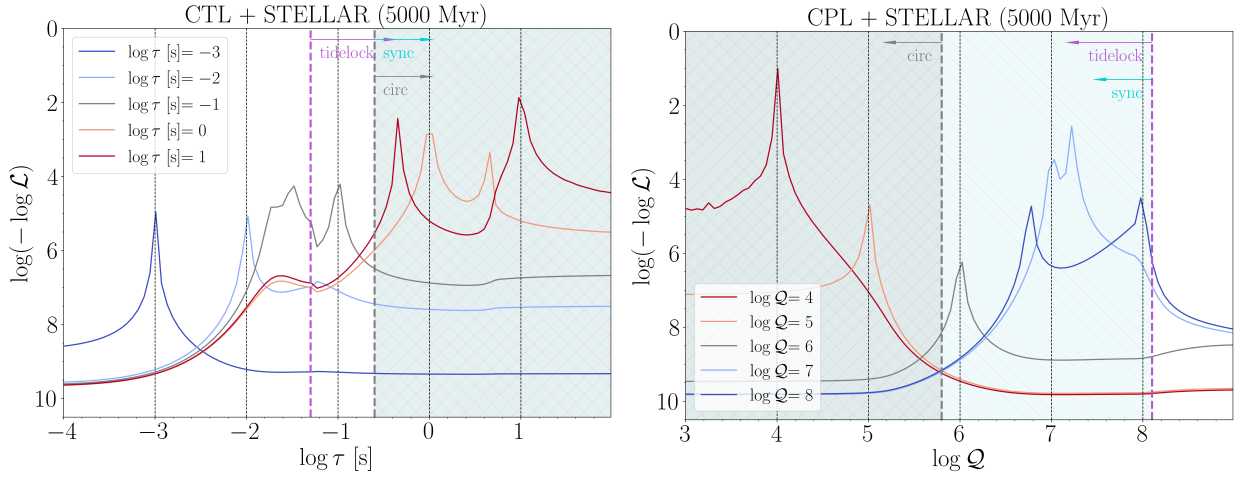


Figure 3.20: Same as Figure 3.18, but for 5 Gyr.

In the case of initial $P_{\text{rot}} < P_{\text{orb}}$, tidal locking would cause angular momentum to be transferred from the rotation period to the orbit, until the two approach synchronization. At the point of tidal locking, the magnetic braking acting to slow the rotation of the star is balanced by tidal forces that speed up the rotation of the star (Section 2.2). The angular momentum lost due to magnetic braking must come from the angular momentum of the orbit, causing orbital decay. Thus, we find that coupling magnetic braking with the constant time lag model results in a “turn-over” in orbital period evolution (Figure 3.21), which results in degenerate solutions (Figures 3.18–3.20).

CTL+STELLAR			
Age	tidelock	sync	circ
50 Myr	$\tau > -1.1$	$\tau > 1.2$	$\tau > 1.2$
500 Myr	$\tau > -1.2$	$\tau > 0.5$	$\tau > 0.5$
5000 Myr	$\tau > -1.3$	$\tau > -0.6$	$\tau > -0.6$
CPL+STELLAR			
Age	tidelock	sync	circ
50 Myr	$\mathcal{Q} < 5.2$	$\mathcal{Q} < 5.2$	$\mathcal{Q} < 3.8$
500 Myr	$\mathcal{Q} < 7.3$	$\mathcal{Q} < 7.3$	$\mathcal{Q} < 4.6$
5000 Myr	$\mathcal{Q} < 8.1$	$\mathcal{Q} < 8.1$	$\mathcal{Q} < 5.8$

Table 3.5: Range of tidal values in which tides are strong enough to cause tidal locking, synchronization, or circularization at a given age. In Figures 3.18–3.20, the tidelock range is shown in magenta, the synchronization range is shown in cyan, and the circularization range is shown in gray.

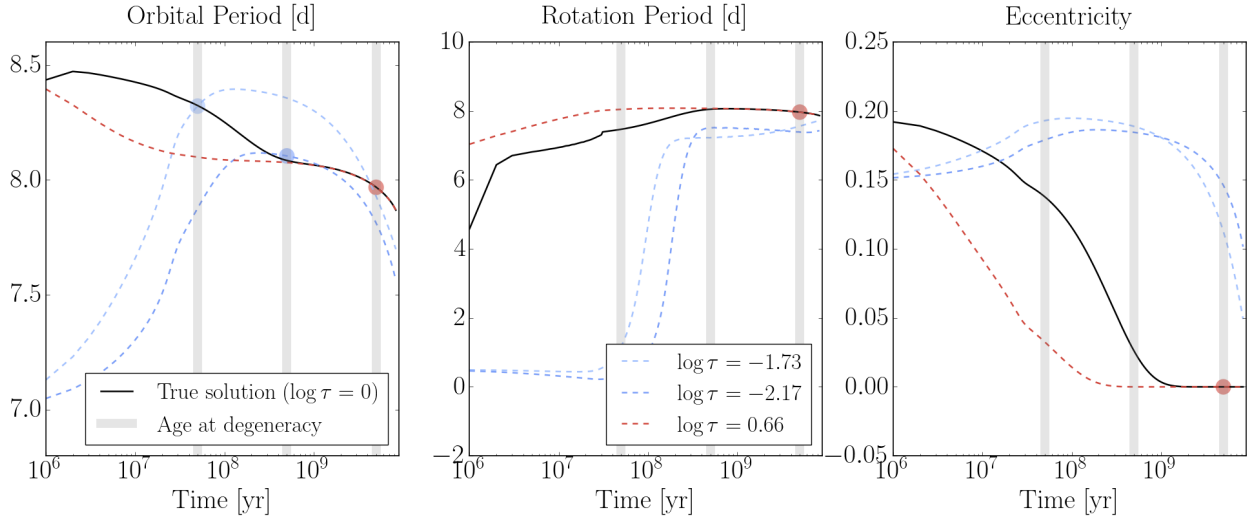


Figure 3.21: Examples showing the how solutions with different $\log \tau$ are degenerate at different ages. The black line highlights the solution for $\log \tau = 0$. The thick grey vertical lines mark the ages 50 Myr, 500 Myr, and 5 Gyr as used in Figures 3.18–3.20. The colored dashed lines show the solutions for the secondary likelihood peak at each age: at 50 Myr the likelihood has peaks at $\log \tau = 0$ and -1.73 ; at 500 Myr the likelihood has peaks at $\log \tau = 0$ and -2.17 ; and at 5 Gyr the likelihood has peaks at $\log \tau = 0$ and 0.66 . The color scale is consistent with Figures 3.1–3.2 and 3.18–3.20, where darker red represents stronger tides and darker blue represents weaker tides. The colored dots highlight where the evolutions intersect.

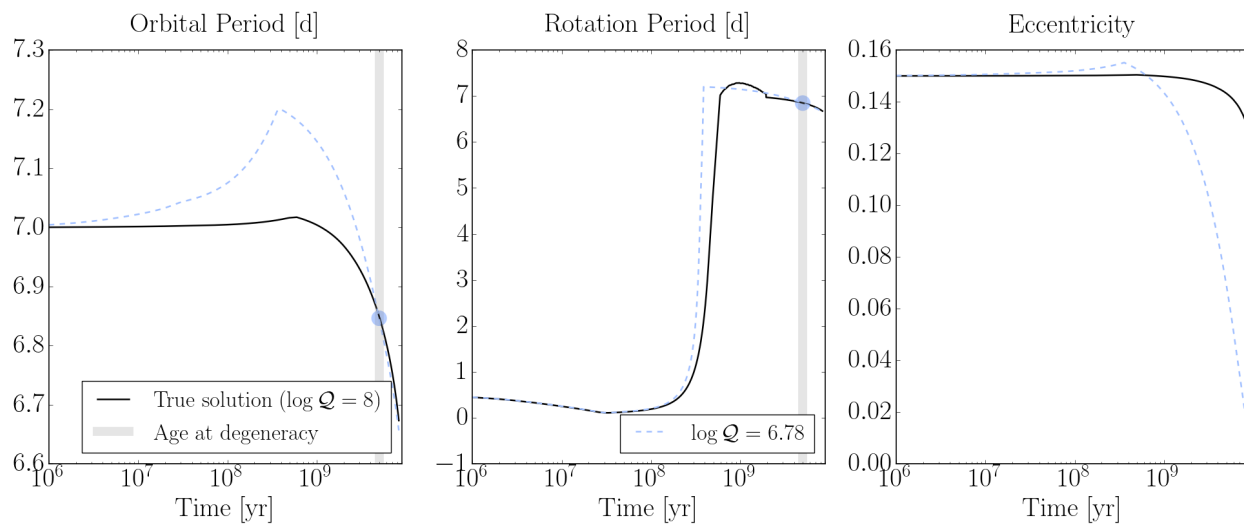


Figure 3.22: Same as Figure 3.21, but for the CPL model. In Figure 3.20, the solution for $\log \mathcal{Q} = 8$ shows a double peak in likelihood. The black line highlights the solution for $\log \mathcal{Q} = 8$. The blue dashed line shows the solution for the secondary likelihood peak: at 5 Gyr the likelihood has peaks at $\log \mathcal{Q} = 8$ and $\log \mathcal{Q} = 6.78$.

5 Discussion

The previous sections revealed that the path to constraining tidal evolution due to equilibrium tides faces fundamental challenges:

1. The evolution is governed by first-order ordinary differential equations that are high dimensional (18-dimensional phase space) and nonlinear. Thus, it is important to systematically understand how inference is sensitive to each model unknown and uncertainty.
2. We cannot observationally constrain the full phase space of a system: we can measure a system’s present-day configuration (*i.e.* final P_{orb} , P_{rot} , e , R), but the time derivatives (*i.e.* dP_{orb}/dt , dP_{rot}/dt , de/dt , dR/dt) are significantly more difficult to constrain.
3. There are strong degeneracies between the input parameters, particularly between $\mathcal{Q}$ and initial orbital conditions (P_{orb} , P_{rot} , e).
4. The observable constraints (final P_{orb} , P_{rot} , e , etc.) are more sensitive to the initial state conditions than tidal $\mathcal{Q}$.
5. For old systems (older than 5 Gyrs), the ratio $P_{\text{orb}}/P_{\text{rot}}$ is more sensitive to tidal $\mathcal{Q}$ than initial conditions. We find that populations of old systems may be useful for constraining upper or lower limits on tidal $\mathcal{Q}$.
6. Our analysis shows that for fixed tidal $\mathcal{Q}$ for all stars that tidal $\mathcal{Q}$ cannot be inferred to order-of-magnitude precision for individual systems by Bayesian methods, even when considering idealized uncertainties on present day orbital states and perfect priors (fixed at true values) for system masses and age.

In addition, when models are applied to real data (without inspection on simulated data), there is the confounding factor that the model may not accurately represent the true physics of the system. Although the ultimate goal is to make meaningful inferences on real data, having a poor understanding of the systematic biases of the model itself (as stated above) makes it difficult to unambiguously interpret inferences and disentangle which factors can be improved. Thus, in this paper we have investigated such systematic model biases in the context of statistical inference. Given these systematic challenges when it comes to inference of tidal $\mathcal{Q}$, we caution the community to consider model biases when interpreting inferred values of tidal $\mathcal{Q}$, particularly based on inference of individual binary systems.

A recent study by [Patel & Penev \(2022\)](#) reports a modified tidal $\mathcal{Q}'$ values for 41 low-mass ($0.4 < M_* < 1.2M_{\odot}$) eclipsing binary systems from *Kepler*. That study does not publish the full posteriors or summary statistics for their individual objects, but reports an overall constraint of $\log \mathcal{Q}' = 7.818 \pm 0.035$ based on the joint posterior of the 41 combined objects. However, it is unclear in the methodology of [Patel & Penev \(2022\)](#) (and similar studies, including [Penev et al. 2018](#)) how they handle model degeneracies in their inference. Their paper states that constraints on the initial orbital state come from finding “the initial orbital period and initial eccentricity, which when evolved to the sampled age of the system matches the orbital period and eccentricity sampled” (Section 4.1; [Patel & Penev, 2022](#)). This

assumption certainly contradicts our findings that the initial orbital period and eccentricity are not unique to a given fit of final values (Section 4.2 and Figures 3.16–3.17). Their paper also states that constraints on mass and age come from fitting stellar isochrones (Paxton et al., 2010) to the estimated effective temperatures, surface gravity, and metallicity from the Mathur et al. (2017) catalog of combined photometric and spectroscopic stellar parameters of *Kepler* targets. The study does not report uncertainties on their mass or age estimates, although isochrone fitting is known to be an unreliable age metric, particularly for low-mass stars, which may be on the main sequence (Soderblom, 2010). It is furthermore unclear whether the catalog estimates from Mathur et al. (2017) represent only the primary, or are biased by the blended photometry or spectroscopy of the secondary (Mathur et al. 2017 notes in Section 5.3 that the catalog is intended for single stars, and is likely biased for multistar systems). We argue that more work is necessary in this field to identify a reliable constraint on tidal $\mathcal{Q}$ than what is present in the existing literature.

5.1 Physical Origin of Short-period Binaries

In addition to the statistical challenge of constraining the physical parameters of equilibrium tides, there is also the fundamental question of what the constraints on the initial orbital parameters tell us about the origin of short-period binaries, which remains an open question in the field. The formation of stars is expected to result from a hierarchical collapse of gas within molecular clouds (Vázquez-Semadeni et al., 2019). However, the amount of thermal pressure expected during the initial collapse precludes the formation of systems at very short separations (with $a \lesssim 0.1$ AU or $P_{\text{orb}} \lesssim 10$ days), implying that additional mechanisms of dissipation must take place to explain the existence of short-period binaries (Sterzik et al., 2003; Tokovinin, 2021).

Results from the simulated posteriors (Figures 3.12–3.15) suggest that present-day short period systems (with orbital periods $P_{\text{orb}} < 10$ days) evolved from an initial configuration that started with a short orbital period, as seen by the high density of points with orbital periods between ~ 5 and 10 days. In other words, equilibrium tides (even for $\log \tau$ as high as 1.0, or $\log \mathcal{Q}$ as low as 4.0) are not an efficient enough dissipation mechanism to drive significant orbital migration.

A number of explanations have been proposed and investigated, including evolution due to a tertiary companion (Fabrycky & Tremaine, 2007), or dissipation after collapse due to interactions with primordial gas (Moe & Kratter, 2018). From a formation standpoint, it is worth mentioning that the plausibility of this conclusion (that short-period systems initially start out with a short orbital period) relies on the assumption that orbits undergo additional dissipation, causing inward migration to orbital periods of < 10 days within the first ~ 5 Myrs of their formation. For now it is understood that young, short-period systems may be plausible, but more work is needed to understand the dominant formation mechanisms of short-period binaries (Moe & Kratter, 2018; Tokovinin & Moe, 2020), as well as the role of dissipation due to dynamical tides, which has been shown to be more efficient than equilibrium tides during pre main-sequence evolution (Zanazzi & Wu, 2021).

5.2 Outlook

Our analysis (as well as recent work by [Mirouh et al. 2023](#)) suggests that constraining tidal dissipation, even using the best possible observational constraints (eclipsing binaries in open clusters with precise ages, orbital parameters, and rotation periods; *e.g.*, [Southworth & Clausen, 2006](#); [David et al., 2015, 2016](#); [Gillen et al., 2017](#); [Torres et al., 2018](#)), is difficult to near-impossible. While the prospects of constraining tides based on existing metrics (*e.g.*, measuring the tidal circularization period, or performing Bayesian inference of $\mathcal{Q}$ based on present-day orbital state) may not be promising, there may be alternative methods worth exploring.

5.2.1 Constraints on Orbital Period Decay

In particular, it is worth exploring alternative methods to constrain the derivative states of the evolutions. There are a few cases in which orbital period decay has been measured, including the planetary systems WASP-12b ([Patra et al., 2017](#); [Yee et al., 2019](#); [Patra et al., 2020](#)) and Kepler-1658b ([Chontos et al., 2019](#); [Vissapragada et al., 2022](#)) with implications for stellar tidal $\mathcal{Q}$'s, though orbital decay has not been widely measured for many stellar systems due to the long baseline of observation time required. Moreover, in general, we expect that it would be difficult to disambiguate how tidal $\mathcal{Q}$ influences orbital decay, apart from other factors (*e.g.*, magnetic braking, presence of unobserved companion) that could influence the angular momentum of the system. This limitation is significant when it comes to inferring timescales related to parameters such as tidal $\mathcal{Q}$ or tidal τ , which characterize tidal energy dissipation.

5.2.2 Understanding the Attractor Space

An alternative approach would seek to gain constraints based on a population of systems. The full dynamical phase space of this problem is 18-dimensional, including the nine state variables and their derivatives (Table 3.1). However, as predicted from tidal models, the evolution of a binary system is dissipative (*i.e.* flows contract in volume in phase space), due to the energy being lost to the interior of the stars in the form of heat due to tides.

For the CTL model the predicted equilibrium rotation period occurs for $d\omega/dt = 0$, or when

$$\frac{\omega_i}{n} = \frac{2 \cos \psi}{1 + \cos^2 \psi} \frac{N(e)}{\Omega(e)} \approx \frac{N(e)}{\Omega(e)}. \quad (3.15)$$

For obliquities $\psi < 30$ deg, the factor of (Eq. 3.15) containing ψ terms is insignificant to the equilibrium period by less than 1%. Since the CPL model we used is only applicable to low eccentricities ([Greenberg, 2009](#)), and can only predict pseudosynchronous ratios of 1:1 and 3:2 ([Fleming et al., 2019](#)), for now we apply only the equilibrium state analysis to the CTL model.

Next we consider the limiting behaviors when tides are combined with stellar evolution and magnetic braking. Synchronization occurs when the net torque acting on the rotation of a star becomes zero and the rotation period locks with the orbital period. This happens when the torque due to magnetic braking balances with the torque due to tides ($\frac{dJ}{dt}|_{tides} + \frac{dJ}{dt}|_{MB} = 0$). Finally, circularization occurs when ($de/dt = 0$ and $e = 0$).

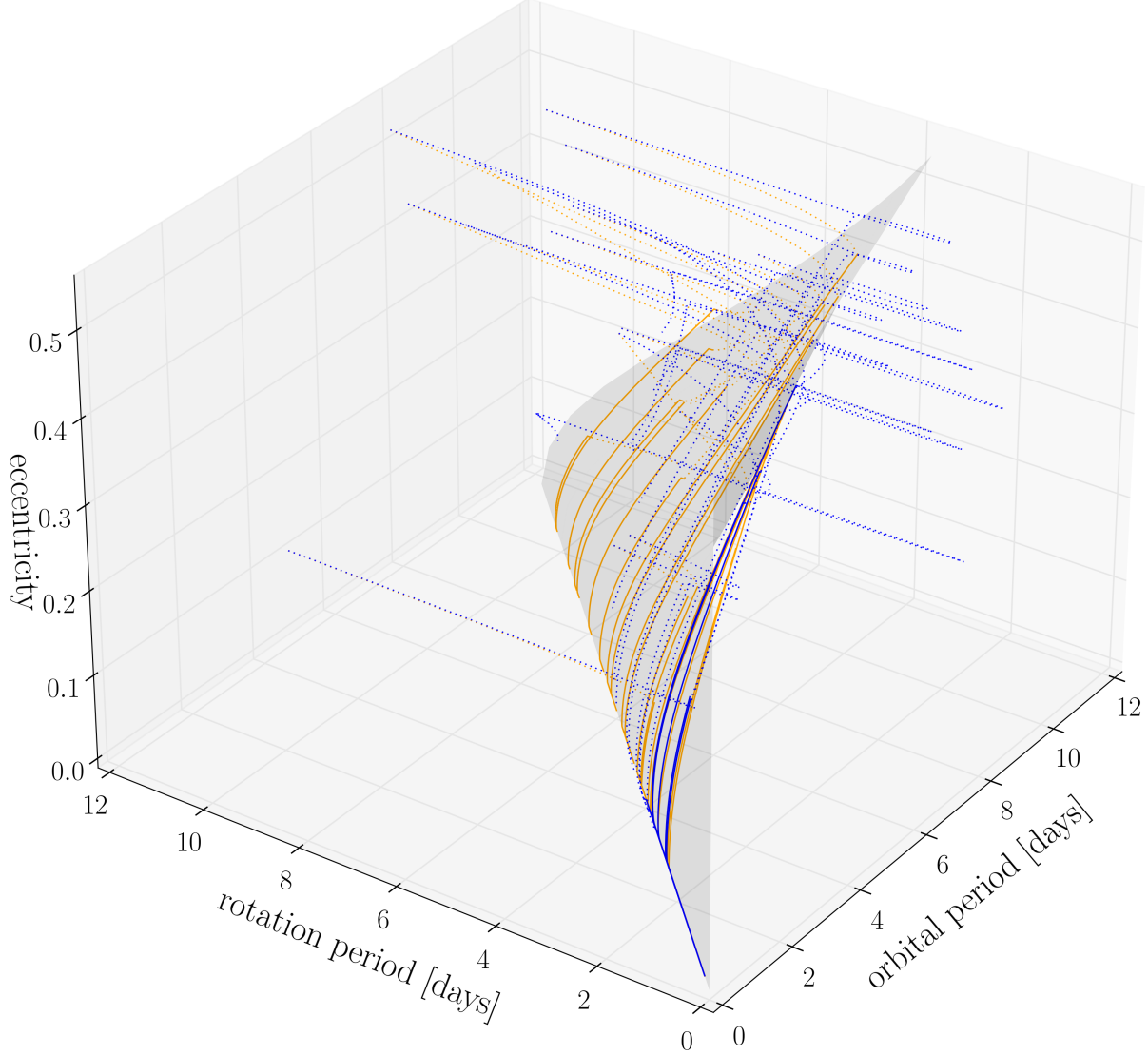


Figure 3.23: Orbital and rotational evolution for randomized initial conditions with strong tides (orange, $\log \tau = 0$) and weak tides (blue, $\log \tau = -2$). Axes plot the rotation period (days), orbital period (days), and eccentricity. The dashed portion of the lines show the trajectory before rotational torque balance, and the solid portion shows after. At rotational torque equilibrium (*i.e.* when tidal torque balances with magnetic braking), evolution trajectories converge to the shaded manifold, in which all $(P_{\text{orb}}, P_{\text{rot}}, e)$ solutions are constrained to a very precise 2D subspace of 3D.

While stellar evolution plays a large role in early dynamical evolution (as seen by trajectories in Figure 3.23 that initially “blow past” torque equilibrium), due to small variations in stellar radius and moment of inertia during the main sequence, the limiting behavior is driven by the equilibrium tide, which reduces the potential parameter space of our analysis. For old systems (on the main sequence), this implies that the model-observation comparison of the dominant variable space reduces to $(P_{\text{orb}}, P_{\text{rot},i}, e)$, or equivalently (n, ω_i, e) . That means the predicted equilibrium states lie along a very precise, low-dimensional manifold of (n, ω_i, e) space as shown in Fig. 3.23. Hence more work is needed to precisely constrain the (n, ω_i, e) space of observed binaries. In particular (as discovered in observational work by Lurie et al. 2017; Hobson-Ritz et al. 2025, and suggested by simulation work by Mirouh et al. 2023), more work is needed to characterize subsynchronously rotating binaries (systems near synchronization, but with spin-orbit ratios less than 1:1), which may provide key evidence of the precise balance between tidal forces and magnetic braking.

5.2.3 Model Reparameterization

We can also consider the possibility of reparameterizing the model as a way to reduce the dimensionality of the problem and degenerate parameters. Considering only the effect of tides, it is possible to reduce the number of orbital states from (n, ω_i, e) to just (ω_i, e) . This is because the total angular momentum of the system is given by

$$J_{\text{tot}} = J_{\text{orb}} + J_{\text{rot},1} + J_{\text{rot},2} \quad (3.16)$$

$$= \frac{\alpha}{n} \sqrt{1 - e^2} + \mathcal{I}_1 \omega_1 + \mathcal{I}_2 \omega_2, \quad (3.17)$$

so the mean motion can be written in terms of the rotation frequencies of each star and the eccentricity:

$$n = \left(\frac{\alpha \sqrt{1 - e^2}}{J_{\text{tot}} - \mathcal{I}_1 \omega_1 + \mathcal{I}_2 \omega_2} \right) \quad (3.18)$$

where $\alpha = M_1 M_2 G^{2/3} / (M_1 + M_2)^{1/3}$.

Figure 3.24 shows the evolution of the ratio ω_i/n (or $P_{\text{orb}}/P_{\text{rot}}$) for the the primary star in a binary system evolved according to CTL tides. For systems with initial conditions of $P_{\text{rot}} < P_{\text{orb}}$, the rate at which the system synchronizes depends linearly on $\log \tau$, without crossing degeneracies. However, this model reparameterization can only be done if the system conserves angular momentum. When magnetic braking is coupled with tides (Section 2.2), angular momentum is taken away from the system from stellar winds, introducing nonlinearities and crossing trajectories to the evolution of $P_{\text{orb}}/P_{\text{rot}}$ (Figure 3.25). Notably, the trajectories in Figure 3.25 exhibit a kink near $P_{\text{orb}}/P_{\text{rot}} \approx 0.75$, where the evolution temporarily stalls.

In this case an additional parameter ΔJ_{tot} (representing the amount of angular momentum lost from the system) would need to be sampled. This would not reduce the total number of parameters needed to be sampled, but would allow for the initial eccentricity and orbital period to be fit simultaneously and consistently, rather than independently. Further could test whether different parameterizations would be more efficient at sampling the parameter space.

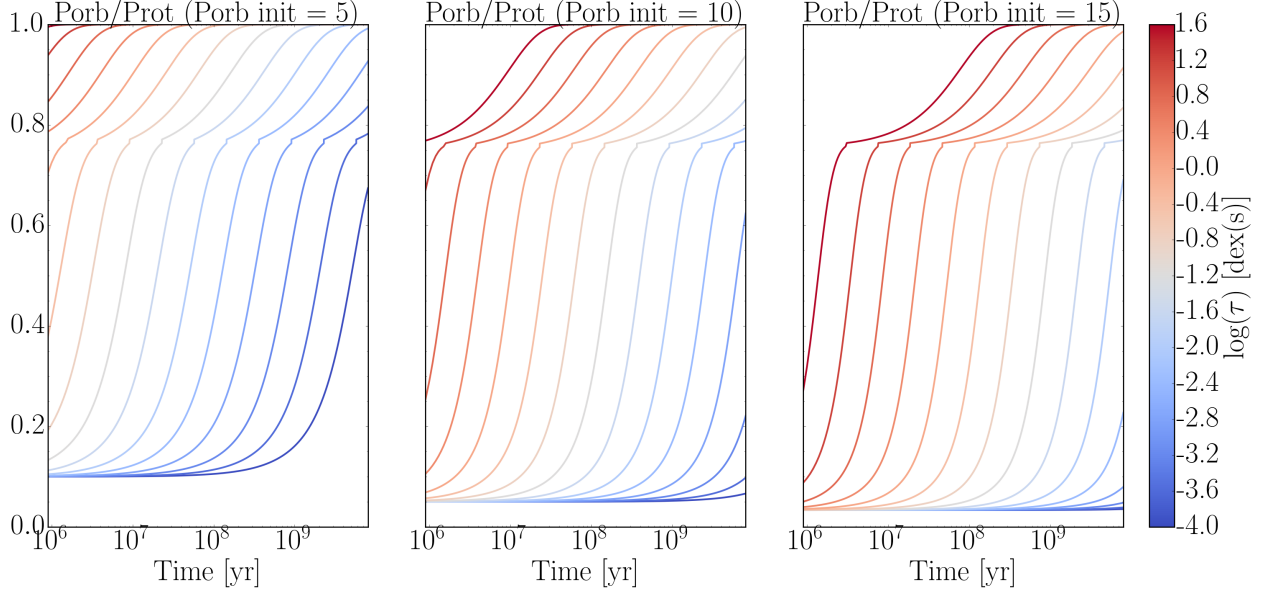


Figure 3.24: Evolution of the $P_{\text{orb}}/P_{\text{rot}}$ ratio for the primary star in a $1M_{\odot} - 1M_{\odot}$ binary system assuming only CTL tides. The system is evolved with an initial eccentricity of $e = 0.2$ and initial rotation periods of $P_{\text{rot},1} = P_{\text{rot},2} = 0.5\text{d}$ for different initial orbital periods: 5 days (left), 10 days (middle) and 15 days (right).

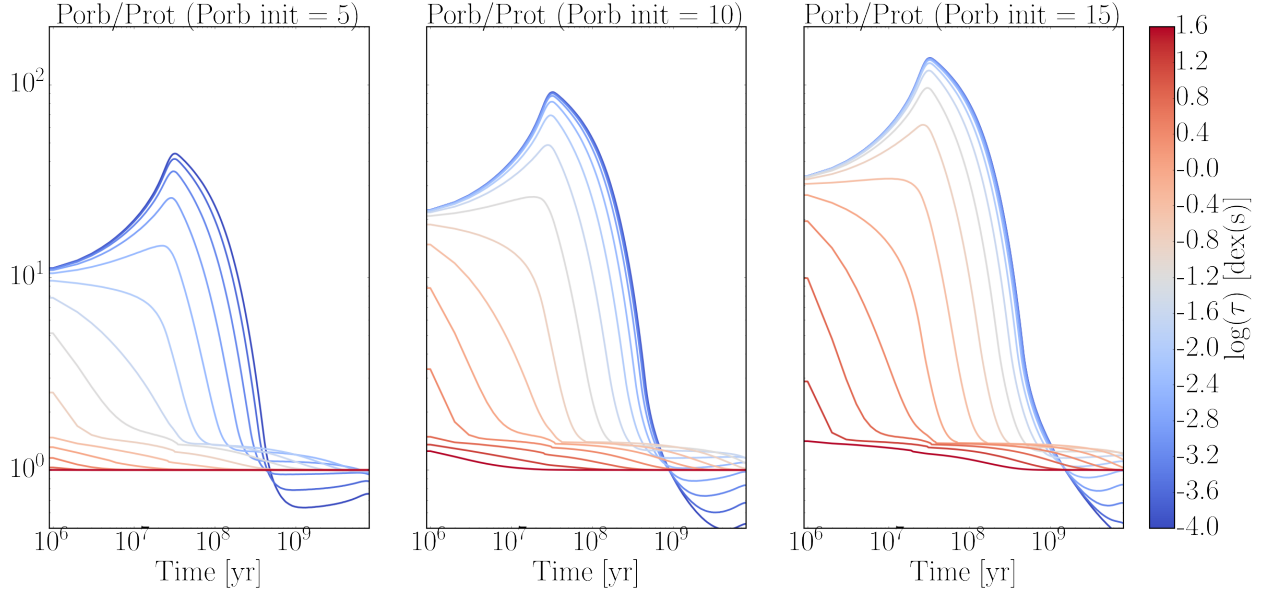


Figure 3.25: Same system as Figure 3.24, but modeled with CTL tides coupled with stellar evolution.

5.3 Limitations

The scope of this paper focuses on the limitations of equilibrium tidal dissipation theory (which is currently ill-constrained to orders of magnitude), not limitations of stellar evolution or magnetic braking (which we assume are correct by factors within an order of magnitude). In this study we only consider equilibrium tides, which only account for tidal dissipation when a hydrostatic bulge is raised on a star. We also assume a simplified tidal response, in which the mis-alignment of the bulge components are parameterized by either a time lag or phase lag that is constant in time. This is a significant simplification, as the efficiency of tidal dissipation may vary with changes in stellar internal structure over evolutionary timescales, or on shorter timescales due to magnetic activity cycles (*e.g.*, [Ogilvie, 2013](#)). The CTL and CPL mechanisms do not account for all of the complexities of the stellar interior structures that have been proposed, including turbulent viscosity acting on equilibrium tides in the convective envelope ([Barker, 2020](#)) or the fluid motions of dynamical tide prescriptions ([Zahn & Bouchet, 1989](#)).

We also assume that stellar evolution influences tides, but we make no assumptions about how tides may influence stellar evolution. Studies such as [Casey et al. \(2019\)](#) suggest that tides may influence stellar evolution, including lithium production in red-giant stars. However, modeling the effects on stellar evolution would be a significant endeavor requiring the use of a stellar evolution code such as MESA (Modules for Experiments in Stellar Astrophysics; [Paxton et al., 2010](#)), rather than grids of evolution models ([Baraffe et al., 2015](#)). Furthermore, we only model low-mass systems through their main-sequence evolution phase with separations larger than the Roche radius, hence we do not model or make assumptions about any mass transfer between stars in the system.

We also make the assumption that orbital evolution occurs for an isolated binary system, in which there are no additional perturbers (such as planet or nearby star) to influence the angular momentum of the system. The presence of additional bodies could introduce a variety of different effects, such as Kozai-Lidov effects, resonances, or chaos ([Toonen et al., 2016](#); [Fabrycky & Tremaine, 2007](#); [Naoz, 2016](#)). Accounting for the presence of triple (or other higher order systems) would be an important role in future studies when comparing binary evolution models to observations, as unobserved companions could account for differences between the observed and simulated distributions of orbits.

6 Conclusion

In this paper, we investigated the prospects and limitations for constraining equilibrium tides in low-mass binary stars based on ideal uncertainties. We used the package **VPLanet** to simulate the coupled effects of stellar evolution, magnetic braking, and equilibrium tides. We then applied statistical methods including sensitivity analysis and simulated Bayesian inference to systematically analyze how model degeneracies and observational uncertainties limit constraints on tidal dissipation.

Sensitivity analysis allowed us to systematically search high-dimensional and non-linear parameter spaces to identify which types of individual systems are most promising for con-

straining tidal dissipation, thus informing observational follow-up efforts. We used sensitivity analysis to assess which input parameters dominate the final orbital and rotational states of the system. We found to first order that the final orbital state (in particular the final orbital period) is dominated by the initial orbital period of the system. We further investigated whether the final eccentricity and rotation states would be sensitive enough to gain meaningful constraints on tidal dissipation. To perform this assessment, we simulated inference using the package `alabi` (Chapter 2) to trace out 5-dimensional posteriors in a computationally efficient manner.

Our analysis suggests that even when employing implausibly precise orbital parameters from *Kepler*/*TESS*, constraining tidal Q remains ill-posed because of strong degeneracies with initial conditions and because we cannot determine how quickly a system evolves (*i.e.* we have no constraints on dP_{orb}/dt , dP_{rot}/dt , de/dt , etc. for any given system). An alternate approach would take advantage of the constraints from a large population of binaries. However, different binary systems can originate from a wide range of initial conditions, which are degenerate and dependent on complex (and very ill-constrained, uncertain) formation mechanisms.

An alternative approach worth investigating would seek to characterize the attractor state dynamics of tides. Stellar evolution and magnetic braking dominate in early evolution; however, during the main-sequence tidal dissipation dominates, even for weak tides. Considering the case of constant radius and radius of gyration (which is approximately the case when stars reach the main sequence), the attractor phase space under torque equilibrium reduces to 6 dimensions (e , a , ω , de/dt , da/dt , $d\omega/dt$). Thus, overdensities in $(P_{\text{orb}}, e, P_{\text{rot}})$ of observed binaries may be a promising way to constrain and validate theoretical tidal models, regardless of how fast these systems are evolving.

In summary this work critically examines our understanding of equilibrium tides in binary systems. While there is certainly still much work left to realistically model the tidal response of stars and to understand how stellar structure plays a role in the gravitational interactions of stars, the impact of work in this subfield has many implications for broader astrophysics. Stars play a fundamental role on a broad range of scales from the long-term evolution of planets and planetary systems, to the dynamics of stellar clusters and galactic dynamics. Thus the dynamical outcomes on the orbits of binary stars due to tidal evolution has the potential to give insight to many fundamental physical processes involving gravitational interactions with stars.

1 Equilibrium Tide Equations

In this appendix, we present the equilibrium tide models used in this study.

1.1 Governing Equations

The average energy dissipated $\langle \dot{E}_{\text{orb}} \rangle_i$ and torque on each star in the system $\langle \dot{J}_{\text{rot}} \rangle_i$ can be derived from the potential and an assumption for the phase lag dependence (e.g. constant time, or constant phase). For the CTL model, the tidal energy and torque terms are given by (Leconte et al., 2010, Equation A.21 & Equation 9):

$$\langle \dot{E}_{\text{orb}} \rangle_i = 2K_i \tau_i \left[N(e) \cos \psi_i \frac{\omega_i}{n} - N_a(e) \right] \quad (.19)$$

$$\langle \dot{J}_{\text{rot}} \rangle_i = -\frac{K_i \tau_i}{n} \left[(1 + \cos^2 \psi_i) \Omega(e) \frac{\omega_i}{n} - 2 \cos \psi_i N(e) \right] \quad (.20)$$

where the eccentricity functions $N(e)$, $N_a(e)$, $\Omega(e)$ from Leconte et al. (2010) are given by:

$$N(e) = \frac{1 + \frac{15}{2}e^2 + \frac{45}{8}e^4 + \frac{5}{16}e^6}{(1 - e^2)^6} \quad (.21)$$

$$N_a(e) = \frac{1 + \frac{31}{2}e^2 + \frac{255}{8}e^4 + \frac{185}{16}e^6 + \frac{25}{64}e^8}{(1 - e^2)^{15/2}} \quad (.22)$$

$$\Omega_e(e) = \frac{1 + \frac{3}{2}e^2 + \frac{1}{8}e^4}{(1 - e^2)^5} \quad (.23)$$

$$N_e(e) = \frac{1 + \frac{15}{4}e^2 + \frac{15}{8}e^4 + \frac{5}{64}e^6}{(1 - e^2)^{13/2}} \quad (.24)$$

$$\Omega(e) = \frac{1 + 3e^2 + \frac{3}{8}e^4}{(1 - e^2)^{9/2}} \quad (.25)$$

and the intermediate coefficient variable K_i is given by:

$$K_i = \frac{3}{2} G k_2 n^2 \left(\frac{M_j^2 R_i^5}{a^6} \right). \quad (.26)$$

Here k_2 is known as the tidal Love number. The Love number k_l describes the degree to which a stellar body deforms due to the tidal forces in the radial direction, where k_2 is the Love number due to the quadrupole order potential. In this study we adopt a value of $k_2 = 0.5$, but note that uncertainty in k_2 is degenerate with that of $\mathcal{Q}$ and can also be written as the effective quality factor $\mathcal{Q}' = 3\mathcal{Q}/2k_2$, where $k_2 = 3/2$ and $\mathcal{Q}' = \mathcal{Q}$ for a homogeneous fluid sphere (Jackson et al., 2008).

For the alternative linear equilibrium tide model, CPL, the tidal energy dissipation and

torque terms are given by (Ferraz-Mello et al., 2008, Equation 48–49):

$$\langle \dot{E}_{\text{orb}} \rangle_i = \frac{K_i}{4n} \left[4\varepsilon_0 + e^2(-20\varepsilon_0 + \frac{147}{2}\varepsilon_1 + \frac{1}{2}\varepsilon_2 - 3\varepsilon_5) - 4\sin^2\psi(\varepsilon_0 - \varepsilon_8) \right] \quad (.27)$$

$$\langle \dot{J}_{\text{rot}} \rangle_i = -\frac{K_i}{4n^2} \left[4\varepsilon_0 + e^2(-20\varepsilon_0 + 49\varepsilon_1 + \varepsilon_2) + 2\sin^2\psi(-2\varepsilon_0 + \varepsilon_8 + \varepsilon_9) \right] \quad (.28)$$

Under the constant phase lag assumption, the amplitude of the phase lags are proportional to a constant $\mathcal{Q} \approx 1/\varepsilon$. For quadrupole order potential ($l = 2$), we consider the dominant forcing frequencies $\varepsilon = kn - m\omega$ up to $m = \pm 2$, where the amplitude of each phase lag term corresponds to the following the frequencies (Table 1, Ferraz-Mello et al., 2008):

$$\begin{aligned} \varepsilon_{0,i} &= 2\omega_i - 2n \\ \varepsilon_{1,i} &= 2\omega_i - 3n \\ \varepsilon_{2,i} &= 2\omega_i - n \\ \varepsilon_{5,i} &= n \\ \varepsilon_{6,i} &= 2n \\ \varepsilon_{7,i} &= 2n \\ \varepsilon_{8,i} &= \omega_i - 2n \\ \varepsilon_{9,i} &= \omega_i \end{aligned}$$

1.2 Secular Evolution of Keplerian Elements

This section elaborates on the equations for the total evolution of the Keplerian components: semi-major axis, eccentricity, rotation rate, and obliquity for both the constant-time-lag and constant-phase-lag equilibrium tidal models. For any tidal model, the orbit-averaged evolution of the Keplerian orbital elements due to each component is given by:

$$\text{semi-major axis: } \left\langle \frac{da}{dt} \right\rangle = -\frac{a}{E_{\text{orb}}} \left\langle \frac{dE_{\text{orb}}}{dt} \right\rangle, \quad (.29)$$

$$\text{eccentricity: } \left\langle \frac{de}{dt} \right\rangle = \left(\frac{1}{a} \frac{da}{dt} - \frac{2}{J_{\text{orb}}} \frac{dJ_{\text{orb}}}{dt} \right) \cdot \frac{1 - e^2}{2e}, \quad (.30)$$

$$(.31)$$

where $\mu_{\text{tot}} = M_1 M_2 / (M_1 + M_2)$ is the reduced mass, and $\mathcal{I}_i = M_i r_{g,i}^2 R^2$ is the moment of inertia for each body. The spin evolution for each body is given by:

$$\text{rotation frequency: } \left\langle \frac{d\omega_i}{dt} \right\rangle = \frac{1}{\mathcal{I}_i} \left(\frac{dJ_{\text{tide}}}{dt} + \frac{dJ_{\text{mb}}}{dt} - \omega_i \frac{d\mathcal{I}_i}{dt} \right), \quad (.32)$$

where $i = 1$ is the primary, $i = 2$ is the secondary. The terms bracketed by $\langle \cdot \rangle$ denote quantities that are orbit-averaged.

1.2.1 Semi-major Axis

The equation for the semi-major axis evolution when coupled with stellar evolution and magnetic braking remains the same as when only considering equilibrium tide. The total semi-major axis derivative is the sum of the contributions from the primary ($i = 1$) and secondary ($i = 2$):

$$\left\langle \frac{da}{dt} \right\rangle_{\text{tot}} = \sum_{i=1}^2 \left\langle \frac{da_i}{dt} \right\rangle_{\text{tide}}. \quad (.33)$$

For the constant time lag model, the equation for the orbit-averaged derivative of semi-major axis can be derived from substituting Equations (.29) and (.19):

$$\left\langle \frac{da_i}{dt} \right\rangle_{\text{CTL}} = \frac{4a^2}{GM_1M_2} K_i \tau_i \left[N(e) \cos \psi_i \frac{\omega_i}{n} - N_a(e) \right]. \quad (.34)$$

Similarly, the orbit-averaged derivative of semi-major axis for the constant phase lag assumption is found from substituting Equations (.29) and (.27):

$$\left\langle \frac{da_i}{dt} \right\rangle_{\text{CPL}} = \frac{a^2}{GM_1M_2} \frac{K_i}{2nQ_i} \left[4\hat{\varepsilon}_0 + e^2(-20\hat{\varepsilon}_0 + \frac{147}{2}\hat{\varepsilon}_1 + \frac{1}{2}\hat{\varepsilon}_2 - 3\hat{\varepsilon}_5) - 4\sin^2 \psi(\hat{\varepsilon}_0 - \hat{\varepsilon}_8) \right]. \quad (.35)$$

Applying the constant phase lag assumption ($\varepsilon = \hat{\varepsilon}/Q$ where Q is a constant) lets us factor out the magnitude of the phase lag Q on the denominator, and each $\hat{\varepsilon}$ term corresponds to the sign of each phase lag term (Heller et al., 2011), $\hat{\varepsilon}_i = \text{sign}(\varepsilon_i)$.

1.2.2 Eccentricity

The eccentricity evolution can be obtained by differentiating the orbital angular momentum and rearranging to find:

$$\left\langle \frac{de}{dt} \right\rangle = -\frac{1}{E_{\text{orb}}} \left[\frac{dE_{\text{orb}}}{dt} - \frac{n}{\sqrt{1-e^2}} \frac{dJ_{\text{orb}}}{dt} \right] \cdot \frac{1-e^2}{2e}. \quad (.36)$$

From conservation of angular momentum, the change in orbital angular momentum is the negative of the magnitude of the change in rotational angular momentum, which is the sum of the torques due to tides and magnetic braking:

$$\frac{dJ_{\text{orb}}}{dt} = -\sum_{i=1}^2 \frac{dJ_{\text{rot},i}}{dt} = -\sum_{i=1}^2 \left(\frac{dJ_{\text{tide}}}{dt} + \frac{dJ_{\text{mb}}}{dt} \right)_i. \quad (.37)$$

Equation (.36) can thus be broken into components, $\langle de/dt \rangle = \langle de/dt \rangle_{\text{tide}} + \langle de/dt \rangle_{\text{mb}}$. Following the assumption of Repetto & Nelemans (2014) and Fleming et al. (2019), we assume $\langle de/dt \rangle_{\text{mb}}$ is relatively small compared to $\langle de/dt \rangle_{\text{tide}}$, and thus just consider the effect of tides on eccentricity evolution.

The contribution to eccentricity due to tides for the constant time lag model, can be

obtained by substituting Eqns (.19) and (.20) into Eqn. (.30) to get:

$$\left\langle \frac{de_i}{dt} \right\rangle_{\text{CTL}} = \frac{ae}{GM_1 M_2} K_i \tau_i \left[11 \cos \psi_i \Omega_e(e) \frac{\omega_i}{n} - 18 N_e(e) \right], \quad (.38)$$

and for the constant phase lag model, can be obtained by substituting Eqns (.27) and (.28) into Eqn. (.36):

$$\left\langle \frac{de_i}{dt} \right\rangle_{\text{CPL}} = -\frac{ae}{4GM_1 M_2} \frac{K_i}{n \mathcal{Q}} \left[2\hat{\varepsilon}_0 - \frac{49}{2}\hat{\varepsilon}_1 + \frac{1}{2}\hat{\varepsilon}_2 + 3\hat{\varepsilon}_5 \right]. \quad (.39)$$

1.2.3 Rotation

The secular rotational evolution for each individual body i in the system is solved from the sum of torques due to both tides and magnetic braking acting on the body:

$$\left\langle \frac{d\omega_i}{dt} \right\rangle = \left\langle \frac{d\omega_i}{dt} \right\rangle_{\text{tide}} + \left\langle \frac{d\omega_i}{dt} \right\rangle_{\text{mb}} = \frac{1}{\mathcal{I}} (\mathcal{T}_{\text{tide}} + \mathcal{T}_{\text{mb}}), \quad (.40)$$

where the total rotational evolution due to CTL tides and stellar evolution is given by the sum of Equation (.20) and (3.4) divided by moment of inertia, $\mathcal{I} = Mr_g^2 R^2$:

$$\left\langle \frac{d\omega_i}{dt} \right\rangle_{\text{CTL}} = -\frac{K_i \tau_i}{n \mathcal{I}_i} \left[(1 + \cos^2 \psi_i) \Omega(e) \frac{\omega_i}{n} - 2 \cos \psi_i N(e) \right]. \quad (.41)$$

Similarly, the total rotational evolution due to CPL tides and stellar evolution is given by the sum of Equation (.28) and (3.4) divided by moment of inertia:

$$\left\langle \frac{d\omega_i}{dt} \right\rangle_{\text{CPL}} = -\frac{K_i}{4n^2 \mathcal{Q}_i \mathcal{I}_i} \left[4\hat{\varepsilon}_0 + e^2(-20\hat{\varepsilon}_0 + 49\hat{\varepsilon}_1 + \hat{\varepsilon}_2) + 2 \sin^2 \psi (-2\hat{\varepsilon}_0 + \hat{\varepsilon}_8 + \hat{\varepsilon}_9) \right]. \quad (.42)$$

1.2.4 Obliquity

Finally, the secular evolution of the obliquity for each star in the system for the CTL model was derived in [Leconte et al. \(2010\)](#):

$$\left\langle \frac{d\psi_i}{dt} \right\rangle = \frac{K_i \sin \psi_i}{\mathcal{I}_i n} \left[(\cos \psi_i - \eta_i) \Omega(e) \frac{\omega_i}{n} - 2N(e) \right], \quad (.43)$$

where $\eta_i = J_{\text{rot},i}/J_{\text{orb}}$ is the ratio of rotational over orbital angular momentum:

$$\eta_i = \frac{M_1 + M_2}{M_1 M_2} \frac{\mathcal{I}_i \omega_i}{a^2 n \sqrt{1 - e^2}}. \quad (.44)$$

For the CPL model, we use the obliquity evolution derived in [Ferraz-Mello et al. \(2008\)](#):

$$\left\langle \frac{d\psi_i}{dt} \right\rangle = \frac{K_i \sin \psi_i}{2\omega_i \mathcal{I}_i n^2 \mathcal{Q}} \left[(1 - \xi_i) \hat{\varepsilon}_0 + (1 + \xi_i) (\hat{\varepsilon}_8 - \hat{\varepsilon}_9) \right], \quad (.45)$$

where ξ_i is defined as

$$\xi_i = \frac{\omega_i \mathcal{I}_i a n}{GM_i M_j}. \quad (.46)$$

CHAPTER 4

A PHYSICALLY MOTIVATED GAUSSIAN PROCESS KERNEL FOR STELLAR ROTATION

1	Introduction	118
2	Lightcurve Model	121
2.1	Single-Spot Projected Area	121
2.2	The Emergence and Decay Profile of Spots	123
2.3	Lattitudinal and Longitudinal Evolution of Spots	124
2.4	Stellar Limb Darkening	126
2.5	Summary of Parameters	126
3	Gaussian Process	128
3.1	Motivation for Gaussian Processes	128
3.2	Mathematical Formulation of Gaussian Processes	129
3.3	The Covariance Matrix and Kernel Function	131
4	Deriving the Mean Function	132
5	Deriving the Analytical Kernel for Starspot Variability	134
5.1	Fourier Transform Notation and Identities	134
5.2	Approach to Derivation	134
5.3	Fourier Transform: The Rotation Term $\mathcal{V}(t)$	137
5.4	Fourier Transform: Spot Evolution Envelope Term $\Gamma(t)$	139
5.5	Combined Kernel for a Single Spot	141
5.6	Averaging Over Random Spot Parameters	142
5.7	Multi-Spot Kernel	144
5.8	Power Spectral Density of the Full Kernel	147
6	Physical Interpretation of Kernel Hyperparameters	149
6.1	Information Content of the Mean and Covariance	149
6.2	Kernel Components and Hyperparameters	149
7	GP Solver Implementation	161
7.1	Naive Cholesky Solver	161
7.2	Banded Cholesky Solver	162
7.3	Software Architecture	163

8	MCMC Sampling the Kernel Hyperparameters	164
8.1	Log-Posterior and Gradients	166
8.2	Prior	166
8.3	Mass Matrix Estimation	167
8.4	Sampler Backends	167
9	Numerical Results	168
9.1	Analytic vs. Numerical Kernel	168
9.2	Comparison to Kernels from Literature	173
9.3	Comparison to Other GP Packages	174
9.4	Inference Conditions	177
9.5	Testing the Cranmer–Rao Bound of Different Parameter Regimes . .	178
9.6	MCMC Inference	186
10	Case Study: Application to Kepler-63	188
10.1	System Geometry	188
10.2	Fitting Procedure	191
10.3	Posterior Results	192
11	Discussion	202
11.1	Summary of Results	202
11.2	Applications and Future Work	203
11.3	Limitations	204
12	Conclusion	204
1	Fourier Coefficients of the Visibility Function	205
1.1	Setup	205
1.2	Evaluating the Integral	206
2	Fourier Transform of the Trapezoidal Envelope	207
2.1	Piecewise Definition of $\Gamma(t)$	208
2.2	Direct Integration	208
3	Autocorrelation of the Squared Envelope	209
3.1	Setup	209
3.2	Result	210

Abstract

e present a novel, physically motivated Gaussian process (GP) kernel designed to model starspot variability from photometric lightcurves. Using an analytic flux model, we derive an analytic kernel function that describes the temporal correlations induced by rotating starspots as a function of spot properties and stellar rotation: equatorial rotation period (P_{eq}), differential rotation (κ), inclination angle (I), spot lifetimes (ℓ_{spot}), spot emergence and decay timescales (τ_{spot}), maximum spot size (α_{max}), spot contrast (f_{spot}), and number of spots (N_{spot}). We thoroughly validate our derivation using SymPy, compare it to numerical simulations, and interpret how each parameter affects the kernel structure. We show that α_{max} , f_{spot} , and N_{spot} are degenerate parameters controlling the amplitude of the correlations (which can be re-parameterized into a single parameter σ_k), while P_{eq} , κ , and I control the periodicity and harmonic structure, and ℓ_{spot} , τ_{em} , and τ_{dec} control the decay of the correlations over time. Our implementation is available in the package `spotgp`, which uses JAX for efficient automatic differentiation and GPU acceleration. Furthermore, our implementation makes use of a sparse banded matrix solver (with bandwidth b), reducing the computational complexity of GP inference from $\mathcal{O}(N^3)$ to $\mathcal{O}(Nb^2)$ for N data points. Using simulated lightcurves, we use Markov chain Monte Carlo (MCMC) and Cramer-Rao bound analysis to systematically explore how inference performs for different parameter regimes, and observational conditions (cadence, baseline, and noise level). Finally, we apply our GP model to Kepler-63 and demonstrate that we can infer an inclination consistent with $v_* \sin I$ measurements from spectroscopy and spot parameters with differential rotation and spot lifetimes consistent with transit spot analysis. This study is the first to demonstrate that differential rotation can be inferred from 1-dimensional lightcurves with a physically interpretable Gaussian process.

1 Introduction

High-precision photometric surveys such as *Kepler* (Borucki et al., 2010), *K2* (Howell et al., 2014), and *TESS* (Ricker et al., 2014) have been transforming our understanding of fundamental stellar physics, revealing the effects of stellar variability and dynamics for stars across a wide range of masses, ages, and Galactic environments. In particular, precise time-domain data have enabled constraints on the rotation periods of thousands of stars (McQuillan et al., 2014; Gordon et al., 2021; Boyle et al., 2026), which has been crucial for understanding the evolution of stellar angular momentum (Weber & Davis, 1967; Skumanich, 1972), the surface variability and magnetic activity of stars (Santos et al., 2017; Forgács-Dajka et al., 2021), and the stellar dynamo process (Brun et al., 2017b; Käpylä et al., 2023; Noraz et al., 2022c).

Over time, stars lose angular momentum through magnetized stellar winds, causing them to spin down as they age. The rate of this spin-down depends on the star’s mass, age, and magnetic field strength, and is driven by the torque from stellar winds flowing along open magnetic field lines – a process known as *magnetic braking*. Early work established that $P_{\text{rot}} \propto t^{-1/2}$ (Skumanich, 1972), opening up the field of gyrochronology for inferring stellar ages from rotation periods (Barnes, 2003a, 2007). Further work has developed empirical relations for how magnetic braking torque depends on stellar properties (Kawaler, 1988; Sills

et al., 2000; Reiners & Mohanty, 2012; van Saders & Pinsonneault, 2013; Matt et al., 2015; Amard & Matt, 2020; Ireland et al., 2022). However open questions remain: recent studies have discovered evidence for weakened magnetic braking in rapidly rotating old field stars (van Saders et al., 2016; Saunders et al., 2024), suggesting that the braking process may be more complex than originally thought.

Much of our knowledge of stellar rotation comes from photometric variability driven by starspots. Low-mass stars ($M \lesssim 1.5 M_{\odot}$) possess outer convective envelopes in which turbulent motions interact with rotation and magnetic fields to drive a stellar dynamo (Brun et al., 2017b). Differential rotation (the latitudinal variation in angular velocity) plays a fundamental role in this process: it shears poloidal magnetic field lines into a toroidal configuration (the Ω -effect), which is a key mechanism driving the activity cycle (Noraz et al., 2022c). Numerical simulations of rotating convection predict that the magnitude and geometry of differential rotation depend on the stellar mass, rotation rate, and convective structure (Augustson et al., 2012; Guerrero et al., 2013b; Käpylä et al., 2023). In particular, these simulations predict a transition from solar-like differential rotation (fast equator, slow poles) to anti-solar differential rotation (fast poles, slow equator) as the Rossby number $Ro = P_{\text{rot}}/\tau_{\text{conv}}$ increases past $Ro \sim 1$. Many starspot properties – including spot lifetimes (Giles et al., 2017), latitude distributions, and emergence patterns – are also governed by Ro , making it a highly important dimensionless parameter for characterizing the stellar dynamo regime (Gastine et al., 2013; Noraz et al., 2022a). However, this predicted transition in differential rotation remains largely untested observationally, because measuring differential rotation for large samples of stars has been extremely challenging.

Existing techniques for measuring stellar differential rotation include: Doppler imaging (Collier Cameron, 2007), the Fourier transform method for spectral line profiles (Reiners & Mohanty, 2012), asteroseismology (Benomar et al., 2018), and transit spot-crossing (Morris et al., 2017). However, each have each been limited to samples of $\sim$ tens of stars, and are restricted to specific stellar types (e.g., fast rotators, bright stars, or transiting systems). On the other hand, the high-precision, long-baseline photometric datasets from *Kepler* and *TESS* encode information on the spot variability for hundreds of thousands of stars, which has yet to be fully leveraged. However, extracting differential rotation from disk-integrated photometry is notoriously difficult: Aigrain et al. (2015) conducted a blind hare-and-hounds exercise on 1,000 simulated *Kepler*-like lightcurves and found *little correlation between the reported and simulated values of the differential rotation shear*. Morris et al. (2019b) reinforced this conclusion using the Sun itself as a benchmark, finding no detectable differential rotation signal even in a noise-free solar lightcurve reconstructed from archival sunspot records. These results suggest that the stochastic nature of spot emergence at a broad range of latitudes causes the measured period to reflect an average over active latitudes, rather than revealing the latitude-dependent rotation law.

Modeling the magnetohydrodynamical processes that accurately reproduces the convection, magnetic fields, and dynamo process of low-mass stars is inherently complex and computationally expensive (Miesch et al., 2000, 2008; Charbonneau, 2020). As an alternative, analytic spot models (Kipping, 2012; Luger et al., 2019) treat the stellar surface as hosting an arbitrary number of circular spots that emerge and decay at different latitudes, each modulating the disk-integrated flux according to its projected area and contrast. These models have successfully matched the statistical properties of observed lightcurves (Aigrain et al.,

2015), while providing a physically interpretable framework for the spot geometry.

Although approaches for forward modeling surfaces with spots exist, inferring information about *individual* star spots is an ill-posed problem with many degeneracies (Luger et al., 2019; Basri & Shah, 2020). Basri & Shah (2020) showed that the single-dip ratio (a metric for whether a lightcurve shows one or two dips per rotation) is only weakly sensitive to differential rotation shear and is strongly confounded by spot lifetime, concluding that differential rotation may be “nearly impossible to accurately infer from light curves alone unless spots live for many rotations.” A Gaussian process (GP; Rasmussen & Williams, 2006) offers an alternative approach that treats starspot variability as a stochastic process and models the *distribution* of spot properties rather than individual spots. GPs have become a standard tool for modeling stellar rotation (Angus et al., 2018b; Foreman-Mackey et al., 2017; Gordon et al., 2021), but the most widely used quasi-periodic (QP) kernel has no parameter that encodes differential rotation, meaning that any differential rotation signal is absorbed into the other hyperparameters in an uncontrolled way. (Nicholson & Aigrain, 2022) systematically tested the mapping between QP kernel hyperparameters and physical spot properties, finding that while the rotation period and evolution timescale are reliably recovered, the harmonic complexity parameter has no intuitive physical interpretation. These results collectively motivate the development of a physically grounded GP kernel—one whose hyperparameters map directly to measurable stellar properties including differential rotation.

In this paper, we derive a physically motivated GP kernel for stellar rotation from first principles. Starting from an analytic starspot model (Kipping, 2012), we obtain a closed-form kernel function that naturally encodes the effects of spot geometry, evolution, and differential rotation. We validate the analytic kernel against numerical Monte Carlo simulations, demonstrate that it captures harmonic structure and non-sinusoidal periodicity absent in standard quasi-periodic and simple harmonic oscillator kernels, and show that it provides significant computational speedups over numerical approaches. The paper is organized as follows: Section 2 describes the lightcurve model; Section 3 describes the formalisms for defining the Gaussian process model; Section 5 derives the GP kernel; Section 6 discusses the physical interpretation of the kernel hyperparameters; Section 7 presents the GP solver implemented in JAX; Section 8 describes how the GP hyperparameters can be sampled in an MCMC inference framework; Section 9 validates the analytic kernel against numerical results, compares our kernel to other quasi-periodic kernels used in literature, and discusses the conditions for precisely inferring starspot parameters; Section 10 demonstrates a case study on real data, in which we compare our methodology to parameters derived in literature for Kepler-63; Section 11 summarizes the main findings of our work and discusses future work and known limitations of our model. Finally, we discuss concluding remarks in Section 12 and provide detailed derivations in the Appendix.

2 Lightcurve Model

We use the analytic starspot model defined in [Kipping \(2012\)](#). The flux model for N_{spot} spots with contrast f_k is:

$$F(t) = 1 - \frac{1}{\pi} \sum_{k=1}^{N_{\text{spot}}} A_k(t) \cdot (1 - f_{\text{spot},k}). \quad (4.1)$$

where $A_k(t)$ is the projected area of spot k at time t and $f_{\text{spot},k}$ is the contrast of the spot ($0 =$ black spot, $1 =$ no contrast). The flux deficit at a given time is the sum of the contributions from all visible spots on the surface of the star, where each contribution is determined by the projected area of the spot and its contrast relative to the stellar photosphere.

2.1 Single-Spot Projected Area

For each individual spot, the flux deficit is determined by the projected area of the spot on the visible hemisphere of the star. For ease of notation we drop the spot index k and write the projected area of a single spot as $A(t)$, however these equations are applied to each spot k with its own time-dependent area $A_k(t)$ and angle $\beta_k(t)$. This projected area of starspot k at time t is given by:

$$A_k(t) = \mathbb{R} \left[\cos^{-1}[\cos \alpha_k \csc \beta_k] + \cos \beta_k \sin \alpha_k \Xi_k - \cos \alpha_k \sin \beta_k \Psi_k \right] \quad (4.2)$$

where the intermediate terms are

$$\Xi_k = \sin \alpha_k \cos^{-1}[-\cot \alpha_k \cot \beta_k], \quad (4.3)$$

$$\Psi_k = \sqrt{1 - \cos^2 \alpha_k \csc^2 \beta_k} \quad (4.4)$$

Expanded out, $A(t)$ becomes a piecewise expression with three cases:

$$A_k(t) = \begin{cases} \pi \sin^2 \alpha_k \cos \beta_k & (0 < \beta_k < \pi/2 - \alpha_k) \\ \cos^{-1}[\cos \alpha_k \csc \beta_k] + \cos \beta_k \sin^2 \alpha_k \cos^{-1}[-\cot \alpha_k \cot \beta_k] \\ \quad - \cos \alpha_k \sin \beta_k \sqrt{1 - \cos^2 \alpha_k \csc^2 \beta_k} & (\pi/2 - \alpha_k < \beta_k < \pi/2 + \alpha_k) \\ 0 & (\pi/2 + \alpha_k < \beta_k < \pi), \end{cases} \quad (4.5)$$

where α_k is the angular radius of the spot and β_k is the angle between the spot normal and the line of sight. The three cases correspond to distinct geometric configurations of the spot on the stellar disk: fully visible, partially occulted by the limb, and fully hidden on the far side, as shown in [Figure 4.1](#).

The angle between spot normal and the line of sight for a single spot at a given time is:

$$\cos \beta(t) = \cos I \sin \Phi + \sin I \cos \Phi \cos \Lambda(t), \quad (4.6)$$

where I is the inclination of the stellar rotation axis, Φ is the latitude of the spot, and $\Lambda(t)$

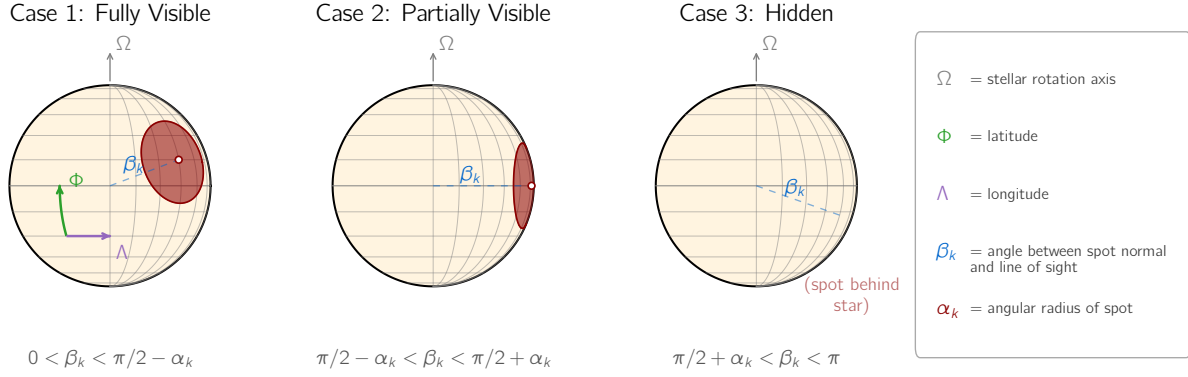


Figure 4.1: Illustration of the three geometric regimes for the projected area of a circular starspot on a rotating star. Each panel shows the stellar disk as seen by the observer, with latitude and longitude grid lines, the spot (dark red), the angular radius α_k , and the angle β_k between the spot normal and the line of sight. **Left:** The spot is fully visible on the near hemisphere. **Center:** The spot straddles the stellar limb and is partially occulted. **Right:** The spot is on the far side and hidden from view.

is the longitude of the spot as a function of time. The time dependence of $\beta(t)$ arises from the rotation of the star, which causes the spot to move in longitude and therefore change its angle relative to the line of sight. The projected area $A(t)$ is then determined by both $\alpha(t)$ and $\beta(t)$ according to the piecewise function in Eqn 4.5.

When the spot angular radius α_k is small compared to 1 radian (typical for Sun-like stars where spots subtend $\lesssim 6^\circ$), we can make the approximation $\sin \alpha_k \approx \alpha_k$, so $\sin^2 \alpha_k \approx \alpha_k^2$. In this limit:

- The **fully visible** case simplifies to $A_k \approx \pi \alpha_k^2 \cos \beta_k$.
- The **partial-visibility window** has width $2\alpha_k$ and becomes vanishingly narrow, so the spot transitions nearly instantaneously from fully visible to fully hidden as it crosses the limb.
- The **geometric projection effect** is purely $\cos \beta$ —the same projection factor as a flat disk viewed at angle β . Higher-order corrections from spherical curvature (the Ξ_k , Ψ_k terms) vanish at leading order in α_k .

This simplification allows the entire three-case piecewise function to be replaced with a single expression:

$$A(t) \approx \underbrace{\pi \alpha^2(t)}_{\Gamma(t)} \cdot \underbrace{\max\{\cos \beta(t), 0\}}_{\nu(t)}, \quad (4.7)$$

where $\alpha(t)$ is the angular radius of the spot (whose time evolution is described in Section 2.2). Figure 4.2 illustrates this approximation. The left panel shows the size of spots with three different angular radii $\alpha_{\max}$ on the stellar disk at $\beta = 0$ (disk center), providing a visual reference for the spot-to-star size ratio. The right panel compares the exact projected area $A(\beta)$ from Eq. (4.5) (solid lines) with the small-spot approximation $\max\{\cos \beta, 0\}$ (dashed

line). For small spots ($\alpha_{\max} \lesssim 15^\circ$), the exact and approximate curves are nearly indistinguishable, validating the factorization. For larger spots ($\alpha_{\max} \sim 30^\circ$), the transition at the limb ($\beta = \pi/2$) becomes more gradual in the exact solution, but the approximation remains reasonable.

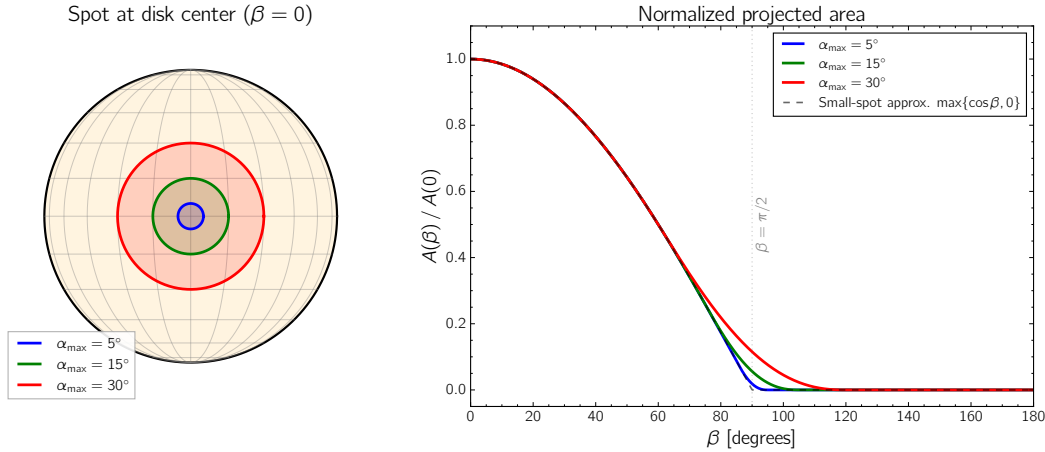


Figure 4.2: *Left:* Stellar disk viewed face-on, with circles indicating the projected size of spots with angular radii $\alpha_{\max} = 5^\circ$, 15° , and 30° centered at $\beta = 0$ (disk center). *Right:* Normalized projected area $A(\beta)/A(0)$ as a function of the angle β between the spot normal and the line of sight. Solid colored lines show the exact piecewise expression (Eq. 4.5) for each $\alpha_{\max}$; the dashed black line shows the small-spot approximation $\max\{\cos \beta, 0\}$ (Eq. 4.7). The two converge for $\alpha_{\max} \lesssim 15^\circ$.

The key consequence for the kernel derivation is that the projected area now *factorizes* into a product of two functions:

$$A(t) \propto \Gamma(t) \cdot \mathcal{V}(t), \quad (4.8)$$

where we define $\Gamma(t) \equiv \alpha^2(t)$ as the squared spot-size envelope, and $\mathcal{V}(t) \equiv \max\{\cos \beta(t), 0\}$ as the spot visibility function. Figure 4.3 illustrates this decomposition. Without this factorization, the time dependence of α and β would be too entangled for analytic Fourier analysis. The small-spot limit makes $A(t)$ decomposable, enabling the convolution-based approach developed in the following sections.

2.2 The Emergence and Decay Profile of Spots

In this work, we assume that the angular size of the spot, $\alpha(t)$, evolves over time as the spot emerges and decays. Here we test two different parameterizations of the spot evolution: a piecewise linear function and an exponential growth/decay function. In both cases, the spot reaches its maximum angular size $\alpha_{\max}$ at time t_{ref} , which is the reference time for the spot. The spot emerges over a timescale τ_{em} and decays over a timescale τ_{dec} .

Trapezoidal emergence/decay profile: Following Kipping (2012), we define a piecewise linear function defined by two timescales: Δt_1 for the emergence phase and Δt_2 for the decay phase. The spot reaches its maximum angular size $\alpha_{\max}$ at time t_{ref} , which is the

reference time for the spot. The Heaviside step function $H(x)$ is used to ensure that the spot only contributes to the flux deficit during its emergence and decay phases. The spot angular radius is given by:

$$\frac{\alpha_{\text{trap}}(t)}{\alpha_{\text{max}}} = \frac{1}{\tau_{\text{em}}} [\Delta t_1 H(\Delta t_1) - \Delta t_2 H(\Delta t_2)] - \frac{1}{\tau_{\text{dec}}} [\Delta t_3 H(\Delta t_3) - \Delta t_4 H(\Delta t_4)], \quad (4.9)$$

where

$$\begin{aligned} \Delta t_1 &= t - t_{\text{ref}} + \frac{\ell_{\text{spot}}}{2} + \tau_{\text{em}}, \\ \Delta t_2 &= t - t_{\text{ref}} + \frac{\ell_{\text{spot}}}{2}, \\ \Delta t_3 &= t - t_{\text{ref}} - \frac{\ell_{\text{spot}}}{2}, \\ \Delta t_4 &= t - t_{\text{ref}} - \frac{\ell_{\text{spot}}}{2} - \tau_{\text{dec}}. \end{aligned}$$

The emergence and decay durations (τ_{em} and τ_{dec}) are normalized such that $\tau_{\text{em}} = \int_{-\infty}^{\infty} \Delta t_1 H(\Delta t_1) - \Delta t_2 H(\Delta t_2) dt$ and $\tau_{\text{dec}} = \int_{-\infty}^{\infty} \Delta t_3 H(\Delta t_3) - \Delta t_4 H(\Delta t_4) dt$ ensure that the peak of the trapezoidal function is normalized to 1, so that $\alpha_{\text{trap}}(t)$ reaches α_{max} at its maximum. The trapezoidal profile is a simple, piecewise linear model that captures the general shape of spot emergence and decay, with a linear rise to maximum size followed by a linear decay back to zero.

Exponential emergence/decay profile: An alternative, model for the spot-size evolution (similar to [Aigrain et al. 2015](#)) uses exponential emergence and decay rather than linear ramps. The spot angular radius is:

$$\alpha_{\text{exp}}(t) = \alpha_{\text{max}} \times \begin{cases} e^{(t-t_1)/\tau_{\text{em}}} & t < t_1 \quad (\text{exponential rise}) \\ 1 & t_1 \leq t \leq t_2 \quad (\text{plateau}) \\ e^{-(t-t_2)/\tau_{\text{dec}}} & t > t_2 \quad (\text{exponential decay}) \end{cases} \quad (4.10)$$

where t_1 and $t_2 = t_1 + \tau_{\text{max}}$ mark the start and end of the plateau phase. The three parameters are: τ_{em} (emergence e -folding time), τ_{max} (plateau duration), and τ_{dec} (decay e -folding time). Without loss of generality we center the envelope by setting $t_1 = 0$, $t_2 = \tau_{\text{max}}$.

2.3 Latitudinal and Longitudinal Evolution of Spots

The position of a starspot on the surface is described by its latitude (Φ) and longitude (Λ). Similar to the Sun, the latitude of the spot is assumed to be constant over time, while the

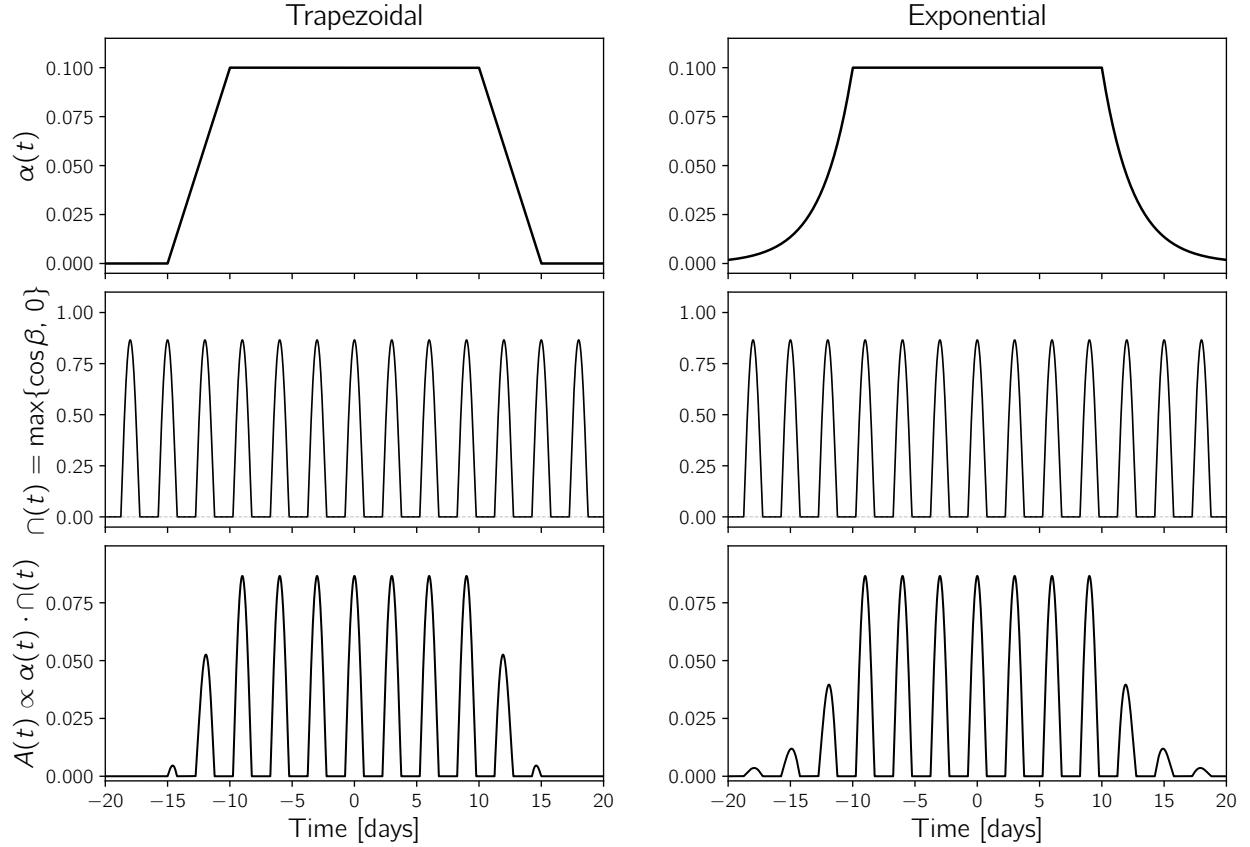


Figure 4.3: Decomposition of the single-spot projected area signal $A(t)$ into the spot-size evolution envelope $\alpha(t)$ (top) and the visibility function $\mathcal{V}(t) = \max\{\cos \beta(t), 0\}$ (middle), with their product $A(t) \propto \alpha(t) \cdot \mathcal{V}(t)$ (bottom). *Left:* trapezoidal envelope (Eq. 4.9). *Right:* exponential envelope (Eq. 4.10). Both use the same rotation period $P = 3$ d, spot latitude $\Phi = 30^\circ$, and edge-on inclination $I = 90^\circ$.

longitude evolves due to the rotation of the star:

$$\Phi = \Phi_{\text{ref}}, \quad (4.11)$$

$$\Lambda(t) = \Lambda_{\text{ref}} + \frac{2\pi(t - t_{\text{ref}})}{P(\Phi)}, \quad (4.12)$$

We assume a Sun-like differential rotation profile, where the rotation period P at latitude Φ_k depends on the latitudinal shear κ and the rotation period at the equator of the star P_{eq} :

$$P(\Phi_k) = \frac{P_{\text{eq}}}{1 - \kappa \sin^2 \Phi_k}, \quad (4.13)$$

where $\kappa = 0$ corresponds to solid body rotation and $\kappa > 0$ corresponds to solar-like differential rotation (equator rotates faster than poles), while $\kappa < 0$ corresponds to anti-solar differential rotation (poles rotate faster than equator). Or, in other words, κ quantifies the

relative difference in rotation rate between the equator and the poles, $\kappa = (\Omega_{\text{eq}} - \Omega_{\text{pole}})/\Omega_{\text{eq}}$ for equatorial rotation rate $\Omega_{\text{eq}} = 2\pi/P_{\text{eq}}$ and polar rotation rate $\Omega_{\text{pole}} = 2\pi/P_{\text{pole}}$. Figure 4.4 illustrates the effect of differential rotation on the rotational velocity at different latitudes on the stellar surface.

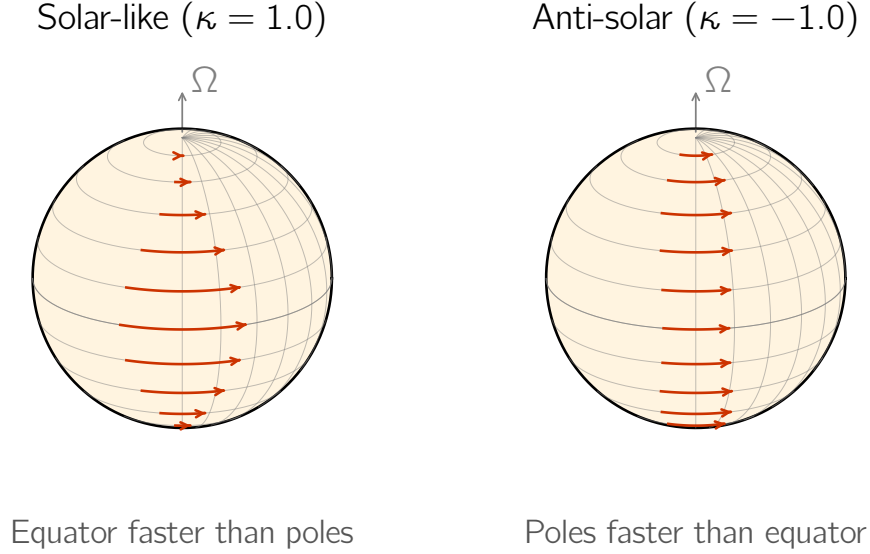


Figure 4.4: Illustration of differential rotation on a stellar sphere viewed at inclination $I = 70^\circ$. Arrows show the rotational velocity at each latitude, with length proportional to the angular velocity, $\omega(\Phi) = 2\pi/P(\Phi)$. *Left:* Solar-like differential rotation ($\kappa = 1.0$), where the equator rotates faster than the poles. *Right:* Anti-solar differential rotation ($\kappa = -1.0$), where the poles rotate faster than the equator.

2.4 Stellar Limb Darkening

Following Kipping (2012), we also add a quadratic limb darkening law for the stellar limb darkening:

$$F(t) = 1 - \sum_{n=0}^4 \left(\frac{nc_n}{n+4} \right) - \frac{1}{\pi} \sum_{k=1}^{N_{\text{spot}}} A_k(t) \cdot (1 - f_{\text{spot},k}). \quad (4.14)$$

In this work we implement this term into the numerical code, but not in the analytic kernel derivation. However, as we show in Section 9, the limb darkening term has a negligible effect on the kernel, and so we ignore including it in the analytic derivation for simplicity.

2.5 Summary of Parameters

Tables 4.1 and 4.2 summarize the physical parameters of the lightcurve model. Table 4.1 lists the stellar parameters that are held fixed for all spots on the star, while Table 4.2 lists the parameters that can vary for each individual spot.

<i>Star variables</i>	
P_{eq}	rotation period at equator
κ	latitudinal differential rotation shear
I	inclination angle
N_{spot}	total number of spots observed within the simulated time T_{obs}
l_{spot}	lifetime of each spot
τ_{spot}	emergence and decay timescale of each spot

Table 4.1: Stellar parameters and their definitions.

<i>Individual spot variables</i>	
k	index number for each spot
$A_k(t)$	projected area of star spot
$\beta_k(t)$	angle between spot normal and line of sight
$\alpha_k(t)$	angular size of spot (emerges and decays)
$\Lambda_k(t)$	starspot longitude ($0, 2\pi$)
$\Lambda_{\text{ref},k}$	reference longitude of spot ($0, 2\pi$)
$\Phi_{\text{ref},k}$	reference latitude of spot ($0, \pi$)
$t_{\text{ref},k}$	time when spot reaches maximum size
$\alpha_{\text{max},k}$	maximum angular size of spot
$f_{\text{spot},k}$	spot-star contrast ($0 = \text{black spot}, 1 = \text{no contrast}$)

Table 4.2: Individual spot parameters and their definitions.

Figures 4.5 and 4.6 show examples of lightcurves generated from the starspot model, with contributions from each individual spot shown in color and the cumulative flux deficit shown in black.

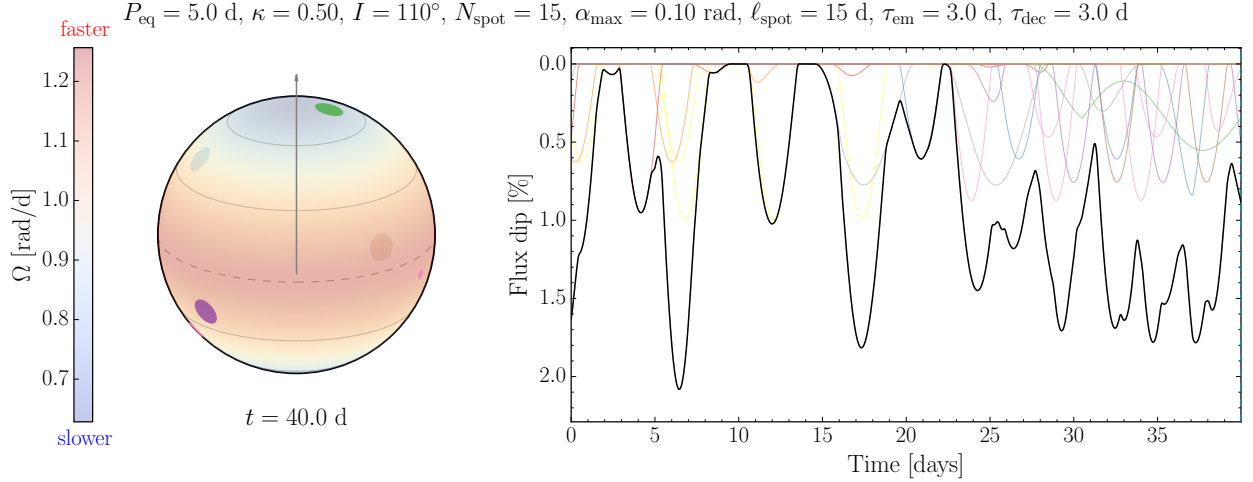


Figure 4.5: Examples of simulated lightcurves generated from the starspot model from the Kipping (2012) model. The top panel shows the case with solar-like differential rotation ($\kappa = 0.5$), while the bottom panel shows the case with anti-solar differential rotation ($\kappa = -0.5$). The contributions from each individual spot are shown in color, while the cumulative flux deficit is shown in black. The lightcurve exhibits quasi-periodic variability due to the rotation of the star and the emergence and decay of spots.

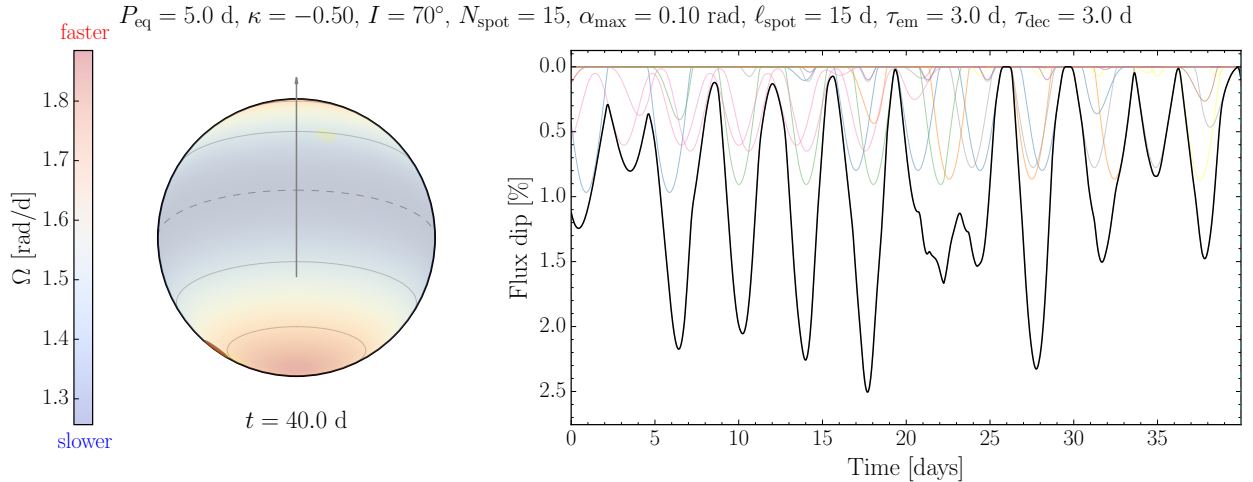


Figure 4.6: Same as Figure 4.5, but showing anti-solar differential rotation ($\kappa = -0.5$).

3 Gaussian Process

3.1 Motivation for Gaussian Processes

In practice, extracting the physical parameters of a star from its lightcurve is challenging due to the complex and stochastic nature of starspot-induced variability. The observed lightcurve is a superposition of signals from multiple spots, each emerging at different times

and locations on the stellar surface. Even with very high quality data, performing model inversion to infer the underlying spot properties is inherently challenging for at least several reasons:

- **Degeneracies:** Multiple combinations of spot properties can produce similar lightcurves, leading to degeneracies in the inferred parameters (Figure 4.7).
- **Non-uniqueness:** The same lightcurve can be explained by different spot configurations, making it difficult to uniquely determine the spot properties (Figure 4.8).
- **Noise and systematics:** Real lightcurves are affected by instrumental noise, stellar granulation, and other sources of variability that can obscure the starspot signal and further complicate the inference of spot properties.

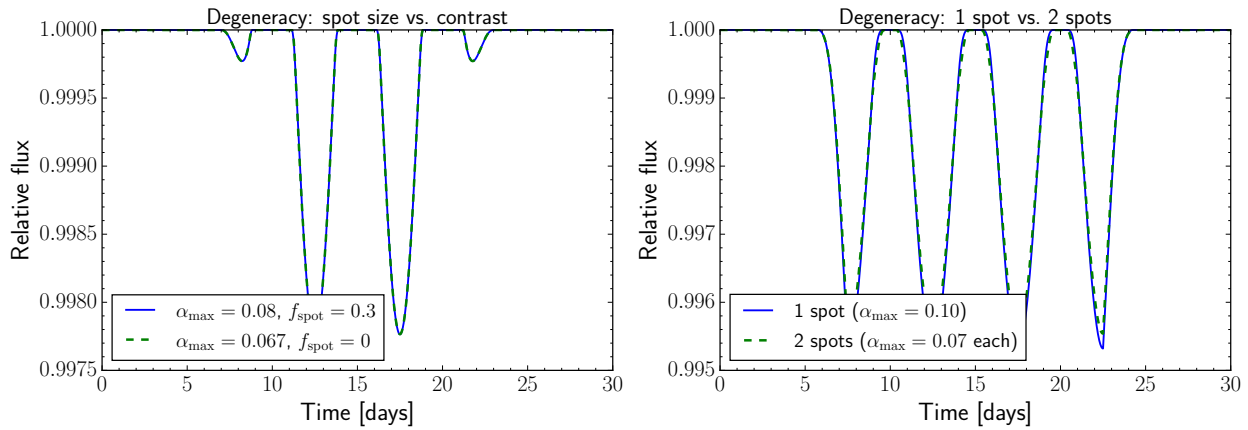


Figure 4.7: Illustration of degeneracies in starspot lightcurve modeling.

Given these challenges, a more practical approach is to not just model the properties of *individual* spots, but to instead model the *statistical properties* of the starspot population as a whole. This is where Gaussian processes (GPs) come in: they provide a flexible and powerful framework for modeling the covariance structure of the lightcurve without needing to explicitly model each individual spot. By deriving a physically motivated kernel function that captures the key features of starspot variability, we can use GPs to infer the underlying physical parameters of the star and its spots from the observed lightcurve, while marginalizing over the complex and degenerate spot configurations. This approach allows the extraction of meaningful information about the star’s rotation period, differential rotation, spot lifetimes, and emergence/decay timescales as a distribution over the entire spot population, rather than the properties of each individual spot.

3.2 Mathematical Formulation of Gaussian Processes

Formally, a Gaussian process is defined by a multivariate Gaussian:

$$\mathcal{GP}(\mathbf{x}) \sim \mathcal{N}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (4.15)$$

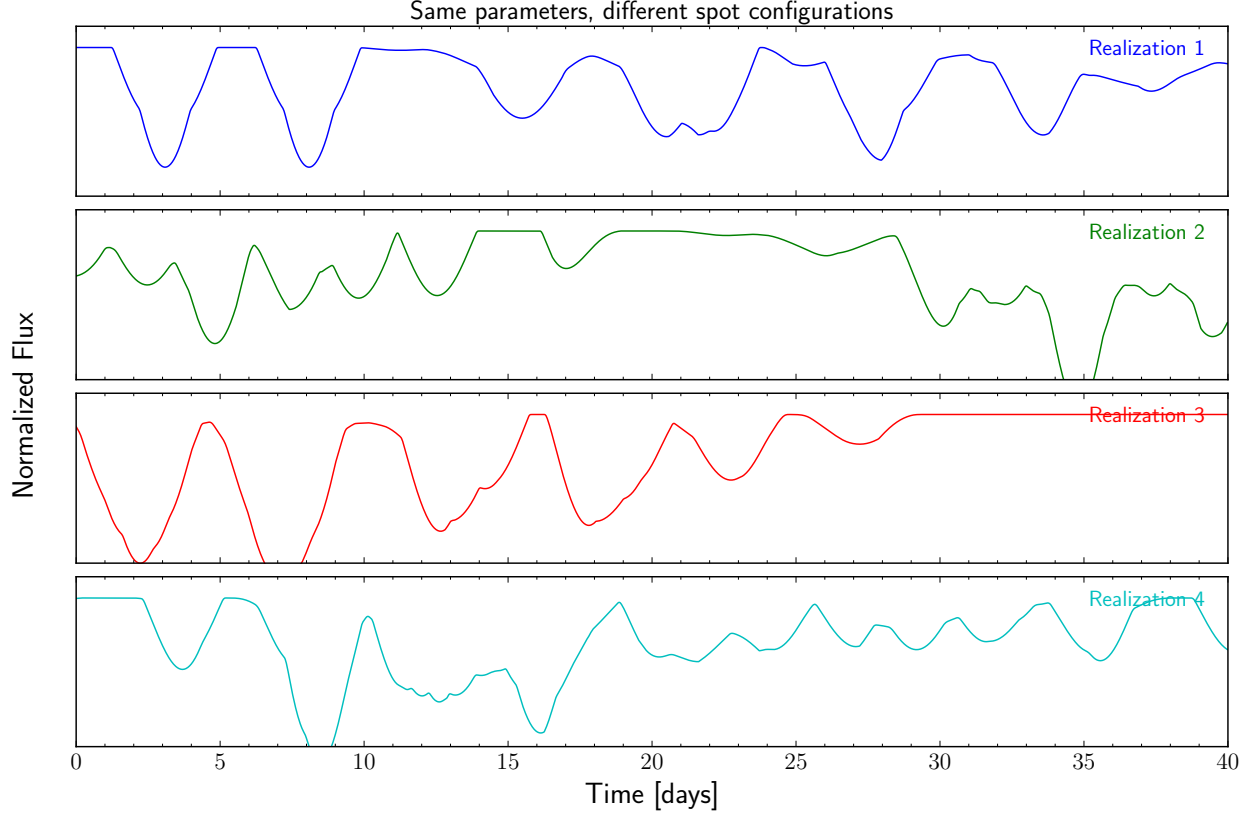


Figure 4.8: Illustration of non-uniqueness in starspot lightcurve modeling. All lightcurves have the same parameters: $P_{\text{eq}} = 5 \text{ d}$, $\kappa = 0$, $I = 80^\circ$, $N_{\text{spot}} = 10$, $l_{\text{spot}} = 8 \text{ d}$, $\tau_{\text{spot}} = 3 \text{ d}$, yet the resulting lightcurves differ due to the random positions and emergence times of individual spots.

where the mean function, $m(\mathbf{x})$, represents the expected value of the function at each input point $\mathbf{x}$, while the covariance function, $k(\mathbf{x}, \mathbf{x}')$ models the correlation between function values at different input points $\mathbf{x}$ and $\mathbf{x}'$, quantifying the smoothness and spatial properties of the underlying function.

This GP, defined by a mean function and a covariance function, represents a probability distribution over functions and can be used to model the lightcurve data. Given some lightcurve data $(\mathbf{x}, \mathbf{y})$, the log marginal likelihood (or evidence) of a GP is given by:

$$\log P(\mathbf{y}|\mathbf{x}, \theta) = -\frac{1}{2}\mathbf{y}^T K^{-1}\mathbf{y} - \frac{1}{2}\log |K| - \frac{n}{2}\log 2\pi \quad (4.16)$$

This distribution enables predictions at new test points, where the predictive mean and variance at a given test point x^* can be computed as:

$$\mu = k(x_*, \mathbf{x})k(\mathbf{x}, \mathbf{x})^{-1}\mathbf{y} \quad (4.17)$$

$$\sigma^2 = k(x_*, x_*) - k(x_*, \mathbf{x})k(\mathbf{x}, \mathbf{x})^{-1}k(\mathbf{x}, x_*) \quad (4.18)$$

For this particular application, the mean function $m(t)$ is determined by the expected flux

deficit from the average spot coverage, which we derive analytically in Section 4. In practice, we subtract the mean from the observed data and set $m(t) = 0$. The covariance function $k(\mathbf{x}, \mathbf{x}')$ derived in Section 5 encodes the physical properties of the star and its spots. By maximizing the log marginal likelihood with respect to the hyperparameters θ of the kernel, we can infer the underlying physical parameters of the star and its spots from the observed lightcurve data. For simplification we assume that the distribution of starspot properties does not evolve with time. In other words, the kernel is *stationary* and depends only the time lag τ between two points. With these basic assumptions, we move on to how to compute the kernel function numerical and analytically. However, for a more comprehensive introduction to Gaussian processes and their properties, see [Rasmussen & Williams \(2006\)](#).

3.3 The Covariance Matrix and Kernel Function

A brute-force approach to modeling the covariance structure of the lightcurve is to directly compute the covariance matrix K from a large sample of simulated lightcurves generated from the starspot model. Assuming we have 1-dimensional training data, where $\mathbf{x} = t_1, \dots, t_n$ represent the times and $\mathbf{y} = F_1, \dots, F_n$ represent the fluxes of the lightcurve, then we want to solve for covariance matrix:

$$K = \begin{bmatrix} \text{Var}(F_1) & \cdots & \text{Cov}(F_1, F_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(F_n, F_1) & \cdots & \text{Var}(F_n) \end{bmatrix}. \quad (4.19)$$

Each element of the covariance matrix K is the equivalent to the covariance between F_i and F_j :

$$K_{ij} = \text{Cov}(F_i, F_j). \quad (4.20)$$

The first term is the expectation of the flux product at two different times t and $t + \tau$. Flux at a given time is given by the [Kipping \(2012\)](#) lightcurve model (Eqn. 4.1). The random variables of this process are the reference time, latitude, and longitude of each spot ($t_{\text{ref},k}$, $\Lambda_{\text{ref},k}$, and $\Phi_{\text{ref},k}$, respectively).

The covariance matrix can then be estimated over a large sample of simulations with random variables (e.g., a uniform distribution with $t_{\text{ref},k} \in [0, T]$, $\Lambda_{\text{ref},k} \in [-\pi/2, \pi/2]$, and $\Phi_{\text{ref},k} \in [0, 2\pi]$). The covariance can then be estimated as the sample covariance of the fluxes at two different times t and $t + \tau$:

$$\mathcal{K}_{\text{num}}(\tau) = \text{Cov}[F(t), F(t + \tau)] \quad (4.21)$$

$$= \frac{1}{N-1} \sum_{i=1}^N (F(t_i) - \mathbb{E}[F])(F(t_i + \tau) - \mathbb{E}[F]). \quad (4.22)$$

In practice, the numerical approach is computationally expensive, since it requires generating a large number of lightcurves and computing the covariance for each pair of time points. Moreover, it does not provide a closed-form expression for the kernel function, which limits its interpretability and makes it difficult to use in a Gaussian process regression framework.

In the following sections, we first derive the mean function analytically (Section 4) and then derive an analytical expression for the kernel $\mathcal{K}(\tau)$ (Section 5), using the numerical estimate for validation.

4 Deriving the Mean Function

A Gaussian process is fully specified by its mean function $m(t) = \mathbb{E}[F(t)]$ and covariance function $\mathcal{K}(\tau)$. Before deriving the kernel in Section 5, we compute the mean function analytically from the flux model.

We model spot emergence as a stationary process with rate $\dot{N}_{\text{spot}}$ [spots day⁻¹], so that each spot contributes a finite-duration flux pulse. For identical spots with contrast f_{spot} , the expected flux deficit at any instant is the emergence rate times the time-integrated flux deficit of a single spot:

$$\mathbb{E}[\delta F] = \frac{\dot{N}_{\text{spot}} (1 - f_{\text{spot}})}{\pi} \int_{-\infty}^{\infty} \mathbb{E}_{\Lambda, \Phi}[A(t)] dt. \quad (4.23)$$

The integral converges because each spot has finite support through $\Gamma(t)$, and t_{ref} does not appear: a shift in spot emergence time is equivalent to a phase rotation of the longitude, which is already uniformly distributed.

Using the small-spot factorization (Eq. 4.7), $A(t) = \pi \alpha_{\text{max}}^2 \Gamma(t) \mathcal{V}(t; \Lambda_{\text{ref}}, \Phi)$, we average over Λ_{ref} inside the time integral. Since Λ_{ref} is uniform on $[0, 2\pi]$, the longitude-averaged visibility is constant in time:

$$\bar{\mathcal{V}}(\Phi, I) \equiv \langle \mathcal{V}(t) \rangle_{\Lambda} = \frac{1}{2\pi} \int_0^{2\pi} \max\{b_0 + b_1 \cos(\omega_0 t - \Lambda_{\text{ref}}), 0\} d\Lambda_{\text{ref}} = \frac{b_0 \theta_{\text{vis}} + b_1 \sin \theta_{\text{vis}}}{\pi}, \quad (4.24)$$

where $b_0 = \cos I \sin \Phi$, $b_1 = \sin I \cos \Phi$, and $\theta_{\text{vis}} = \arccos(-b_0/b_1)$. This is exact: averaging over the initial longitude phase is equivalent to averaging over one full rotation period, so $\bar{\mathcal{V}}$ is independent of t regardless of the spot lifetime. In the Fourier analysis of Section 5.3, $\bar{\mathcal{V}}$ will be identified as the zeroth Fourier coefficient $c_0(\Phi)$ of the visibility function.

Since $\bar{\mathcal{V}}$ is constant, it factors out of the time integral, and the remaining integral acts only on the spot-size envelope:

$$\int_{-\infty}^{\infty} \langle A(t) \rangle_{\Lambda} dt = \pi \alpha_{\text{max}}^2 \bar{\mathcal{V}}(\Phi, I) \int_{-\infty}^{\infty} \Gamma(t) dt. \quad (4.25)$$

Averaging over the latitude distribution $p(\Phi)$ and combining, the expected flux for a general envelope $\Gamma(t)$ is:

$$\mathbb{E}[F] = 1 - \dot{N}_{\text{spot}} (1 - f_{\text{spot}}) \alpha_{\text{max}}^2 \left(\int_{-\infty}^{\infty} \Gamma(t) dt \right) \int_{-\pi/2}^{\pi/2} \bar{\mathcal{V}}(\Phi, I) p(\Phi) d\Phi \quad (4.26)$$

This is an intrinsic stellar property, independent of both the observing baseline T_{obs} and

the rotation parameters P_{eq} and κ , since the time-averaged visibility $\bar{\mathcal{V}}$ depends only on the inclination and latitude geometry. The only envelope-dependent quantity is $\int \Gamma(t) dt$, which has a closed-form expression for each profile considered in this work. Note the contrast with the kernel amplitude (Section 5.5), which involves $R_{\Gamma}(0) = \int \Gamma^2(t) dt$: the mean depends on the first moment of the envelope, while the variance depends on the second.

4.0.1 Reparameterization in terms of a_{spot}

We define the **spot contrast-area product**:

$$a_{\text{spot}} \equiv (1 - f_{\text{spot}}) \alpha_{\text{max}}^2, \quad (4.27)$$

which collapses the degenerate pair $(f_{\text{spot}}, \alpha_{\text{max}})$ into a single parameter controlling the flux deficit per spot. In terms of $(\dot{N}_{\text{spot}}, a_{\text{spot}})$, the mean flux (Eq. 4.26) becomes:

$$\mathbb{E}[F] = 1 - \dot{N}_{\text{spot}} a_{\text{spot}} \left(\int_{-\infty}^{\infty} \Gamma(t) dt \right) \int_{-\pi/2}^{\pi/2} \bar{\mathcal{V}}(\Phi, I) p(\Phi) d\Phi. \quad (4.28)$$

As we show in Section 5.7, the kernel amplitude can also be written in terms of these same parameters as $\sigma_k^2 = \dot{N}_{\text{spot}} a_{\text{spot}}^2$. The mean depends on the product $\dot{N}_{\text{spot}} a_{\text{spot}}$ while the kernel depends on $\dot{N}_{\text{spot}} a_{\text{spot}}^2$, so fitting both the mean and kernel simultaneously constrains $\dot{N}_{\text{spot}}$ and a_{spot} independently—breaking the amplitude degeneracy that σ_k alone cannot resolve.

For the **trapezoidal** profile, $\Gamma(t) = [\alpha(t)/\alpha_{\text{max}}]^2$ has parabolic ramps, giving

$$\int_{-\infty}^{\infty} \Gamma(t) dt = \ell_{\text{spot}} + \frac{2\tau_{\text{spot}}}{3}, \quad (4.29)$$

where the plateau contributes ℓ_{spot} and each ramp contributes $\int_0^{\tau_{\text{spot}}} (s/\tau_{\text{spot}})^2 ds = \tau_{\text{spot}}/3$. For the **exponential** profile (Eq. 4.10), the corresponding integral is $\tau_{\text{em}}/2 + \tau_{\text{max}} + \tau_{\text{dec}}/2$.

For edge-on viewing ($I = \pi/2$) with a uniform latitude distribution $p(\Phi) = 1/\pi$, the mean visibility reduces to $\bar{\mathcal{V}} = \cos \Phi/\pi$ and the latitude integral evaluates to $2/\pi^2$. Specializing to the trapezoidal envelope:

$$\mathbb{E}[F] = 1 - \frac{2 \dot{N}_{\text{spot}} a_{\text{spot}}}{\pi^2} \left(\ell_{\text{spot}} + \frac{2\tau_{\text{spot}}}{3} \right) \quad (I = \pi/2). \quad (4.30)$$

The mean flux deficit is typically small ($\lesssim 1\%$ for Sun-like stars), and observed lightcurves are generally mean-subtracted before fitting. We therefore set the GP mean function to $m(t) = 0$ and model the mean-subtracted residuals. The covariance function $\mathcal{K}(\tau)$ derived in the following section is the autocovariance of these residuals.

5 Deriving the Analytical Kernel for Starspot Variability

The goal of this section is to derive an analytical expression for the kernel function $\mathcal{K}(\tau)$ that describes the covariance of the flux at two different times t and $t + \tau$ due to starspot variability. We first start with the analytic flux model defined in Equation 4.1 and then derive the kernel function by computing the autocorrelation of the flux signal.

The result is a physically motivated kernel function that captures the key features of starspot-induced variability, such as the rotation period, spot lifetime, and emergence/decay timescales. The kernel function depends on the underlying physical properties of the star and its spots, which are encapsulated in the hyperparameters:

$$\theta = \{P_{\text{eq}}, \kappa, I, N_{\text{spot}}, l_{\text{spot}}, \tau_{\text{spot}}, \alpha_{\text{max}}\}.$$

This kernel can be used in a Gaussian process regression framework to model and predict stellar variability, while also allowing for inference on the physical properties of the star and its spots through the hyperparameters. In the following sections, we give an overview of the Fourier transform notation and identities used in the derivation, and then we present the step-by-step derivation of the kernel function $\mathcal{K}(\tau)$ from the flux model.

5.1 Fourier Transform Notation and Identities

We use the Fourier transform convention $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$ with inverse $f(t) = (2\pi)^{-1} \int \hat{f}(\omega) e^{i\omega t} d\omega$ throughout. Table 4.3 summarizes the notation.

The derivation relies on three standard results. First, a periodic function with period T and Fourier coefficients $c_n = T^{-1} \int_0^T f(t) e^{-in\omega_0 t} dt$ (where $\omega_0 = 2\pi/T$) has a Fourier transform given by the weighted Dirac comb $\hat{f}(\omega) = 2\pi \sum_n c_n \delta(\omega - n\omega_0)$. Second, the convolution theorem: convolution in the time domain corresponds to multiplication in the frequency domain (FT4), and a product in the time domain becomes a convolution in the frequency domain (FT5). Third, the Wiener–Khinchin theorem: the autocorrelation $R_f(\tau) = \int f(t) f(t + \tau) dt$ is the inverse Fourier transform of the power spectral density $S_f(\omega) = |\hat{f}(\omega)|^2$:

$$R_f(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 e^{i\omega\tau} d\omega. \quad (4.31)$$

Together, these definitions allow us to obtain the autocovariance kernel $\mathcal{K}(\tau)$ from the Fourier transform of the single-spot signal $A(t)$. Table 4.4 lists the Fourier transform properties and Table 4.5 lists the specific transform pairs used in this paper.

5.2 Approach to Derivation

The autocorrelation function of function $A(t)$ is equivalent to taking the convolution between $A(t)$ and a lagged version of the function. By the convolution property (FT4), convolution in the time domain is multiplication in the frequency domain. Let $\mathcal{F}\{A(t)\} = \hat{A}(\omega)$ denote the Fourier transform of a function $A(t)$, and $\mathcal{F}^{-1}\{\hat{A}(\omega)\} = A(t)$ denote the inverse Fourier

Table 4.3: Notation used in Section 5.

Symbol	Description
t	Time
τ	Time lag
ω	Angular frequency
ω_0	Fundamental angular frequency, $2\pi/T$
T	Period of a periodic function
n	Harmonic index (integer)
$f(t), g(t)$	Functions in the time domain
$\hat{f}(\omega), \hat{g}(\omega)$	Fourier transforms of f and g
c_n	Fourier series coefficients
$\delta(\cdot)$	Dirac delta function
$(f * g)(t)$	Convolution of f and g
$R_f(\tau)$	Autocorrelation function of f
$\varepsilon(\omega)$	Energy spectral density, $ \hat{f}(\omega) ^2$
$k(\tau)$	Autocovariance kernel
$A(t)$	Single-spot flux signal
$\text{rect}_W(t)$	Rectangular pulse of width W
$\text{sinc}(x)$	Sinc function, $\sin(x)/x$
$\text{sgn}(\cdot)$	Sign function
a, b	Arbitrary constants
t_0	Fixed time shift

Table 4.4: Properties of the Fourier transform. We use the convention $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$. Here a, b are constants, t_0 and ω_0 are fixed shifts, and f, g are functions with transforms $\hat{f}, \hat{g}$.

No.	Property	Time domain $\longleftrightarrow$ Frequency domain
FT1	Linearity	$a f(t) + b g(t) \longleftrightarrow a \hat{f}(\omega) + b \hat{g}(\omega)$
FT2	Time shifting	$f(t - t_0) \longleftrightarrow e^{-i\omega t_0} \hat{f}(\omega)$
FT3	Frequency shifting	$e^{i\omega_0 t} f(t) \longleftrightarrow \hat{f}(\omega - \omega_0)$
FT4	Convolution	$(f * g)(t) \longleftrightarrow \hat{f}(\omega) \cdot \hat{g}(\omega)$
FT5	Multiplication	$f(t) \cdot g(t) \longleftrightarrow \frac{1}{2\pi} \hat{f}(\omega) * \hat{g}(\omega)$
FT6	Even symmetry	$f(t)$ real and even $\implies \hat{f}(\omega)$ purely real
FT7	Odd symmetry	$f(t)$ real and odd $\implies \hat{f}(\omega)$ purely imaginary

transform. So the Fourier transform of the kernel function is:

$$\mathcal{F}\{\mathcal{K}(\tau)\} = A(t) * A(t - \tau) \quad (4.32)$$

$$= \hat{A}(\omega) \cdot \hat{A}(\omega) \quad (4.33)$$

$$= |\hat{A}(\omega)|^2 \quad (4.34)$$

$$= S_A(\omega) \quad (4.35)$$

Table 4.5: Fourier transform pairs used in this work. We use the convention $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$ and $\text{sinc}(x) \equiv \sin(x)/x$.

No.	Function	$f(t)$	$\hat{f}(\omega)$
P1	Rectangular pulse	$\text{rect}_W(t) = \begin{cases} 1 & t \leq W/2 \\ 0 & \text{otherwise} \end{cases}$	$W \text{sinc}\left(\frac{\omega W}{2}\right)$
P2	Cosine	$\cos(\omega_0 t + \phi)$	$\pi [e^{i\phi} \delta(\omega - \omega_0) + e^{-i\phi} \delta(\omega + \omega_0)]$
P3	Complex exponential	$e^{i\omega_0 t}$	$2\pi \delta(\omega - \omega_0)$
P4	Square wave	$\text{sgn}(\cos \omega_0 t)$	$2\pi \sum_{n \text{ odd}} \frac{(-1)^{(n-1)/2}}{n} \delta(\omega - n\omega_0)$
P5	Dirac delta	$\delta(t - t_0)$	$e^{-i\omega t_0}$
P6	Dirac comb	$\sum_n \delta(t - nT)$	$\frac{2\pi}{T} \sum_n \delta(\omega - n\omega_0), \quad \omega_0 = 2\pi/T$

The term $S_A(\omega)$ is the *power spectral density* of the single-spot signal. By the Wiener–Khinchin theorem (Eq. 4.31), the autocorrelation function is the inverse Fourier transform of the power spectral density:

$$\mathcal{K}(\tau) \xleftrightarrow{\mathcal{F}} S_A(\omega) \quad (4.36)$$

$$\mathcal{K}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_A(\omega) e^{i\omega \tau} d\omega. \quad (4.37)$$

Hence to solve for $\mathcal{K}(\tau)$, we need to first solve for the Fourier transform of a single spot, $\hat{A}(\omega) = \mathcal{F}\{A(t)\}$. Referring back to Equation 4.7, ignoring the edge cases (for small spot sizes), the drop in brightness for a single spot is the product of the squared spot-size envelope $\Gamma(t) = \alpha^2(t)$ and the visibility function $\mathcal{V}(t) = \max\{\cos \beta(t), 0\}$. The visibility function is a truncated cosine, which can be written as the product of a square wave (P4) and cosine wave (P2): $\mathcal{V}(t) = \Pi(t) \cdot C(t)$. By the multiplication property (FT5), a product in the time domain becomes a convolution in the frequency domain:

$$\mathcal{F}\{A(t)\} = \mathcal{F}\{\Gamma(t) \cdot \mathcal{V}(t)\} \quad (4.38)$$

$$= \mathcal{F}\{\Gamma(t)\} * \mathcal{F}\{\mathcal{V}(t)\} \quad (4.39)$$

$$\hat{A}(\omega) = \hat{\Gamma}(\omega) * \hat{\mathcal{V}}(\omega) \quad (4.40)$$

As we show in the next section, the Fourier transform of the spot evolution $\hat{A}(\omega)$ can be solved analytically for both of the components $\hat{\Gamma}(\omega)$ and $\hat{\mathcal{V}}(\omega)$. In Section 5.3 we derive the Fourier transform of the truncated cosine $\hat{\mathcal{V}}(\omega)$ (representing the variability as the star rotates), and in Section 5.4 we derive the Fourier transform of the squared spot-size envelope $\hat{\Gamma}(\omega)$ (representing the growth and decay of the spot). Finally, in Section 5.5 we combine these results to obtain the final expression for the kernel $\mathcal{K}(\tau)$.

Figure ?? provides an overview of the derivation strategy. The analytic path (top row)

derives $\mathcal{K}(\tau)$ through Fourier analysis, while a parallel numerical path (bottom row) estimates the kernel via Monte Carlo simulation. Finally, we compare the numerical and analytical kernels as a validation test of the derivation.

5.3 Fourier Transform: The Rotation Term $\mathcal{V}(t)$

As given in Equation 4.6, the drop in flux due to a spot rotating into and out of view is determined by the angle β between the spot normal and the line of sight. The spot is visible when $\cos \beta > 0$, and the visibility function is $\mathcal{V}(t) = \max\{\cos \beta(t), 0\}$. The angle β depends on the spot latitude Φ , the stellar inclination I , and the rotation period P of the star.

Figure 4.9 shows how the signal varies as a function of spot latitude Φ and stellar inclination I . For a star viewed equator-on ($I = 90^\circ$, left panel of Figure 4.9), the visibility function looks like a half-truncated cosine wave, where the spot rotates into view for half a rotation period and then rotates out of view for the other half. The different colors show the relative change in flux for spots at different latitudes Φ , with spots at higher latitudes having smaller amplitudes due to their smaller projected areas.

Figure 4.9 also shows how the signal varies as a function of stellar inclination I for a fixed spot latitude $\Phi = 60^\circ$ (right panel). For a star viewed equator-on ($I = 90^\circ$), the visibility function is a half-truncated cosine wave (red). For a star viewed at intermediate inclination ($60^\circ < I < 90^\circ$), the visibility function is a semi-truncated cosine wave (blue). For a star viewed closer to pole-on ($I = 30^\circ$), the visibility function is a full cosine wave which never rotates out of view, but varies in amplitude due to the changing projected area of the spot (orange). Finally, for a star viewed pole-on ($I = 0^\circ$), the visibility function is a constant (blue) since the spot is always visible with the same projected area.

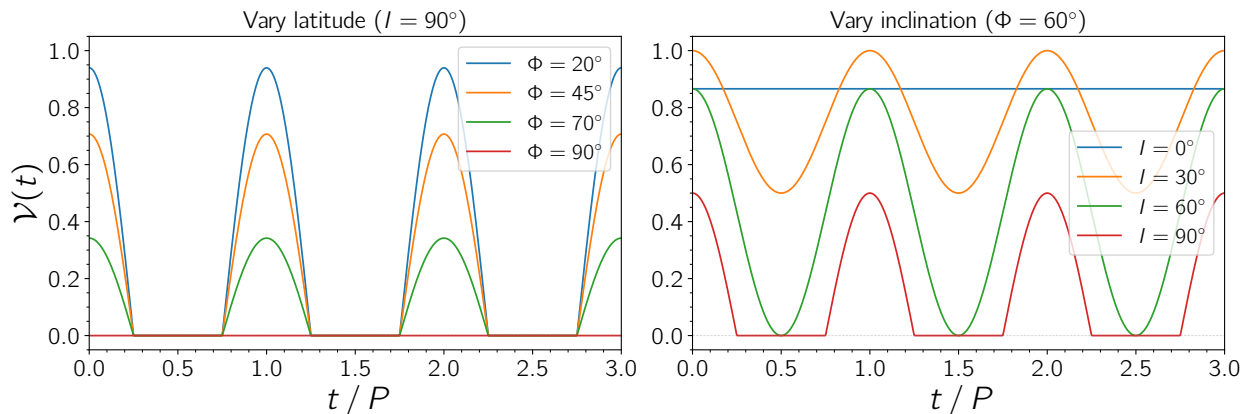


Figure 4.9: Visibility function $\mathcal{V}(t) = \max\{\cos \beta(t), 0\}$ over three rotation periods. *Left:* varying spot latitude Φ at fixed edge-on inclination $I = 90^\circ$. *Right:* varying stellar inclination I at fixed latitude $\Phi = 60^\circ$. The amplitude and duty cycle of $\mathcal{V}(t)$ set the Fourier coefficients $|c_n|^2$ that enter the kernel.

The visibility function $\mathcal{V}(t) = \max\{\cos \beta(t), 0\}$ is periodic with the rotation period $P =$

$2\pi/\omega_0$, so it has a discrete Fourier series (Section 5):

$$\mathcal{V}(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega_0 t}, \quad (4.41)$$

and its Fourier transform is a summation over Dirac functions (P3):

$$\hat{\mathcal{V}}(\omega) = 2\pi \sum_{n=-\infty}^{\infty} c_n \delta(\omega - n\omega_0). \quad (4.42)$$

Solution for the edge-on case: For the edge-on case ($I = 90^\circ$), $\cos \beta(t) = \cos \Phi \cos(\omega_0 t)$, and the visibility function reduces to $\mathcal{V}(t) = \cos \Phi \cdot \max\{\cos(\omega_0 t), 0\}$. The Fourier coefficients of the half-rectified cosine $\max\{\cos \theta, 0\}$ are (see Appendix 1 for the full derivation):

$$c_n = \begin{cases} \cos \Phi / \pi & n = 0 \\ \cos \Phi / 4 & n = \pm 1 \\ \frac{(-1)^{|n|/2+1} \cos \Phi}{\pi(n^2 - 1)} & n = \pm 2, \pm 4, \pm 6, \dots \\ 0 & n = \pm 3, \pm 5, \pm 7, \dots \end{cases} \quad (4.43)$$

All coefficients are real and satisfy $c_{-n} = c_n$, as expected for a real, even function.

Solution for general inclinations: For general inclination, $\cos \beta(t) = b_0 + b_1 \cos(\omega_0 t)$ where $b_0 = \cos I \sin \Phi$ and $b_1 = \sin I \cos \Phi$. The spot is visible when $|\omega_0 t| < \theta_{\text{vis}} = \arccos(-b_0/b_1)$, and the general Fourier coefficients (derived in Appendix 1) are:

$$c_n = \frac{1}{2\pi} \int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} (b_0 + b_1 \cos \theta) e^{-in\theta} d\theta = \frac{1}{\pi} \left[\frac{b_0 \sin(n\theta_{\text{vis}})}{n} + \frac{b_1}{2} \left(\frac{\sin((n-1)\theta_{\text{vis}})}{n-1} + \frac{\sin((n+1)\theta_{\text{vis}})}{n+1} \right) \right] \quad (4.44)$$

with $c_0 = (b_0 \theta_{\text{vis}} + b_1 \sin \theta_{\text{vis}})/\pi$ and $c_{\pm 1} = (b_0 \sin \theta_{\text{vis}} + b_1[\theta_{\text{vis}} + \sin \theta_{\text{vis}} \cos \theta_{\text{vis}}]/2)/(2\pi)$ obtained by taking the appropriate limits.

In general, it is difficult to solve for a closed-form solution to the integral of the full kernel function in Equation 4.70 due to the complexity of the latitude-dependent terms. However, for specialized cases (e.g., a star with face on inclination, solid body rotation, and uniform latitude spots), the kernel reduces to a simple, closed-form solution. Truncating to second harmonic order, the Fourier transform of the visibility function is:

$$\hat{\mathcal{V}}(\omega) \approx 2\pi [c_0 \delta(\omega) + c_1 [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] + c_2 [\delta(\omega - 2\omega_0) + \delta(\omega + 2\omega_0)]] . \quad (4.45)$$

As we will show in Section 5.5, the kernel $\mathcal{K}(\tau)$ is dominated by the first two harmonics, so this approximation is sufficient for our purposes (though in our `spotgp` implementation, this Fourier series can be expanded to an arbitrary number of terms as specified by the user).

5.4 Fourier Transform: Spot Evolution Envelope Term $\Gamma(t)$

In this section we derive the Fourier transform of the spot evolution envelope, which captures the growth and decay of the spot over time. Here we present several common spot evolution profiles, but the same approach can be applied to any arbitrary profile $\alpha(t)$ by performing the appropriate integral. These profiles include: (1) squared trapezoidal (Kipping, 2012), (2) exponential (Aigrain et al., 2015), and (3) skewed normal (Forgács-Dajka et al., 2021). The trapezoidal and exponential profiles are simple functions that produce an analytic kernel function. On the other hand, the skewed normal profile is a more flexible profile that can capture asymmetric growth and decay timescales and reproduce the observed spot lifetimes of the Sun (Forgács-Dajka et al., 2021).

In the Results and Discussion (Sections 9–11) we show a detailed analysis of how the properties of the trapezoidal profile translate to the structure of the kernel function, as its analytic solution is easily interpretable. While perhaps not the most realistic model, the trapezoidal parameterization provides an intuitive insight into how the kernel behaves in different limits. We include the implementations for all three kernels in our `spotgp` implementation (as well as the flexibility to define custom envelope functions) so that users can experiment with different assumptions for spot lifetimes and decay timescales.

(1) Squared trapezoidal profile for $\Gamma(t)$: In the first example, we consider the trapezoidal profile (Equation 4.9) from Kipping (2012). Appendix 2, shows how we solve for the Fourier transform of the normalized squared envelope $\Gamma(t) = \alpha^2(t)/\alpha_{\max}^2$: Since $\alpha(t)$ is a trapezoid with peak $\alpha_{\max}$, the normalized squared envelope $\Gamma(t) = \alpha^2(t)/\alpha_{\max}^2$ has unit peak height, the same support $\ell_{\text{spot}} + 2\tau_{\text{spot}}$, the same plateau duration ℓ_{spot} , but parabolic (rather than linear) ramps. Since $\Gamma(t)$ is real and even, its Fourier transform is purely real and can be computed via the cosine transform $\hat{\Gamma}(\omega) = 2 \int_0^\infty \Gamma(t) \cos(\omega t) dt$, splitting into a plateau integral and a parabolic-ramp integral with two integrations by parts (see Appendix 2):

$$\hat{\Gamma}(\omega) = \frac{4}{\tau_{\text{spot}}^2 \omega^3} \left[\tau_{\text{spot}} \omega \cos\left(\frac{\omega \ell_{\text{spot}}}{2}\right) + \sin\left(\frac{\omega \ell_{\text{spot}}}{2}\right) - \sin\left(\frac{\omega \ell_{\text{spot}}}{2} + \omega \tau_{\text{spot}}\right) \right]. \quad (4.46)$$

A visualization of this profile is shown in the left panel of Figure 4.10.

(2) Exponential profile for $\Gamma(t)$: For the piecewise exponential envelope (Eq. 4.10), setting $t_1 = 0$ and $t_2 = \tau_{\max}$, the squared envelope is

$$\Gamma(t) = \alpha_{\max}^2 \times \begin{cases} e^{2t/\tau_{\text{em}}} & t < 0 \\ 1 & 0 \leq t \leq \tau_{\max} \\ e^{-2(t-\tau_{\max})/\tau_{\text{dec}}} & t > \tau_{\max} \end{cases} \quad (4.47)$$

with doubled exponential rates compared to $\alpha(t)$. Because $\Gamma(t)$ is not symmetric when $\tau_{\text{em}} \neq \tau_{\text{dec}}$, the Fourier transform is complex. Splitting $\hat{\Gamma}(\omega) = \hat{\Gamma}_R(\omega) - i \hat{\Gamma}_I(\omega)$ and evaluating each piece integral analytically:

$$\hat{\Gamma}_R(\omega) = \alpha_{\max}^2 \left[\frac{2\tau_{\text{em}}}{4 + \omega^2 \tau_{\text{em}}^2} + \frac{\sin(\omega \tau_{\max})}{\omega} + \frac{2\tau_{\text{dec}} \cos(\omega \tau_{\max})}{4 + \omega^2 \tau_{\text{dec}}^2} + \frac{\omega \tau_{\text{dec}}^2 \sin(\omega \tau_{\max})}{4 + \omega^2 \tau_{\text{dec}}^2} \right] \quad (4.48)$$

$$\hat{\Gamma}_I(\omega) = \alpha_{\max}^2 \left[\frac{\omega \tau_{\text{em}}^2}{4 + \omega^2 \tau_{\text{em}}^2} + \frac{1 - \cos(\omega \tau_{\text{max}})}{\omega} + \frac{\omega \tau_{\text{dec}}^2 \cos(\omega \tau_{\text{max}})}{4 + \omega^2 \tau_{\text{dec}}^2} - \frac{2 \tau_{\text{dec}} \sin(\omega \tau_{\text{max}})}{4 + \omega^2 \tau_{\text{dec}}^2} \right] \quad (4.49)$$

The full derivation, including the direct convolution for $R_\Gamma(\tau)$, is analogous to the trapezoidal case (Appendix 2), and a figure of the function is shown in the middle panel of Figure 4.10.

(3) Skewed Normal profile for $\Gamma(t)$: Forgács-Dajka et al. (2021) modeled sunspot group area evolution using a skew-normal function (their Eq. 1), which naturally captures the observed asymmetry between rapid emergence and slower decay phases of sunspot lifetimes. We adopt this model as a third envelope option. The normalized-squared envelope is

$$\Gamma(t) \propto \exp\left(-\frac{t^2}{2\sigma_{\text{sn}}^2}\right) \left[1 + \text{erf}\left(\frac{n_{\text{sn}} t}{\sigma_{\text{sn}} \sqrt{2}}\right)\right], \quad (4.50)$$

rescaled so that the peak value equals unity. The two free parameters are σ_{sn} (the scale parameter in days, controlling the overall spot lifetime) and n_{sn} (the dimensionless skewness parameter). When $n_{\text{sn}} = 0$ the envelope reduces to a symmetric Gaussian; $n_{\text{sn}} < 0$ gives rapid rise and slow decay (typical of solar spots), while $n_{\text{sn}} > 0$ gives the opposite asymmetry. Unlike the trapezoidal and exponential profiles, the skew-normal $\Gamma(t)$ does not yield a closed-form autocorrelation $R_\Gamma(\tau)$ or Fourier transform $\hat{\Gamma}(\omega)$, so we therefore compute both numerically via FFT-based methods in our `spotgp` implementation. This function is shown in the right panel of Figure 4.10.

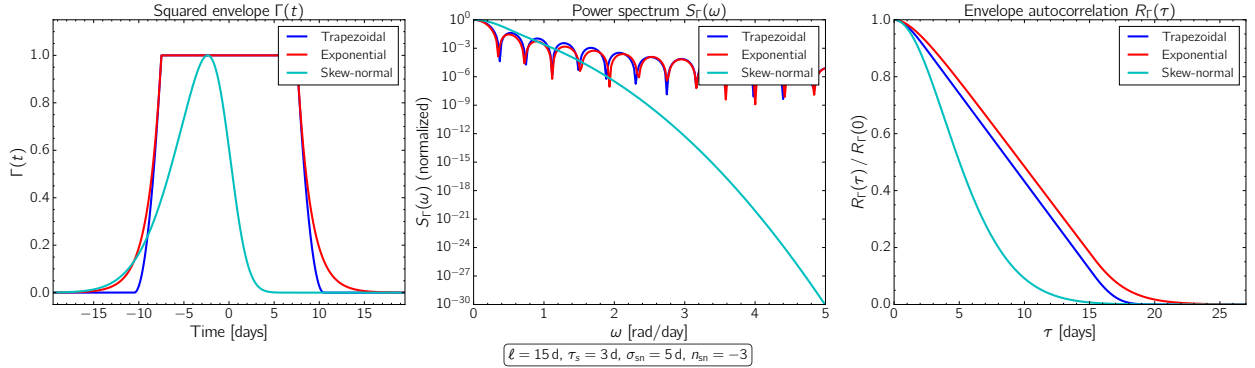


Figure 4.10: Comparison of the three spot envelope models: trapezoidal (blue), exponential (green), and skew-normal (red). *Left:* Normalized squared envelopes $\Gamma(t)$. *Middle:* Energy spectral densities $S_\Gamma(\omega)$ (normalized, log scale), showing the different high-frequency decay rates set by envelope smoothness. *Right:* Normalized envelope autocorrelation $R_\Gamma(\tau)/R_\Gamma(0)$. All three implementations are included in our `spotgp` package, along with the functionality to define custom envelope profiles.

5.5 Combined Kernel for a Single Spot

Substituting the delta train (4.42) into Eq. (4.40) and using the sifting property of the Dirac delta (P5, FT3), $f(\omega) * \delta(\omega - a) = f(\omega - a)$ yields:

$$\hat{A}(\omega) = \sum_n c_n \hat{\Gamma}(\omega - n\omega_0) \quad (4.51)$$

Explicitly to second harmonic order, using linearity (FT1):

$$\boxed{\hat{A}(\omega) = c_0 \hat{\Gamma}(\omega) + c_1 [\hat{\Gamma}(\omega - \omega_0) + \hat{\Gamma}(\omega + \omega_0)] + c_2 [\hat{\Gamma}(\omega - 2\omega_0) + \hat{\Gamma}(\omega + 2\omega_0)]} \quad (4.52)$$

where c_0, c_1, c_2 are the Fourier coefficients of the visibility function given by Eq. (4.44) (all real, with $c_{-n} = c_n$). Each term has a clear physical interpretation in the time domain:

- The $n = 0$ (DC “direct current”) term, $c_0 \Gamma(t)$, is the non-oscillatory component: the mean flux deficit that tracks the spot-size envelope. It is largest when the spot is visible for most of the rotation (e.g., a high-latitude spot on a low-inclination star), contributing a smooth, aperiodic dimming.
- The $n = 1$ (fundamental) term, $2c_1 \Gamma(t) \cos(\omega_0 t)$, is the dominant periodic component at the stellar rotation frequency. It produces the familiar once-per-rotation brightness dip as the spot rotates in and out of view.
- The $n = 2$ (second harmonic) term, $2c_2 \Gamma(t) \cos(2\omega_0 t)$, captures the non-sinusoidal shape of the visibility function—the flattened profile near disk center and sharp limb transitions—which a pure cosine cannot reproduce.

Figure 4.11 shows the Fourier decomposition of the single-spot signal $A(t) = \Gamma(t) \mathcal{V}(t)$ for three stellar inclinations at a fixed latitude $\Phi = 30^\circ$, with the same spot decay envelope $\Gamma(t)$. Here we see that terms up to second harmonic order are sufficient for capturing the shape of the signal for varying I .

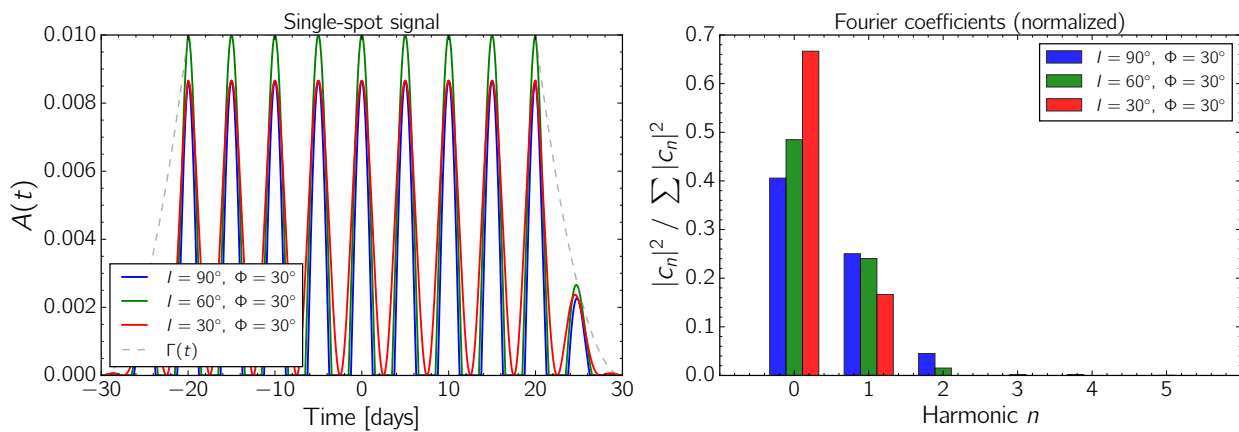


Figure 4.11: Fourier decomposition of the single-spot signal $A(t) = \Gamma(t) \cdot \mathcal{V}(t)$ for a spot at latitude $\Phi = 30^\circ$. *Left:* Time-domain signal for three stellar inclinations, with the squared envelope $\Gamma(t) = \alpha^2(t)$ shown as a dashed gray line. *Right:* Normalized squared Fourier coefficients $|c_n|^2 / \sum |c_n|^2$ of the visibility function for each inclination.

Figure 4.12 illustrates how the individual $n = 0$, $n = 1$, and $n = 2$ harmonic components combine to produce the full kernel shape for a single spot at a fixed latitude $\Phi = 30^\circ$. Each panel decomposes the normalized kernel $\mathcal{K}(\tau)/\mathcal{K}(0)$ (Eq. 4.64) for a different stellar inclination. The fractional power in each harmonic is defined as

$$f_n = \frac{(2 - \delta_{n0}) |c_n|^2}{\sum_{m=0}^N (2 - \delta_{m0}) |c_m|^2}, \quad (4.53)$$

where the factor $(2 - \delta_{n0})$ accounts for the positive and negative frequency contributions at each harmonic. At low inclination ($I = 30^\circ$), the DC term dominates ($f_0 \sim 89\%$), producing a kernel that decays nearly monotonically with only weak periodic modulation. As the inclination increases toward edge-on ($I = 90^\circ$), the fundamental becomes the leading component ($f_1 \sim 50\%$), and the kernel develops pronounced oscillations at the rotation period. The second harmonic remains a minor but non-negligible contribution ($f_2 \sim 9\%$ at $I = 90^\circ$), sharpening the peaks and introducing structure at half the rotation period. The shaded region under the $n = 0$ component highlights the aperiodic “floor” of the kernel set by the DC visibility power.

5.6 Averaging Over Random Spot Parameters

The expression for $\hat{A}(\omega)$ in Eq. (4.51) describes a *single* spot with a particular reference time t_{ref} , reference longitude Λ_{ref} , and latitude Φ_{ref} . To obtain the kernel for a *population* of spots, we must average the power spectral density $|\hat{A}(\omega)|^2$ over these random variables. We treat t_{ref} and Λ_{ref} as uniformly distributed, and Φ_{ref} as drawn from some latitude distribution $p(\Phi)$.

5.6.1 Effect of t_{ref} : time shift of the envelope

Shifting the envelope by t_{ref} gives $\Gamma(t - t_{\text{ref}})$, whose Fourier transform picks up a phase by the time-shifting property (FT2, P5):

$$\hat{\Gamma}(\omega; t_{\text{ref}}) = \hat{\Gamma}(\omega) e^{-i\omega t_{\text{ref}}}. \quad (4.54)$$

This phase propagates to $\hat{A}(\omega)$, but since $|e^{-i\omega t_{\text{ref}}}|^2 = 1$, it **drops out** of the power spectral density $|\hat{A}(\omega)|^2$. The kernel is therefore independent of t_{ref} – any spot reference time yields the same autocorrelation.

5.6.2 Effect of Λ_{ref} : longitude phase

Shifting the initial longitude Λ_{ref} is equivalent to a time shift of $\Lambda_{\text{ref}}/\omega_0$ in the periodic visibility function. By the time-shifting property (FT2), the Fourier series coefficients acquire a harmonic-dependent phase:

$$c_n(\Lambda_{\text{ref}}) = c_n e^{in\Lambda_{\text{ref}}}, \quad (4.55)$$

where $c_n \equiv c_n(\Lambda_{\text{ref}} = 0)$. Substituting into Eq. (4.51), the full single-spot spectrum is:

$$\hat{A}(\omega) = e^{-i\omega t_{\text{ref}}} \sum_n c_n e^{in\Lambda_{\text{ref}}} \hat{\Gamma}(\omega - n\omega_0). \quad (4.56)$$

The power spectral density is:

$$|\hat{A}(\omega)|^2 = \sum_n \sum_m c_n c_m^* e^{i(n-m)\Lambda_{\text{ref}}} \hat{\Gamma}(\omega - n\omega_0) \hat{\Gamma}^*(\omega - m\omega_0). \quad (4.57)$$

Averaging over Λ_{ref} uniformly distributed on $[0, 2\pi]$:

$$\langle e^{i(n-m)\Lambda_{\text{ref}}} \rangle_{\Lambda_{\text{ref}}} = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)\Lambda_{\text{ref}}} d\Lambda_{\text{ref}} = \delta_{nm}. \quad (4.58)$$

This **cancels out all cross terms** between different harmonics. Only the diagonal ($n = m$) terms survive:

$$\boxed{\langle |\hat{A}(\omega)|^2 \rangle_{t_{\text{ref}}, \Lambda_{\text{ref}}} = \sum_{n=-\infty}^{\infty} |c_n|^2 S_{\Gamma}(\omega - n\omega_0),} \quad (4.59)$$

where $S_{\Gamma}(\omega) \equiv |\hat{\Gamma}(\omega)|^2$ is the power spectral density of the squared envelope, and each harmonic of the visibility function contributes independently to the power spectrum, with weight $|c_n|^2$.

5.6.3 Obtaining the kernel from the averaged power spectral density

Applying the Wiener–Khinchin theorem (Eq. 4.31) to the averaged power spectral density yields:

$$\mathcal{K}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle |\hat{A}(\omega)|^2 \rangle e^{i\omega\tau} d\omega = \sum_n |c_n|^2 \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\Gamma}(\omega - n\omega_0) e^{i\omega\tau} d\omega. \quad (4.60)$$

Substituting $u = \omega - n\omega_0$:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\Gamma}(u) e^{i(u+n\omega_0)\tau} du = e^{in\omega_0\tau} \cdot \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\Gamma}(u) e^{iu\tau} du}_{R_{\Gamma}(\tau)}, \quad (4.61)$$

where $R_{\Gamma}(\tau)$ is the autocorrelation of the squared envelope $\Gamma(t) = \alpha^2(t)$. Therefore:

$$\mathcal{K}(\tau) = R_{\Gamma}(\tau) \sum_{n=-\infty}^{\infty} |c_n|^2 e^{in\omega_0\tau}. \quad (4.62)$$

Since the c_n are real with $c_{-n} = c_n$, the complex exponentials combine into cosines:

$$\mathcal{K}(\tau) = \alpha_{\max}^4 R_{\Gamma}(\tau) \left[|c_0|^2 + 2 \sum_{n=1}^{\infty} |c_n|^2 \cos(n\omega_0\tau) \right] \quad (4.63)$$

The kernel factorizes into two interpretable components:

- $R_{\Gamma}(\tau)$: the **normalized envelope autocorrelation** of $\Gamma(t) = \alpha^2(t)/\alpha_{\max}^2$, which sets the overall decay timescale of correlations (governed by spot lifetime ℓ_{spot} and rise/decay time τ_{spot}).
- The **cosine sum**: a quasi-periodic modulation at the rotation frequency and its harmonics, with amplitudes set by the squared Fourier coefficients $|c_n|^2$ of the visibility function.

Truncating to second harmonic order yields:

$$\mathcal{K}(\tau) \approx \alpha_{\max}^4 R_{\Gamma}(\tau) \left[|c_0|^2 + 2|c_1|^2 \cos(\omega_0\tau) + 2|c_2|^2 \cos(2\omega_0\tau) \right], \quad (4.64)$$

which represents the kernel function for a *single* spot.

5.7 Multi-Spot Kernel

The observed flux deficit is the sum over N_{spot} independent spots (Eq. 4.1):

$$\delta F(t) = \frac{1}{\pi} \sum_{k=1}^{N_{\text{spot}}} A_k(t) (1 - f_{\text{spot},k}). \quad (4.65)$$

For independent spots with identical contrast f_{spot} , the covariance of δF is the sum of the individual spot covariances:

$$\text{Cov}[\delta F(t), \delta F(t + \tau)] = \frac{(1 - f_{\text{spot}})^2}{\pi^2} \sum_{k=1}^{N_{\text{spot}}} \langle R_{A_k}(\tau) \rangle, \quad (4.66)$$

where $\langle R_{A_k}(\tau) \rangle$ denotes the autocorrelation of spot k averaged over its random parameters. Since all spots are drawn from the same distribution, each contributes equally, so:

$$\mathcal{K}(\tau) = \frac{N_{\text{spot}} (1 - f_{\text{spot}})^2}{\pi^2} \langle \mathcal{K}(\tau) \rangle_{\Phi}, \quad (4.67)$$

where $\langle \mathcal{K}(\tau) \rangle_{\Phi}$ is the single-spot kernel (Eq. 4.63) averaged over the latitude distribution $p(\Phi)$.

5.7.1 Averaging over latitude

The latitude Φ_{ref} affects both the Fourier coefficients c_n (through $b_0 = \cos I \sin \Phi$ and $b_1 = \sin I \cos \Phi$) and, in the presence of differential rotation, the rotation frequency $\omega_0(\Phi) =$

$2\pi(1 - \kappa \sin^2 \Phi)/P_{\text{eq}}$. Averaging Eq. (4.63) over $p(\Phi)$:

$$\langle \mathcal{K}(\tau) \rangle_{\Phi} = \alpha_{\text{max}}^4 \int_{-\pi/2}^{\pi/2} R_{\Gamma}(\tau) \left[|c_0(\Phi)|^2 + 2 \sum_{n=1}^N |c_n(\Phi)|^2 \cos(n\omega_0(\Phi) \tau) \right] p(\Phi) d\Phi. \quad (4.68)$$

In general, this integral must be evaluated numerically.

5.7.2 Solid-body rotation ($\kappa = 0$)

When $\kappa = 0$, the rotation frequency ω_0 is the same at all latitudes. The $\cos(n\omega_0\tau)$ factors come out of the latitude integral, and only the squared coefficients need to be averaged:

$$\langle \mathcal{K}(\tau) \rangle_{\Phi} = \alpha_{\text{max}}^4 R_{\Gamma}(\tau) \left[\langle |c_0|^2 \rangle_{\Phi} + 2 \sum_{n=1}^N \langle |c_n|^2 \rangle_{\Phi} \cos(n\omega_0\tau) \right], \quad (\kappa = 0), \quad (4.69)$$

where $\langle |c_n|^2 \rangle_{\Phi} = \int |c_n(\Phi)|^2 p(\Phi) d\Phi$. The latitude averaging simply replaces each $|c_n|^2$ with its population mean.

5.7.3 Full kernel

Combining Eqs. (4.67) and (4.68), the complete GP kernel for the flux variability due to a population of independent spots is:

$$\mathcal{K}(\tau) = \frac{N_{\text{spot}} (1 - f_{\text{spot}})^2 \alpha_{\text{max}}^4}{\pi^2} \int R_{\Gamma}(\tau) \left[\sum_n |c_n(\Phi)|^2 \cos(n\omega_0(\Phi) \tau) \right] p(\Phi) d\Phi \quad (4.70)$$

where $R_{\Gamma}(\tau)$ is the normalized autocorrelation of the squared spot-size envelope $\Gamma(t) = \alpha^2(t)/\alpha_{\text{max}}^2$, $c_n(\Phi)$ are the Fourier coefficients of the visibility function (Eq. 4.44), $\omega_0(\Phi) = 2\pi(1 - \kappa \sin^2 \Phi)/P_{\text{eq}}$ is the latitude-dependent rotation frequency, and $p(\Phi)$ is the spot latitude distribution. The full hyperparameters of this kernel are:

$$\theta = \{P_{\text{eq}}, \kappa, I, N_{\text{spot}}, \ell_{\text{spot}}, \tau_{\text{spot}}, \alpha_{\text{max}}, f_{\text{spot}}\}. \quad (4.71)$$

The parameter N_{spot} in Eq. (4.70) represents the total number of spots active during the observation. In practice, spots emerge stochastically at a rate $\dot{N}_{\text{spot}}$ [spots day⁻¹], so for a time series of duration T_{obs} the expected number of spots is $N_{\text{spot}} = \dot{N}_{\text{spot}} T_{\text{obs}}$. Parameterizing in terms of the emergence rate $\dot{N}_{\text{spot}}$ rather than the total count N_{spot} has the advantage that the kernel amplitude becomes independent of the observing baseline: the amplitude prefactor becomes $\dot{N}_{\text{spot}} T_{\text{obs}} (1 - f_{\text{spot}})^2 \alpha_{\text{max}}^4 / \pi^2$, which is an intrinsic stellar property. In our `spotgp` implementation, users may specify either `nspot` (total count) or `nspot_rate` (emergence rate in spots day⁻¹); the latter is the recommended parameterization for inference.

The amplitude prefactor is degenerate in practice (only the product can be constrained from the lightcurve variance) so it can be absorbed into a single amplitude hyperparameter

σ_k^2 . Using the spot contrast-area product $a_{\text{spot}} = (1 - f_{\text{spot}}) \alpha_{\text{max}}^2$ defined in Section 4:

$$\sigma_k^2 \equiv \dot{N}_{\text{spot}} a_{\text{spot}}^2, \quad (4.72)$$

and the reduced hyperparameter set becomes:

$$\theta = \{P_{\text{eq}}, \kappa, I, \ell_{\text{spot}}, \tau_{\text{spot}}, \sigma_k\}. \quad (4.73)$$

When fitting the kernel alone (with mean-subtracted data), only σ_k is constrained. However, jointly fitting the mean function (Eq. 4.28) and the kernel constrains $\dot{N}_{\text{spot}} a_{\text{spot}}$ and $\dot{N}_{\text{spot}} a_{\text{spot}}^2$ independently, breaking the degeneracy and recovering the individual values of $\dot{N}_{\text{spot}}$ and a_{spot} . The full shared parameter set is then:

$$\theta = \{P_{\text{eq}}, \kappa, I, \ell_{\text{spot}}, \tau_{\text{spot}}, \dot{N}_{\text{spot}}, a_{\text{spot}}\}. \quad (4.74)$$

Truncating to second harmonic order ($n \leq 2$) and specializing to solid-body rotation ($\kappa = 0$), the cosines factor out of the latitude integral and the kernel becomes:

$$\mathcal{K}(\tau) = \sigma_k^2 R_{\Gamma}(\tau) [\langle |c_0|^2 \rangle_{\Phi} + 2 \langle |c_1|^2 \rangle_{\Phi} \cos(\omega_0 \tau) + 2 \langle |c_2|^2 \rangle_{\Phi} \cos(2\omega_0 \tau)], \quad (4.75)$$

where $\omega_0 = 2\pi/P_{\text{eq}}$ and the latitude-averaged squared coefficients are $\langle |c_n|^2 \rangle_{\Phi} = \int_0^{\pi} |c_n(\Phi)|^2 p(\Phi) d\Phi$, with $c_n(\Phi)$ given by Eq. (4.44).

The envelope autocorrelation $R_{\Gamma}(\tau)$ is the *normalized* autocorrelation of the squared spot-size envelope $\Gamma(t) = \alpha^2(t)/\alpha_{\text{max}}^2$, obtained via the Wiener–Khinchin theorem:

$$R_{\Gamma}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\Gamma}(\omega) e^{i\omega\tau} d\omega = \frac{1}{\pi} \int_0^{\infty} S_{\Gamma}(\omega) \cos(\omega\tau) d\omega, \quad (4.76)$$

where $S_{\Gamma}(\omega) \equiv |\hat{\Gamma}(\omega)|^2$ is the power spectral density of the squared envelope and $\hat{\Gamma}(\omega)$ is given by Eq. (4.46). Since $\hat{\Gamma}(\omega)$ is real and even, $S_{\Gamma} = \hat{\Gamma}^2$ and $R_{\Gamma}(\tau)$ reduces to a cosine transform. Equivalently, $R_{\Gamma}(\tau)$ can be computed directly as the convolution $\int \Gamma(s)\Gamma(s+\tau) ds$; since $\Gamma(t)$ is piecewise polynomial, this yields a closed-form piecewise polynomial (see Appendix 3, Eq. .131).

For the special case of edge-on viewing ($I = \pi/2$) and a uniform latitude distribution $p(\Phi) = 1/\pi$ on $[0, \pi]$, all Fourier coefficients are proportional to $\cos \Phi$: $c_n(\Phi) = g_n \cos \Phi$, where $g_0 = 1/\pi$, $g_1 = 1/4$, and $g_2 = 1/(3\pi)$. The latitude integrals evaluate to $\langle |c_n|^2 \rangle_{\Phi} = g_n^2/2$, giving the fully closed-form kernel:

$$\mathcal{K}(\tau) = \sigma_k^2 R_{\Gamma}(\tau) \left[\frac{1}{2\pi^2} + \frac{1}{16} \cos(\omega_0 \tau) + \frac{1}{9\pi^2} \cos(2\omega_0 \tau) \right] \quad (I = \pi/2, \kappa = 0). \quad (4.77)$$

For general inclination I , the latitude integrals $\langle |c_n|^2 \rangle_{\Phi}$ must be evaluated numerically, since c_0 and c_1 depend on $\theta_{\text{vis}}(\Phi) = \arccos(-\cot I \tan \Phi)$, which prevents elementary integration. However, the numerical quadrature over Φ is inexpensive (a single 1D integral per harmonic). This simplified form (4.77) is applicable when a system is viewed edge-on ($I = \pi/2$) with no differential rotation ($\kappa = 0$) and under the assumption that spots are

uniformly distributed across all latitudes, however it is flexible enough to work for different envelope terms $R_\Gamma(\tau)$.

5.8 Power Spectral Density of the Full Kernel

By the Wiener–Khinchin theorem (Eq. 4.31), the kernel $\mathcal{K}(\tau)$ and its power spectral density (PSD) $S_{\mathcal{K}}(\omega)$ form a Fourier transform pair:

$$S_{\mathcal{K}}(\omega) = \int_{-\infty}^{\infty} \mathcal{K}(\tau) e^{-i\omega\tau} d\tau. \quad (4.78)$$

Since $\mathcal{K}(\tau)$ is real and even, $S_{\mathcal{K}}(\omega)$ is real, non-negative, and even in ω . To evaluate the transform, we use the frequency-shift property: for a real, even function $R_\Gamma(\tau)$ with transform $S_\Gamma(\omega) = \mathcal{F}\{R_\Gamma\}$,

$$\mathcal{F}\{R_\Gamma(\tau) \cos(n\omega_0\tau)\} = \frac{1}{2} [S_\Gamma(\omega - n\omega_0) + S_\Gamma(\omega + n\omega_0)], \quad (4.79)$$

which follows from writing $\cos(n\omega_0\tau) = \frac{1}{2}(e^{in\omega_0\tau} + e^{-in\omega_0\tau})$ and applying the frequency-shift property (FT3).

Applying this equation to the solid-body rotation kernel (Eq. 4.75) yields:

$$S_{\mathcal{K}}(\omega) = \sigma_k^2 \left[\langle |c_0|^2 \rangle_\Phi S_\Gamma(\omega) + \sum_{n=1}^N \langle |c_n|^2 \rangle_\Phi [S_\Gamma(\omega - n\omega_0) + S_\Gamma(\omega + n\omega_0)] \right], \quad (\kappa = 0), \quad (4.80)$$

where $S_\Gamma(\omega) = |\hat{\Gamma}(\omega)|^2$ is the energy spectral density of the squared envelope and $\omega_0 = 2\pi/P_{\text{eq}}$. The PSD is a sum of shifted copies of the envelope spectrum $S_\Gamma(\omega)$, each centered on a harmonic $n\omega_0$ and weighted by the corresponding latitude-averaged squared Fourier coefficient $\langle |c_n|^2 \rangle_\Phi$. The $n = 0$ term produces a low-frequency component centered at $\omega = 0$, while each $n \geq 1$ term contributes a pair of symmetric peaks at $\omega = \pm n\omega_0$, whose width is set by the envelope spectral width (inversely proportional to the spot lifetime).

For the general case with differential rotation ($\kappa \neq 0$), the latitude-dependent rotation frequency $\omega_0(\Phi)$ prevents factoring the cosine out of the latitude integral. Taking the Fourier transform of the full kernel (Eq. 4.70):

$$S_{\mathcal{K}}(\omega) = \sigma_k^2 \int_{-\pi/2}^{\pi/2} \left[|c_0(\Phi)|^2 S_\Gamma(\omega) + \sum_{n=1}^N |c_n(\Phi)|^2 [S_\Gamma(\omega - n\omega_0(\Phi)) + S_\Gamma(\omega + n\omega_0(\Phi))] \right] p(\Phi) d\Phi. \quad (4.81)$$

Differential rotation causes the harmonic peaks to broaden: spots at different latitudes contribute power at slightly different frequencies $n\omega_0(\Phi)$, smearing each peak over the range $n\omega_0(0)$ to $n\omega_0(\pi/2) = n\omega_0(0)(1 - \kappa)$. This smearing is the signal of DR that can be extracted from a lightcurve.

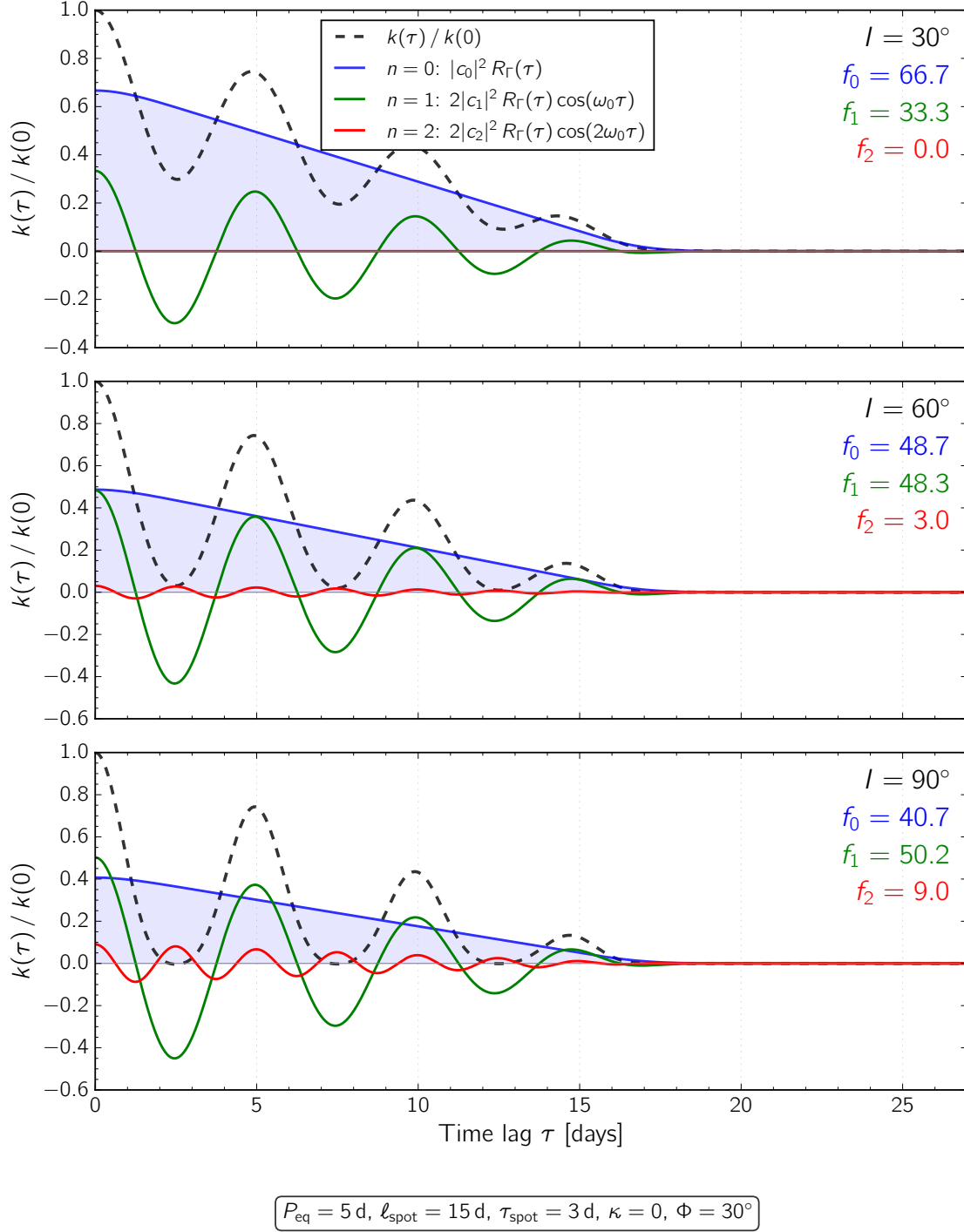


Figure 4.12: Harmonic decomposition of the kernel $\mathcal{K}(\tau)/\mathcal{K}(0)$ (Eq. 4.64) for a single spot at latitude $\Phi = 30^\circ$ at three stellar inclinations ($I = 30^\circ, 60^\circ, 90^\circ$), with fixed spot parameters ($P_{\text{eq}} = 5 \text{ d}$, $\ell_{\text{spot}} = 15 \text{ d}$, $\tau_{\text{spot}} = 3 \text{ d}$, $\kappa = 0$). The black curve is the full kernel; colored curves show the individual $n = 0$ (DC, blue), $n = 1$ (fundamental, orange), and $n = 2$ (second harmonic, green) components. The shaded blue region marks the DC contribution. Inset annotations give the fractional power f_n (Eq. 4.53) in each harmonic. For pole-on viewing the kernel is dominated by the aperiodic DC term, while for edge-on viewing the fundamental produces strong periodic oscillations.

6 Physical Interpretation of Kernel Hyperparameters

6.1 Information Content of the Mean and Covariance

Before discussing the individual hyperparameters, we illustrate how the single-spot signal connects to the two statistical quantities modeled by the GP: the mean function (first moment) and the kernel (second moment). Figure 4.13 shows 80 realizations of the single-spot signal $A(t) = \Gamma(t) \mathcal{V}(t)$ for random initial longitudes Λ_{ref} , at fixed inclination $I = 60^\circ$ and latitude $\Phi = 30^\circ$. The ensemble mean over all longitudes is the smooth, aperiodic curve $\langle A(t) \rangle_\Lambda = c_0 \Gamma(t)$: the spot-size envelope scaled by the zeroth Fourier coefficient of the visibility function. This is the first moment of the process – the quantity that enters the GP mean function (Section 4). It depends linearly on c_0 and on the first moment of the envelope, $\int \Gamma dt$.

The lower panel shows the residuals after subtracting this mean. These oscillatory residuals (produced by the periodic rotation of the spot in and out of view) are what the kernel (second moment) describes. Their variance depends quadratically on the Fourier coefficients ($|c_n|^2$) and on the autocorrelation of the envelope, $R_\Gamma(\tau) = \int \Gamma(t) \Gamma(t + \tau) dt$.

6.2 Kernel Components and Hyperparameters

Now we will summarize the physical meaning of each kernel hyperparameter and its effect on the shape of the kernel. The shape is governed by the interplay of two components: the envelope autocorrelation $\Gamma(t)$, which sets the overall decay timescale of correlations (due to spot evolution), and the visibility function $\mathcal{V}(t)$, which encodes the periodic structure at the rotation frequency and its harmonics.

The hyperparameters can be grouped into three categories:

- **Rotation parameters:** P_{eq} (equatorial rotation period), κ (differential rotation shear), I (stellar inclination). These parameters determine the rotation frequencies of spots at different latitudes and the shape of the visibility function, which in turn affect the periodic structure of the kernel (e.g., the location and relative amplitudes of peaks in the ACF and PSD).
- **Spot evolution parameters:** ℓ_{spot} (spot lifetime), τ_{spot} (spot rise/decay timescale). These parameters govern the shape of the envelope autocorrelation $\Gamma(t)$, which sets the overall decay timescale of correlations. Longer lifetimes and slower evolution lead to a more slowly decaying ACF, while shorter lifetimes and faster evolution produce a more rapidly decaying ACF.
- **Amplitude parameters:** $\dot{N}_{\text{spot}}$ (spot emergence rate) and $a_{\text{spot}} = (1 - f_{\text{spot}}) \alpha_{\text{max}}^2$ (spot contrast-area product). These parameters affect the overall amplitude of the kernel and the mean flux level, but do not significantly change the shape of the kernel. When fitting the kernel alone, only the product $\sigma_k^2 = \dot{N}_{\text{spot}} a_{\text{spot}}^2$ is constrained. However, jointly fitting the mean function (Section 4) and the kernel constrains $\dot{N}_{\text{spot}}$ and a_{spot} independently, as discussed below.

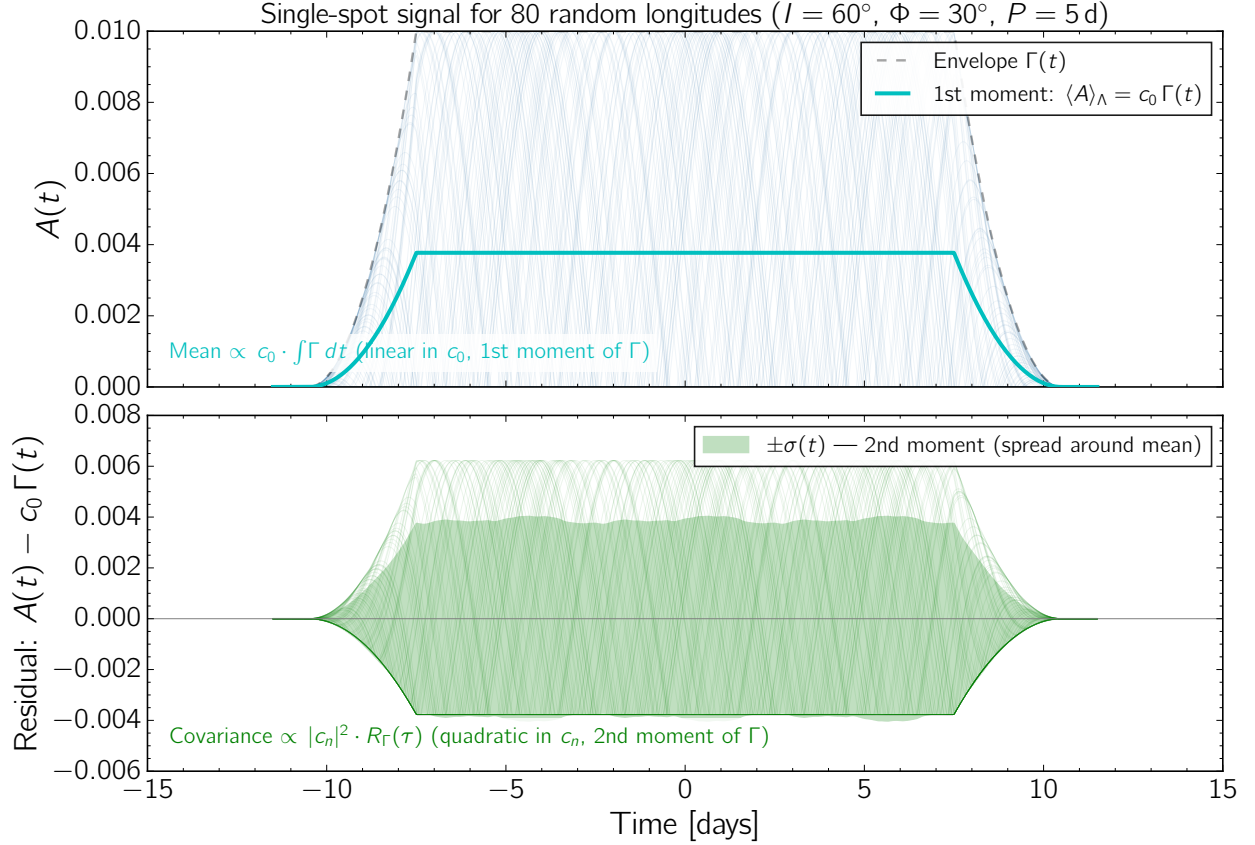


Figure 4.13: Decomposition of the single-spot signal into first and second moments of the GP, for a spot at latitude $\Phi = 30^\circ$ with $I = 60^\circ$, $P_{\text{eq}} = 5$ d, $\ell_{\text{spot}} = 15$ d, and $\tau_{\text{spot}} = 3$ d. *Top:* 80 realizations of $A(t) = \Gamma(t) \mathcal{V}(t)$ for random longitudes (blue), with the ensemble mean $\langle A \rangle_\Lambda = c_0 \Gamma(t)$ (red). The mean depends linearly on c_0 and on $\int \Gamma dt$ (first moment of the envelope). *Bottom:* Residuals after subtracting the mean, with $\pm\sigma(t)$ envelope (orange). This spread is what the kernel describes; it depends quadratically on $|c_n|^2$ and on the envelope autocorrelation $R_\Gamma(\tau) = \int \Gamma(t) \Gamma(t + \tau) dt$ (second moment of the envelope).

6.2.1 Lifetime of the Kernel

Figure 4.14 illustrates how the kernel of the envelope term $R_\Gamma(\tau)$ depends on the spot lifetime parameters ℓ_{spot} and τ_{spot} . For simplification, we show the simplified symmetric trapezoid case with equal spot emergence and decay timescales (where $\tau_{\text{spot}} = \tau_{\text{em}} = \tau_{\text{dec}}$), although the non-symmetric case is conceptually similar. The four colors show the shape profiles for different values of ℓ_{spot} and τ_{spot} , but same total spot lifetime $\ell_{\text{spot}} + 2\tau_{\text{spot}} = 20$ d. As shown in blue, when the spot lifetime is dominated by the plateau duration ℓ_{spot} , the envelope autocorrelation shallowly decays to zero at $\tau = \ell_{\text{spot}} + 2\tau_{\text{spot}}$, with a piecewise linear derivative. On the other hand, when the spot lifetime is dominated by the rise/decay timescale τ_{spot} (red), the envelope autocorrelation decays more rapidly and smoothly to zero, with a piecewise quadratic derivative. The intermediate case (purple) show a combination of these behaviors.

This difference qualitatively makes sense: for time lags $\tau > \ell_{\text{spot}} + 2\tau_{\text{spot}}$, the autocorre-

lation is effectively zero, reflecting the finite lifetime of the spots (which is consistent with the piecewise polynomial form of $R_\Gamma(\tau)$ derived in Appendix 3). The more time that a spot spends at its max size $\alpha_{\max}$ (i.e., the longer the plateau ℓ_{spot}), the more slowly the autocorrelation decays, since the spot remains visible for a longer time. Conversely, if the spot spends most of its lifetime in the rise/decay phases (i.e., short ℓ_{spot} and long τ_{spot}), the autocorrelation decays more quickly, since the spot is only visible for a short time before it disappears.

For $\tau < \ell_{\text{spot}} + 2\tau_{\text{spot}}$, the shape of the autocorrelation reflects the relative contributions of the plateau and rise/decay phases of the spot evolution. The derivative plots in the bottom row of Figure 4.14 show how the slope changes at the breakpoints $\tau = \tau_s$, ℓ , $\ell + \tau_s$, and $\ell + 2\tau_s$, which correspond to transitions between different phases of the spot evolution.

6.2.2 Amplitude of the Kernel

The kernel amplitude—the variance of the GP at zero lag—is obtained by evaluating Eq. (4.75) at $\tau = 0$:

$$\mathcal{K}(0) = \sigma_k^2 R_\Gamma(0) [\langle |c_0|^2 \rangle_\Phi + 2\langle |c_1|^2 \rangle_\Phi + 2\langle |c_2|^2 \rangle_\Phi]. \quad (4.82)$$

This factorizes into three independent contributions:

1. **Amplitude prefactor** σ_k^2 : Recall that $\sigma_k^2 = \dot{N}_{\text{spot}} a_{\text{spot}}^2$ (Eq. 4.72) absorbs the spot emergence rate $\dot{N}_{\text{spot}}$ and the contrast-area product $a_{\text{spot}} = (1 - f_{\text{spot}}) \alpha_{\max}^2$ into a single quantity. When fitting the kernel alone, $\dot{N}_{\text{spot}}$ and a_{spot} are fully degenerate—only their product σ_k^2 is constrained by the data.
2. **Envelope energy** $R_\Gamma(0)$: The zero-lag value of the normalized envelope autocorrelation equals the integrated energy of the squared spot-size envelope, $R_\Gamma(0) = \int \Gamma^2(t) dt$. For the trapezoidal profile (Appendix 3), this evaluates to

$$R_\Gamma(0) = \ell_{\text{spot}} + \frac{2\tau_{\text{spot}}}{5}, \quad (4.83)$$

which is basically the effective duration over which the spot contributes significant flux. The plateau phase contributes ℓ_{spot} (the spot is at full size), while each ramp phase contributes $\tau_{\text{spot}}/5$ (the $1/5$ factor arises because $\int_0^1 u^4 du = 1/5$ for the squared ramp $\Gamma^2 \propto u^4$). Thus, longer-lived spots or longer plateau durations directly increase the kernel amplitude, independently of the rotation or geometry parameters.

3. **Visibility power** $\sum_n \langle |c_n|^2 \rangle_\Phi$: The bracketed term $\langle |c_0|^2 \rangle_\Phi + 2\langle |c_1|^2 \rangle_\Phi + 2\langle |c_2|^2 \rangle_\Phi$ is the total power in the visibility function, averaged over latitude. This factor depends on the stellar inclination I and the latitude distribution $p(\Phi)$. Perhaps counterintuitively, the total visibility power is *largest* for pole-on viewing ($I \rightarrow 0^\circ$) and smallest for edge-on ($I = 90^\circ$). This is because the DC component $|c_0|^2$ dominates the power budget: for pole-on viewing, nearly all spots are permanently visible and contribute a large, constant flux deficit, which inflates $|c_0|^2$ and hence the total variance. For edge-on viewing, spots are occulted for roughly half the rotation period, reducing the time-averaged contribution and thus the total power. In other words, pole-on stars have a higher photometric variance $\mathcal{K}(0)$ (driven by the mean flux deficit from ever-present spots), while edge-on stars have stronger *periodic modulation* relative to their variance

(i.e., a larger fraction of the power in the $n \geq 1$ harmonics).

In practice, the factorization in Eq. (4.82) means that the observed photometric variance constrains the product $\sigma_k^2 \cdot R_\Gamma(0) \cdot \sum_n \langle |c_n|^2 \rangle_\Phi$, but from the kernel alone the individual values of $\dot{N}_{\text{spot}}$ and a_{spot} remain degenerate. However, jointly fitting the mean function (Eq. 4.28) provides independent leverage: the mean flux deficit constrains $\dot{N}_{\text{spot}} a_{\text{spot}} \cdot \int \Gamma dt \cdot \langle \bar{\mathcal{V}} \rangle_\Phi$, while the variance constrains $\dot{N}_{\text{spot}} a_{\text{spot}}^2 \cdot R_\Gamma(0) \cdot \sum_n \langle |c_n|^2 \rangle_\Phi$. Since the shape factors ($\int \Gamma dt$, $R_\Gamma(0)$, $\langle \bar{\mathcal{V}} \rangle_\Phi$, $\langle |c_n|^2 \rangle_\Phi$) are determined by the five shape hyperparameters, the ratio of the variance to the mean squared isolates a_{spot} , and hence $\dot{N}_{\text{spot}}$, without external information. External constraints (e.g., spectroscopic f_{spot} or transit-mapped α_{max}) can further decompose a_{spot} into its constituent factors.

If the lightcurve is further normalized to unit variance (i.e., $\mathcal{K}(0) = 1$), then σ_k^2 is fully absorbed into the normalization and carries no information. The normalized kernel becomes

$$\hat{\mathcal{K}}(\tau) \equiv \frac{\mathcal{K}(\tau)}{\mathcal{K}(0)} = \frac{R_\Gamma(\tau)}{R_\Gamma(0)} \cdot \frac{\langle |c_0|^2 \rangle_\Phi + 2\langle |c_1|^2 \rangle_\Phi \cos(\omega_0 \tau) + 2\langle |c_2|^2 \rangle_\Phi \cos(2\omega_0 \tau)}{\langle |c_0|^2 \rangle_\Phi + 2\langle |c_1|^2 \rangle_\Phi + 2\langle |c_2|^2 \rangle_\Phi}, \quad (4.84)$$

which depends on only five hyperparameters: P_{eq} , κ , I , ℓ_{spot} , and τ_{spot} . The first factor, $R_\Gamma(\tau)/R_\Gamma(0)$, is the normalized envelope autocorrelation whose *shape* depends on ℓ_{spot} and τ_{spot} individually (through the piecewise breakpoints at $\tau = \tau_{\text{spot}}$, ℓ_{spot} , $\ell_{\text{spot}} + \tau_{\text{spot}}$, and $\ell_{\text{spot}} + 2\tau_{\text{spot}}$), even though the absolute scale $R_\Gamma(0) = \ell_{\text{spot}} + 2\tau_{\text{spot}}/5$ is divided out. The second factor contains the *relative* harmonic amplitudes $\langle |c_n|^2 \rangle_\Phi / \sum_m \langle |c_m|^2 \rangle_\Phi$, which constrain I through the ratio of the DC power to the oscillatory power (Section 6). Thus, even with a fully normalized lightcurve, the kernel shape alone constrains the rotation period, differential rotation, inclination, spot lifetime, and spot rise/decay timescale—all parameters that affect the temporal *structure* of the variability. What is lost is all information about the absolute *amplitude*: $\dot{N}_{\text{spot}}$ and a_{spot} become unconstrained by the kernel shape. If the lightcurve is *not* mean-subtracted, the mean flux level (Eq. 4.28) provides an independent constraint on $\dot{N}_{\text{spot}} a_{\text{spot}}$, which—combined with the unnormalized variance constraint on $\dot{N}_{\text{spot}} a_{\text{spot}}^2$ —recovers both amplitude parameters even without external information.

6.2.3 The DC Term and the Mean Function

The $n = 0$ (DC) harmonic plays a dual role in the GP: it contributes to both the kernel (through $|c_0|^2$) and the mean function (through $c_0 = \bar{\mathcal{V}}$). Understanding this connection clarifies how to handle mean subtraction and why the kernel is strictly non-negative.

The kernel (Eq. 4.63) is strictly non-negative for all lags: each spot’s flux deficit is one-signed (always dimming, never brightening), so the autocorrelation of a non-negative pulse is non-negative, and the kernel—a sum of such autocorrelations—inherits this property. The $|c_0|^2 R_\Gamma(\tau)$ term provides the positive floor that prevents the kernel from crossing zero, representing the covariance of the non-oscillatory flux deficit as spots stochastically emerge and decay. The oscillatory ($n \geq 1$) terms modulate around this floor but cannot push the total below it (except negligibly at high harmonics; cf. the lower envelope, Eq. 4.90).

In contrast, the sample autocorrelation function (ACF) of observed lightcurves typically oscillates between positive and negative values at the rotation period. This discrepancy arises because sample ACFs are computed by subtracting the *time mean* of the lightcurve,

which removes the DC flux level—the same quantity encoded by c_0 . More precisely, the sample ACF estimates

$$\hat{C}(\tau) = \frac{1}{N} \sum_i (y_i - \bar{y})(y_{i+\tau} - \bar{y}), \quad (4.85)$$

which is biased relative to the true autocovariance $\mathcal{K}(\tau)$ by a term proportional to $\bar{\mathcal{K}} \equiv T^{-1} \int_0^T \mathcal{K}(s) ds$. Since $|c_0|^2$ dominates $\bar{\mathcal{K}}$, the sample ACF is shifted downward by approximately $|c_0|^2 \bar{R}_\Gamma$, allowing it to cross zero at half-period lags where the oscillatory terms are at their most negative.

The DC power fraction,

$$f_0 = \frac{\langle |c_0|^2 \rangle_\Phi}{\langle |c_0|^2 \rangle_\Phi + 2\langle |c_1|^2 \rangle_\Phi + 2\langle |c_2|^2 \rangle_\Phi}, \quad (4.86)$$

quantifies the severity of this mismatch. For pole-on viewing ($I \rightarrow 0^\circ$), $f_0 \rightarrow 1$ and the kernel is dominated by DC power that is removed by mean subtraction, making the mismatch large. For edge-on viewing ($I = 90^\circ$), $f_0 \approx 0.08$ and the effect is minor.

This has a practical consequence for GP modeling: when fitting mean-subtracted data with a zero-mean GP, the $|c_0|^2$ floor in the kernel over-predicts the autocovariance at trough lags, because the kernel models DC variance that the mean subtraction has already removed. The effect is largest for low-inclination stars. Two strategies avoid this:

1. **Model the mean explicitly:** Use the analytic mean function (Eq. 4.28) as the GP mean and retain the full kernel including $|c_0|^2$. The mean absorbs the DC flux level, and the kernel correctly models the residual covariance.
2. **Drop the DC term:** For mean-subtracted data, use a reduced kernel with only the $n \geq 1$ harmonics:

$$\mathcal{K}_{\text{osc}}(\tau) = \sigma_k^2 R_\Gamma(\tau) [2\langle |c_1|^2 \rangle_\Phi \cos(\omega_0 \tau) + 2\langle |c_2|^2 \rangle_\Phi \cos(2\omega_0 \tau)]. \quad (4.87)$$

This oscillatory-only kernel matches the behavior of the sample ACF, but loses the constraint on I from the DC-to-oscillatory power ratio.

The first approach is preferred when amplitude information is available (i.e., the lightcurve is not normalized), since it preserves all constraints including the degeneracy-breaking from the mean (Section 4).

6.2.4 Upper and Lower Envelopes of the Kernel

The kernel (Eq. 4.64) is the product of the slowly varying envelope autocorrelation $R_\Gamma(\tau)$ and a sum of cosines at the rotation frequency and its harmonics. Because the cosine terms oscillate between fixed extrema, we can solve for the curves that bound the kernel from above and below.

From Equation 4.64, the kernel for a single spot (truncated to $N = 2$) is:

$$\mathcal{K}(\tau) = R_\Gamma(\tau) [|c_0|^2 + 2|c_1|^2 \cos(\omega_0 \tau) + 2|c_2|^2 \cos(2\omega_0 \tau)]. \quad (4.88)$$

The **upper envelope** is achieved when all cosine terms equal +1 (i.e., at integer multiples

of the rotation period, $\tau = mP_{\text{eq}}$):

$$k_{\text{upper}}(\tau) = R_{\Gamma}(\tau) [|c_0|^2 + 2|c_1|^2 + 2|c_2|^2]. \quad (4.89)$$

This expression is basically the envelope autocorrelation $R_{\Gamma}(\tau)$ scaled by the total power in all harmonics. The upper envelope is therefore set entirely by the spot evolution parameters ($\ell_{\text{spot}}, \tau_{\text{spot}}$) through $R_{\Gamma}(\tau)$, and does not depend on inclination or latitude (which only affect the *relative* amplitudes $|c_n|^2$, not their sum when normalized).

The **lower envelope** is achieved at odd half-periods, $\tau = (2m+1)P_{\text{eq}}/2$, where $\cos(\omega_0\tau) = -1$ and $\cos(2\omega_0\tau) = +1$:

$$k_{\text{lower}}(\tau) = R_{\Gamma}(\tau) [|c_0|^2 - 2|c_1|^2 + 2|c_2|^2]. \quad (4.90)$$

The lower envelope is controlled by the balance between the DC component $|c_0|^2$, the fundamental $|c_1|^2$, and the second harmonic $|c_2|^2$. The ratio of the lower to upper envelope factors,

$$\frac{k_{\text{lower}}}{k_{\text{upper}}} = \frac{|c_0|^2 - 2|c_1|^2 + 2|c_2|^2}{|c_0|^2 + 2|c_1|^2 + 2|c_2|^2}, \quad (4.91)$$

quantifies the *fractional oscillation depth* of the kernel: a ratio near +1 means negligible oscillation (the kernel decays monotonically), while a ratio near zero means the kernel oscillates down to approximately zero at the troughs. The magnitude and sign of this ratio depend primarily on the stellar inclination I :

- **Edge-on viewing** ($I = 90^\circ$): For edge-on geometry, the visibility angle is $\theta_{\text{vis}} = \pi/2$ at all latitudes, and the DC component vanishes ($c_0 \propto \cos I = 0$). All Fourier coefficients are proportional to $\cos \Phi$, so the envelope ratio becomes independent of latitude: $k_{\text{lower}}/k_{\text{upper}} \approx -0.005$. The kernel oscillates nearly to zero at the troughs between rotation peaks, reflecting the fact that edge-on spots are fully occulted for half the rotation period.
- **Inclined viewing** ($I < 90^\circ$): The DC component $c_0 \propto \cos I \sin \Phi$ grows, since spots near the visible pole are always in view and contribute a non-oscillating flux deficit. This raises the lower envelope well above zero, compressing the oscillation depth. For example, at $I = 60^\circ$ the ratio is $\approx +0.3$, and at $I = 30^\circ$ it reaches $\approx +0.8$. In the limit of pole-on viewing ($I \rightarrow 0^\circ$), $|c_0|^2$ dominates all harmonics and the kernel becomes nearly monotonically decaying with negligible oscillation.

The differential rotation parameter κ also affects the envelope structure indirectly. When $\kappa \neq 0$, spots at different latitudes rotate at different frequencies, which broadens and reduces the coherent peaks of the cosine sum. This effect raises the lower envelope toward zero, since the destructive interference at $\cos(\omega_0\tau) = -1$ is diluted across a range of frequencies. The effect is most pronounced at large lags, where the accumulated phase differences between latitudes are largest.

Figure 4.15 illustrates the upper and lower envelopes for three representative inclinations. For edge-on viewing, the kernel oscillates nearly to zero at the troughs (lower/upper ≈ -0.005), while for $I = 30^\circ$ the oscillations are shallow and the lower envelope remains close to the upper envelope (lower/upper $\approx +0.78$).

6.2.5 Number of Visible Peaks

A practical observable in any autocorrelation analysis is the number of distinct rotation peaks visible before the signal decays into the noise. In our kernel, this is controlled by the ratio of the spot lifetime to the rotation period. The envelope autocorrelation $R_{\Gamma}(\tau)$ has compact support: it vanishes identically for $\tau > \ell_{\text{spot}} + 2\tau_{\text{spot}}$ (Eq. .131). Since the kernel peaks occur at integer multiples of the rotation period, the number of visible peaks is approximately

$$N_{\text{peaks}} \approx \left\lfloor \frac{\ell_{\text{spot}} + 2\tau_{\text{spot}}}{P_{\text{eq}}} \right\rfloor. \quad (4.92)$$

This is the starspot kernel’s analog of the quality factor $Q = \pi f_0 \tau_{\text{damp}}$ in a damped harmonic oscillator kernel: a larger ratio of coherence time to period yields more peaks. For example, with $\ell_{\text{spot}} = 15 \text{ d}$, $\tau_{\text{spot}} = 3 \text{ d}$, and $P_{\text{eq}} = 5 \text{ d}$, we expect $N_{\text{peaks}} \approx 4$ visible rotation peaks, consistent with Figures 4.14 and 4.15.

Because $R_{\Gamma}(\tau)$ has compact support (unlike the exponential or Gaussian decays used in quasi-periodic kernels, which are strictly positive for all τ), the kernel exactly vanishes beyond the spot lifetime. This means that any observed correlation at lags $\tau > \ell_{\text{spot}} + 2\tau_{\text{spot}}$ must arise from other sources (e.g., activity cycles or instrumental systematics), providing a clean diagnostic for the spot lifetime.

6.2.6 Harmonic Content and Peak Shape

The relative amplitudes of the fundamental and second harmonic control whether the kernel oscillation looks sinusoidal or has sharper, non-sinusoidal peaks. Defining the harmonic ratio

$$\mathcal{H} \equiv \frac{\langle |c_2|^2 \rangle_{\Phi}}{\langle |c_1|^2 \rangle_{\Phi}}, \quad (4.93)$$

we can characterize two limiting regimes:

- **Fundamental-dominated** ($\mathcal{H} \ll 1$): The kernel oscillation is approximately sinusoidal, $\mathcal{K}(\tau) \approx R_{\Gamma}(\tau) [|c_0|^2 + 2|c_1|^2 \cos(\omega_0 \tau)]$. This occurs for low inclinations ($I \lesssim 30^\circ$), where the visibility function is nearly sinusoidal and the second harmonic is negligible. The ACF peaks are broad and symmetric.
- **Significant second harmonic** ($\mathcal{H} \gtrsim 0.1$): The second harmonic sharpens the peaks and introduces a secondary structure at half the rotation period. This occurs for edge-on viewing ($I \approx 90^\circ$), where the spot visibility function is a truncated cosine with significant power beyond the fundamental. In the lightcurve, this manifests as “double-dipping”—two flux minima per rotation period when two spots (or one spot visible over a broad longitude range) produce distinct ingress and egress features.

The harmonic ratio $\mathcal{H}$ is therefore a diagnostic of the stellar inclination, complementary to the envelope ratio (Eq. 4.91). For edge-on viewing with a uniform latitude distribution, $\mathcal{H} = g_2^2/g_1^2 = (4/(3\pi))^2 \approx 0.18$, while for $I = 30^\circ$ the ratio drops to $\mathcal{H} \approx 0.02$. Importantly, $\mathcal{H}$ is a property of the normalized kernel shape and is therefore measurable even from a variance-normalized lightcurve.

6.2.7 Peak Height Decay

A commonly measured ACF diagnostic is the height of the first rotation peak relative to the zero-lag value. In our kernel, this ratio is

$$\frac{\mathcal{K}(P_{\text{eq}})}{\mathcal{K}(0)} = \frac{R_{\Gamma}(P_{\text{eq}})}{R_{\Gamma}(0)}, \quad (4.94)$$

since all cosine terms equal +1 at $\tau = P_{\text{eq}}$. This simplification is remarkable: the first peak height ratio depends *only* on the spot evolution parameters ℓ_{spot} and τ_{spot} (through R_{Γ}), and is completely independent of the stellar inclination I , the differential rotation κ , or the visibility geometry.

More generally, the peak heights at all integer multiples of P_{eq} trace out the upper envelope $R_{\Gamma}(\tau)/R_{\Gamma}(0)$, which decays from 1 to 0 over the spot lifetime. The rate of this decay encodes the spot lifetime parameters. For the trapezoidal profile, if mP_{eq} falls in the linear regime of R_{Γ} (i.e., $\tau_{\text{spot}} < mP_{\text{eq}} < \ell_{\text{spot}}$; Appendix 3, Interval 2), the peak heights decay linearly:

$$\frac{\mathcal{K}(mP_{\text{eq}})}{\mathcal{K}(0)} = \frac{\ell_{\text{spot}} + 2\tau_{\text{spot}}/3 - mP_{\text{eq}}}{\ell_{\text{spot}} + 2\tau_{\text{spot}}/5}, \quad \tau_{\text{spot}} < mP_{\text{eq}} < \ell_{\text{spot}}. \quad (4.95)$$

This linear decay of peak heights is a distinctive prediction of the trapezoidal spot profile and could be tested observationally. Exponential spot profiles, by contrast, produce exponentially decaying peak heights.

6.2.8 Effect of Differential Rotation on Peak Broadening

When $\kappa \neq 0$, spots at different latitudes rotate at different angular velocities $\omega_0(\Phi) = 2\pi(1 - \kappa \sin^2 \Phi)/P_{\text{eq}}$. At lag τ , a spot at latitude Φ accumulates a phase $\omega_0(\Phi)\tau$ relative to the equatorial rate. The spread in phase across latitudes is

$$\Delta\phi(\tau) = \frac{2\pi |\kappa| \tau}{P_{\text{eq}}} \langle \sin^2 \Phi \rangle_{\Phi}, \quad (4.96)$$

where $\langle \sin^2 \Phi \rangle_{\Phi}$ is the mean of $\sin^2 \Phi$ over the latitude distribution. For a uniform distribution on $[0, \pi/2]$, $\langle \sin^2 \Phi \rangle = 1/2$.

This phase spread has two observable consequences:

1. **Peak broadening:** At small lags ($\Delta\phi \ll 2\pi$), the kernel peaks remain sharp but become slightly asymmetric—skewed toward longer lags for solar-like differential rotation ($\kappa > 0$), since high-latitude spots rotate more slowly and contribute power at slightly longer periods. At large lags ($\Delta\phi \gtrsim \pi$), the peaks broaden substantially and eventually wash out entirely, as spots at different latitudes fall out of phase with each other.
2. **Coherence lag:** The lag at which the periodic signal is effectively destroyed can be estimated by setting $\Delta\phi \approx \pi$ (half-cycle dephasing):

$$\tau_{\text{coh}} \approx \frac{P_{\text{eq}}}{2 |\kappa| \langle \sin^2 \Phi \rangle_{\Phi}}. \quad (4.97)$$

For $\kappa = 0.2$ (solar-like) and $P_{\text{eq}} = 5$ d, this gives $\tau_{\text{coh}} \approx 25$ d. Beyond this lag, the kernel oscillations are strongly suppressed even if $R_{\Gamma}(\tau)$ has not yet decayed to zero. In effect, differential rotation imposes a second coherence timescale that competes with the spot lifetime: the kernel decays on a timescale of $\min(\ell_{\text{spot}} + 2\tau_{\text{spot}}, \tau_{\text{coh}})$.

In the PSD, differential rotation broadens the sharp peaks at ω_0 and $2\omega_0$ into bands of width $\Delta\omega \approx 2\pi |\kappa| \langle \sin^2 \Phi \rangle / P_{\text{eq}}$, as noted in Section 9.2. The degree of broadening is a direct diagnostic of $|\kappa|$, though the sign of κ (prograde vs. retrograde differential rotation) can only be inferred from the direction of peak asymmetry in the ACF at intermediate lags.

Figure 4.16 illustrates this decoherence effect. For each value of $\kappa \in \{0, 0.5, 1\}$, three GP lightcurve samples are drawn from the kernel and shown alongside the normalized ACF. With solid-body rotation ($\kappa = 0$), the lightcurves are highly periodic and the ACF retains sharp peaks out to the spot lifetime. As $|\kappa|$ increases, the periodic pattern progressively loses coherence: at $|\kappa| = 0.5$ the signal becomes noticeably aperiodic beyond the first two peaks ($\tau_{\text{coh}} \approx 10$ d from Eq. 4.97), and at $|\kappa| = 1$ only the first peak survives ($\tau_{\text{coh}} \approx 5$ d). The ACF peaks simultaneously broaden and diminish, while the $R_{\Gamma}(\tau)$ envelope (dashed) remains unchanged—confirming that differential rotation destroys the periodic coherence without affecting the spot lifetime. As expected from Eq. (4.96), the decoherence depends on $|\kappa|$: the ACFs for $\kappa = +0.5$ and $\kappa = -0.5$ (and likewise ± 1) are nearly identical, since the phase spread is symmetric in κ . The sign of κ affects only the direction of peak asymmetry (skewness toward longer or shorter lags), which is subtle and most apparent at intermediate lags.

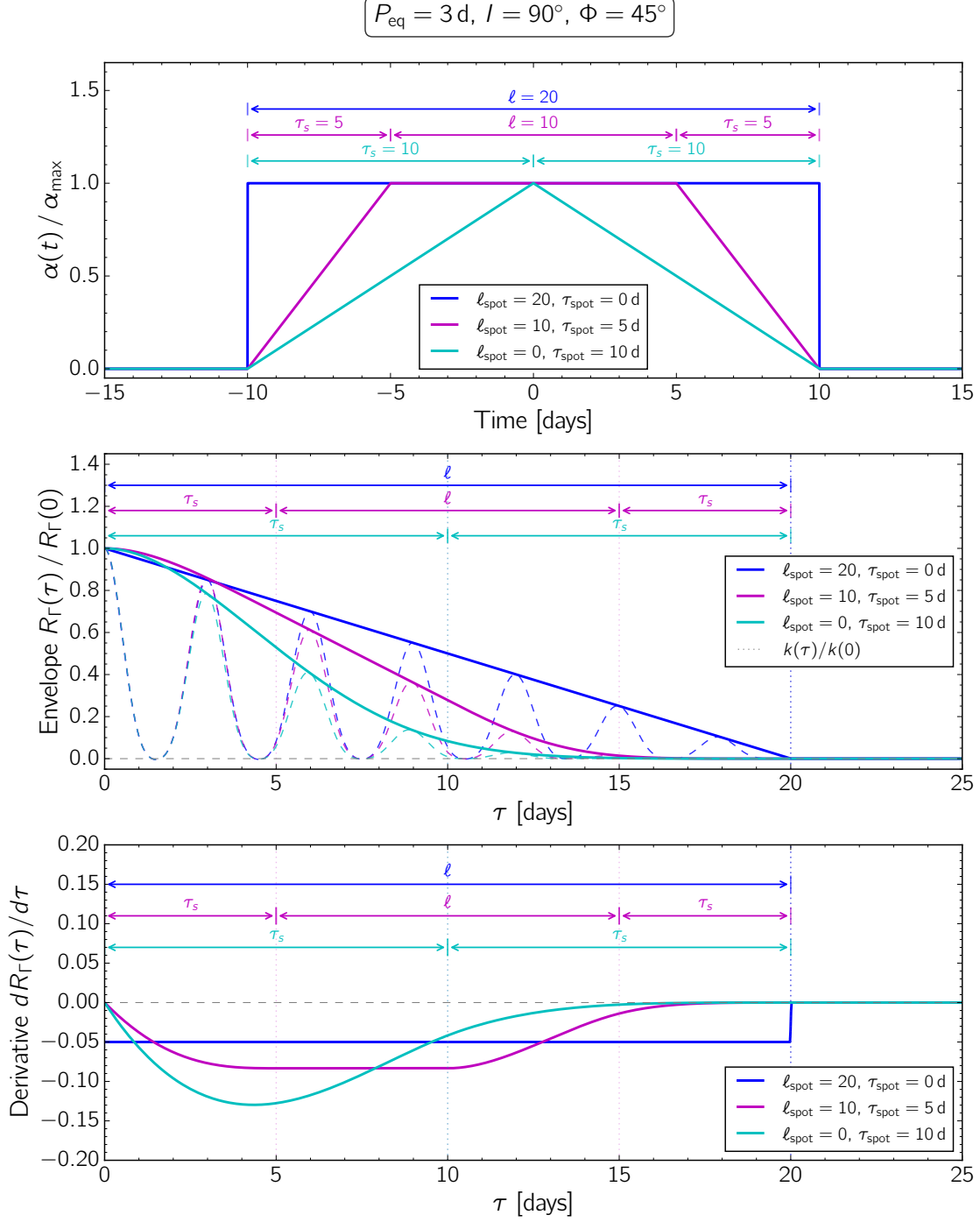


Figure 4.14: Effect of varying ℓ_{spot} and τ_{spot} on the kernel envelope and autocorrelation. Left column: vary ℓ_{spot} with τ_{spot} fixed. Right column: vary τ_{spot} with ℓ_{spot} fixed. Top row: trapezoidal envelope $\alpha(t)/\alpha_{\text{max}}$ with horizontal annotations marking ℓ and τ_s . Middle row: normalized autocorrelation $R_\Gamma(\tau)/R_\Gamma(0)$ (solid) with normalized kernel $\mathcal{K}(\tau)/\mathcal{K}(0)$ overlaid (dashed). Bottom row: derivative $dR_\Gamma(\tau)/d\tau$. In the middle and bottom rows, dotted vertical lines mark the piecewise breakpoints of R_Γ at $\tau = \tau_s, \ell, \ell + \tau_s$, and $\ell + 2\tau_s$ (Equation D50), and horizontal annotations indicate the corresponding interval widths.

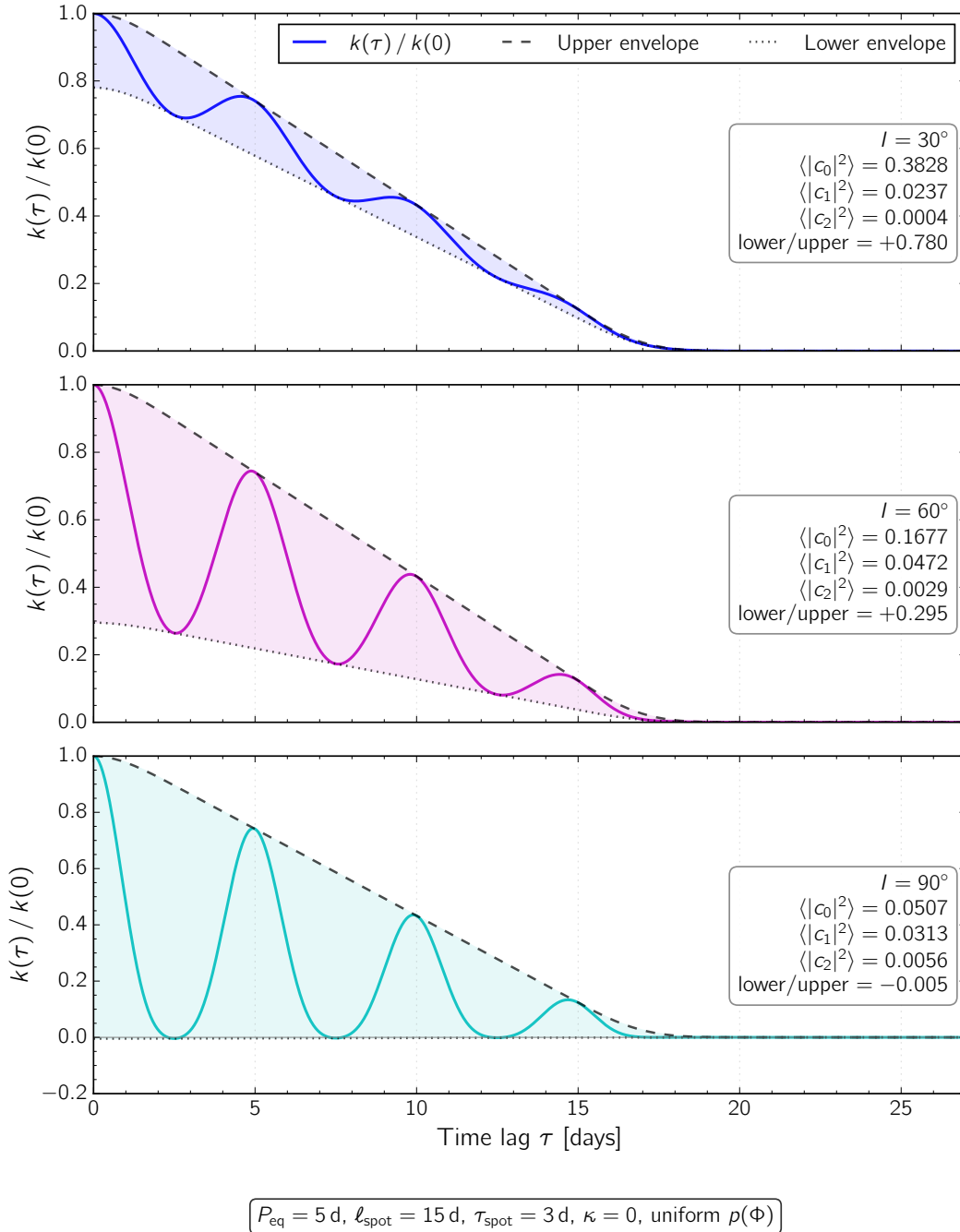


Figure 4.15: Upper and lower envelopes of the kernel for three stellar inclinations ($I = 30^\circ$, 60° , 90°), with fixed spot parameters ($P_{\text{eq}} = 5 \text{ d}$, $\ell_{\text{spot}} = 15 \text{ d}$, $\tau_{\text{spot}} = 3 \text{ d}$, $\kappa = 0$, uniform $p(\Phi)$). The upper envelope (dashed) is $R_{\Gamma}(\tau) [|c_0|^2 + 2|c_1|^2 + 2|c_2|^2]$, identical in shape for all inclinations. The lower envelope (dotted) is $R_{\Gamma}(\tau) [|c_0|^2 - 2|c_1|^2 + 2|c_2|^2]$. For edge-on viewing, the kernel oscillates nearly to zero at the troughs (lower/upper ≈ -0.005); as inclination decreases, the growing DC component $|c_0|^2$ raises the lower envelope toward the upper, compressing the oscillation depth. Shaded regions show the range between envelopes. Inset annotations list the latitude-averaged squared Fourier coefficients and the lower/upper ratio (Eq. 4.91).

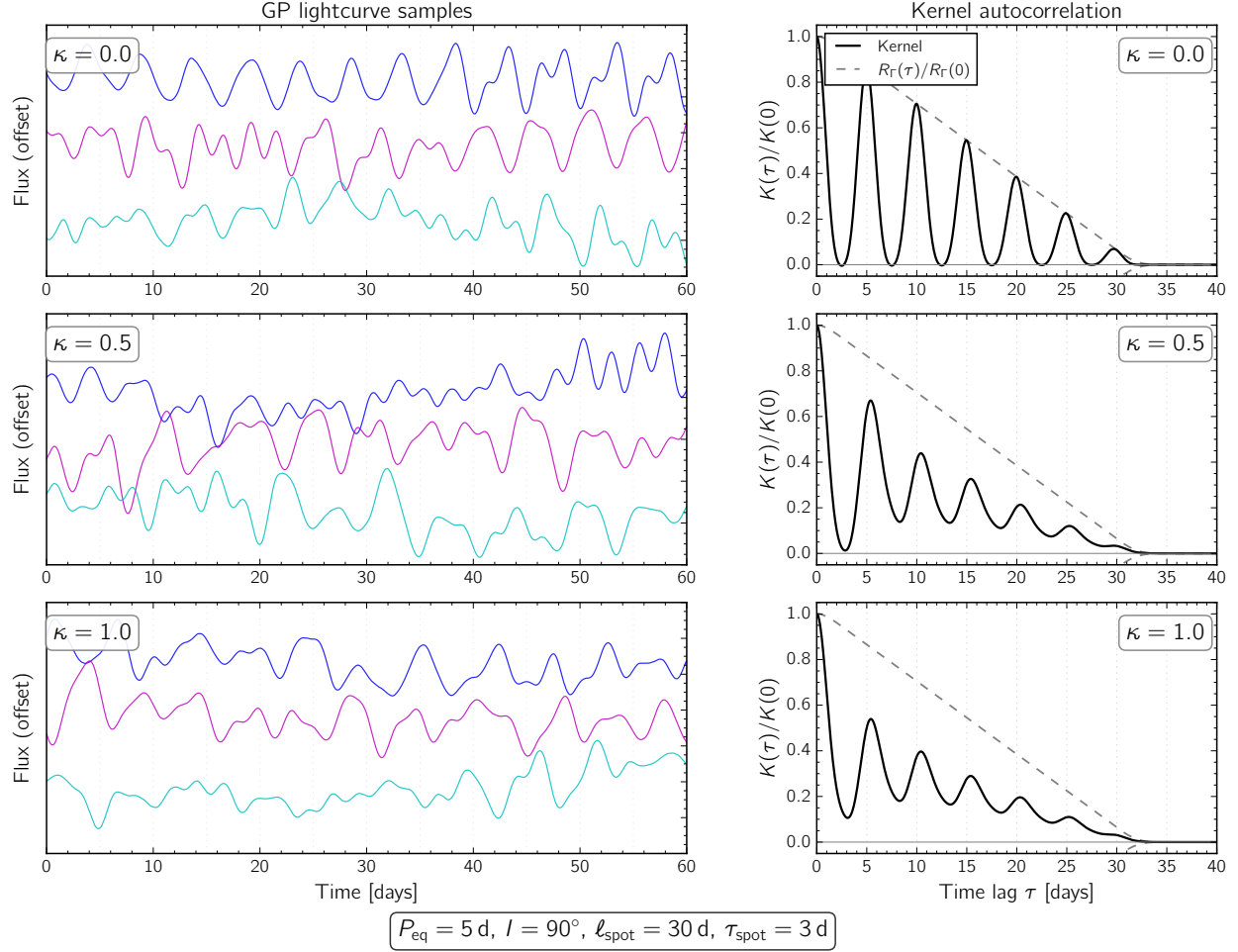


Figure 4.16: Decoherence from differential rotation. *Left:* Three GP lightcurve samples drawn from the kernel for each κ value. *Right:* Normalized kernel ACF (solid) with the envelope autocorrelation $R_\Gamma(\tau)/R_\Gamma(0)$ (dashed). For $\kappa = 0$ (solid-body rotation), the lightcurves are highly periodic and the ACF peaks remain sharp. As $|\kappa|$ increases, spots at different latitudes rotate at different rates, causing the periodic signal to decohere: ACF peaks broaden and wash out, and the lightcurves become irregular. The decoherence is symmetric in $|\kappa|$: the ACFs for $\kappa = \pm 0.5$ and $\kappa = \pm 1$ are nearly identical. The envelope R_Γ is unchanged since it depends only on spot lifetime parameters. Parameters: $P_{\text{eq}} = 5 \text{ d}$, $I = 90^\circ$, $\ell_{\text{spot}} = 30 \text{ d}$, $\tau_{\text{spot}} = 3 \text{ d}$.

7 GP Solver Implementation

In order to solve for the GP posterior and perform inference on the hyperparameters θ , we need to evaluate the GP marginal log-likelihood (Eq. 4.98), which requires computing the kernel matrix inversion. In this section, we show how the structure of our kernel allows for efficient computation using a banded Cholesky solver, which reduces the computational cost from $\mathcal{O}(N^3)$ to $\mathcal{O}(Nb^2)$, where b is the bandwidth determined by the spot lifetime and observing cadence. We implement this solver in JAX, enabling fast automatic differentiation for MCMC sampling of the hyperparameters.

7.1 Naive Cholesky Solver

We implement a JAX-accelerated GP solver that uses the analytic kernel for Bayesian inference on stellar lightcurves. The GP marginal log-likelihood for N observations $\mathbf{y}$ at times $\mathbf{x}$ is

$$\ln p(\mathbf{y} \mid \theta) = -\frac{1}{2} [(\mathbf{y} - \boldsymbol{\mu})^\top K_\theta^{-1} (\mathbf{y} - \boldsymbol{\mu}) + \ln |K_\theta| + N \ln 2\pi], \quad (4.98)$$

where K_θ is the $N \times N$ covariance matrix with entries $[K_\theta]_{ij} = k(|x_i - x_j|; \theta) + (\sigma_{\text{err},i}^2 + \sigma_n^2) \delta_{ij}$, $\boldsymbol{\mu}$ is the mean vector, and σ_n is an optional white-noise amplitude.

Evaluating Eq. (4.98) involves two computationally distinct steps:

1. *Kernel evaluation*: computing $k(|x_i - x_j|; \theta)$ for all pairs, which requires the latitude integral in Eq. (4.70). We perform this integral using Gauss-Legendre quadrature with $n_{\text{lat}} = 32$ nodes. To avoid materializing all latitude contributions simultaneously (which would require an $n_{\text{lat}} \times M$ intermediate tensor, where M is the number of unique lags), we accumulate the weighted latitude contributions sequentially, reducing peak memory from $\mathcal{O}(n_{\text{lat}} \times M)$ to $\mathcal{O}(M)$. We further reduce memory by exploiting the symmetry of the covariance matrix, evaluating the kernel only on the $N(N+1)/2$ upper-triangular lags rather than the full N^2 matrix.
2. *Cholesky factorization and solve*: decomposing $K_\theta = LL^\top$ and solving for $L^{-1}(\mathbf{y} - \boldsymbol{\mu})$, which dominates at $\mathcal{O}(N^3)$.

Our implementation provides two interfaces. The `GPSolver` class pre-computes and caches the Cholesky factorization at construction time, enabling fast repeated queries (predictions, log-likelihood evaluation) for fixed hyperparameters. For MCMC sampling, where hyperparameters change at every step, we provide a pure-functional just-in-time (JIT) compiled log-likelihood function that recomputes the full pipeline (kernel evaluation, Cholesky, and solve) from scratch on each call. JAX’s just-in-time compilation and automatic differentiation provide efficient gradient computation for Hamiltonian Monte Carlo (HMC) and related samplers.

Table 4.6 benchmarks the wall-clock time and peak memory usage of the GP solver as a function of the number of data points N . The “Build” column measures the `GPSolver` construction (kernel evaluation + Cholesky), while the “Log-likelihood” column measures the JIT-compiled pure-functional evaluation used during MCMC. After JIT warmup, the log-likelihood evaluation is faster than the class-based build because XLA reuses pre-allocated buffers. At $N = 1000$, the full log-likelihood evaluation takes ~ 0.1 s on a single CPU core,

Table 4.6: GPSolver wall-clock time and peak memory as a function of the number of data points N . *Build*: covariance matrix + Cholesky factorization. *Log-likelihood*: JIT-compiled evaluation (kernel + Cholesky + solve) as called during MCMC. *Predict*: predictive mean at 200 test points. Times are the median of 5 runs after JIT warmup. Memory is the peak RSS increase during each operation, measured in an isolated subprocess.

N	Build		Log-likelihood		Predict	
	Time [s]	Memory	Time [s]	Memory	Time [s]	Memory
10	0.0810	17 MB	0.0001	0.1 MB	0.1749	29 MB
100	0.0900	17 MB	0.0057	0.1 MB	0.2116	29 MB
1000	0.3546	137 MB	0.1416	0.1 MB	0.2892	38 MB
5000	8.4480	3.57 GB	3.0122	503 MB	0.7084	246 MB
10000	31.9826	14.16 GB	12.3798	1.97 GB	1.7840	882 MB

making MCMC chains of $\sim 10^4$ samples feasible within hours. The $\mathcal{O}(N^2)$ memory scaling of the covariance matrix is the primary limitation for larger datasets.

7.2 Banded Cholesky Solver

The compact support of the spot-lifetime envelope $R_\Gamma(\tau)$ (Section 6) guarantees that the kernel vanishes exactly for lags $|\tau| > \tau_{\max} \equiv \ell_{\text{spot}} + 2\tau_{\text{spot}}$. For uniformly sampled data with cadence Δt , this means the covariance matrix K_θ is banded with bandwidth

$$b = \left\lceil \frac{\ell_{\text{spot}} + 2\tau_{\text{spot}}}{\Delta t} \right\rceil, \quad (4.99)$$

i.e., $[K_\theta]_{ij} = 0$ whenever $|i - j| > b$. For typical stellar monitoring campaigns where the observing baseline greatly exceeds the spot lifetime ($T_{\text{obs}} \gg \tau_{\max}$), the bandwidth satisfies $b \ll N$, and the covariance matrix is sparse.

We exploit this banded structure to reduce the cost of the Cholesky factorization from $\mathcal{O}(N^3)$ to $\mathcal{O}(Nb^2)$ and the memory from $\mathcal{O}(N^2)$ to $\mathcal{O}(Nb)$ (Golub & Van Loan, 2013, Section 4.3.5). The lower Cholesky factor L inherits the banded structure of K_θ (i.e., $L_{ij} = 0$ for $i - j > b$), so both the factorization and the subsequent forward/backward substitution can be restricted to within the band.

We store L in compact banded form as a $(b + 1) \times N$ array $\tilde{L}$, where

$$\tilde{L}_{d,j} = L_{j+d,j}, \quad d = 0, 1, \dots, b, \quad (4.100)$$

so that the diagonal occupies row $d = 0$ and the d -th sub-diagonal occupies row d . The kernel is evaluated directly into this compact storage, computing only the $\mathcal{O}(Nb)$ entries within the band rather than the full N^2 matrix. The log-determinant is then $\ln |K_\theta| = 2 \sum_j \ln \tilde{L}_{0,j}$.

We implement the banded Cholesky factorization using `jax.lax.scan` over columns: at each column j , the diagonal and the b sub-diagonal entries are updated using only the $(b \times b)$ block of previously computed entries within the band. The entire factorization is

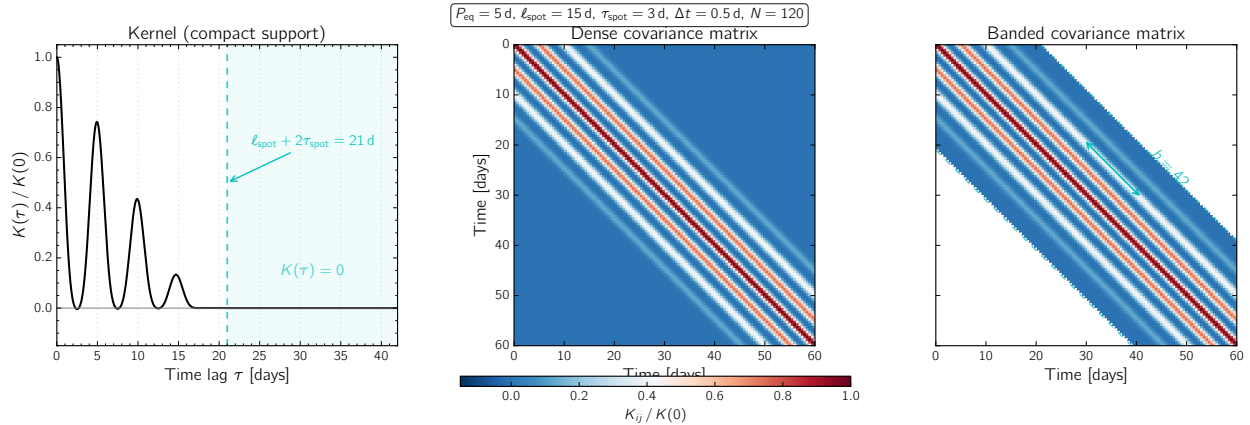


Figure 4.17: Scaling of the GP solver with bandwidth b for fixed $N = 1000$. The banded Cholesky solver scales as $\mathcal{O}(Nb^2)$, while the full Cholesky scales as $\mathcal{O}(N^3)$. For typical spot lifetimes and observing cadences, the bandwidth is small ($b \ll N$), leading to significant speedups.

Table 4.7: GPSolver wall-clock time: `cholesky_full` vs. `cholesky_banded` (JAX backend). *Build*: kernel evaluation + Cholesky factorization. *Predict*: GP predictive mean at 200 test points. Times are the median of 3 runs after JIT warmup. The full Cholesky solver scales as $\mathcal{O}(N^3)$ and becomes impractical for $N \gtrsim 10^4$; the banded solver scales as $\mathcal{O}(Nb^2)$ and handles $N = 10^5$ in seconds.

$x_{\max}$ [d]	N	Full Cholesky		Banded Cholesky		Speedup
		Build	Predict	Build	Predict	
10	100	333 ms	493 ms	535 ms	589 ms	0.6×
50	500	458 ms	669 ms	587 ms	639 ms	0.8×
100	1,000	732 ms	761 ms	595 ms	734 ms	1.2×
500	5,000	16.9 s	2.66 s	810 ms	773 ms	20.8×
1,000	10,000	36.2 s	9.55 s	957 ms	1.02 s	37.8×
5,000	50,000	—	—	2.52 s	2.52 s	—
10,000	100,000	—	—	4.47 s	4.56 s	—

JIT-compiled and end-to-end differentiable via JAX’s automatic differentiation, enabling gradient-based sampling with no additional implementation effort. The forward and backward triangular solves ($L\mathbf{y} = \mathbf{r}$ and $L^\top\mathbf{x} = \mathbf{y}$) similarly operate on the compact storage, restricting each dot product to at most b terms.

7.3 Software Architecture

We implement the kernel derivation and GP inference framework described above in the open-source Python package `spotgp`. The package is designed with a modular architecture so that each physical assumption (envelope shape, amplitude parameterization, latitude distribution) can be swapped independently without modifying the kernel or solver code.

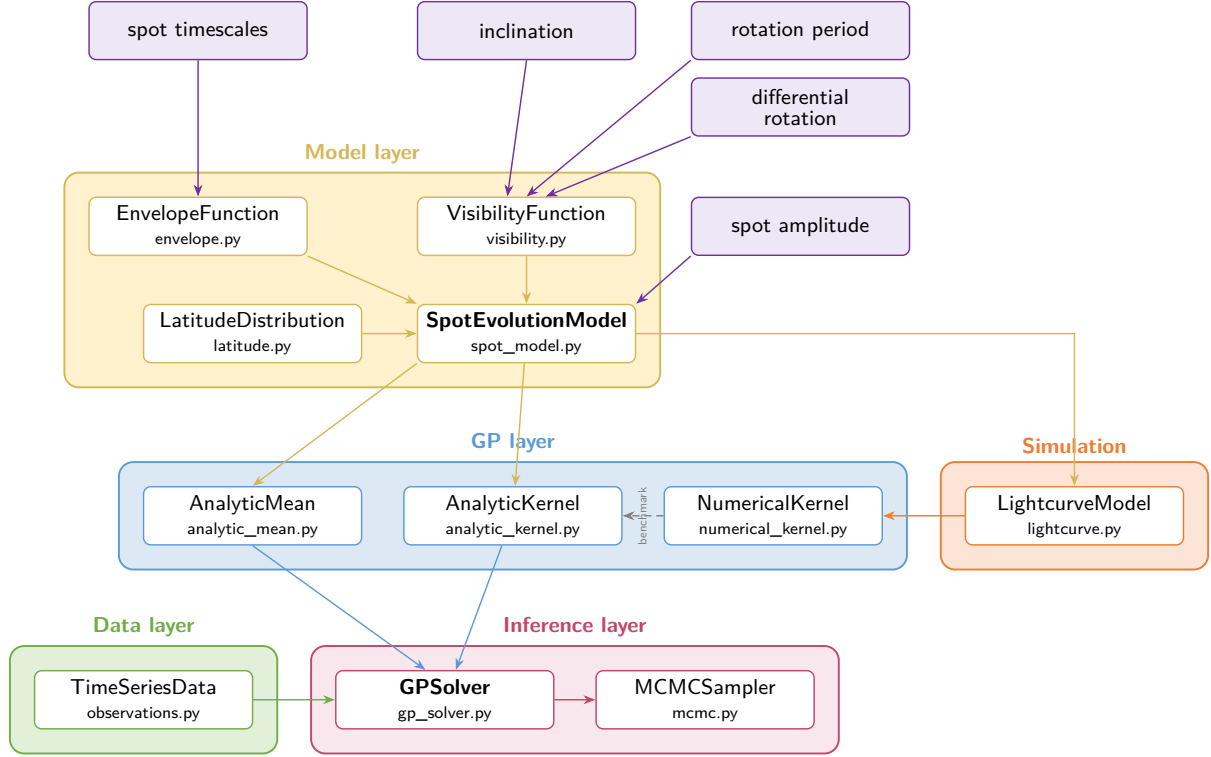


Figure 4.18: Architecture of the `spotgp` package. The core components are the `Kernel` class, which computes the covariance matrix from the kernel function; the `GPSolver` class, which performs Cholesky factorization and solves for the GP posterior; and the MCMC sampling module, which uses NUTS to sample the kernel hyperparameters. The kernel is constructed from modular components: the envelope autocorrelation $R_{\Gamma}(\tau)$, the amplitude parameterization (e.g., Fourier coefficients), and the latitude distribution $p(\Phi)$. Each component can be independently modified or extended without affecting the overall architecture.

Figure 4.18 illustrates the data flow through the main components, which we describe below.

8 MCMC Sampling the Kernel Hyperparameters

Figure 4.19 summarizes the probabilistic structure of the model, showing how the physical parameters determine the photometric kernel and how the flux observations derive from it.

Given a set of observations $\{t_i, y_i, \sigma_i\}_{i=1}^N$, we sample the posterior distribution of the kernel hyperparameters $\theta = (P_{\text{eq}}, \kappa, I, N_{\text{spot}}, \ell_{\text{spot}}, \tau_{\text{spot}}, \alpha_{\text{max}}, \sigma_k)$ using the No-U-Turn Sampler (NUTS; Hoffman & Gelman, 2014), a variant of Hamiltonian Monte Carlo (HMC) that automatically tunes the trajectory length. We implement this algorithm using the BlackJAX library (Lao et al., 2020), which provides a JAX-native NUTS kernel that is fully compatible with the JIT-compiled, autodifferentiable log-posterior described below.

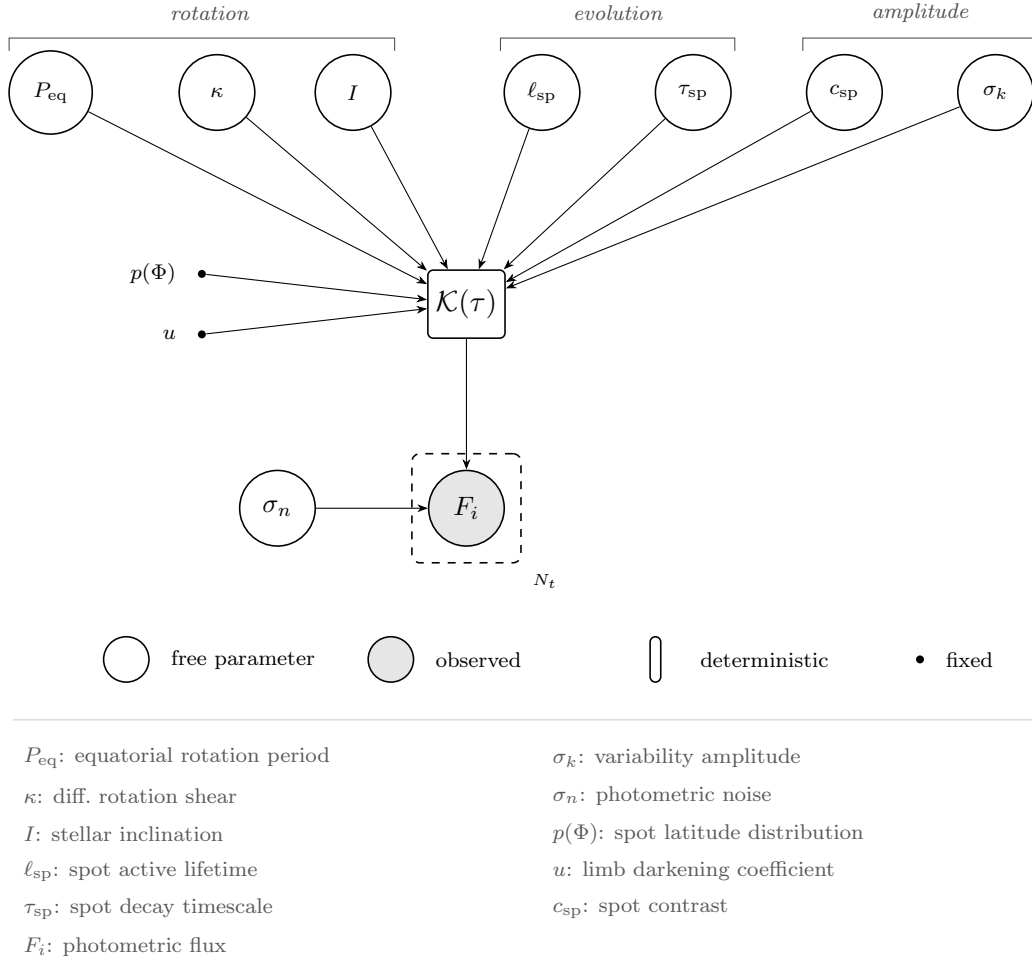


Figure 4.19: Probabilistic graphical model for the photometric starspot GP. Free parameters (open circles) determine the photometric kernel $\mathcal{K}(\tau)$ (Equation 4.70), which governs the temporal covariance of the flux observations F_i across N_t epochs. The rotation parameters (P_{eq} , κ , I) control the harmonic structure, the evolution parameters (ℓ_{sp} , τ_{sp}) control the temporal envelope $R_{\text{T}}(\tau)$, and the amplitude parameters (c_{sp} , σ_k) control the overall variability level. The spot latitude distribution $p(\Phi)$ and limb darkening coefficient u (filled dots) enter the kernel as fixed inputs from stellar models.

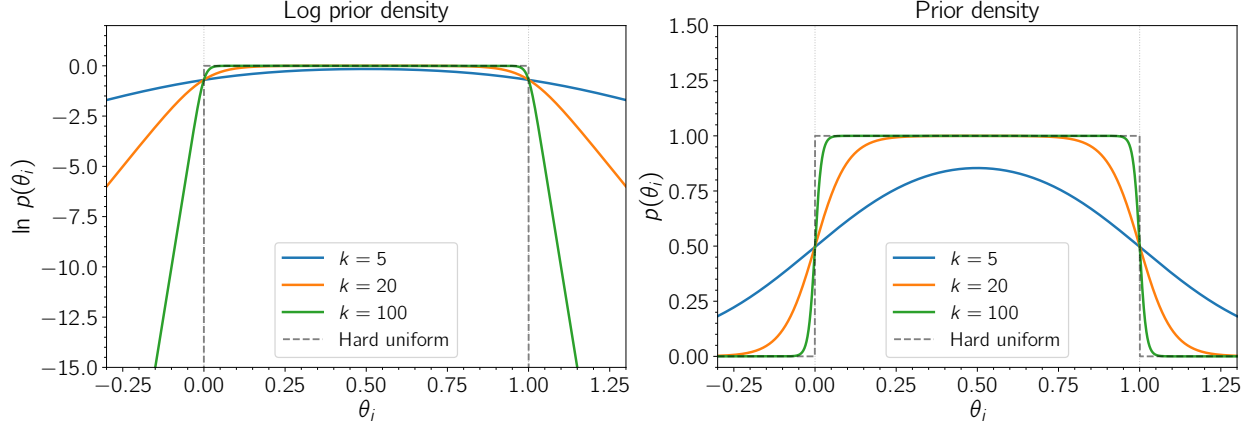


Figure 4.20: Soft uniform prior for a single parameter θ with bounds $[a, b]$. The sigmoid-based formulation provides a smooth transition at the boundaries while maintaining an approximately uniform density within the bounds.

8.1 Log-Posterior and Gradients

The GP marginal log-likelihood is

$$\ln \mathcal{L}(\boldsymbol{\theta}) = -\frac{1}{2} [\mathbf{r}^\top K_\theta^{-1} \mathbf{r} + \ln |K_\theta| + N \ln(2\pi)], \quad (4.101)$$

where $\mathbf{r} = \mathbf{y} - \boldsymbol{\mu}$ is the residual vector, $K_\theta = K(\boldsymbol{\theta}) + \text{diag}(\boldsymbol{\sigma}^2 + \sigma_n^2) + \epsilon \mathbf{I}$ is the noise-augmented covariance matrix, and $\epsilon = 10^{-8}$ is a small jitter for numerical stability. The Cholesky factorization $K_\theta = LL^\top$ is used for efficient computation of both the data-fit term and the log-determinant: $\ln |K_\theta| = 2 \sum_i \ln L_{ii}$.

The entire computational graph—from hyperparameters $\boldsymbol{\theta}$ through the kernel evaluation (Eq. 4.70), covariance matrix assembly, Cholesky decomposition, and log-likelihood—is implemented in JAX and is end-to-end differentiable. The gradient $\nabla_{\boldsymbol{\theta}} \ln p(\boldsymbol{\theta}|\mathbf{y})$ is computed via reverse-mode automatic differentiation, providing exact gradients at roughly the cost of a single likelihood evaluation.

8.2 Prior

We adopt a soft uniform prior that enforces parameter bounds while maintaining differentiability. For each parameter θ_i with bounds $[a_i, b_i]$:

$$\ln p(\boldsymbol{\theta}) = \sum_i [\ln \sigma(k(\theta_i - a_i)) + \ln \sigma(k(b_i - \theta_i)) - \ln(b_i - a_i)], \quad (4.102)$$

where $\sigma(\cdot)$ is the sigmoid function and $k = 100$ controls the steepness of the boundary (Figure 4.20). This provides effectively uniform support within the bounds while keeping the gradient finite everywhere, which is critical for HMC-based samplers.

8.3 Mass Matrix Estimation

The efficiency of NUTS depends on the choice of mass matrix M , which preconditions the Hamiltonian dynamics to account for differences in the scale and correlation of the parameters. An identity mass matrix performs poorly when the hyperparameters span vastly different scales (e.g., $P_{\text{eq}} \sim 5$ d vs. $\alpha_{\text{max}} \sim 0.1$ rad). We provide three methods for estimating the inverse mass matrix M^{-1} :

1. **Hessian at the MAP.** We first find the maximum a posteriori estimate θ_{MAP} via L-BFGS-B using JAX-computed gradients, then compute the Hessian of the negative log-likelihood at that point using `jax.hessian`:

$$M^{-1} = \left[\nabla_{\theta}^2 (-\ln \mathcal{L}) \Big|_{\theta_{\text{MAP}}} \right]^{-1}. \quad (4.103)$$

The Hessian is regularized by clamping its eigenvalues to be at least 10^{-6} , ensuring positive definiteness.

2. **Fisher information.** The expected Fisher information matrix for a GP with covariance K_{θ} is

$$\mathcal{I}_{ij} = \frac{1}{2} \text{tr} \left(K_{\theta}^{-1} \frac{\partial K_{\theta}}{\partial \theta_i} K_{\theta}^{-1} \frac{\partial K_{\theta}}{\partial \theta_j} \right). \quad (4.104)$$

The kernel derivatives $\partial K_{\theta} / \partial \theta_i$ are computed via forward-mode autodiff (`jax.jacfwd`), which is memory-efficient for the case of few parameters mapping to a large covariance matrix. The inverse mass matrix is then $M^{-1} = \mathcal{I}^{-1}$.

3. **Laplace approximation.** This is equivalent to the Hessian method but additionally provides a Gaussian approximation to the posterior, $p(\theta|\mathbf{y}) \approx \mathcal{N}(\theta_{\text{MAP}}, M^{-1})$, which can be used for rapid approximate inference or as a proposal distribution.

In practice, for NUTS we extract the diagonal of the full inverse mass matrix and use a diagonal mass matrix, which is more numerically robust. Extreme values are clamped to within a factor of 10^4 of the median diagonal element to prevent instability.

8.4 Sampler Backends

Because the `spotgp` likelihood and its gradients are implemented as pure JAX functions, they can be readily coupled to any sampler that accepts a callable log-probability (and, optionally, its gradient). We currently provide built-in interfaces to two complementary backends:

- **BlackJAX NUTS.** The default sampler is the NUTS kernel from **BlackJAX** (Lao et al., 2020), which exploits the exact gradients provided by JAX autodifferentiation to efficiently explore the posterior. This is the recommended choice when the posterior is unimodal and the mass matrix estimation described in Section 8.3 provides adequate preconditioning.

- **Dynesty nested sampling.** For problems where the posterior may be multimodal or where one additionally requires a Bayesian evidence estimate, `spotgp` also supports the `dynesty` dynamic nested sampler (Speagle, 2020b). Because nested sampling does not require gradient information, this backend uses only the log-likelihood and prior transform, making it a useful complement to NUTS when gradient-based sampling encounters difficulties such as highly correlated or multimodal posteriors.

Users may select the sampler backend at runtime, and the interface is designed so that additional samplers can be integrated with minimal effort by supplying the appropriate log-probability and prior functions.

9 Numerical Results

9.1 Analytic vs. Numerical Kernel

In this Section, we compare the analytic kernel derived in Section 5 to a numerical kernel computed by simulating many random spot realizations and empirically estimating their autocorrelation (Eq. 4.22). For the numerical kernel, we simulate $N_{\text{sim}} = 500$ realizations of the flux deficit $\delta F(t)$, each with $N_{\text{spot}} = 10$ spots drawn from the same parameter distribution. We then compute the sample autocorrelation function across all simulations to obtain an empirical estimate of the kernel. For the analytic kernel, we evaluate Eq. (4.70) numerically for the same set of parameters.

Table 4.8 summarizes the parameter combinations tested and the mean squared error (MSE) between the analytic and numerical autocorrelation functions. For each row, a single parameter is varied from its baseline value ($P_{\text{eq}} = 5$ d, $\kappa = 0$, $I = 90^\circ$, $\ell_{\text{spot}} = 15$ d, $\tau_{\text{spot}} = 3$ d, $\alpha_{\text{max}} = 0.1$ rad), with $N_{\text{spot}} = 10$ held fixed throughout.

Figures 4.21–4.25 show the analytic vs. numerical kernel comparison for each parameter variation. In all cases, the analytic kernel closely matches the numerical kernel, with MSE values on the order of 10^{-3} or smaller. Overall, these results validate our analytic kernel against numerical simulations across a range of physically relevant parameters.

Table 4.9 benchmarks the computational speed of evaluating the analytic kernel vs. computing the numerical kernel. The analytic kernel is at least an order of magnitude faster to evaluate than the numerical kernel, which requires simulating many spot realizations and computing their autocorrelation. This speed advantage makes the analytic kernel practical for use in GP regression on large datasets, where repeated kernel evaluations are required.

Table 4.8: Mean squared error (MSE) of the analytic kernel ACF compared to numerical Monte Carlo simulations, with and without stellar limb darkening (LD).

Varied param.	Value	MSE (no LD)	MSE (w/ LD)
P_{eq}	3	6.25×10^{-4}	9.38×10^{-4}
	5	7.17×10^{-4}	7.11×10^{-4}
	10	7.24×10^{-4}	7.64×10^{-4}
κ	-1.0	7.34×10^{-4}	9.71×10^{-4}
	0.0	6.70×10^{-4}	8.66×10^{-4}
	1.0	5.10×10^{-4}	6.14×10^{-4}
I	30°	1.31×10^{-3}	7.37×10^{-4}
	60°	1.31×10^{-3}	1.41×10^{-3}
	90°	8.76×10^{-4}	7.53×10^{-4}
ℓ_{spot}	5	1.36×10^{-4}	1.53×10^{-4}
	15	8.55×10^{-4}	1.03×10^{-3}
	30	1.13×10^{-3}	1.35×10^{-3}
τ_{spot}	1	3.78×10^{-4}	5.98×10^{-4}
	3	5.11×10^{-4}	4.73×10^{-4}
	8	1.15×10^{-3}	1.02×10^{-3}

Table 4.9: A comparison showing the kernel computation time (in seconds) using the analytic and numerical approaches. The analytic kernel is evaluated at varying lag-array sizes (with N simulated datapoints); the numerical (Monte Carlo) kernel uses 200 and 500 simulations over 100 d.

N	Analytic [s]	$N_{\text{sim}} = 200$		$N_{\text{sim}} = 500$	
		Numerical [s]	Speedup	Numerical [s]	Speedup
50	0.0053	0.23	42×	0.33	62×
100	0.0048	0.22	46×	0.34	71×
500	0.0050	0.26	53×	0.47	94×
1000	0.0054	0.36	66×	0.48	89×
5000	0.0075	0.50	67×	0.86	114×

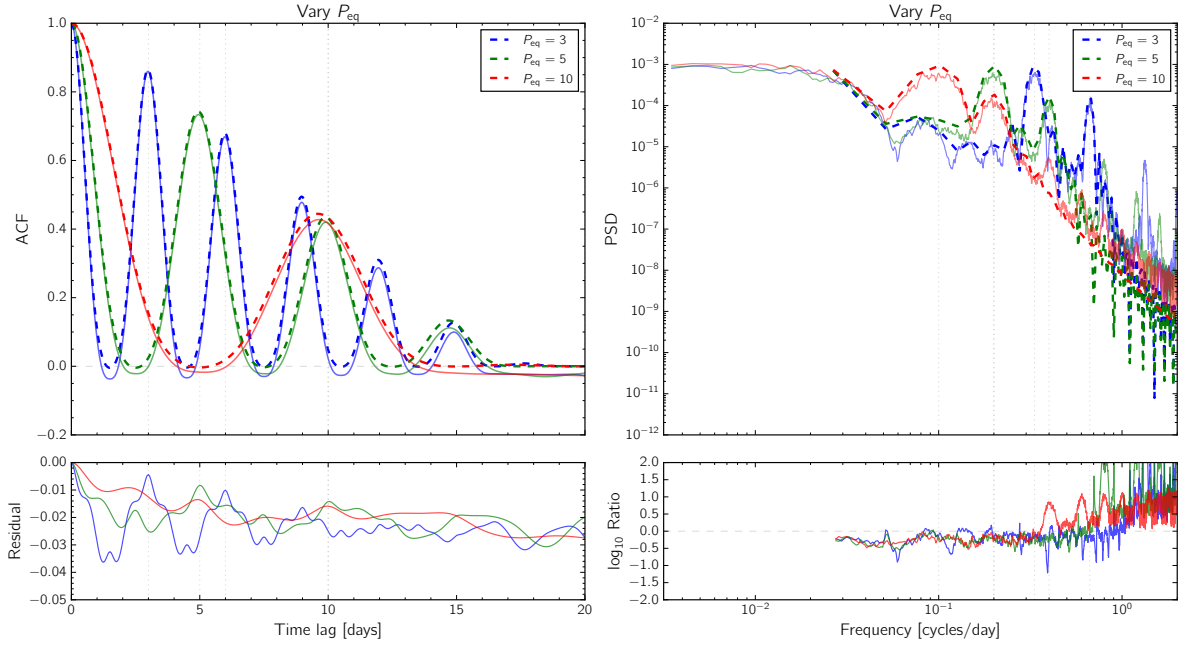


Figure 4.21: Analytic (dashed) vs. numerical Monte Carlo (solid) kernel comparison, varying $P_{\text{eq}} \in \{3, 5, 10\}$ d. The numerical kernel is computed using $N_{\text{sim}} = 500$ simulations. *Left:* Autocorrelation function. *Right:* Power spectral density. Vertical dotted lines mark the rotation frequency and its first harmonic.

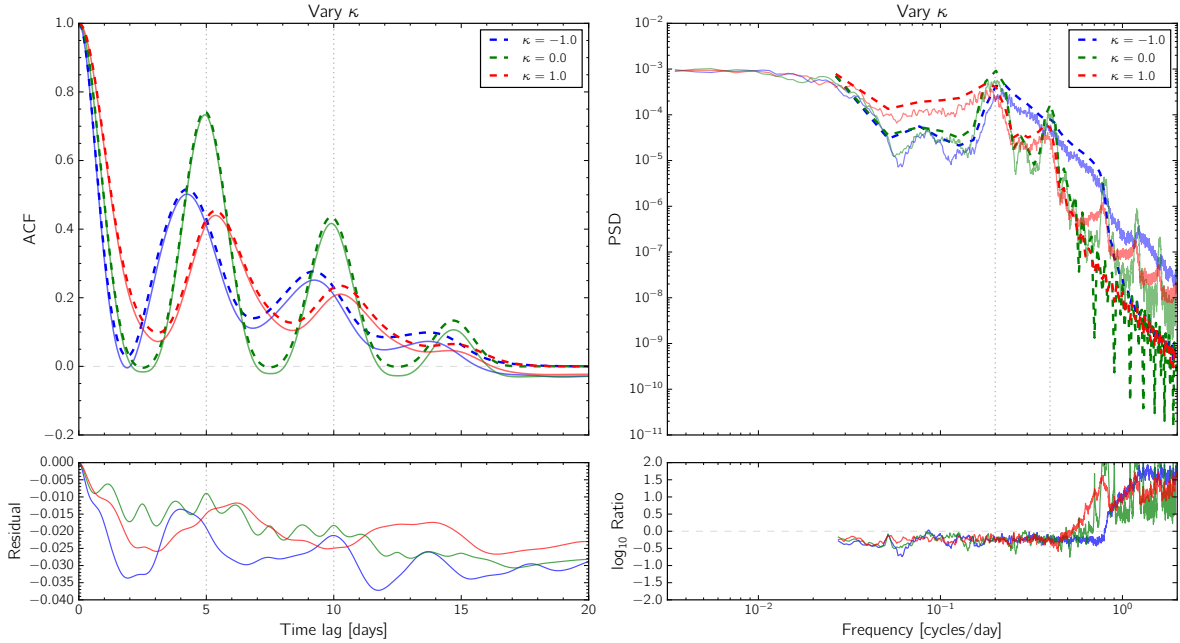


Figure 4.22: Same as Fig. 4.21, varying differential rotation shear $\kappa \in \{-1.0, 0.0, 1.0\}$.

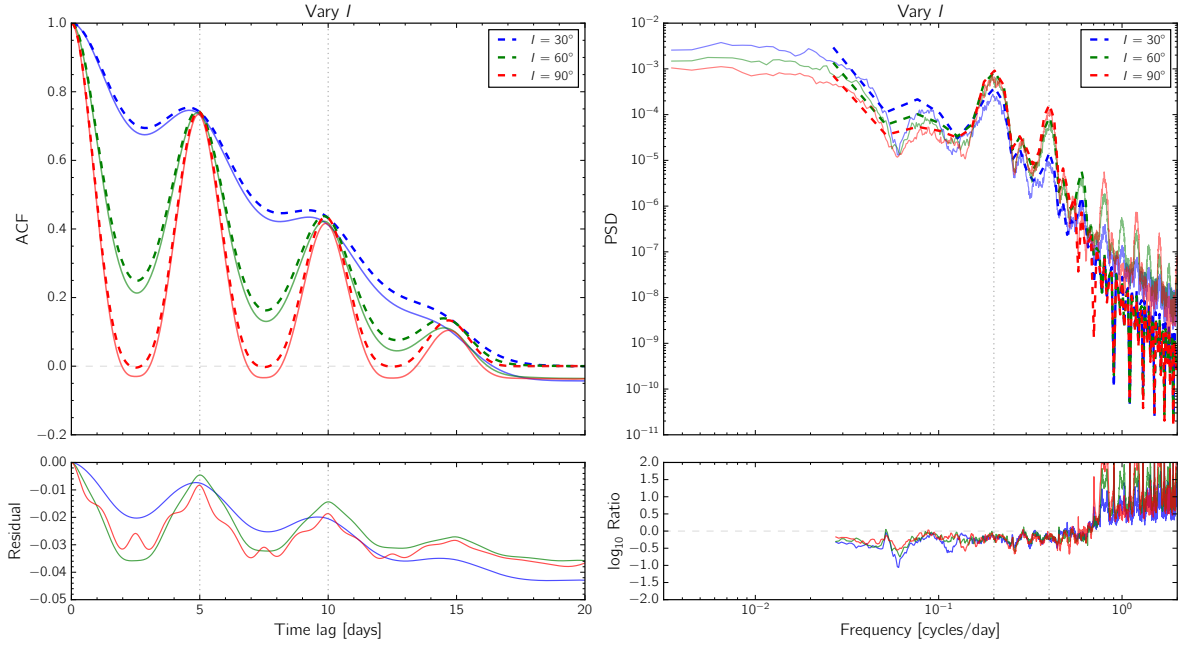


Figure 4.23: Same as Fig. 4.21, varying stellar inclination $I \in \{30^\circ, 60^\circ, 90^\circ\}$.

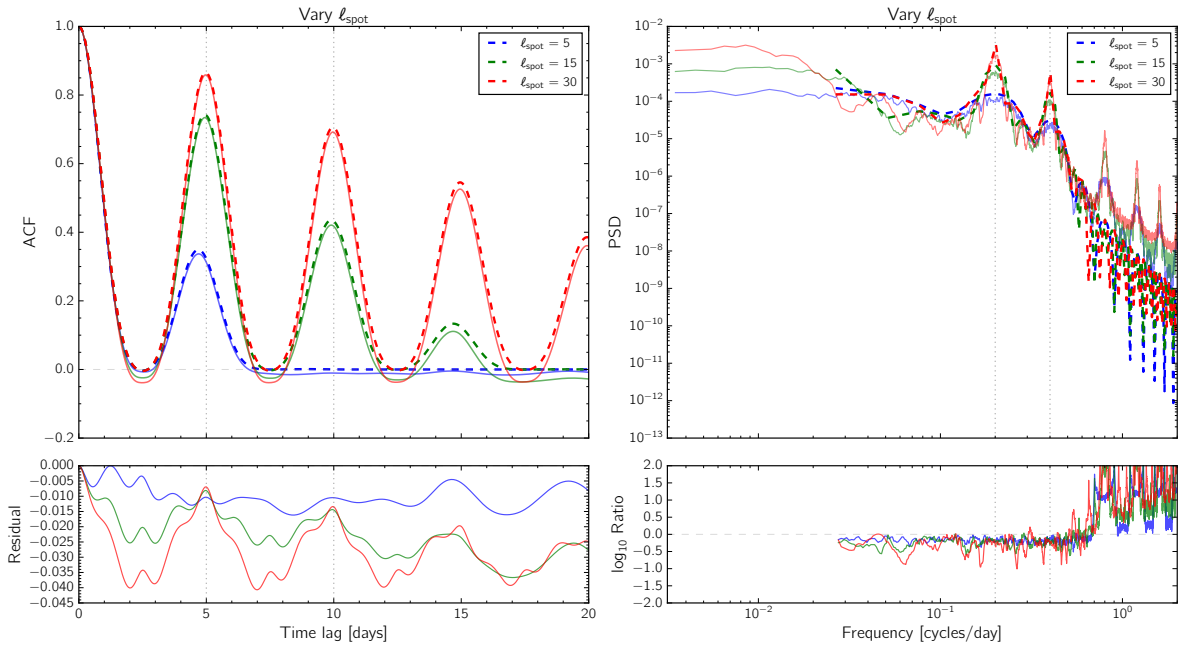


Figure 4.24: Same as Fig. 4.21, varying spot plateau duration $\ell_{\text{spot}} \in \{5, 15, 30\}$ d.

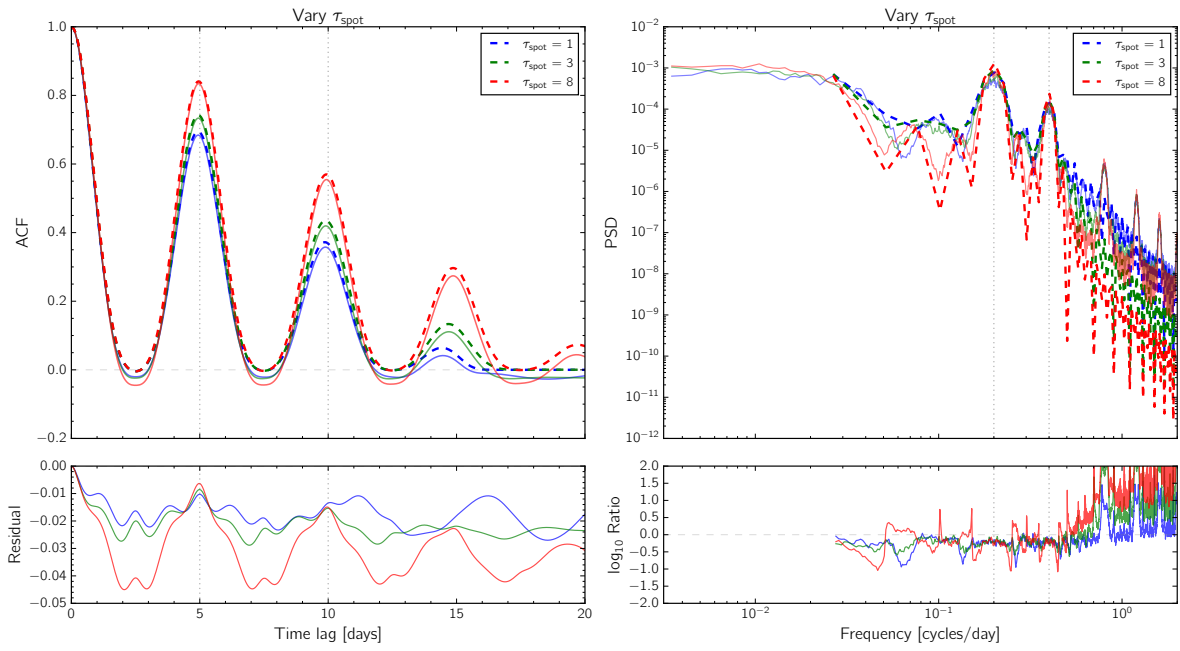


Figure 4.25: Same as Fig. 4.21, varying spot rise/decay timescale $\tau_{\text{spot}} \in \{1, 3, 8\}$ d.

9.2 Comparison to Kernels from Literature

In this Section, we compare our physically motivated kernel to commonly used kernels in the literature, such as the quasi-periodic (QP) kernel and the simple harmonic oscillator (SHO) kernel. We show that our kernel can reproduce the behavior of these kernels in certain limits, but also captures additional features such as harmonic structure and non-sinusoidal periodicity that are not present in the QP or SHO kernels.

Quasi-periodic kernels typically have the form:

$$k_{\text{QP}}(\tau) = A^2 \exp\left(-\frac{\tau^2}{2\lambda^2}\right) \exp\left(-\frac{2 \sin^2(\pi\tau/P)}{\omega^2}\right), \quad (4.105)$$

where P is the period, λ is the decay timescale, and ω controls the smoothness of the periodic component. In the limit of large ω , the periodic component approaches a pure cosine, and the kernel resembles a product of an exponential decay and a cosine.

The simple harmonic oscillator kernel has the form:

$$k_{\text{SHO}}(\tau) = A^2 \exp\left(-\frac{\tau}{\lambda}\right) \cos\left(\frac{2\pi\tau}{P}\right), \quad (4.106)$$

which is a damped cosine with a single period and an exponential decay.

Figure 4.26 compares the normalized autocorrelation function (left) and power spectral density (right) of our starspot kernel against the QP and SHO kernels for three different parameter regimes. The top panel shows a case with short-lived spots and edge-on geometry. Here, the starspot kernel exhibits a rapidly decaying ACF with prominent harmonic peaks in the PSD, while the QP and SHO kernels show a more sinusoidal ACF and a single dominant peak at the rotation frequency. The second panel shows a case with long-lived spots and inclined geometry, where the starspot kernel has a more complex ACF shape and multiple harmonic peaks in the PSD, which are not captured by the QP or SHO kernels.

The third panel of Figure 4.26 shows a case with strong solar-like differential rotation ($\kappa = 1$). Here, the starspot kernel is able to capture the “skewed” behavior around the ACF peaks (i.e., slower moving spots at higher latitudes have a longer period, thus skewing the ACF around the equatorial period peaks to slightly longer lags). For the QP and SHO kernels, shifts in the ACF peaks can only be made by adjusting the single period parameter, which cannot account for the range of rotation periods due to differential rotation. Differential rotation in the starspot kernel also produces a broadened peak in the PSD around the rotation frequency, reflecting the range of rotation frequencies across different latitudes. In contrast, the QP and SHO kernels cannot capture these features, as they assume a single period and do not account for differential rotation.

Overall, this comparison demonstrates that our physically motivated starspot kernel can capture a richer variety of behaviors in the autocorrelation and power spectrum than commonly used kernels, making it a more flexible, interpretable (and potentially more accurate) choice for modeling stellar variability due to starspots.

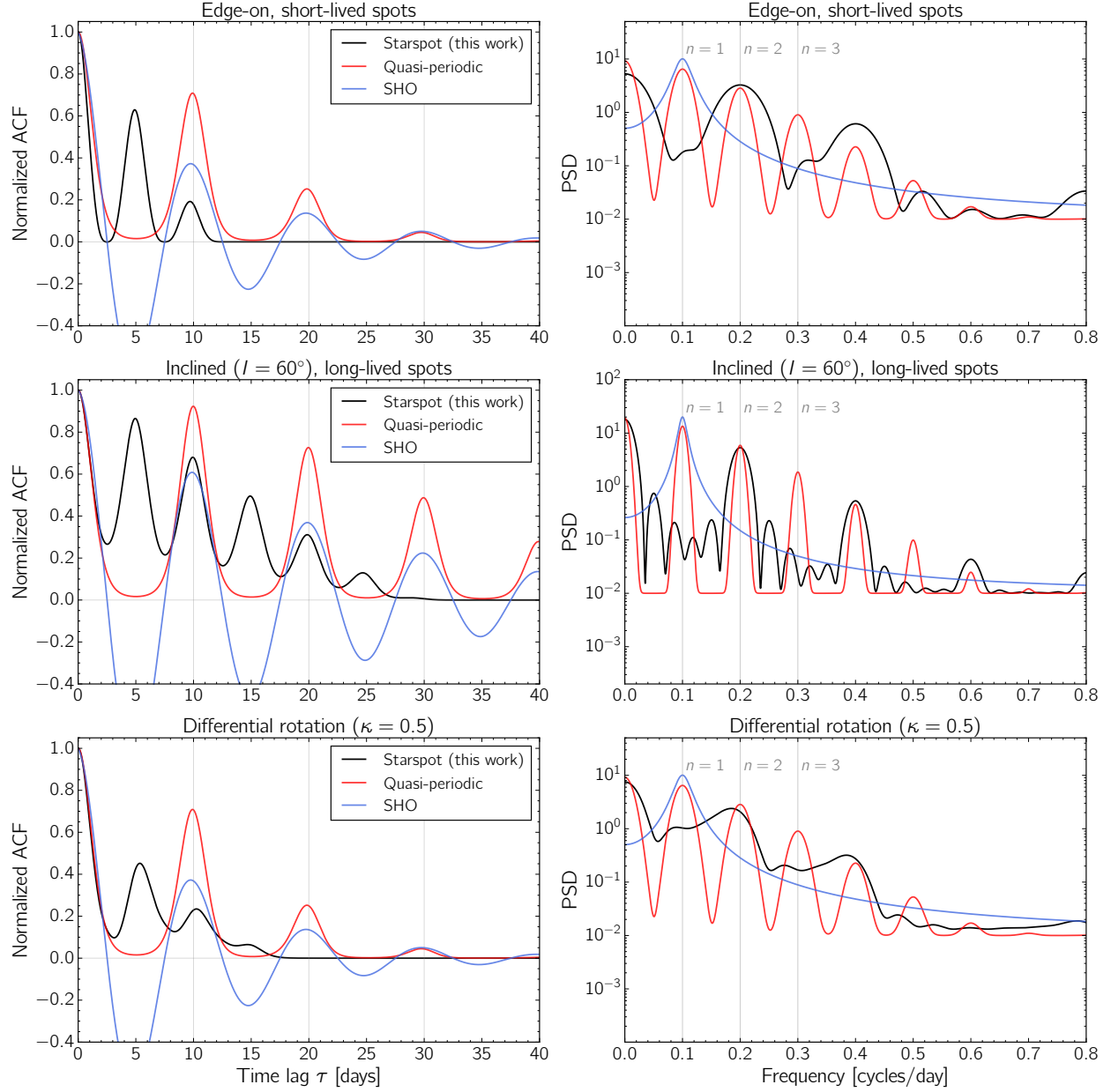


Figure 4.26: Comparison of the starspot kernel (this work) with the quasi-periodic (QP) and simple harmonic oscillator (SHO) kernels. **Left:** Normalized autocorrelation functions. **Right:** Power spectral densities. **Top row:** Edge-on geometry with short-lived spots ($P_{\text{eq}} = 5$ d, $I = 90^\circ$, $\ell = 10$ d, $\tau = 3$ d). **Bottom row:** Inclined geometry with long-lived spots ($P_{\text{eq}} = 5$ d, $I = 60^\circ$, $\ell = 25$ d, $\tau = 5$ d). Vertical dotted lines in the PSD panels mark integer harmonics of the rotation frequency. The starspot kernel naturally produces harmonic peaks and non-sinusoidal ACF shapes that the QP and SHO kernels cannot capture.

9.3 Comparison to Other GP Packages

While our physically motivated kernel provides richer structure than standard QP or SHO kernels, practical adoption depends on computational performance. We benchmark the

spotgp GP solver against two widely used packages: **tinygp** (Foreman-Mackey, 2024), which implements the **celerite** $\mathcal{O}(N)$ semi-separable solver (Foreman-Mackey et al., 2017), and **george** (Ambikasaran et al., 2015), which uses a standard $\mathcal{O}(N^3)$ Cholesky factorization.

For a fair comparison, we use a kernel that is representable in all three frameworks: the one-sided exponential envelope $\Gamma(t) = e^{-t/\tau_{\text{spot}}} H(t)$, which gives $R_\Gamma(\tau) = (2a)^{-1} e^{-a|\tau|}$ with $a = 1/\tau_{\text{spot}}$. With $N_{\text{harm}} = 2$ harmonics, this decomposes into $J = 3$ **celerite** terms (one real + two complex), making it compatible with the $\mathcal{O}(N)$ solver while also being evaluable by the standard dense and banded solvers.

Figure 4.27 shows the wall-clock time for a single log-likelihood evaluation as a function of the number of datapoints N , with fixed cadence $\Delta t = 1$ d. We compare four solver configurations: (1) **tinygp** with the **celerite** $\mathcal{O}(N)$ solver, (2) **george** with a dense $\mathcal{O}(N^3)$ Cholesky, (3) **spotgp** with a banded $\mathcal{O}(Nb^2)$ Cholesky, and (4) **spotgp** with a full $\mathcal{O}(N^3)$ Cholesky. All JAX-based solvers (**tinygp** and **spotgp**) are JIT-compiled and run on the GPU, while **george** uses compiled C/LAPACK and can only run on the CPU.

The top panel of Figure 4.27 shows that **tinygp**’s **celerite** solver scales linearly with N and has the lowest overhead at small dataset sizes. The **george** solver scales as N^3 but benefits from highly optimized C/LAPACK routines, making it competitive for $N \lesssim 500$. The **spotgp** banded solver carries higher per-evaluation overhead than **tinygp** at small N , but because the bandwidth b is set by the kernel correlation length and does not grow with N , its $\mathcal{O}(Nb^2)$ scaling is effectively linear for long baselines. At $N = 10,000$, the banded solver reaches wall-clock times comparable to **tinygp** (~ 0.5 s), while **george** exceeds 30 s. The banded solver overtakes the full $\mathcal{O}(N^3)$ Cholesky around $N \sim 500$.

The middle panel shows matrix storage requirements. The **celerite** solver stores only $\mathcal{O}(NJ)$ elements (where $J = 2N_{\text{harm}} + 1$ is the number of semi-separable terms), keeping memory below a few MB even at $N = 10,000$. The banded solver stores the $(2b + 1) \times N$ band, which grows linearly with N for fixed b . In contrast, the dense solvers (**george** and **spotgp** full) store the full $N \times N$ covariance matrix, reaching ~ 800 MB at $N = 10,000$. For the full **spotgp** solver run on the GPU, the max matrix storage size is limited to the memory capacity of the GPU, which is typically 16–32 GB for consumer hardware (in this case, an Nvidia RTX 5070 TI with 16 GB of memory), thus limiting the maximum N to a few thousand. Similarly for the full **george** solver, the max matrix size is limited by the RAM capacity of the CPU, which is typically 16–64 GB for consumer hardware, thus also limiting the maximum N to a few thousand. In comparison, the banded solver’s linear memory scaling allows it to handle much larger datasets without hitting memory limits.

The bottom panel shows the sparsity fraction $(2b + 1)/N$ of the banded covariance matrix. Because the bandwidth is determined by the kernel’s correlation length (a few rotation periods), it remains fixed as the observation baseline grows. The sparsity fraction therefore drops as $1/N$, falling below 10% for $N \gtrsim 1000$, which is the regime where the banded solver’s advantage is most pronounced.

The **celerite** solver’s $\mathcal{O}(N)$ scaling is only available for kernels that decompose into a sum of (damped) exponentials and cosines. Our trapezoidal and skew-normal envelopes do not admit this decomposition, so the banded Cholesky solver is the only option for the general kernel. For applications where speed is critical and the one-sided exponential envelope is a sufficient approximation (e.g., for noise modeling, or applications like lightcurve detrending where the user does not care about interpreting the hyperparameters), the **celerite**

decomposition provides a practical fast path.

9.3.1 Comparison with `starry_process`

An alternative GP framework for stellar variability is the `starry_process` of Luger et al. (2021b,c), which models the stellar surface as a linear combination of spherical harmonics $Y_{\ell m}$ and derives the covariance of the disk-integrated flux analytically. Our approach differs in several fundamental respects; we summarize the key distinctions here.

Surface representation: `starry_process` represents the stellar surface intensity as a truncated spherical harmonic expansion, $I(\theta, \phi) = \sum_{\ell, m} y_{\ell m} Y_{\ell m}(\theta, \phi)$, with the flux computed via the linear design-matrix relation $\mathbf{f} = \mathbf{A} \mathbf{y}$, where $\mathbf{A}$ encodes the rotational geometry and limb darkening (Luger et al., 2019). This yields a non-parametric surface model: the coefficients $y_{\ell m}$ are treated as Gaussian random variables, and the GP covariance of the flux follows from the prior covariance of the harmonic coefficients, $\Sigma_f = \mathbf{A} \Sigma_y \mathbf{A}^\top$. By contrast, our model parameterizes the surface in terms of individual circular spots (each described by a latitude Φ , longitude evolution $\lambda(t)$, and trapezoidal size envelope $\alpha(t)$), and derives the kernel by taking the autocorrelation of the resulting flux analytically.

Temporal evolution: A key distinction is the treatment of time dependence. The `starry_process` covariance is *non-stationary*: it depends on absolute time through the rotation matrix $\mathbf{A}(t)$, since the surface pattern is treated as fixed (or drawn once) and the star merely rotates. There is no built-in mechanism for spot emergence, evolution, or decay; the surface is effectively a static random field viewed from different angles. Our kernel, by contrast, is *stationary* (depending only on the time lag $\tau = |t_2 - t_1|$) and explicitly encodes spot lifetimes through the envelope autocorrelation $R_\Gamma(\tau)$, which models the birth, plateau, and decay of each spot. This stationarity arises because we marginalize over the random arrival times of spots (Section 5.7), treating the spot process as a temporal Poisson process. As a consequence, our kernel naturally captures the finite coherence time of the variability signal: the autocorrelation decays to zero at $\tau = \ell_{\text{spot}} + 2\tau_{\text{spot}}$, a feature that has no analog in `starry_process`.

Differential rotation: Our model includes differential rotation through the shear parameter κ , which allows spots at different latitudes to rotate at different angular velocities (Section 6). This produces observable effects including peak broadening, decoherence, and the coherence lag τ_{coh} (Figure 4.16). The `starry_process` framework assumes solid-body rotation.

Spot parameterization: In `starry_process`, the surface statistics are controlled by the angular power spectrum of the spherical harmonic coefficients (parameterized by a characteristic angular scale and latitude-dependent modulation). These parameters describe the *spatial* statistics of the surface but do not map directly onto physical spot properties. In our model, the hyperparameters have direct physical interpretations: P_{eq} is the equatorial rotation period, ℓ_{spot} is the spot plateau lifetime, τ_{spot} is the emergence/decay timescale, α_{max} is the peak angular spot radius, and κ is the differential rotation shear.

Computational structure: The `starry_process` covariance matrix Σ_f is dense and generally $O(N^3)$ to factorize, as it is for any GP. The computational advantage of that framework lies in the analytic evaluation of the matrix entries (avoiding numerical integration over the stellar surface). Our kernel shares this advantage—its evaluation reduces to sums of cosines weighted by tabulated Fourier coefficients—but additionally benefits from *compact support*:

$K(\tau) = 0$ for $|\tau| > \ell_{\text{spot}} + 2\tau_{\text{spot}}$, which produces a banded covariance matrix with bandwidth $b = (\ell_{\text{spot}} + 2\tau_{\text{spot}})/\Delta t$. This enables $O(Nb^2)$ Cholesky factorization via banded linear algebra, a significant speedup for long time series with fine cadence.

Degeneracies and information content: [Luger et al. \(2021c\)](#) showed that there exist fundamental degeneracies in mapping stellar surfaces from disk-integrated photometry alone: infinitely many surface maps can produce identical lightcurves. These degeneracies persist in both frameworks. In `starry_process`, they manifest as a null space of the design matrix $\mathbf{A}$; in our model, they appear as the well-known degeneracy between N_{spot} , f_{spot} , and α_{max} (which are absorbed into the single amplitude parameter σ_k^2 ; Eq. 4.82), and the fact that the kernel constrains only the latitude-averaged squared Fourier coefficients $\langle |c_n|^2 \rangle_\Phi$, not the surface map itself. However, our kernel provides direct constraints on *temporal* properties (spot lifetime, evolution timescale, differential rotation) that are not accessible from the static surface model of `starry_process`.

Summary: The two approaches are complementary. `starry_process` is well suited for modeling single-epoch or short-baseline observations where the surface can be treated as approximately static and the goal is to characterize the spatial statistics of surface features. Our kernel is designed for longer time baselines where spot evolution, finite lifetimes, and differential rotation are important, and where a stationary GP with physically interpretable hyperparameters and compact support is desirable.

9.4 Inference Conditions

Before applying our GP model to real data, we first test its ability to recover known parameters from synthetic datasets. We generate mock lightcurves by simulating the flux using the same spot model (Equation 4.1) that underlies our kernel, with known hyperparameters θ_{true} . We then add Gaussian noise (with a standard deviation σ_n) to the simulated fluxes to create realistic observations.

9.4.1 Kernel Convergence and Data Requirements

As a first test, we estimate the minimum observation duration (T_{obs}) and sampling cadence (Δt) required for the ACF to converge within a certain error of the true kernel function. To perform this test, we generate a grid of synthetic timeseries with varying T_{obs} , Δt , and σ_n , and numerically compute the ACF for each individual lightcurve. We then compare the empirical ACF to the true kernel function (computed using the known parameters) and compute the mean squared error (MSE) as a function of T_{obs} , Δt , and σ_n . This analysis identifies the regimes in which the ACF provides a reliable estimate of the kernel, which is critical for informing the design of observing campaigns and setting expectations for parameter recovery.

Figure 4.28 shows the results of this test for two sampling cadences: $\Delta t = 0.5$ d (representative of space-based photometry such as *Kepler* or *TESS*) and $\Delta t = 5.0$ d (representative of ground-based monitoring campaigns). In each panel, the x -axis gives the number of datapoints in the lightcurve $N = T_{\text{obs}}/\Delta t$, and the y -axis gives the MSE between the sample ACF (computed from a single simulated lightcurve) and the analytic kernel, averaged over 10 random realizations with 1σ errorbars. Different colored lines correspond to different Gaussian noise levels σ_{noise} , spanning from 10^{-4} (photon-noise-limited space pho-

tometry) to 10^{-1} (heavily contaminated data). The spot emergence rate is held fixed at $\dot{N}_{\text{spot}} = 0.25 \text{ spots day}^{-1}$, so that longer baselines contain proportionally more spots and the kernel amplitude is independent of T_{obs} (Section 5.7).

Several trends are apparent. First, for low noise levels ($\sigma_{\text{noise}} \leq 10^{-3}$), the MSE decreases steadily with increasing N , roughly as a power law, reflecting the improved statistical convergence of the sample ACF as more independent spot events contribute to the estimate. At the highest noise levels ($\sigma_{\text{noise}} = 10^{-1}$), the noise dominates the signal and the MSE saturates at $\sim 10^{-1}$ regardless of the number of datapoints, indicating that the spot-induced correlations are buried in the noise floor.

Second, comparing the two cadences, the coarser sampling ($\Delta t = 5.0 \text{ d}$) yields systematically higher MSE at fixed N for the intermediate noise levels ($\sigma_{\text{noise}} = 10^{-2}$), because fewer samples per rotation period degrade the ACF estimate at short lags. For the fine cadence ($\Delta t = 0.5 \text{ d}$), the low-noise curves ($\sigma_{\text{noise}} \leq 10^{-3}$) converge to $\text{MSE} \lesssim 10^{-3}$ with $N \gtrsim 10^3$ datapoints (corresponding to $T_{\text{obs}} \gtrsim 500 \text{ d}$), while the coarse cadence requires $N \gtrsim 10^2$ – 10^3 points ($T_{\text{obs}} \gtrsim 500$ – 5000 d) to reach comparable accuracy.

These results suggest that for typical spot parameters, reliable kernel estimation from a single lightcurve requires observation baselines spanning at least ~ 10 – 20 spot lifetimes ($\ell_{\text{spot}} + 2\tau_{\text{spot}} = 21 \text{ d}$ in our test case), with noise levels $\sigma_{\text{noise}} \lesssim 10^{-2}$ in relative flux units.

9.5 Testing the Cranmer–Rao Bound of Different Parameter Regimes

Even though we are only modeling a six parameter kernel model, there are many possible combinations of parameters that can produce similar kernel shapes, leading to degeneracies and non-identifiability. For example, if the spot lifetime is very short compared to the rotation period, the kernel will be dominated by the envelope autocorrelation and may not contain enough information about the periodic structure to constrain P_{eq} or κ . Conversely, if the spot lifetime is relatively long compared to the rotation period or observational baseline, the kernel may be dominated by the periodic structure and have a very slowly decaying ACF, which could also make it difficult to constrain the envelope parameters ℓ_{spot} and τ_{spot} .

Before performing MCMC sampling (which is computationally intensive to complete across a wide parameter space), we can first test the theoretical identifiability of the parameters by computing the Fisher information matrix (Eq. 4.104) across a grid of parameters. The inverse of the Fisher information matrix gives the Cramér–Rao bound (CRB; Cramér, 1946; Rao, 1945) on the covariance of any unbiased estimator of the parameters, which provides a theoretical limit on how well we can constrain the parameters from data generated by our model. By analyzing the eigenvalues and eigenvectors of the Fisher information matrix, we can identify which combinations of parameters are most and least identifiable, and how this depends on the underlying spot parameters. This analysis will inform our expectations for MCMC sampling and help us interpret the results in terms of parameter degeneracies and uncertainties.

To systematically explore parameter space, we draw $N = 4096$ parameter combinations from a scrambled Sobol quasi-random sequence spanning the bounds listed in Table 4.10. For each sample θ_k , we construct a synthetic lightcurve of duration $T_{\text{sim}} = 200 \text{ d}$ sampled at a cadence of $\Delta t = 0.25 \text{ d}$ (i.e., $N_t = 800$ observations) with white noise $\sigma_n = 10^{-4}$ and a mean spot emergence rate of $\dot{N}_{\text{spot}} = 0.25 \text{ d}^{-1}$. We then evaluate the expected Fisher

information matrix $\mathcal{I}(\boldsymbol{\theta}_k)$ via Eq. (4.104) and invert it to obtain the CRB covariance matrix $\Sigma_{\text{CRB}} = \mathcal{I}^{-1}$. Cases where any eigenvalue of $\mathcal{I}$ falls below a regularization floor of 10^{-6} are flagged as numerically singular and excluded from the analysis, as the corresponding CRB is unreliable. From the CRB covariance we extract the marginal standard deviations $\sigma_{\text{CRB}}(\theta_i) = \sqrt{\Sigma_{ii}}$ and the pairwise Pearson correlation coefficients $r_{ij} = \Sigma_{ij}/(\sigma_i\sigma_j)$ for all parameter pairs.

Figures 4.29–4.32 show the results of this analysis for the key parameters κ , I , P_{eq} , and σ_k . Each panel shows a corner plot of the parameter space, with the color indicating the average percent uncertainty on κ (or the parameter of interest) as predicted by the CRB. The plots reveal that certain regions of parameter space allow for much tighter constraints on κ than others. For example, systems with short periods ($P_{\text{eq}} \lesssim 20$ d), long spot lifetimes ($\ell_{\text{spot}} + 2\tau_{\text{spot}} \gtrsim 10P_{\text{eq}}$), and high inclinations ($I \gtrsim 45^\circ$) can theoretically constrain κ to within $\lesssim 10\%$ uncertainty, while other regions (e.g., pole-on inclinations or short spot lifetimes) yield much looser constraints (in some cases $\gtrsim 100\%$ uncertainty).

Table 4.10: Parameter bounds used for the Sobol quasi-random sweep of the CRB analysis.

Parameter	Lower	Upper
P_{eq} [d]	0.1	50
κ	−1.0	2.0
I [deg]	−1	91
ℓ_{spot} [d]	0.1	50
τ_{spot} [d]	0.1	50
σ_k	10^{-3}	1

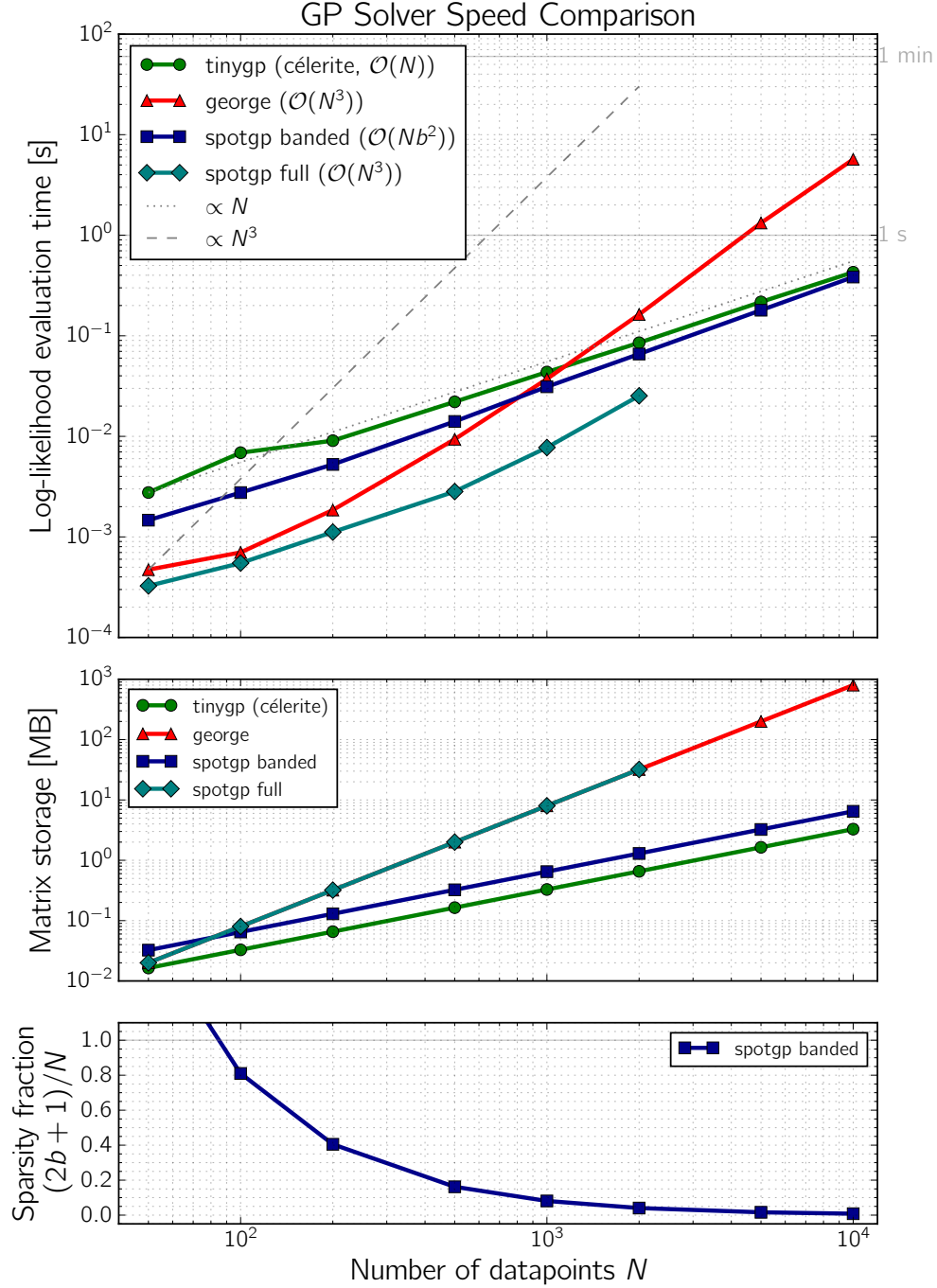


Figure 4.27: Benchmark of GP solver performance as a function of the number of datapoints N , using a one-sided exponential envelope with $N_{\text{harm}} = 2$ harmonics and cadence $\Delta t = 1$ d. *Top:* Wall-clock time for a single log-likelihood evaluation. **tinygp** (célerite, green) scales as $\mathcal{O}(N)$; **george** (red) and **spotgp** full Cholesky (teal) scale as $\mathcal{O}(N^3)$; **spotgp** banded (dark blue) scales as $\mathcal{O}(Nb^2)$. Horizontal gray lines mark 1 s and 1 min. *Middle:* Matrix storage in MB. The **célerite** and banded solvers scale linearly with N , while the dense solvers scale as N^2 . *Bottom:* Sparsity fraction $(2b+1)/N$ of the banded covariance matrix, showing that the bandwidth b remains fixed as N grows. All JAX solvers are JIT-compiled and warmed up before timing.

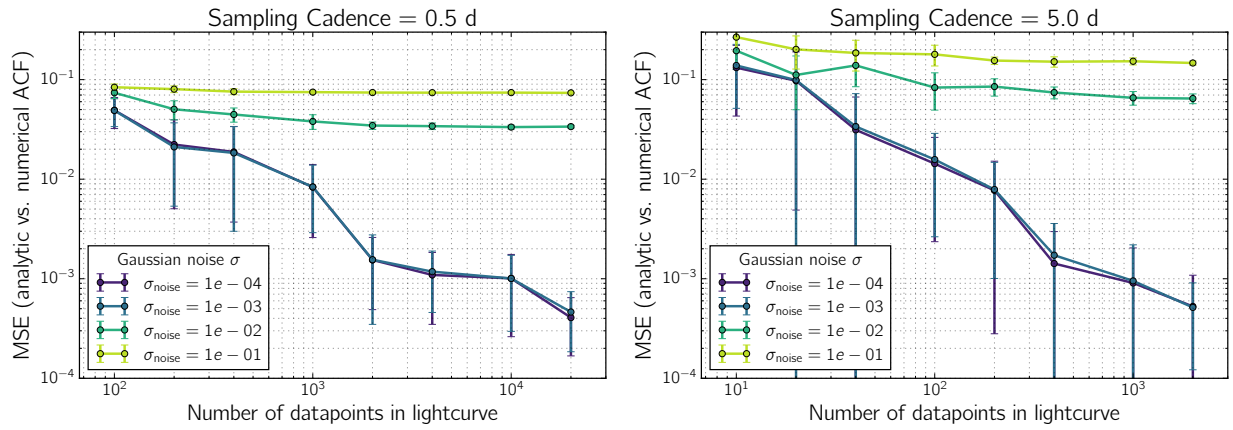


Figure 4.28: MSE between the analytic kernel and the sample ACF computed from individual simulated lightcurves, as a function of the number of datapoints N , for a sampling cadence of $\Delta t = 0.5$ d (left) and $\Delta t = 5.0$ d (right). Each line corresponds to a different Gaussian noise level σ_{noise} added to the simulated flux. Points show the mean over 10 random realizations; errorbars indicate the 1σ scatter. The spot emergence rate is $\dot{N}_{\text{spot}} = 0.25 \text{ spots day}^{-1}$, so N_{spot} scales with the observation duration.

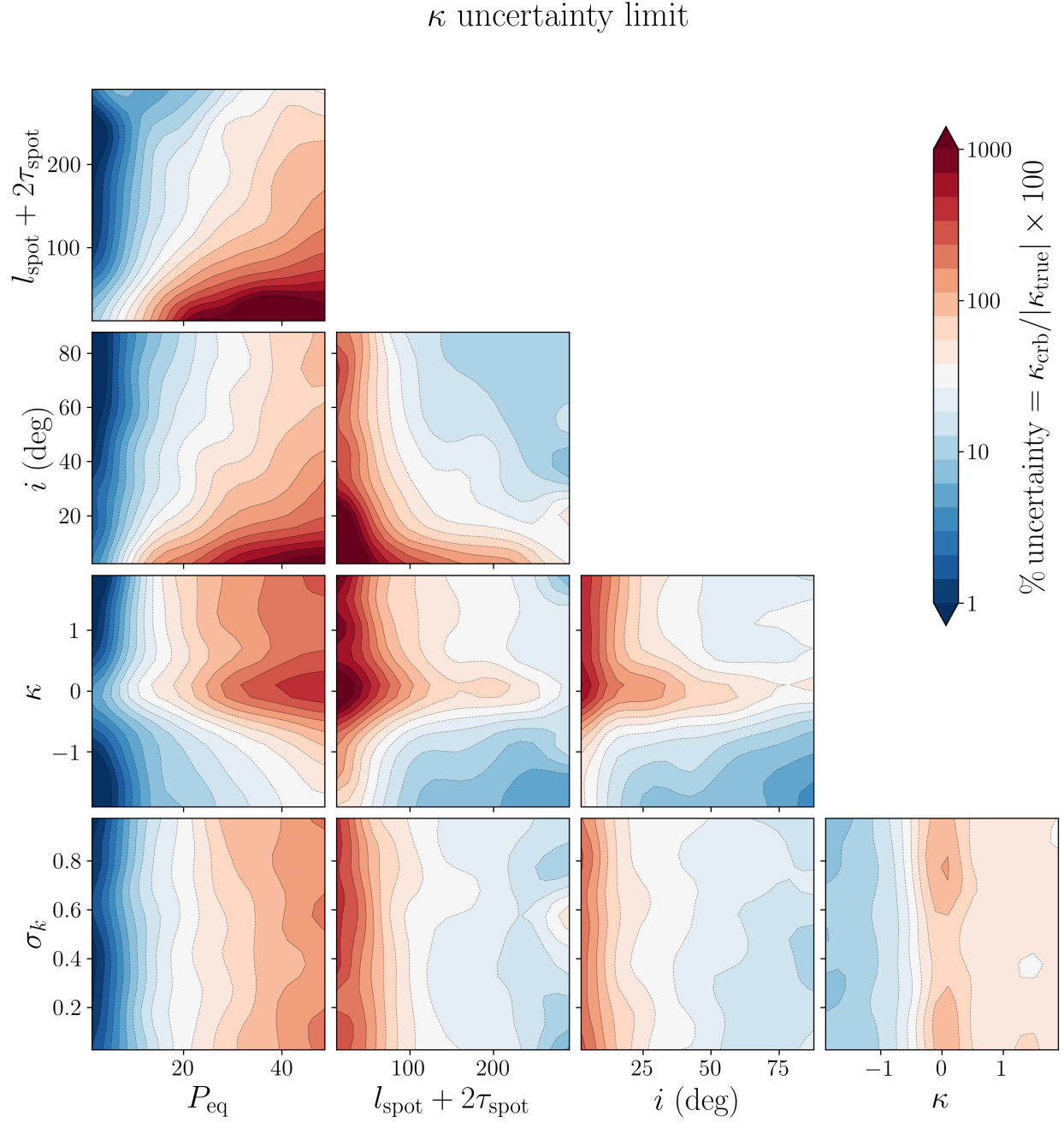


Figure 4.29: Cramér–Rao bound parameter sweep corner plot showing the theoretical lower bounds for κ uncertainty across different parameter regimes. To compute the average percent uncertainty for each parameter region, we computed 4096 simulations within the bounds of Table 4.10 using Sobol quasi-random sampling. The average percent uncertainty on κ is indicated by the color scale, with dark blue indicating tighter constraints and dark red indicating looser constraints. For some regions, with short periods ($P_{\text{eq}} \lesssim 20$), long spot lifetimes ($l_{\text{spot}} + 2\tau_{\text{spot}} \gtrsim 10P_{\text{eq}}$), and high inclinations ($i \gtrsim 45^\circ$), the differential rotation shear κ can theoretically be constrained within $\lesssim 10\%$ uncertainty.

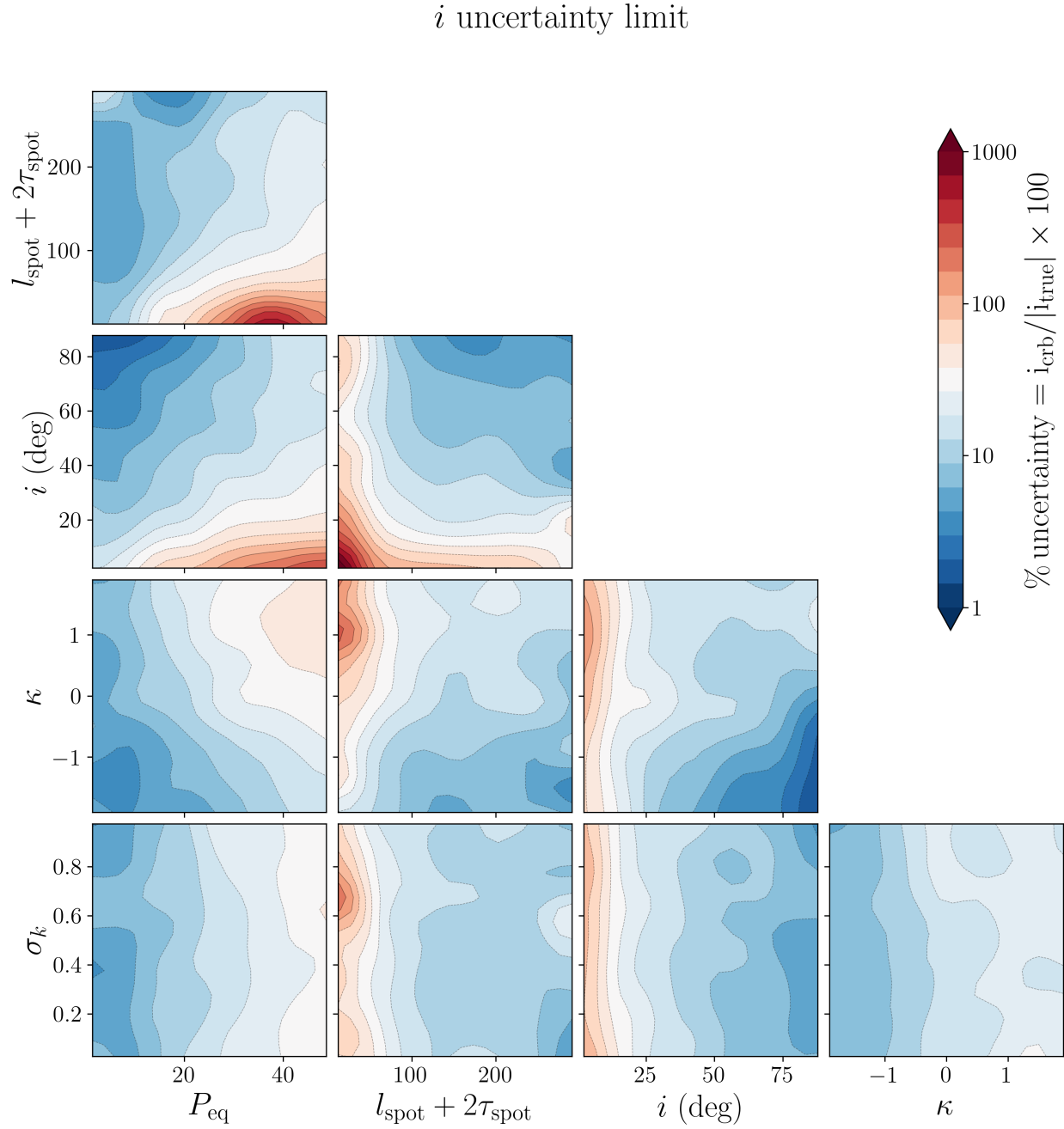


Figure 4.30: Same as Fig. 4.29, but colored by inclination I .

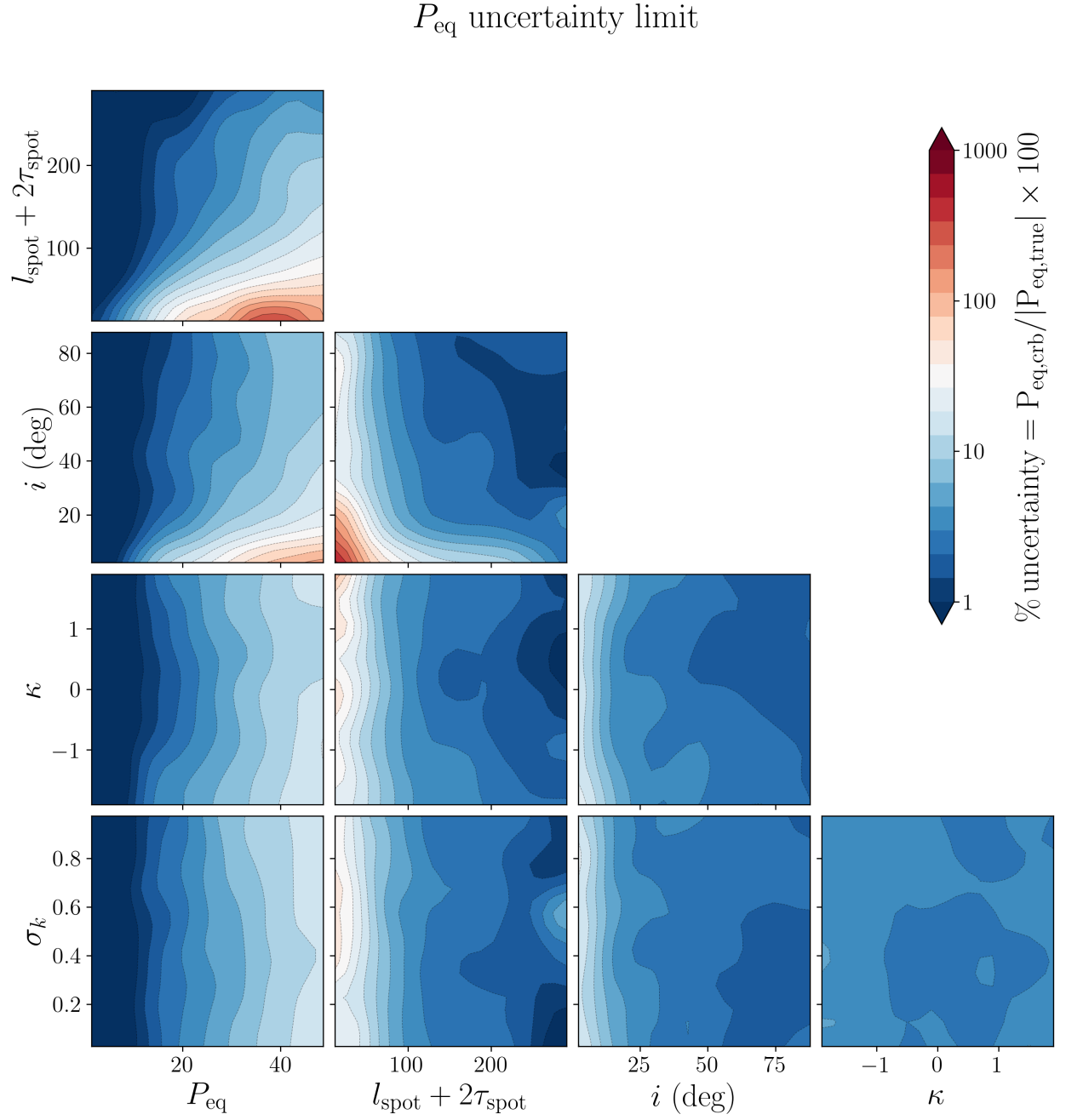


Figure 4.31: Same as Fig. 4.29, but colored by P_{eq} . P_{eq} is by far the easiest of the parameters to constrain. With the exception of a few regions, e.g., systems with short spot lifetimes relative to their period ($l_{\text{spot}} + 2\tau_{\text{spot}} \lesssim P_{\text{eq}}$), at pole-on inclinations ($i \lesssim 30$ deg), the equatorial period can generally be constrained to better than 10% uncertainty, and in many cases better than 1% uncertainty.

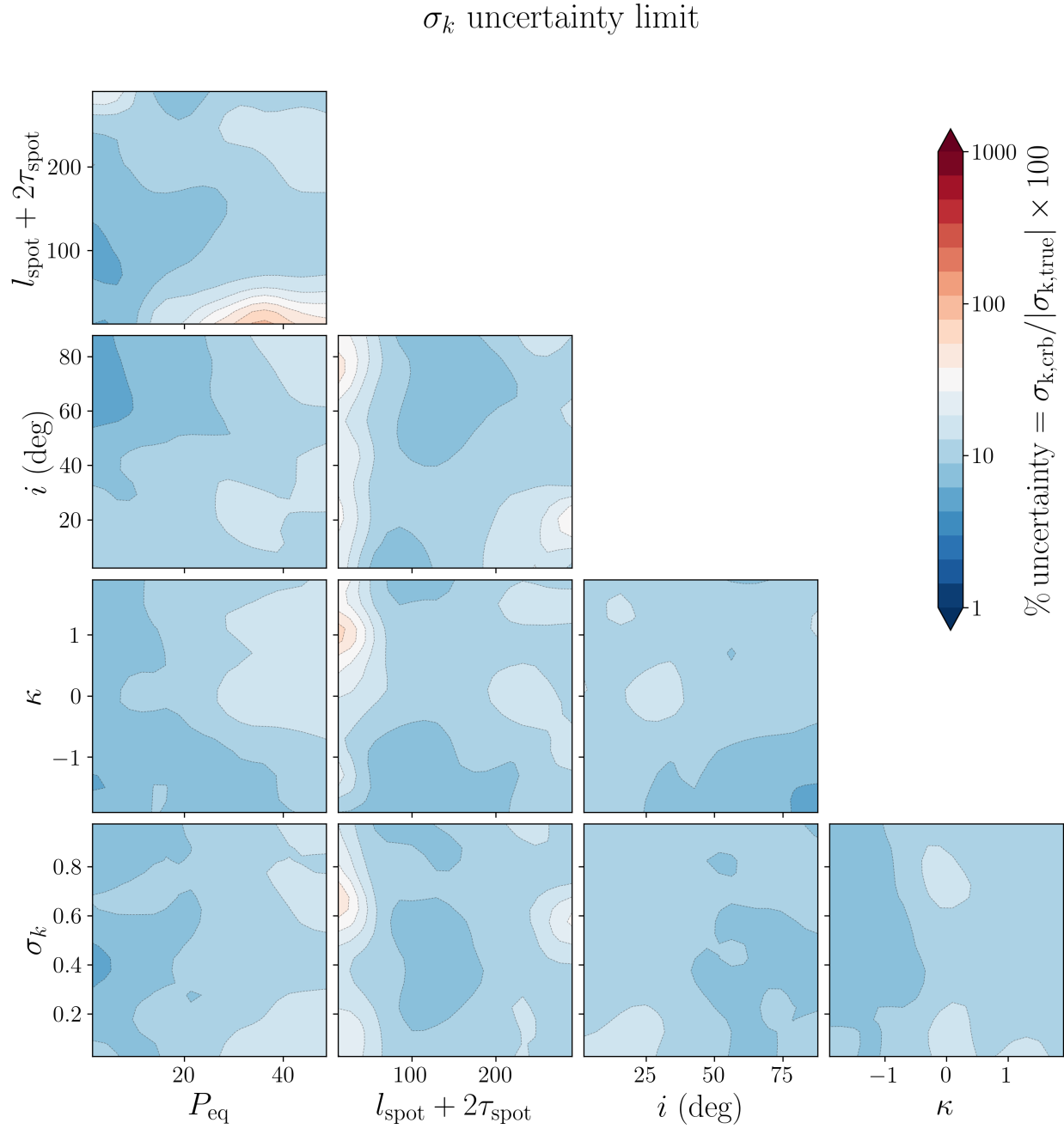


Figure 4.32: Same as Fig. 4.29, but colored by σ_k .

9.6 MCMC Inference

While Section 9.4 gives a sense of the minimum observation requirements for reliable ACF estimation, it doesn’t directly test the ability of our GP model to recover the underlying parameters θ and their uncertainties from noisy data. To assess this issue, we perform MCMC sampling of the posterior distribution $p(\theta|\mathbf{y})$ using the sampling methods described in Section 8. More specifically, we investigate how the noise level and observation duration affect the ability to recover the true parameters, and how this compares to the theoretical limits set by the Cramér–Rao bound. We also analyze the posterior distributions to identify any degeneracies or correlations between parameters, and to understand how well the model can constrain each parameter under different conditions.

For each synthetic lightcurve, we ran both NUTS (with **BlackJAX**) and nested sampling (with **dynesty**) to obtain posterior samples of the parameters θ and compare the results to the true parameters θ_{true} used to generate the data. In practice, we found that the NUTS sampler tended to get stuck in small local modes of the posterior and was less effective at exploring the full parameter space, and so we primarily report results from the nested sampling runs, which were more robust to multimodality and provided better coverage of the posterior.

Table 4.11: Posterior summary statistics from nested sampling (**dynesty**) on simulated lightcurves. For each combination of observation duration (T_{sim}) and true inclination (I_{true}), we report the posterior mean $\pm$ standard deviation alongside the true (injected) parameter value. The inclination is converted from radians to degrees for readability. All simulations use a sampling cadence of $\Delta t = 1.0$ d.

I_{true}	Parameter	True Value	Prior (min, max)	$T_{\text{sim}} = 365$ d		$T_{\text{sim}} = 1460$ d	
				Mean	Std	Mean	Std
30°	P_{eq} [d]	5.0	(4.0, 6.0)	5.005	0.028	4.996	0.014
	κ	0.3	(−1.0, 1.0)	0.304	0.047	0.331	0.022
	I [deg]	30.0	(0.0, 60.0)	90.6	61.1	29.2	0.9
	ℓ_{spot} [d]	20.0	(10.0, 30.0)	20.2	0.3	20.0	0.2
	τ_{spot} [d]	5.0	(0.0, 10.0)	4.95	0.48	5.04	0.28
	σ_k	0.005	(0.001, 0.01)	0.00516	0.00038	0.00487	0.00017
60°	P_{eq} [d]	5.0	(4.0, 6.0)	4.969	0.038	5.001	0.015
	κ	0.3	(−1.0, 1.0)	0.312	0.028	0.298	0.013
	I [deg]	60.0	(30.0, 90.0)	90.7	28.7	58.9	1.2
	ℓ_{spot} [d]	20.0	(10.0, 30.0)	20.2	0.8	20.1	0.3
	τ_{spot} [d]	5.0	(0.0, 10.0)	4.89	1.02	5.11	0.46
	σ_k	0.005	(0.001, 0.01)	0.00490	0.00027	0.00510	0.00014
90°	P_{eq} [d]	5.0	(4.0, 6.0)	5.007	0.035	5.002	0.014
	κ	0.3	(−1.0, 1.0)	0.308	0.028	0.303	0.014
	I [deg]	90.0	(60.0, 120.0)	90.1	3.6	89.8	3.2
	ℓ_{spot} [d]	20.0	(10.0, 30.0)	20.6	0.9	20.0	0.4
	τ_{spot} [d]	5.0	(0.0, 10.0)	4.65	1.01	5.09	0.56
	σ_k	0.005	(0.001, 0.01)	0.00476	0.00025	0.00506	0.00014

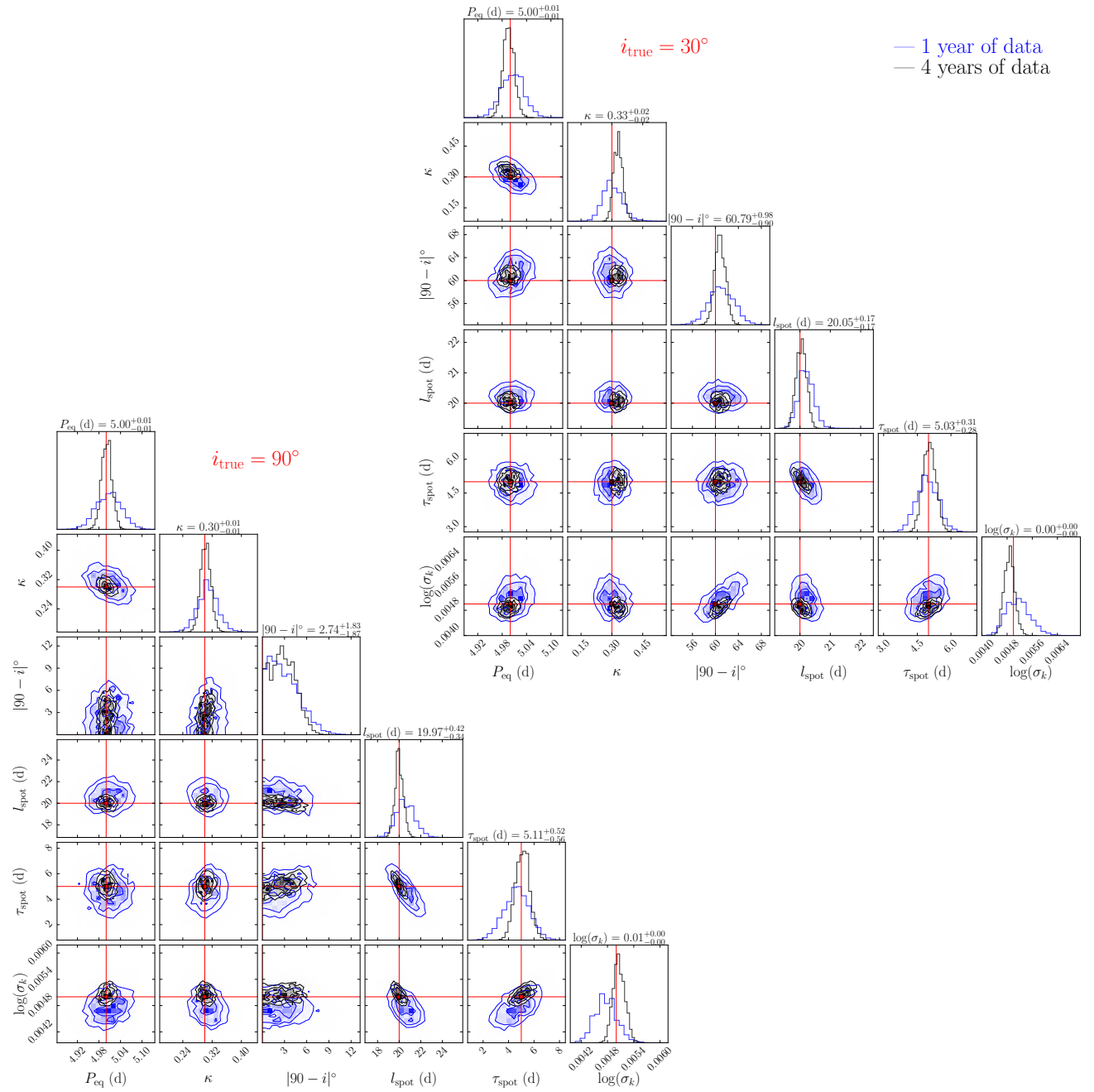


Figure 4.33: Posterior distributions for different inclinations: $I = 90$ (bottom left) and 30 (upper right) degrees based on nested sampling runs on synthetic lightcurves. The blue posterior outlines the constraints using 1 year of data, while the black posterior outlines the constraints using 4 years of data (comparable to the Kepler mission duration), and the true parameters are indicated by the red lines. The posterior distributions show that the model can recover the true parameters with tight precision given a typical noise level of $\sigma_n = 10^{-4}$ and a spot emergence rate of $\dot{N}_{\text{spot}} = 0.25 \text{ d}^{-1}$.

10 Case Study: Application to Kepler-63

Kepler-63 is a young Sun-like star (with a stellar mass and effective temperature of $M_* = 0.984^{+0.035}_{-0.040} M_\odot$ and $T_{\text{eff}} = 5576 \pm 50$ K; [Sanchis-Ojeda et al., 2013b](#)) hosting a transiting hot Jupiter on a polar orbit. Kepler-63 also has a fast rotation period of $P_{\text{rot}} = 5.400 \pm 0.009$ days ([Sanchis-Ojeda et al., 2013b](#)). This system is unique in that the planet’s orbit is highly misaligned relative to the star’s spin axis, so it transits across a wide range of stellar latitudes. Because of this favorable geometry, Kepler-63 has provided a rare opportunity to probe the latitudinal distribution of starspots using transit mapping techniques ([Netto & Valio, 2020](#); [Sagynbayeva & Farr, 2026](#); [Estrela & Valio, 2016](#)).

In this section we discuss the system geometry and how it affects the visibility of starspots. We then apply our GP model to the Kepler-63 lightcurve to infer the spot parameters and compare them to previous results from transit mapping. Finally, we show that our GP approach is able to constrain spot parameters that are consistent with those inferred from transit mapping and spectroscopy. Beyond the specific results for Kepler-63, this case study serves as a demonstration of how our physically motivated GP kernel can be applied to real data to extract information about starspot properties and stellar rotation. Moreover, because our GP method relies solely on the rotational variability in a lightcurve, it can be applied to a much larger samples of stars than those amenable to transit mapping, making it a powerful tool for characterizing starspots across a wide range of stellar types.

10.1 System Geometry

The inclination of Kepler-63 ($I = 138^\circ$) means the observer views the star from below its equatorial plane, with a sub-observer latitude of $90^\circ - I = -48^\circ$. Figure 4.34 illustrates this geometry. The left panel shows the side-view orientation of the rotation axis relative to the line of sight. The middle panel shows the observer’s projected view of the stellar disk, with latitude and longitude grid lines. The right panel maps the visibility $\mu = \cos \beta$ as a function of rotational phase and latitude for a spot at a reference longitude. The black contour marks the boundary between visible ($\mu > 0$) and hidden regions. Latitudes between roughly -90° and $+35^\circ$ are visible at some point during a rotation, while latitudes above $\sim +35^\circ$ remain permanently hidden.

10.1.1 Role of Always-Visible Spots in the GP Fit

An important consideration when interpreting the kernel hyperparameters is the role of spots that remain permanently visible – i.e., spots at latitudes where they never rotate behind the stellar limb. A spot at latitude Φ on a star with inclination I is always visible when $b_0 > b_1$, i.e., $\cos I \sin \Phi > \sin I \cos \Phi$, which simplifies to $\Phi > I$ (measuring Φ from the equator). For Kepler-63, with $I \approx 138^\circ$ (equivalently 42° from pole-on), all spots at latitudes above $\sim 42^\circ$ satisfy this condition – a substantial fraction of the visible hemisphere.

For these always-visible spots, the visibility window spans the full rotation ($\theta_{\text{vis}} = \pi$), and the Fourier coefficients c_n become strongly DC-dominated: $c_0 = b_0 = \cos I \sin \Phi$ is large, while the higher harmonics c_n ($n \geq 1$) are suppressed. In the limiting case of a pole-on star

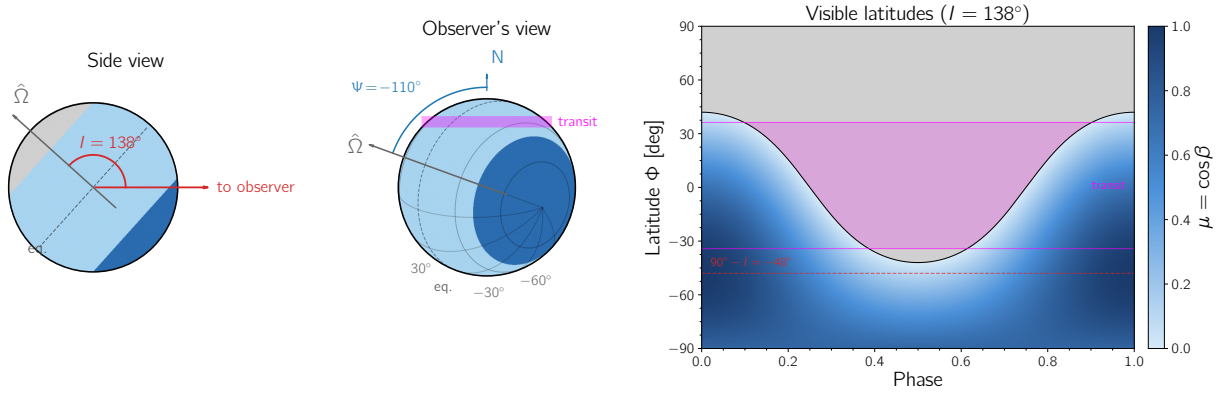


Figure 4.34: System geometry for Kepler-63 ($I = 138^\circ$). In the left and middle panels, the stellar disk is shaded by visibility zone: dark blue for latitudes that are always visible (never rotate behind the limb), light blue for latitudes that are sometimes visible, and gray for latitudes that are never visible. *Left:* Side-view schematic showing the rotation axis $\hat{\Omega}$, the line of sight to the observer, and the inclination angle I . The dashed line marks the stellar equator. *Middle:* Observer’s projected view of the stellar disk with latitude and longitude grid lines, rotated by the sky-plane position angle $\Psi = -110^\circ$ (measured east of celestial north). Because $I > 90^\circ$, the south pole is tilted toward the observer. The magenta band shows the transit chord of Kepler-63b ($i_{\text{orb}} = 87.8^\circ$, $R_p/R_\star = 0.064$, $a/R_\star = 19.35$), with width equal to the planet diameter. Note that Ψ does not affect disk-integrated photometry and therefore does not enter the GP kernel. *Right:* Visibility map $\mu(\phi, \Phi)$ showing the cosine of the center-to-limb angle β as a function of rotational phase and latitude for a spot at reference longitude $\Lambda_0 = 0$. The black contour marks the $\mu = 0$ boundary between the visible and hidden hemispheres. The dashed red line indicates the sub-observer latitude $90^\circ - I = -48^\circ$. The magenta shaded region shows the range of stellar latitudes crossed by the planet during transit; because the system is highly oblique, the transit chord spans a wide range of latitudes ($\sim -40^\circ$ to $\sim +30^\circ$).

($I = 0$), $b_1 = 0$, the kernel contribution from such spots reduces to

$$K_{\text{always-vis}}(\tau) \approx \sigma_k^2 R_\Gamma(\tau) |c_0|^2, \quad (4.107)$$

which is a smooth, aperiodic function that decays on the spot lifetime—with no rotational modulation.

This has several consequences for parameter inference:

1. **Effect on kernel amplitude:** always-visible spots contribute to the total kernel variance $\mathcal{K}(0)$ through the DC term $\langle |c_0|^2 \rangle_\Phi$ (Eq. 4.82), but add negligible power at the rotation frequency or its harmonics. Thus we expect that their contribution is largely degenerate with the amplitude hyperparameter σ_k^2 , which absorbs the persistent flux deficit from these spots.
2. **Effect on rotation parameters:** since always-visible spots produce no periodic modulation, they carry essentially no information about P_{eq} or κ . The periodic signal that constrains these parameters comes entirely from spots in the *partially visible* latitude band ($\Phi < I$), which produce strong $n = 1$ and $n = 2$ harmonics as they rotate into and out of view.
3. **Effect on spot lifetime:** always-visible spots do contribute to the aperiodic decay of the kernel through $R_\Gamma(\tau)$, which in principle constrains ℓ_{spot} and τ_{spot} . However, this aperiodic signal is degenerate with the spot emergence rate and amplitude, making it difficult to isolate the spot lifetime from the DC contribution alone. In practice, the strongest constraints on ℓ_{spot} come from the decay of the *periodic* peak heights (Eq. 4.95), which is driven by the partially visible spots.

The fraction of always-visible spots depends sensitively on the stellar inclination: for edge-on viewing ($I = 90^\circ$), no spots are always visible, and all latitudes contribute periodic modulation. For systems like Kepler-63, where the inclination places a significant fraction of spots in the always-visible regime, the GP fit is effectively driven by the minority of spots at low latitudes that produce detectable rotational signatures.

10.1.2 Role of Partially Visible Spots in Constraining Parameters

While always-visible spots contribute primarily to the DC component of the kernel, it is the *partially visible* spots – those at latitudes $\Phi < I$ ($\Phi < 42^\circ$ for Kepler-63) – that generate the periodic modulation responsible for constraining most of the kernel hyperparameters. These spots rotate in and out of view each period, producing a visibility window of finite duty cycle $\Delta = \theta_{\text{vis}}/\pi < 1$ and exciting strong power in the $n = 1$ and $n = 2$ Fourier harmonics. Their contribution to the kernel is

$$K_{\text{partial}}(\tau) = \sigma_k^2 R_\Gamma(\tau) \sum_n \langle |c_n|^2 \rangle_{\Phi < I} \cos(n\omega_0\tau), \quad (4.108)$$

where the harmonic weights $\langle |c_n|^2 \rangle_{\Phi < I}$ are averaged only over the partially visible latitude band.

The periodic structure introduced by these spots is what enables the GP to separately constrain the following parameters:

1. **Rotation period P_{eq} .** The fundamental periodicity of the kernel is set by the $n =$

- 1 harmonic from partially visible spots. Because these spots produce sharp ingress and egress features as they cross the limb, the spacing of the ACF peaks directly measures P_{eq} . This is the best-constrained parameter in every year of our Kepler-63 fits (Table 4.12).
2. **Differential rotation κ .** Spots at different latitudes rotate at different rates, $P(\Phi) = P_{\text{eq}}/(1 - \kappa \sin^2 \Phi)$. The partially visible band spans a range of latitudes, each contributing periodic signal at a slightly different frequency. This produces a characteristic broadening of the PSD peak around ω_0 and a progressive skewing of the ACF peaks toward longer lags—signatures that are unique to differential rotation and cannot be reproduced by a single-period kernel. The width of the partially visible band therefore controls the lever arm for constraining κ : a wider band (more equator-on viewing) samples a greater range of rotation rates and yields tighter constraints.
 3. **Stellar inclination I .** The inclination controls the *relative* power in the DC versus periodic components of the kernel. As I increases toward edge-on, the always-visible zone shrinks and more spots enter the partially visible regime, shifting power from $\langle |c_0|^2 \rangle$ into $\langle |c_1|^2 \rangle$ and higher harmonics. The ratio of the oscillating to non-oscillating kernel components (Eq. 4.91) is therefore a diagnostic of inclination. Additionally, the harmonic ratio $\langle |c_2|^2 \rangle / \langle |c_1|^2 \rangle$ depends on the typical visibility duty cycle, which is set by I and the latitude distribution.
 4. **Spot lifetime $(\ell_{\text{spot}}, \tau_{\text{spot}})$.** The envelope autocorrelation $R_{\Gamma}(\tau)$ modulates the amplitude of successive ACF peaks. For partially visible spots, this effect manifests as a decay of the periodic peak heights with lag (Eq. 4.95). The rate of this decay directly constrains the total spot lifetime $\ell_{\text{spot}} + 2\tau_{\text{spot}}$, while the shape of the decay (linear vs. curved) distinguishes the plateau duration from the rise/decay timescale.

In summary, the always-visible and partially visible spot populations play complementary roles: the former sets the baseline variance and aperiodic structure of the kernel, while the latter imprints the periodic, harmonic, and differential-rotation signatures that enable the GP to disentangle the physical parameters. For Kepler-63, the partially visible band ($\Phi \lesssim 42^\circ$) includes equatorial and low-latitude spots, which are the primary drivers of the inferred rotation period, differential rotation, and spot lifetime constraints reported in Section 10.3.

10.2 Fitting Procedure

We apply the starspot GP model developed in Sections 5–7 to the Kepler-63 lightcurve obtained from the *Kepler* mission. The full lightcurve spans approximately 1470 days (~ 4 years) and contains $\sim 64,000$ observations after masking out the planet transits and outliers. Because our kernel is stationary (and it is possible that stellar activity may vary due to activity cycles) we divide the dataset into four year-long segments and fit each independently, allowing the inferred spot parameters to evolve from year to year. We also compare our results to a single fit of the full dataset, which assumes constant spot properties over the entire baseline.

For each year, we normalize the lightcurve to unit median flux and downsample to a cadence of $\Delta t = 1.0$ days using inverse-variance-weighted binning, reducing each segment to ~ 600 – 700 data points. This cadence is sufficient to resolve the ~ 5.4 -day rotation period

while keeping the covariance matrix tractable for Cholesky factorization.

We model the kernel using the symmetric trapezoidal envelope (Section 5.4) and the full latitude-averaged visibility function (Section 5.6), giving six free hyperparameters: the equatorial rotation period P_{eq} , differential rotation coefficient κ , stellar inclination I , spot plateau duration ℓ_{spot} , spot rise/decay timescale τ_{spot} , and kernel amplitude σ_k . We adopt uniform priors over the following ranges:

$$\begin{aligned} P_{\text{eq}} &\in \mathcal{U}(5.2, 5.5) \text{ d}, & \kappa &\in \mathcal{U}(-0.5, 0.5), \\ I &\in \mathcal{U}(110^\circ, 150^\circ), & \ell_{\text{spot}} &\in \mathcal{U}(10, 100) \text{ d}, \\ \tau_{\text{spot}} &\in \mathcal{U}(0.1, 10) \text{ d}, & \sigma_k &\in \mathcal{U}(10^{-3}, 10^{-2}). \end{aligned} \quad (4.109)$$

The prior on P_{eq} is centered on the literature value of 5.400 ± 0.009 days (Sanchis-Ojeda et al., 2013b). For the inclination, we adopt a wide uniform prior that encompasses the range of published estimates ($I = 138 \pm 7^\circ$, Sanchis-Ojeda et al. 2013b).

We sample the posterior using the **dynesty** nested sampling algorithm (Speagle, 2020b) with 2000 live points, multi-ellipsoidal bounding, and random-slice sampling. The GP log-likelihood (Eq. 4.101) is evaluated using the banded Cholesky solver described in Section 7, compiled via JAX for GPU acceleration. Each year’s run produces weighted posterior samples from which we extract parameter estimates as the median and 68% credible intervals.

10.3 Posterior Results

The resulting posterior distributions are shown in Figure 4.35. For each year, we evaluate the GP posterior predictive mean at the median posterior parameters and compare to the data in three diagnostic views: the lightcurve, the empirical autocorrelation function (ACF), and the Lomb-Scargle power spectral density (PSD). These are shown in Figures 4.36–4.39. In the ACF panels, we show both the full model kernel (solid line) and the kernel with the c_0 (non-oscillating) visibility component removed (dashed line), which isolates the contribution from stellar rotation. In the PSD panels, the model PSD is computed via the Fourier transform of the kernel on a uniform lag grid, normalized so that its integral matches the data variance. Table 4.12 summarizes the posterior statistics for each year and the combined dataset.

Previous studies have analyzed the starspot properties of Kepler-63 using various methods, including spot modeling of the transit lightcurves and spectroscopic measurements of the star’s rotational velocity. Table 4.13 summarizes the key findings from the literature, and we discuss how they compare to the parameters inferred from our GP modeling below. We find that our GP-inferred posteriors (Table 4.12) are broadly consistent with these independent literature measurements, despite the fact that our method uses only the out-of-transit lightcurve autocorrelation structure and does not rely on spectroscopy, transit spot crossings, or detailed spot modeling.

Rotation period: Our inferred equatorial periods ($P_{\text{eq}} \approx 5.34\text{--}5.39$ d across years; 5.361 ± 0.007 d for the combined fit) are consistent with the periodogram-based measurement of $P_{\text{rot}} = 5.400 \pm 0.009$ d from Sanchis-Ojeda et al. (2013b). The slight offset may reflect our model’s ability to distinguish the equatorial period from a latitude-averaged period, since differential rotation causes high-latitude spots to rotate more slowly.

Inclination: The combined-fit inclination of $I = 138.6^\circ \pm 1.4^\circ$ agrees well with the

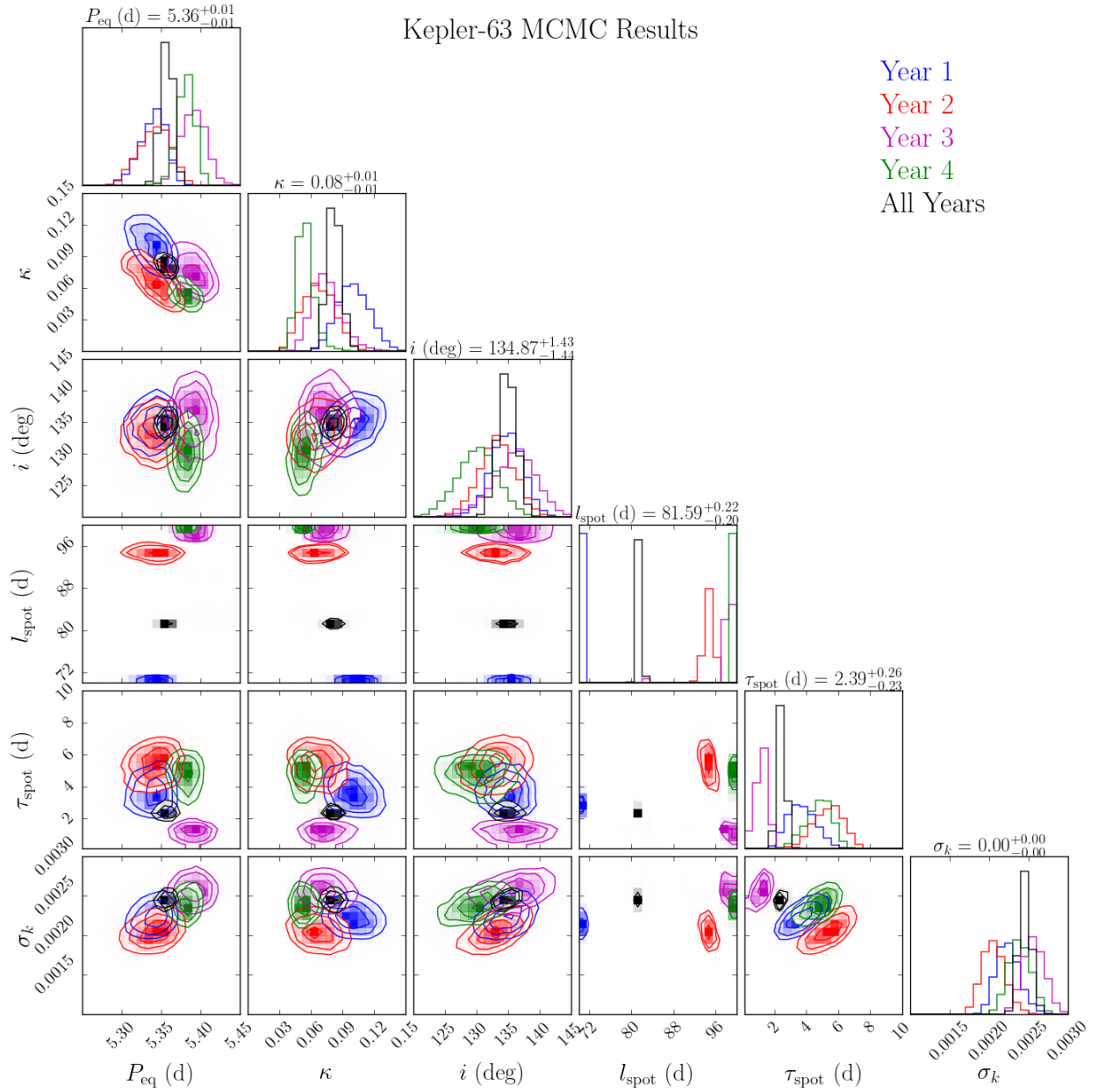


Figure 4.35: Corner plot of the posterior samples for Kepler-63, colored by year.

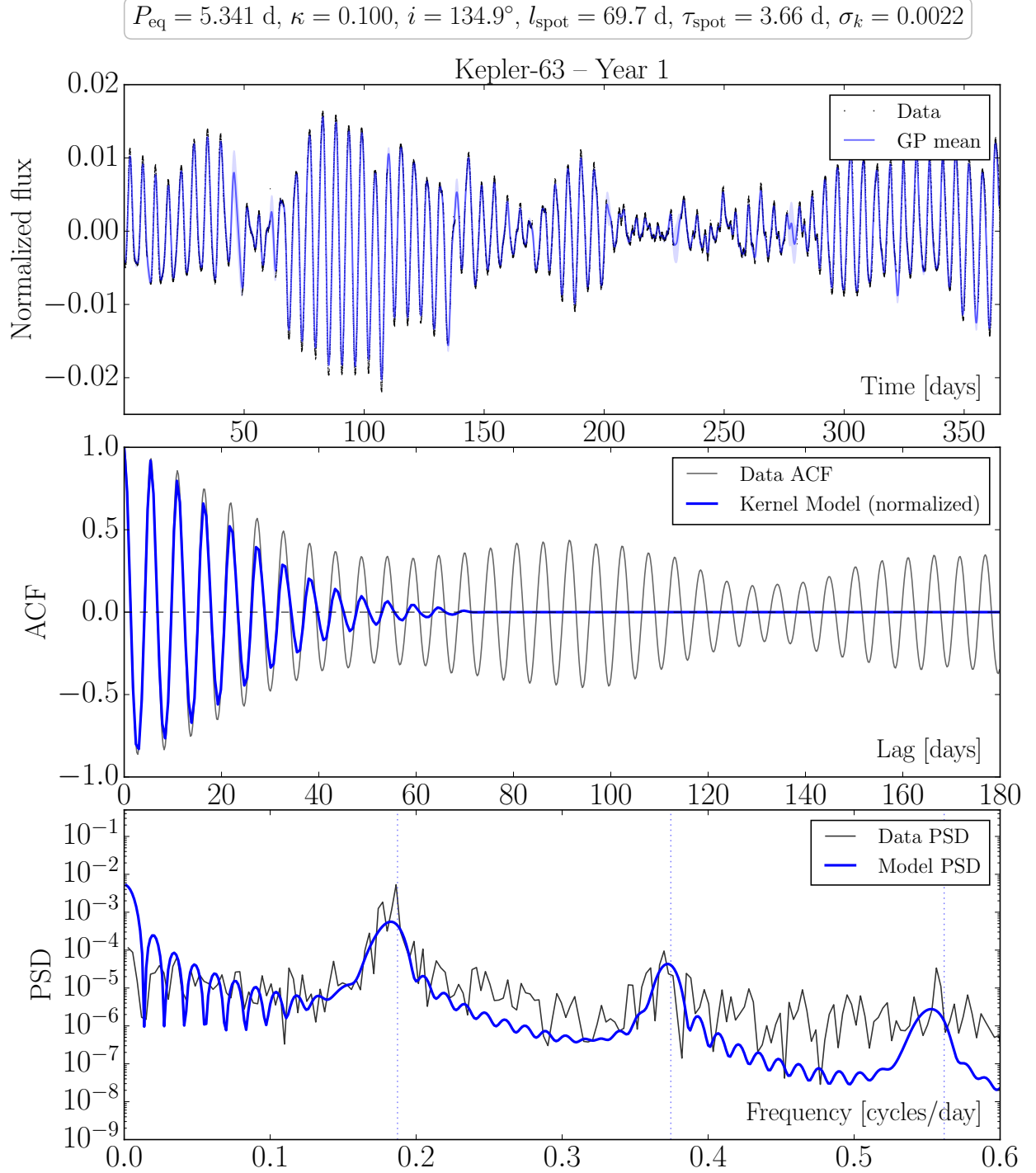


Figure 4.36: Kepler-63 Year 1: lightcurve with GP posterior mean (*top*), autocorrelation function (*middle*), and power spectral density (*bottom*). The colored curves show the model evaluated at the median posterior parameters (listed above each panel); black shows the data. In the ACF panel, the solid line is the full model kernel and the dashed line has the c_0 component removed. Dotted vertical lines in the PSD panel mark the first three rotation harmonics.

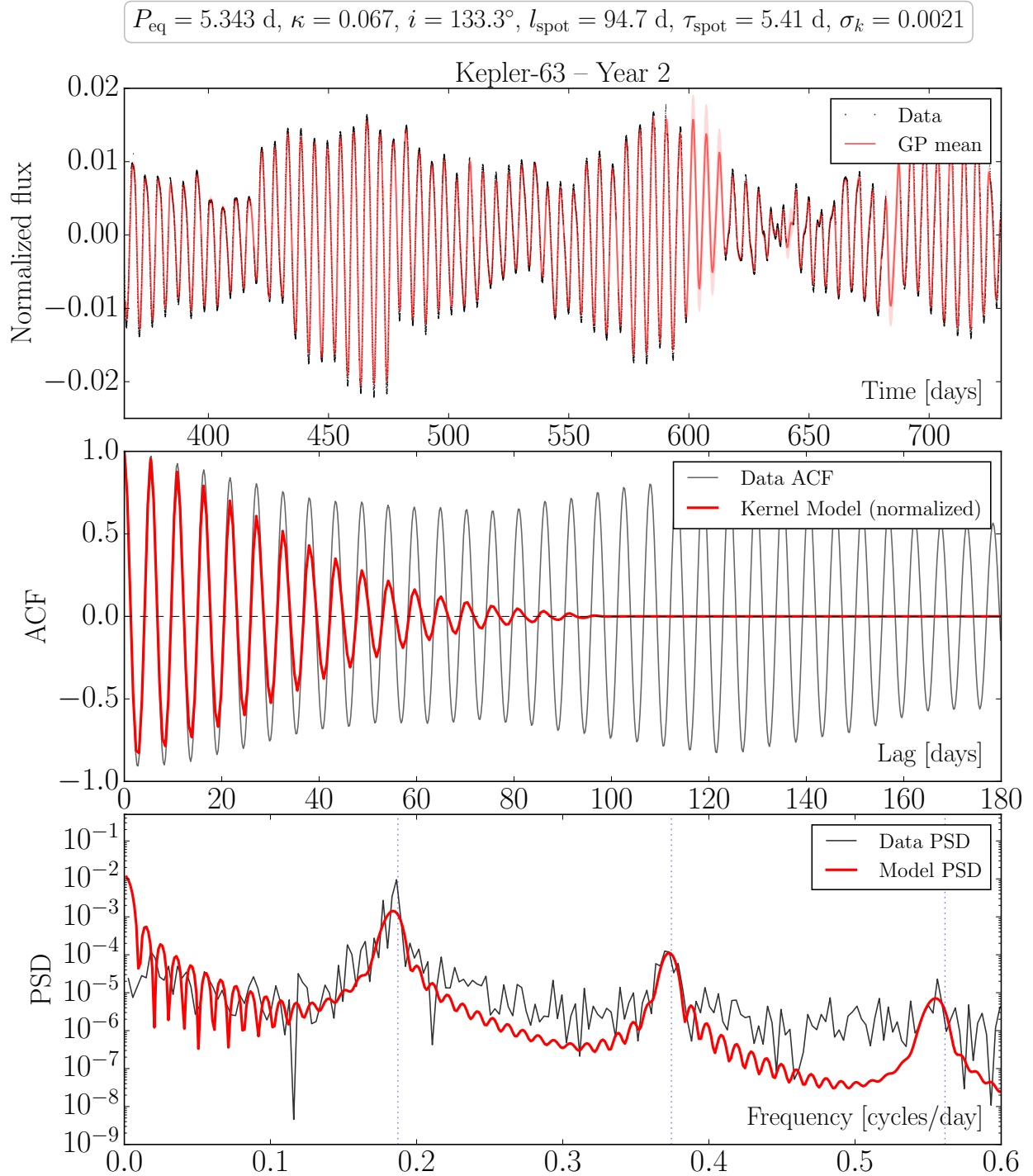


Figure 4.37: Same as Figure 4.36 but for Year 2.

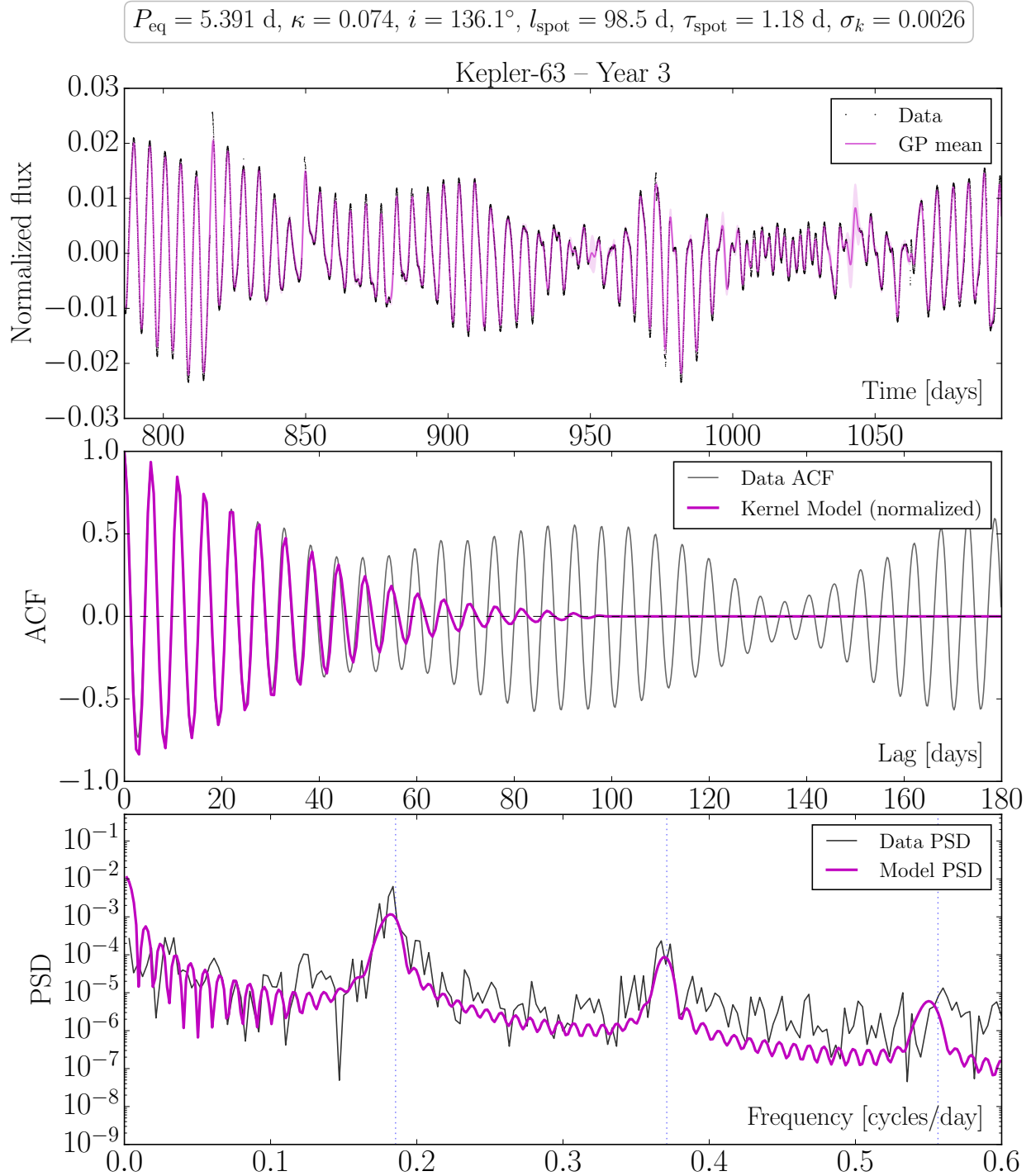


Figure 4.38: Same as Figure 4.36 but for Year 3.

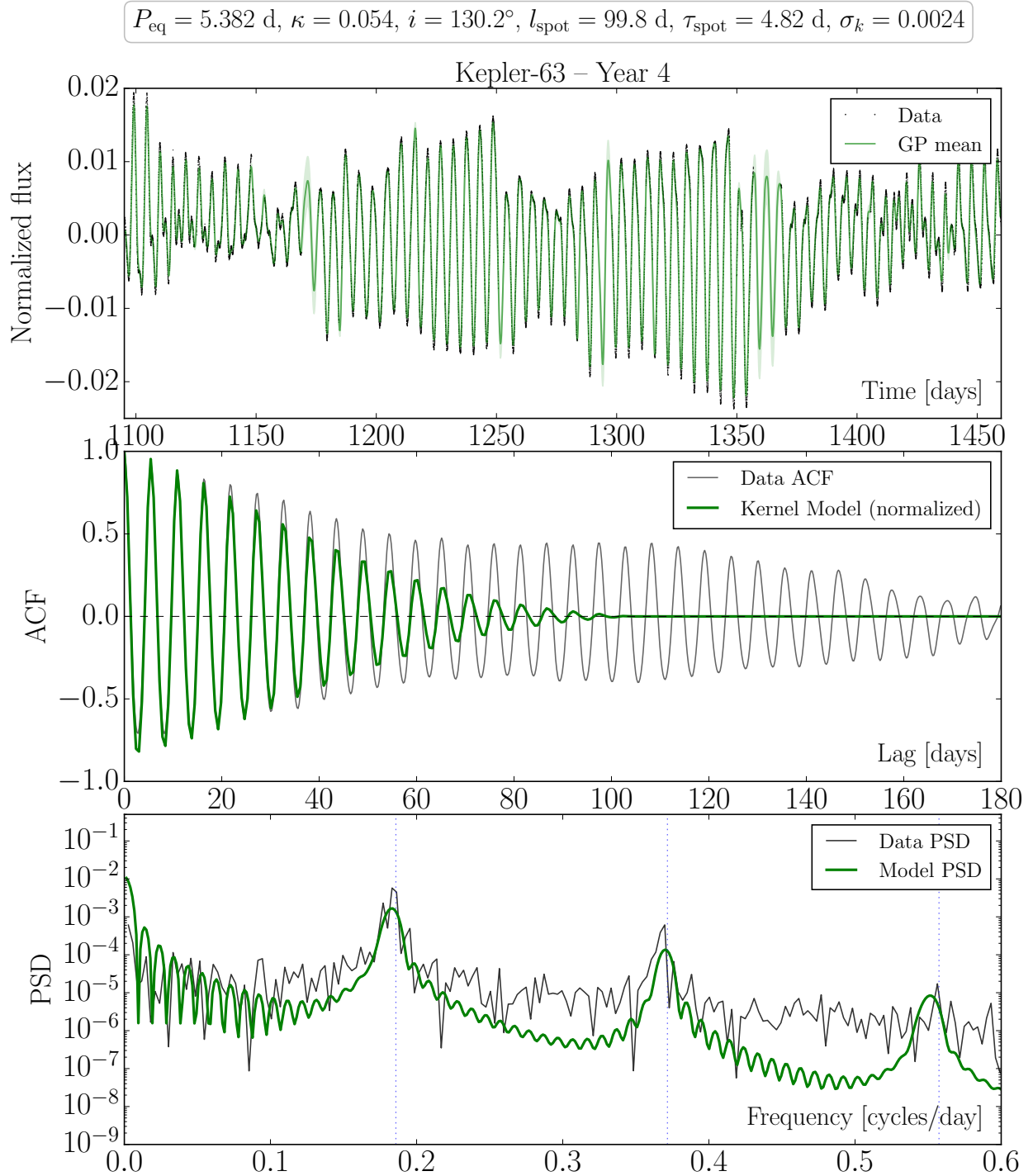


Figure 4.39: Same as Figure 4.36 but for Year 4.

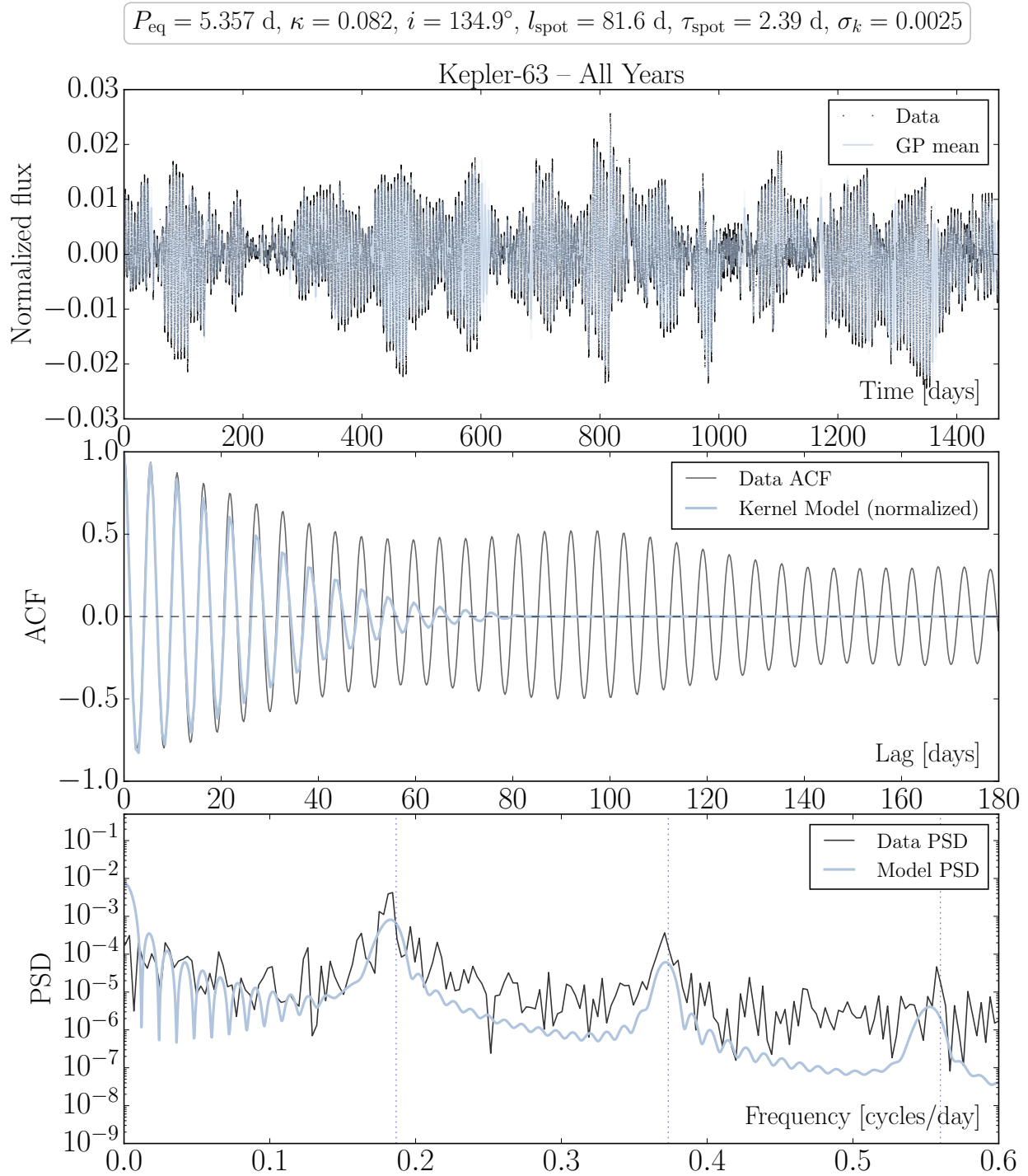


Figure 4.40: Same as Figure 4.36 but fitting the full ~ 4 -year Kepler-63 lightcurve simultaneously with a single set of stationary kernel parameters.

Table 4.12: Posterior summary statistics for the Kepler-63 GP kernel hyperparameters from nested sampling (**dynesty**). For each year of Kepler data (and the combined dataset), we report the posterior mean, standard deviation, median, and 68% credible interval (16th–84th percentiles). Row colors match the corner plot in Figure 4.35.

Dataset	Parameter	Mean	Std	Median	16th %ile	84th %ile
Year 1	P_{eq} (d)	5.357	0.039	5.374	5.309	5.391
	κ	0.085	0.028	0.079	0.056	0.118
	i (deg)	131.6	2.8	131.6	128.8	134.4
	l_{spot} (d)	86.8	5.1	88.1	87.4	88.9
	τ_{spot} (d)	4.09	1.00	4.07	3.04	5.12
	$\dot{N}_{\text{spot}}$	0.1254	0.0942	0.0986	0.0523	0.1935
	a_{spot}	0.0069	0.0022	0.0066	0.0047	0.0091
Year 2	P_{eq} (d)	5.332	0.022	5.327	5.315	5.347
	κ	0.067	0.014	0.069	0.053	0.081
	i (deg)	134.6	3.3	134.8	131.2	137.7
	l_{spot} (d)	78.2	6.0	75.5	74.3	86.4
	τ_{spot} (d)	6.03	0.88	6.04	5.17	6.91
	$\dot{N}_{\text{spot}}$	0.0636	0.0807	0.0299	0.0134	0.1234
	a_{spot}	0.0126	0.0053	0.0129	0.0061	0.0186
Year 3	P_{eq} (d)	5.394	0.017	5.394	5.376	5.410
	κ	0.074	0.015	0.073	0.060	0.089
	i (deg)	132.4	3.6	132.4	128.7	136.1
	l_{spot} (d)	96.7	2.6	97.4	95.1	98.0
	τ_{spot} (d)	1.49	0.46	1.47	1.07	1.90
	$\dot{N}_{\text{spot}}$	0.0954	0.0814	0.0677	0.0335	0.1575
	a_{spot}	0.0094	0.0033	0.0092	0.0060	0.0129
Year 4	P_{eq} (d)	5.356	0.013	5.355	5.343	5.368
	κ	0.064	0.012	0.062	0.052	0.075
	i (deg)	130.0	2.9	130.0	127.2	132.9
	l_{spot} (d)	82.7	3.5	82.0	81.2	82.9
	τ_{spot} (d)	4.10	0.90	4.05	3.21	4.97
	$\dot{N}_{\text{spot}}$	0.1148	0.0983	0.0840	0.0359	0.1916
	a_{spot}	0.0083	0.0033	0.0077	0.0050	0.0117
All	P_{eq} (d)	5.360	0.007	5.360	5.354	5.366
	κ	0.087	0.009	0.087	0.077	0.095
	i (deg)	139.0	1.3	139.0	137.7	140.2
	l_{spot} (d)	83.0	0.3	83.0	82.8	83.3
	τ_{spot} (d)	2.49	0.25	2.47	2.25	2.74
	$\dot{N}_{\text{spot}}$	0.0309	0.0228	0.0234	0.0140	0.0481
	a_{spot}	0.0168	0.0046	0.0167	0.0117	0.0219

[Sanchis-Ojeda et al. \(2013b\)](#) value of $138 \pm 7^\circ$, and is intermediate between the two most precise literature estimates ($138 \pm 7^\circ$ and $151.1^{+0.7}_{-0.6}$ deg from [Sagynbayeva & Farr 2026](#)).

Table 4.13: Literature measurements for Kepler-63 starspot and stellar properties. References: (1) Sanchis-Ojeda et al. (2013b), (2) Sagynbayeva & Farr (2026), (3) Morris et al. (2019b), (4) Rainer et al. (2023), (5) Morris et al. (2019a), (6) Sagynbayeva et al. (2025), (7) Menezes et al. (2024), (8) Araújo & Valio (2023), (9) Netto & Valio (2020), (10) Estrela & Valio (2016), (11) Coffaro et al. (2022).

Property	Value	Method	Ref.
Stellar properties			
I	138 ± 7 deg	Transit + $v_* \sin I$	1
	$151.1^{+0.7}_{-0.6}$ deg	Transit modeling	2
$v_* \sin I$	5.4 ± 0.5 km/s	Spectroscopy	1
	5.43 km/s	Spectroscopy	3
	4.71 ± 0.53 km/s	Spectroscopy	4
P_{rot}	5.400 ± 0.009 d	Periodogram	1
κ	≈ 0.01	Spot period vs. latitude	9
R_*	$0.901^{+0.027}_{-0.022} R_{\odot}$	Isochrone fitting	1
Spot properties			
f_{spot}	0.27 ± 0.06	TiO bands	5
r_{spot}	$10.02^{+0.03}_{-0.02}$ deg	Transit modeling	6
$\langle I_{\text{spot}} \rangle$	0.48 ± 0.16	Spectropolarimetry	7
$\log \langle A_{\text{spot}} \rangle$	$2.71 \pm 0.14 \log[\text{mh}]$	Spectropolarimetry	7
$\langle B_{\text{spot}} \rangle$	3400 ± 110 G	Spectropolarimetry	7
N_{spot}	297 (total)	Spot crossing	8
$\langle A_{\text{spot}} \rangle$	8 ± 5 %	Spot crossing	8
Spot latitude	Polar	Spot crossing	9
Activity cycles			
P_{cyc}	1.27 ± 0.16 yr	Spot modeling	10
	Not detected	X-ray	11

Notably, our constraint is derived entirely from the shape of the kernel rather than from $v_* \sin I$ or transit geometry.

Differential rotation: We infer a positive differential rotation coefficient of $\kappa = 0.09 \pm 0.01$ in all years, indicating near solid-body rotation with a bias towards solar-like differential rotation (equator faster than poles). This measurement is roughly consistent with Netto & Valio (2020): in their study they compare spot latitudes vs spot period for different values of $\kappa = 0.01, 0.1, 1.0$ and find that $\kappa \approx 0.01$ is most consistent with the data (although there is a high scatter to their fit).

Spot lifetimes: We infer typical spot lifetimes of $\ell \sim 70 - 100$ d and $\tau_{\text{spot}} \sim 0 - 8$ d (or total lifetimes of $\ell_{\text{spot}} + 2\tau_{\text{spot}} \sim 85-100$ d). These measurements are consistent with the observations of Netto & Valio (2020), which note that the longest living starspots have lifetimes of at least 75 d. As noted in Section 6, the constraints on the spot lifetime parameters are mostly constrained by the decay timescale of the autocorrelation function. However, looking at the ACF panels in Figures 4.36–4.39, we see that the ACF of the data does not decay monotonically as our kernel with finite spot lifetimes would predict. Instead

the ACF tends to show a decay for the first set of peaks, followed by a flattening or even a slight increase in the peak heights at later lags.

This non-decaying feature of the ACF at high time lags (beyond the initial decay, with $t_{\text{lag}} \gtrsim 50$ d) may be a signature of multiple stationary spot populations, variations in spot properties over time due to activity cycles, long-term evolution of the spot emergence rate, or observational systematics which our kernel cannot capture. We also notice that the ACF of the data tends to increase or modulate for very large time lags ($t_{\text{lag}} \gtrsim 150$ d), however since the baseline of the data used is 1 year, these very large lags are not very well sampled. So in conclusion, the GP appears to be picking up constraints from the longest-lived starspots, however there is likely a non-negligible amount of non-stationarity biasing our spot lifetime estimates. Therefore, the inferred spot lifetime values should be interpreted with caution.

Kernel amplitude and spot properties: Our inferred σ_k values (~ 0.0021 – 0.0026 across years; Table 4.12) encode the degenerate combination $\sigma_k^2 = \dot{N}_{\text{spot}}(1 - f_{\text{spot}})^2 \alpha_{\text{max}}^4 / \pi^2$. While σ_k alone cannot separately determine the spot fraction, spot size, or emergence rate, we can check whether the literature measurements of these quantities are mutually consistent with our inferred amplitude. Taking $f_{\text{spot}} = 0.27 \pm 0.06$ from Morris et al. (2019a) and $r_{\text{spot}} \approx 10^\circ$ from Sagynbayeva et al. (2025) (i.e., $\alpha_{\text{max}} \approx 0.17$ rad), and adopting the spot emergence rate implied by $N_{\text{spot}} = 297$ spots over ~ 1470 d ($\dot{N}_{\text{spot}} \approx 0.2$ spots d $^{-1}$; Araújo & Valio, 2023), we can estimate

$$\sigma_k^{\text{lit}} = \frac{\alpha_{\text{max}}^2}{\pi} \sqrt{\dot{N}_{\text{spot}}} (1 - f_{\text{spot}}) \approx \frac{(0.17)^2}{\pi} \sqrt{0.2} (1 - 0.27) \approx 0.003, \quad (4.110)$$

which is consistent with our GP-inferred values of $\sigma_k \approx 0.002$ – 0.003 to within the substantial uncertainties on each of the input quantities. This order-of-magnitude agreement provides a useful sanity check: even though σ_k is a degenerate combination, its inferred value is compatible with the independently measured spot fraction, spot size, and spot count for Kepler-63. Conversely, if external constraints on any two of $\{f_{\text{spot}}, \alpha_{\text{max}}, \dot{N}_{\text{spot}}\}$ are available, σ_k can be used to estimate the third.

11 Discussion

11.1 Summary of Results

In this work, we derived a physically motivated GP kernel for stellar photometric variability caused by rotating starspots, systematically characterize the performance and limitations of this methodology using simulations, and demonstrated its application to real data. We summarize the main results below.

Analytic kernel derivation (Section 5). Starting from a forward model of the flux deficit produced by circular spots on a differentially rotating star, we decomposed the single-spot signal into a product of an envelope term $\Gamma(t)$ (capturing spot emergence and decay) and a visibility term $\mathcal{V}(t)$ (capturing rotational modulation and limb geometry). By taking the Fourier transform of each component and applying the Wiener–Khinchin theorem, we obtained a closed-form expression for the autocovariance kernel $\mathcal{K}(\tau)$ as a sum over rotation harmonics weighted by latitude-averaged visibility coefficients $\langle |c_n|^2 \rangle_\Phi$ and modulated by the envelope autocorrelation $R_\Gamma(\tau)$ (Eq. 4.70). We provided analytic $R_\Gamma(\tau)$ expressions for trapezoidal, exponential, and skew-normal spot profiles, and validated all derivations symbolically with `SymPy` and numerically against brute-force simulations.

Physical interpretability of hyperparameters (Section 6). The six free kernel hyperparameters (P_{eq} , κ , I , ℓ_{spot} , τ_{spot} , and σ_k) map directly onto stellar and spot properties, unlike the phenomenological parameters of standard quasi-periodic or SHO kernels. We showed that P_{eq} and κ control the periodicity and harmonic structure of the kernel; I controls the balance between the DC and oscillating components via the visibility coefficients; ℓ_{spot} and τ_{spot} set the correlation lifetime; and the amplitude parameters (α_{max} , f_{spot} , N_{spot}) are degenerate and collapse into a single scale parameter σ_k^2 . We further demonstrated that our kernel captures signatures (such as PSD broadening from differential rotation and harmonic ratio variations with inclination) that previously applied kernels cannot reproduce (Section 9.2).

Efficient GP solver (Section 7). Because the kernel $\mathcal{K}(\tau)$ has compact support (set by the spot lifetime $\ell_{\text{spot}} + 2\tau_{\text{spot}}$), the covariance matrix is banded with bandwidth $b \propto (\ell_{\text{spot}} + 2\tau_{\text{spot}})/\Delta t$. Exploiting this structure, we implemented a banded Cholesky solver that reduces the GP inference cost from $\mathcal{O}(N^3)$ to $\mathcal{O}(Nb^2)$. The solver is fully JIT-compiled and autodifferentiable via JAX, enabling gradient-based sampling with no additional implementation effort.

Inference conditions and Cramér–Rao bounds (Sections 9.4–9.5). Using synthetic lightcurves, we showed that reliable kernel estimation requires observation baselines spanning at least ~ 10 – 20 spot lifetimes with noise levels $\sigma_{\text{noise}} \lesssim 10^{-2}$ in relative flux. We computed the Cramér–Rao lower bound across a grid of 4096 parameter combinations and found that the differential rotation coefficient κ is best constrained for short-period stars ($P_{\text{eq}} \lesssim 20$ d) with long-lived spots and moderate-to-high inclinations, where κ can theoretically be measured to $\lesssim 10\%$ uncertainty.

Application to Kepler-63 (Section 10). We applied our GP model to four year-long segments and the full ~ 4 -year lightcurve of Kepler-63, a young, rapidly rotating Sun-like star. Using nested sampling (`dynesty`), we recovered consistent rotation periods of $P_{\text{eq}} \approx 5.35$ – 5.39 d across all years, positive differential rotation ($\kappa \approx 0.06$ – 0.09), and inclinations of

$I \approx 130^\circ\text{--}139^\circ$ (Table 4.12). The inferred inclination from the combined fit ($I = 138.6^\circ \pm 1.4^\circ$) is consistent with the literature value of $138 \pm 7^\circ$ from Sanchis-Ojeda et al. (2013b). The spot lifetime parameters ($\ell_{\text{spot}} \sim 80\text{--}97$ d, $\tau_{\text{spot}} \sim 1.5\text{--}6$ d) suggest long-lived spots with relatively rapid emergence and decay, consistent with the high activity level expected for this young star.

11.2 Applications and Future Work

11.2.1 Direct Applications for this model

There are numerous potential applications for this starspot kernel. Firstly, it can be used as a tool for inferring star spot parameters using only photometric lightcurves, without the need for spectroscopic or transit data. This is particularly useful for stars that do not have transiting planets or for which high-resolution spectroscopy is not available. By fitting the kernel to the lightcurve, one can extract information about the star’s rotation period, differential rotation, inclination, and spot lifetimes, which are important for understanding stellar magnetic activity and its impact on exoplanet habitability.

Additionally, the kernel can be used to model and mitigate stellar variability in the context of exoplanet transit searches and radial velocity measurements, improving the precision of planet detection and characterization. The physically motivated nature of the kernel also allows for more interpretable modeling of stellar variability compared to phenomenological kernels, enabling insights into the underlying stellar physics driving the observed lightcurve features.

11.2.2 Extensions to Spectroscopy

This model can also be used in combination with spectroscopic data to jointly constrain stellar and spot properties. For example, our GP kernel can be used to infer the inclination angle of a star just from the photometric variability. This means that we can combine this information with spectroscopic measurements of $v_* \sin I$, allowing us to infer the equatorial rotation velocity v_{eq} and thus the stellar radius $R_* = v_{\text{eq}} P_{\text{eq}} / 2\pi$. This methodology would allow us to constrain stellar radii independently of stellar evolution/isochrone models.

Another natural extension of this framework is to high-resolution spectroscopic time series, where starspots induce wavelength-dependent distortions in absorption line profiles as they rotate across the stellar disk. In the Doppler imaging framework (e.g., Luger et al., 2021a), a spot at latitude Φ and longitude $\Lambda(t)$ perturbs the line profile at a radial velocity $v(t) = v_{\text{eq}} \sin I \cos \Phi \sin \Lambda(t)$, which is directly related to the same visibility geometry (Eq. 4.6) that underpins our photometric kernel. The Fourier decomposition and Wiener–Khinchin machinery developed in Section 5 should therefore generalize to a two-dimensional spectral-temporal kernel $\mathcal{K}(\tau, \Delta v)$, describing the covariance of spectral perturbations across both time lag and velocity separation. Such an extension could help break degeneracies that are present in photometry alone, in particular, the inclination I and spot latitude Φ enter differently through the line-of-sight velocity than through the projected area, and differential rotation κ produces directly observable velocity-track separations for spots at different latitudes. Joint photometric–spectroscopic inference with a physically motivated kernel could

therefore improve constraints on stellar surface properties, particularly in the context of extreme precision radial velocity (EPRV) surveys where disentangling stellar activity from planetary signals remains a central challenge.

11.2.3 Further Development

We make the kernel implementation publicly available as a Python package, `spotgp`, which provides a user-friendly interface for evaluating the kernel and performing GP regression with it. The package includes functions for computing the kernel given a set of hyperparameters, as well as utilities for sampling the hyperparameters from data using MCMC. We also provide example Jupyter notebooks demonstrating how to use the kernel for modeling stellar lightcurves and inferring starspot properties.

Furthermore, we have implemented the package in a modular way such that users can easily adapt the kernel to their specific needs and assumptions (such as different spot latitude distributions, envelope functions, or spot parameter distributions). We hope that `spotgp` will be a valuable resource for the community and facilitate further research into stellar variability and starspot modeling.

11.3 Limitations

Our kernel model is based on several simplifying assumptions, such as circular spots with uniform contrast, a fixed spot latitude distribution, and independent spot evolution. In reality, starspots can have more complex shapes, non-uniform contrasts, and may interact with each other (e.g., through magnetic interactions or spot merging). Additionally, our model does not currently account for faculae or other types of stellar activity that can also contribute to variability.

Our kernel is also derived under the assumption that all spots are drawn from the same distributions for Φ , ℓ_{spot} , τ_{spot} , and α_{max} . In reality, there may be correlations between these parameters (e.g., larger spots may have longer lifetimes), which could affect the shape of the kernel. We also assume that these starspot properties are constant in time, such that the kernel is stationary. In practice, spot properties may evolve over time (e.g., due to activity cycles), leading to non-stationary behavior in the lightcurve. Future work could explore more complex models that allow for such correlations. However, the current model provides a useful starting point for understanding the key physical parameters that shape the photometric variability of spotted stars and can be extended in future work to relax some of these assumptions. We also provide several examples on the online documentation for how users can test non-stationary kernels and more complex spot models using the modular implementation of `spotgp`.

12 Conclusion

In conclusion, we have derived an analytic, physically motivated Gaussian process kernel for stellar photometric variability driven by rotating starspots. Starting from a forward model

of circular spots on a differentially rotating star, we obtained a closed-form autocovariance function with six hyperparameters (P_{eq} , κ , I , ℓ_{spot} , τ_{spot} , and σ_k) that map directly onto measurable stellar and spot properties. This approach stands in contrast to the phenomenological kernels widely used in the literature, whose parameters lack clear physical interpretation.

Using simulated lightcurves and MCMC/nested sampling, we demonstrated that the kernel can recover injected stellar parameters, including the stellar inclination and differential rotation coefficient κ , given sufficient observational baseline ($\gtrsim 10$ spot lifetimes) and photometric precision ($\sigma_{\text{noise}} \lesssim 10^{-2}$). Using Cramér–Rao analysis we further showed that κ is best constrained for short-period, high-inclination stars with long-lived spots. We validated the method on Kepler-63, recovering a rotation period, differential rotation rate, and stellar inclination consistent with independent literature measurements from transit chord mapping.

The compact support of the kernel enables a banded Cholesky GP solver that scales as $\mathcal{O}(Nb^2)$ rather than $\mathcal{O}(N^3)$, making inference tractable for the long-baseline, densely sampled lightcurves produced by *Kepler* and *TESS*. Combined with a JAX-based implementation that supports automatic differentiation and gradient-based sampling, this framework is well suited for population-level studies of stellar surface properties.

Looking ahead, systematic application of this kernel to large photometric surveys could yield differential rotation and spot lifetime measurements for thousands of stars across a range of spectral types and ages – quantities that are currently available for only a handful of targets. Such measurements would place new empirical constraints on stellar dynamo theory, the evolution of surface magnetic activity, and the impact of stellar variability on exoplanet detection and habitability assessments.

1 Fourier Coefficients of the Visibility Function

Here we derive the Fourier coefficients c_n of the visibility function $\mathcal{V}(t) = \max\{\cos \beta(t), 0\}$ from first principles.¹

1.1 Setup

From Equation 4.6, we have:

$$\cos \beta(t) = \underbrace{\cos I \sin \Phi}_{b_0} + \underbrace{\sin I \cos \Phi}_{b_1} \cos(\Lambda_0 + \omega_0 t). \quad (.111)$$

This expression is a shifted cosine oscillating between $b_0 - b_1$ and $b_0 + b_1$. The spot is visible when $\cos \beta(t) > 0$, which (for $b_1 > 0$, $|b_0| < b_1$) occurs when:

$$|\omega_0 t + \Lambda_0| < \theta_{\text{vis}}, \quad \theta_{\text{vis}} = \arccos\left(-\frac{b_0}{b_1}\right). \quad (.112)$$

¹Symbolic verification: `src/derivations/derive_fourier_cosine.ipynb`

This defines the *duty cycle* $\Delta = 2\theta_{\text{vis}}/(2\pi) = \theta_{\text{vis}}/\pi$ of the visibility window. When $b_0 + b_1 < 0$ the spot is never visible; when $b_0 - b_1 > 0$ the spot is always visible ($\theta_{\text{vis}} = \pi$). Figure .41 illustrates the components of the visibility function $\mathcal{V}(t)$.

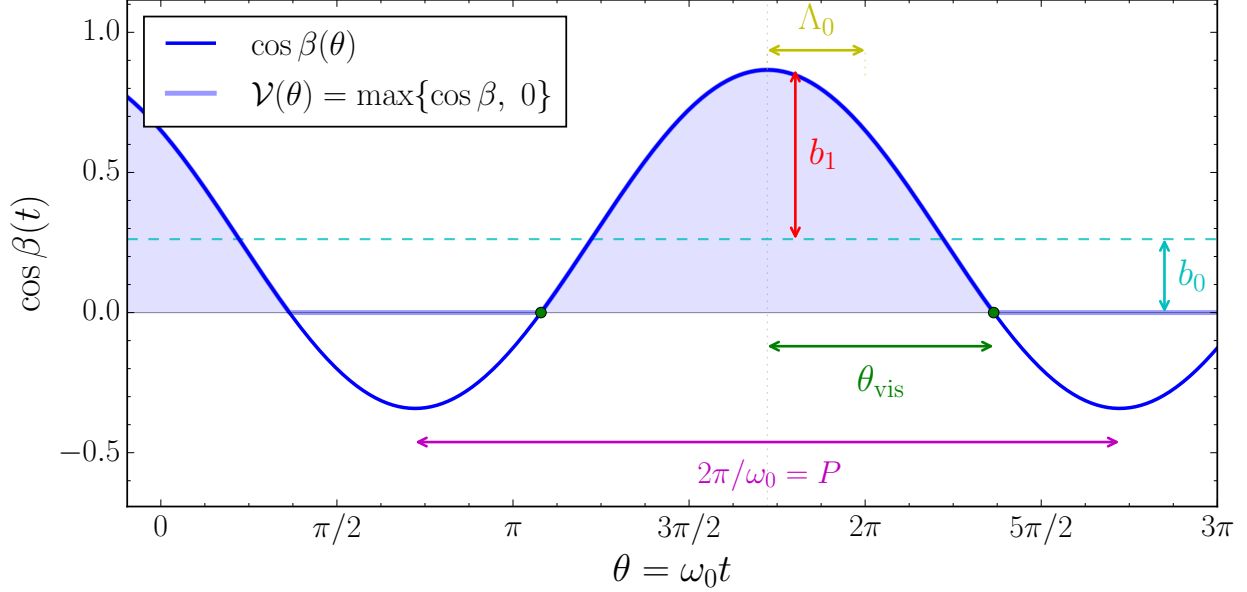


Figure .41: Visibility function $\mathcal{V}(t)$ for a single spot. The spot is visible when $\cos \beta(t) > 0$, which occurs for a fraction $\Delta = 2\theta_{\text{vis}}/(2\pi)$ of the rotation period. The visibility function is a truncated cosine, with amplitude b_1 and offset b_0 .

Since $\mathcal{V}(t)$ is periodic with period $P = 2\pi/\omega_0$, we expand it in a Fourier series:

$$\mathcal{V}(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega_0 t}, \quad (.113)$$

where the Fourier coefficients are:

$$c_n = \frac{1}{T} \int_0^T \mathcal{V}(t) e^{-in\omega_0 t} dt = \frac{1}{2\pi} \int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} (b_0 + b_1 \cos \theta) e^{-in\theta} d\theta, \quad (.114)$$

with the substitution $\theta = \omega_0 t + \Lambda_0$ (and without loss of generality setting $\Lambda_0 = 0$, since a phase shift only affects the complex phase of c_n but not $|c_n|^2$).

1.2 Evaluating the Integral

We split the integrand into the constant and cosine parts:

$$c_n = \frac{1}{2\pi} \left[b_0 \int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} e^{-in\theta} d\theta + b_1 \int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} \cos \theta e^{-in\theta} d\theta \right]. \quad (.115)$$

First integral (constant part): For $n \neq 0$:

$$\int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} e^{-in\theta} d\theta = \left[\frac{e^{-in\theta}}{-in} \right]_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} = \frac{e^{-in\theta_{\text{vis}}} - e^{in\theta_{\text{vis}}}}{-in} = \frac{2 \sin(n \theta_{\text{vis}})}{n}. \quad (.116)$$

For $n = 0$ this gives $2\theta_{\text{vis}}$.

Second integral (cosine part): Using $\cos \theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$:

$$\int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} \cos \theta e^{-in\theta} d\theta = \frac{1}{2} \int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} e^{i(1-n)\theta} d\theta + \frac{1}{2} \int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} e^{-i(1+n)\theta} d\theta. \quad (.117)$$

Each integral is of the same form as the first. For $m \neq 0$:

$$\int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} e^{im\theta} d\theta = \frac{2 \sin(m \theta_{\text{vis}})}{m}. \quad (.118)$$

Therefore, for $n \neq \pm 1$:

$$\int_{-\theta_{\text{vis}}}^{\theta_{\text{vis}}} \cos \theta e^{-in\theta} d\theta = \frac{\sin((1-n)\theta_{\text{vis}})}{1-n} + \frac{\sin((1+n)\theta_{\text{vis}})}{1+n}. \quad (.119)$$

Combining: Substituting back into Eq. (.115), for $n \neq 0, \pm 1$:

$$c_n = \frac{1}{\pi} \left[\frac{b_0 \sin(n \theta_{\text{vis}})}{n} + \frac{b_1}{2} \left(\frac{\sin((n-1)\theta_{\text{vis}})}{n-1} + \frac{\sin((n+1)\theta_{\text{vis}})}{n+1} \right) \right], \quad (.120)$$

where we used $\sin((1-n)x) = -\sin((n-1)x)$ and $1/(1-n) = -1/(n-1)$.

2 Fourier Transform of the Trapezoidal Envelope

Here we derive the Fourier transform and autocorrelation function of the trapezoidal envelope from Section 2.2. Symbolic verification of these equations using `sympy` can be found at: `src/derivations/derive_fourier_gamma.ipynb`

2.1 Piecewise Definition of $\Gamma(t)$

The trapezoidal spot evolution $\alpha(t)$, centered at $t = 0$, has a flat plateau of duration ℓ_{spot} at amplitude α_{max} and linear ramps of duration τ_{spot} on each side:

$$\alpha(t) = \alpha_{\text{max}} \begin{cases} 0 & t < -\frac{\ell_{\text{spot}}}{2} - \tau_{\text{spot}} \\ \frac{t + \frac{\ell_{\text{spot}}}{2} + \tau_{\text{spot}}}{\tau_{\text{spot}}} & -\frac{\ell_{\text{spot}}}{2} - \tau_{\text{spot}} \leq t < -\frac{\ell_{\text{spot}}}{2} \\ 1 & -\frac{\ell_{\text{spot}}}{2} \leq t \leq \frac{\ell_{\text{spot}}}{2} \\ \frac{-t + \frac{\ell_{\text{spot}}}{2} + \tau_{\text{spot}}}{\tau_{\text{spot}}} & \frac{\ell_{\text{spot}}}{2} < t \leq \frac{\ell_{\text{spot}}}{2} + \tau_{\text{spot}} \\ 0 & t > \frac{\ell_{\text{spot}}}{2} + \tau_{\text{spot}} \end{cases} \quad (.121)$$

Squaring and normalizing the trapezoid (.121):

$$\Gamma(t) = \frac{\alpha^2(t)}{\alpha_{\text{max}}^2} = \begin{cases} 0 & |t| > \ell/2 + \tau \\ \left(\frac{|t| - \ell/2 - \tau}{-\tau} \right)^2 & \ell/2 < |t| \leq \ell/2 + \tau \\ 1 & |t| \leq \ell/2 \end{cases} \quad (.122)$$

where $\ell \equiv \ell_{\text{spot}}$ and $\tau \equiv \tau_{\text{spot}}$ for brevity. The ramps are now *parabolic* (concave) rather than linear.

2.2 Direct Integration

Since $\Gamma(t)$ is real and even, $\hat{\Gamma}(\omega)$ is purely real:

$$\hat{\Gamma}(\omega) = 2 \int_0^\infty \Gamma(t) \cos(\omega t) dt = 2 \left[\underbrace{\int_0^{\ell/2} \cos(\omega t) dt}_{I_1} + \underbrace{\int_{\ell/2}^{\ell/2+\tau} \left(\frac{\ell/2 + \tau - t}{\tau} \right)^2 \cos(\omega t) dt}_{I_2} \right]. \quad (.123)$$

Integral I_1 (plateau)::

$$I_1 = \frac{\sin(\omega \ell/2)}{\omega}. \quad (.124)$$

Integral I_2 (parabolic ramp):: With $a = \ell/2$ and $b = \ell/2 + \tau$, we integrate by parts twice. Setting $f(t) = ((b - t)/\tau)^2$:

First integration by parts ($u = f(t)$, $dv = \cos(\omega t) dt$):

$$I_2 = \left[\frac{f(t) \sin(\omega t)}{\omega} \right]_a^b + \frac{2}{\tau^2 \omega} \int_a^b (b - t) \sin(\omega t) dt. \quad (.125)$$

The boundary term gives $0 - f(a) \sin(\omega a)/\omega = -\sin(\omega \ell/2)/\omega$ (since $f(a) = 1$).

Second integration by parts on $J = \int_a^b (b-t) \sin(\omega t) dt$ ($u = b-t$, $dv = \sin(\omega t) dt$):

$$J = \left[-\frac{(b-t) \cos(\omega t)}{\omega} \right]_a^b - \frac{1}{\omega} \int_a^b \cos(\omega t) dt = \frac{\tau \cos(\omega \ell/2)}{\omega} - \frac{\sin(\omega b) - \sin(\omega \ell/2)}{\omega^2}. \quad (.126)$$

Combining:

$$I_2 = -\frac{\sin(\omega \ell/2)}{\omega} + \frac{2 \cos(\omega \ell/2)}{\tau \omega^2} - \frac{2 \left[\sin\left(\frac{\omega \ell}{2} + \omega \tau\right) - \sin\left(\frac{\omega \ell}{2}\right) \right]}{\tau^2 \omega^3}. \quad (.127)$$

Combining $I_1 + I_2$: The $\sin(\omega \ell/2)/\omega$ terms from I_1 and I_2 cancel:

$$I_1 + I_2 = \frac{2 \cos(\omega \ell/2)}{\tau \omega^2} + \frac{2 \sin(\omega \ell/2)}{\tau^2 \omega^3} - \frac{2 \sin\left(\frac{\omega \ell}{2} + \omega \tau\right)}{\tau^2 \omega^3}. \quad (.128)$$

Therefore:

$$\hat{\Gamma}(\omega) = \frac{4}{\tau^2 \omega^3} \left[\tau \omega \cos\left(\frac{\omega \ell}{2}\right) + \sin\left(\frac{\omega \ell}{2}\right) - \sin\left(\frac{\omega \ell}{2} + \omega \tau\right) \right]. \quad (.129)$$

3 Autocorrelation of the Squared Envelope

Next we derive the closed-form autocorrelation $R_\Gamma(\tau) = \int_{-\infty}^{\infty} \Gamma(s) \Gamma(s+\tau) ds$ of the normalized squared trapezoidal envelope $\Gamma(t) = \alpha^2(t)/\alpha_{\max}^2$ defined in Eq. (.122). The `sympy` symbolic verification can be found at: `src/derivations/derive_R_Gamma.ipynb`

3.1 Setup

The function $\Gamma(t)$ has three regions: a plateau of height 1 for $|t| \leq \ell/2$, parabolic ramps for $\ell/2 < |t| \leq S$, and zero otherwise. The autocorrelation integral

$$R_\Gamma(\tau) = \int_{-S}^{S-\tau} \Gamma(s) \Gamma(s+\tau) ds, \quad \tau \geq 0, \quad (.130)$$

must be split into sub-intervals where both $\Gamma(s)$ and $\Gamma(s+\tau)$ have known polynomial forms. The breakpoints of $\Gamma(s)$ are at $s = \pm \ell/2$ and $s = \pm S$, while those of $\Gamma(s+\tau)$ are shifted by $-\tau$. The relevant sub-intervals and the character of the overlap change as τ increases through the four intervals $[0, \tau_s]$, $[\tau_s, \ell]$, $[\ell, \ell + \tau_s]$, and $[\ell + \tau_s, \ell + 2\tau_s]$.

3.2 Result

Evaluating the polynomial integrals in each sub-interval and summing (verified symbolically and numerically), the autocorrelation is:

$$R_{\Gamma}(\tau) = \begin{cases} \ell + \frac{2\tau_s}{5} - \frac{4\tau^2}{3\tau_s} + \frac{2\tau^3}{3\tau_s^2} - \frac{\tau^5}{15\tau_s^4} & 0 \leq \tau \leq \tau_s \\ \ell + \frac{2\tau_s}{3} - \tau & \tau_s \leq \tau \leq \ell \\ P_3(\tau) & \ell \leq \tau \leq \ell + \tau_s \\ \frac{(\ell + 2\tau_s - \tau)^5}{30\tau_s^4} & \ell + \tau_s \leq \tau \leq \ell + 2\tau_s \\ 0 & \tau > \ell + 2\tau_s \end{cases} \quad (.131)$$

where the third piece is the degree-5 polynomial:

$$\begin{aligned} P_3(\tau) = & \frac{\tau^5}{30\tau_s^4} - \frac{(\ell + 2\tau_s)\tau^4}{6\tau_s^4} + \frac{(\ell^2 + 4\ell\tau_s + 2\tau_s^2)\tau^3}{3\tau_s^4} \\ & - \frac{\ell(\ell^2 + 6\ell\tau_s + 6\tau_s^2)\tau^2}{3\tau_s^4} + \frac{(\ell^4 + 8\ell^3\tau_s + 12\ell^2\tau_s^2 - 6\tau_s^4)\tau}{6\tau_s^4} \\ & + \frac{-\ell^5 - 10\ell^4\tau_s - 20\ell^3\tau_s^2 + 30\ell\tau_s^4 + 20\tau_s^5}{30\tau_s^4}. \end{aligned} \quad (.132)$$

Figure .42 illustrates these results. The left panel shows the Fourier transform $\hat{\Gamma}(\omega)$ decomposed into the plateau contribution I_1 and the parabolic ramp contribution I_2 . The right panel shows $R_{\Gamma}(\tau)$ with the four piecewise intervals color-coded, overlaid on a numerical convolution for verification.

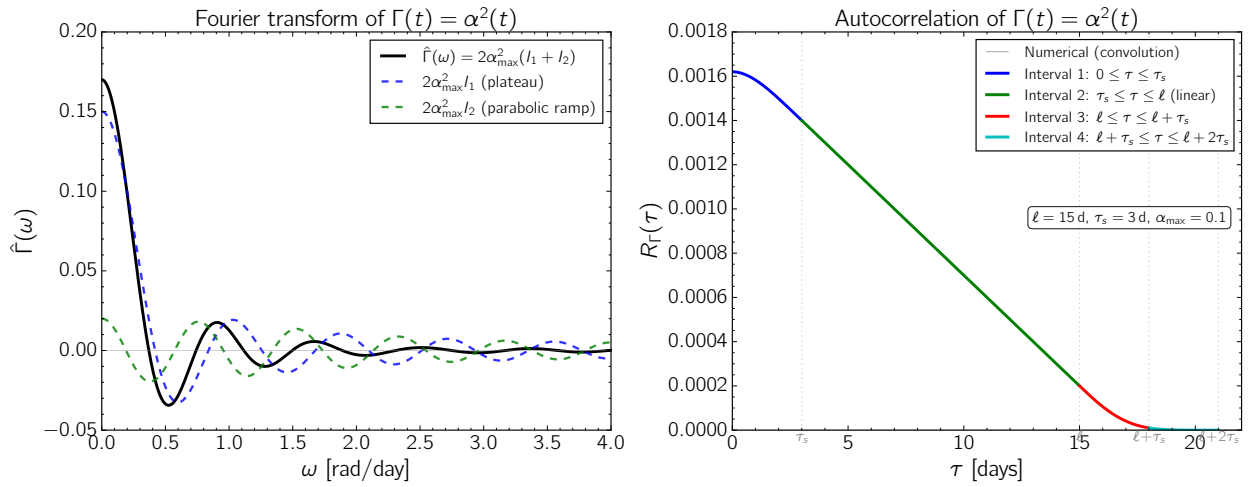


Figure .42: Piecewise structure of the envelope autocorrelation. *Left:* Fourier transform $\hat{\Gamma}(\omega)$ (black) decomposed into the plateau sinc-like term I_1 (blue dashed) and the parabolic ramp term I_2 (orange dashed). *Right:* Autocorrelation $R_\Gamma(\tau)$ from Eq. (.131), with each of the four piecewise intervals shown in a different color. The thin gray curve shows the result of direct numerical convolution, confirming the closed-form expressions. Parameters: $\ell = 15$ d, $\tau_s = 3$ d, $\alpha_{\max} = 0.1$ rad.

CHAPTER 5

CONCLUSION

1	Summary	213
2	Synthesis	214
3	Future Directions	215

1 Summary

This thesis develops and applies novel statistical methods to advance our understanding of stellar rotation, tidal interactions, and starspot variability. The work is motivated by three interconnected questions posed in the Introduction: (1) how can we perform inference with computationally expensive astrophysical models, (2) what can we learn about tidal dissipation from eclipsing binary populations, and (3) how can we infer differential rotation and starspot properties from photometric lightcurves. Each chapter addresses one of these questions, and together they form a progression from general-purpose statistical infrastructure to domain-specific physical models for stellar dynamics.

In Chapter 2, I presented **alabi**, an open-source Python framework for accelerating Bayesian inference when the forward model is computationally expensive. The core idea is to replace the expensive likelihood function with a Gaussian process surrogate trained via active learning, concentrating expensive model evaluations in the regions of parameter space that matter most for posterior inference. Once trained, the surrogate can be evaluated in microseconds, enabling millions of posterior samples at negligible cost.

I validated **alabi** on a suite of benchmark problems spanning unimodal distributions, curved degeneracies, multimodal structure, and dimensionalities up to 64, demonstrating that the framework achieves less than 1% average error with modest training budgets ($\sim 10 \times$ the number of dimensions). For forward models with evaluation times of one second or longer, **alabi** reduces the total number of model evaluations by orders of magnitude compared to direct MCMC, yielding speedups of 10–1000 $\times$. The modular design of the package allows users to swap in different kernels, acquisition functions, and sampling backends, and the entire workflow—from initial sampling through active learning to posterior inference—can be parallelized across computing clusters.

In Chapter 3, I applied **alabi** to the problem of constraining equilibrium tidal dissipation in low-mass binary stars. Using **VPLanet** to simulate the coupled effects of stellar evolution, magnetic braking, and equilibrium tides, I employed global sensitivity analysis and simulated Bayesian inference to systematically investigate how model degeneracies and observational uncertainties limit constraints on the tidal quality factor Q .

The central finding of this work is sobering: even with implausibly precise orbital and rotational parameters from *Kepler* and *TESS*, constraining tidal Q for individual binary systems remains ill-posed. The final orbital state is dominated by the initial orbital period, and strong degeneracies between Q and the unknown initial conditions prevent meaningful inference on a system-by-system basis. Sensitivity analysis revealed that the final orbital period carries little information about tidal dissipation, while the rotation period and eccentricity offer somewhat more leverage—but still insufficient to break the degeneracy with initial conditions in the absence of constraints on time derivatives (dP_{orb}/dt , dP_{rot}/dt , de/dt).

These results suggest that population-level approaches are more promising than individual system analyses. In particular, I identified overdensities in the observed (P_{orb} , e , P_{rot}) distribution of binaries as a potential diagnostic for tidal models, since different formulations of tidal dissipation predict distinct attractor states in this space. This population perspective motivated complementary observational work: in collaboration with undergraduates, I confirmed the existence of a statistically significant population of subsynchronous

binaries in the *TESS* eclipsing binary catalog (Hobson-Ritz et al., 2025), consistent with the earlier *Kepler* discovery by Lurie et al. (2017). The leading physical explanation for this sub-synchronous population—differential rotation combined with preferential spot emergence at higher latitudes—directly motivated the development of the physically motivated GP kernel presented in Chapter 4.

In Chapter 4, I derived a physically motivated Gaussian process kernel for stellar photometric variability from first principles. Starting from an analytic starspot model, I obtained a closed-form autocovariance function whose six hyperparameters—equatorial rotation period (P_{eq}), differential rotation (κ), inclination (I), spot lifetime (ℓ_{spot}), spot emergence/decay timescale (τ_{spot}), and amplitude (σ_k)—map directly onto measurable stellar and spot properties. This stands in contrast to phenomenological kernels such as the quasi-periodic kernel, whose parameters lack clear physical interpretation and which contain no parameter encoding differential rotation.

I validated the analytic kernel against numerical Monte Carlo simulations, demonstrated that it captures harmonic structure and non-sinusoidal periodicity absent in standard kernels, and showed that the compact support of the kernel enables a banded Cholesky solver scaling as $\mathcal{O}(Nb^2)$ rather than $\mathcal{O}(N^3)$, making inference tractable for the long-baseline, densely sampled lightcurves produced by *Kepler* and *TESS*. Using Cramér–Rao bound analysis, I showed that differential rotation is best constrained for short-period, high-inclination stars with long-lived spots, and I used MCMC and nested sampling on simulated lightcurves to demonstrate that the kernel recovers injected parameters when the observational baseline exceeds ~ 10 spot lifetimes.

I applied the kernel to Kepler-63, a young, rapidly rotating Sun-like star, and recovered a rotation period, differential rotation rate, inclination, and spot lifetime consistent with independent literature measurements from transit chord mapping and spectroscopy. This constitutes the first demonstration that differential rotation can be inferred from one-dimensional lightcurves using a physically interpretable Gaussian process.

2 Synthesis

A unifying theme across this thesis is the development and rigorous application of statistical methodology to extract physical insight from astrophysical data, and—equally important—to characterize the limits of what can be inferred.

In Chapter 2, *alabi* provides the computational infrastructure that makes Bayesian inference feasible for expensive forward models. In Chapter 3, applying this infrastructure to tidal evolution reveals that the inference problem itself is fundamentally limited by degeneracies, a conclusion that could not have been reached without the ability to efficiently explore the parameter space. In Chapter 4, the physically motivated GP kernel overcomes a different kind of limitation—the lack of interpretability in phenomenological models—by grounding the kernel in the physics of starspot variability. Together, these contributions illustrate that progress in stellar dynamics requires not only better physical models but also better statistical tools for connecting those models to data.

The three chapters are also connected by the physics of stellar rotation. Tidal interactions govern how binary stars exchange angular momentum and evolve toward synchronization, but

the observed rotation periods of tidally interacting binaries are shaped by differential rotation and the latitude distribution of starspots. The subsynchronous binary population discovered in *Kepler* and confirmed in *TESS* may be a direct consequence of differential rotation, and the GP kernel developed in Chapter 4 provides the first tool capable of measuring this effect from photometric data alone. Conversely, the limitations identified in Chapter 3—that tidal $\mathcal{Q}$ is best constrained through population statistics rather than individual systems—motivate the kind of large-scale, survey-level application that the efficient `spotgp` implementation is designed for.

3 Future Directions

The methods and results presented in this thesis open several avenues for future work.

Population-level differential rotation measurements. The most immediate application of the `spotgp` kernel is to measure differential rotation, spot lifetimes, and stellar inclinations for large samples of stars observed by *Kepler* and *TESS*. Such measurements—currently available for only a handful of targets from Doppler imaging and asteroseismology—would place new empirical constraints on stellar dynamo theory and the dependence of surface shear on stellar mass, age, and rotation rate. The banded GP solver and JAX-based implementation make this computationally tractable.

Connecting tidal evolution and surface activity. A natural next step is to apply `spotgp` to the eclipsing binary samples from *Kepler* (Lurie et al., 2017) and *TESS* (Hobson-Ritz et al., 2025) to directly test whether the subsynchronous population can be explained by differential rotation. If the measured κ values and spot latitude distributions of subsynchronous binaries are consistent with the $\sim 13\%$ period offset, this would confirm the differential rotation hypothesis and provide a new pathway to constraining tidal synchronization physics. Furthermore, Wang et al. (2024) found that starspot lifetimes on tidally interacting binaries differ systematically from those on single stars, an effect that `spotgp` is well suited to characterize with its physically interpretable lifetime parameters.

Extensions to spectroscopy. The Fourier decomposition and Wiener–Khinchin framework underlying the `spotgp` kernel should generalize naturally to spectroscopic time series. Starspots perturb spectral line profiles at radial velocities determined by the same visibility geometry that governs photometric variability, and a two-dimensional kernel $\mathcal{K}(\tau, \Delta v)$ describing covariance across both time lag and velocity separation could break degeneracies present in photometry alone. Joint photometric–spectroscopic inference would be particularly valuable for extreme precision radial velocity surveys, where disentangling stellar activity from planetary signals remains a central challenge.

Broader applications of surrogate-accelerated inference. While `alabi` was applied here to tidal evolution, the framework is general-purpose and applicable to any domain where the likelihood is expensive to evaluate but smooth in parameter space. Potential applications include stellar evolution modeling, chemical abundance analysis with synthetic spectra, and cosmological parameter inference with expensive N -body simulations.

Toward a unified model of stellar rotation and activity. Looking further ahead, the tools developed in this thesis lay the groundwork for a more integrated understanding of stellar rotation. By combining differential rotation measurements from `spotgp` with

tidal evolution models, gyrochronology relations, and activity cycle diagnostics, it may become possible to build a self-consistent picture of how angular momentum is generated, transported, and dissipated over the lifetime of a star—and how these processes shape the environments of orbiting planets.

BIBLIOGRAPHY

- Aigrain, S., Llama, J., Ceillier, T., et al. 2015, *Testing the recovery of stellar rotation signals from Kepler light curves using a blind hare-and-hounds exercise*, Monthly Notices of the Royal Astronomical Society, 450, 3211, doi: [10.1093/mnras/stv853](#)
- Albrecht, S., Reffert, S., Snellen, I., Quirrenbach, A., & Mitchell, D. S. 2007, *The spin axes orbital alignment of both stars within the eclipsing binary system V1143Cyg using the Rossiter-McLaughlin effect*, Astronomy & Astrophysics, 474, 565, doi: [10.1051/0004-6361:20077953](#)
- Albrecht, S. H., Dawson, R. I., & Winn, J. N. 2022, *Stellar obliquities in exoplanetary systems*, arXiv:2203.05460 [astro-ph]. <http://arxiv.org/abs/2203.05460>
- Amard, L., & Matt, S. P. 2020, *The Impact of Metallicity on the Evolution of the Rotation and Magnetic Activity of Sun-like Stars*, The Astrophysical Journal, 889, 108, doi: [10.3847/1538-4357/ab6173](#)
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O’Neil, M. 2015, *Fast Direct Methods for Gaussian Processes*, arXiv:1403.6015 [astro-ph, stat]. <http://arxiv.org/abs/1403.6015>
- Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018a, *Inferring probabilistic stellar rotation periods using Gaussian processes*, Mon Not R Astron Soc, 474, 2094, doi: [10.1093/mnras/stx2109](#)
- . 2018b, *Inferring probabilistic stellar rotation periods using Gaussian processes*, Mon Not R Astron Soc, 474, 2094, doi: [10.1093/mnras/stx2109](#)
- Araújo, A., & Valio, A. 2023, *Dependence of Stellar Differential Rotation on Effective Temperature and Rotation: An Analysis from Starspot Transit Mapping*, The Astrophysical Journal, 956, 141, doi: [10.3847/1538-4357/acf4f0](#)
- Araújo, & Valio. 2023, *Dependence of stellar differential rotation on effective temperature and rotation: an analysis from starspot transit mapping*, arXiv, doi: [10.48550/arXiv.2309.10805](#)
- Augustson, K. C., Brown, B. P., Brun, A. S., Miesch, M. S., & Toomre, J. 2012, *CONVECTION AND DIFFERENTIAL ROTATION IN F-TYPE STARS*, ApJ, 756, 169, doi: [10.1088/0004-637X/756/2/169](#)
- Balandat, M., Karrer, B., Jiang, D. R., et al. 2020, in *Advances in Neural Information Processing Systems* 33. <http://arxiv.org/abs/1910.06403>
- Baliunas, S. L., & Vaughan, A. H. 1985, *Stellar Activity Cycles*, Annual Review of Astronomy and Astrophysics, 23, 379, doi: [10.1146/annurev.aa.23.090185.002115](#)
- Ballot, J., Brun, A. S., & Turck-Chièze, S. 2007, *Simulations of Turbulent Convection in Rotating Young Solarlike Stars: Differential Rotation and Meridional Circulation*, ApJ, 669, 1190, doi: [10.1086/521617](#)
- Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, *New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit*, Astronomy & Astrophysics, 577, A42, doi: [10.1051/0004-6361/201425481](#)
- Barker, A. J. 2020, *Tidal dissipation in evolving low-mass and solar-type stars with predictions for planetary orbital decay*, arXiv:2008.03262 [astro-ph]. <http://arxiv.org/abs/2008.03262>
- . 2022, *Tidal Dissipation Due to Inertial Waves Can Explain the Circularization Periods of Solar-type Binaries*, The Astrophysical Journal Letters, 927, L36, doi: [10.3847/2041-8213/ac5b63](#)
- Barker, A. J., & Ogilvie, G. I. 2009, *On the tidal evolution of Hot Jupiters on inclined orbits*, \mnras, 395, 2268, doi: [10.1111/j.1365-2966.2009.14694.x](#)
- Barnes, J. R., Cameron, A. C., Donati, J.-F., et al. 2005, *The Dependence of Differential Rotation on Temperature and Rotation*, Monthly Notices of the Royal Astronomical Society: Letters, 357, L1, doi: [10.1111/j.1745-3933.2005.08587.x](#)
- Barnes, R. 2017, *Tidal Locking of Habitable Exoplanets*, doi: [10.1007/s10569-017-9783-7](#)
- Barnes, R., do Amaral, L. N. R., Birky, J., et al. 2026, *The Range of Cumulative XUV Flux on GJ 1132 b*. <https://arxiv.org/abs/2605.03939>
- Barnes, R., Mullins, K., Goldblatt, C., et al. 2013, *Tidal Venuses: Triggering a Climate Catastrophe via Tidal Heating*, Astrobiology, 13, 225, doi: [10.1089/ast.2012.0851](#)
- Barnes, R., Luger, R., Deitrick, R., et al. 2020, *VPLANET: The Virtual Planet Simulator*, Publications of the Astronomical Society of the Pacific, 132, 024502, doi: [10.1088/1538-3873/ab3ce8](#)
- Barnes, R., Luger, R., Deitrick, R., et al. 2020, *VPLANET: The Virtual Planet Simulator*, Proc. Astron. Soc. Pac., 132, 024502, doi: [10.1088/1538-3873/ab3ce8](#)
- Barnes, S. A. 2003a, *On the Rotational Evolution of Solar- and Late-Type Stars, Its Magnetic Origins, and the Possibility of Stellar Gyrochronology**, The Astrophysical Journal, 586, 464, doi: [10.1086/367639](#)
- . 2003b, *On the Rotational Evolution of Solar- and Late-Type Stars, Its Magnetic Origins, and the Possibility of Stellar Gyrochronology**, The Astrophysical Journal, 586, 464, doi: [10.1086/367639](#)
- . 2007, *Ages for Illustrative Field Stars Using Gyrochronology: Viability, Limitations, and Errors*, \textbackslashapj, 669, 1167, doi: [10.1086/519295](#)
- Bashi, D., Mazeh, T., & Faigler, S. 2023, *Features of Gaia DR3 Spectroscopic Binaries I. Tidal circularization of Main-Sequence Stars*, Monthly Notices of the Royal Astronomical Society, stad999, doi: [10.1093/mnras/stad999](#)
- Basri, G., & Shah, R. 2020, *The Information Content in Analytic Spot Models of Broadband Precision Light Curves. II. Spot Distributions and Lifetimes, Global and Differential Rotation*, arXiv, doi: [10.48550/arXiv.2008.04969](#)

- Benomar, O., Bazot, M., Nielsen, M. B., et al. 2018, *Asteroseismic detection of latitudinal differential rotation in 13 Sun-like stars*, Science, 361, 1231, doi: [10.1126/science.aao6571](https://doi.org/10.1126/science.aao6571)
- Binois, M., & Wycoff, N. 2022, *A survey on high-dimensional Gaussian process modeling with application to Bayesian optimization*, arXiv, doi: [10.48550/arXiv.2111.05040](https://doi.org/10.48550/arXiv.2111.05040)
- Birky, J., Barnes, R., & Fleming, D. P. 2021a, *Improved Constraints for the XUV Luminosity Evolution of Trappist-1*, RNAAS, 5, 122, doi: [10.3847/2515-5172/ac034c](https://doi.org/10.3847/2515-5172/ac034c)
- . 2021b, *Improved Constraints for the XUV Luminosity Evolution of Trappist-1*, Research Notes of the AAS, 5, 122, doi: [10.3847/2515-5172/ac034c](https://doi.org/10.3847/2515-5172/ac034c)
- Birky, J., Barnes, R. K., & Davenport, J. R. A. 2025, *Prospects of Constraining Equilibrium Tides in Low-Mass Binary Stars*, arXiv, doi: [10.48550/arXiv.2507.12639](https://doi.org/10.48550/arXiv.2507.12639)
- Birky, J., Fleming, D. P., & Barnes, R. K. in prep., *Active Learning for Accelerated Bayesian Inference*
- Bolmont, E., & Mathis, S. 2016, *Effect of the rotation and tidal dissipation history of stars on the evolution of close-in planets*, Celestial Mechanics and Dynamical Astronomy, 126, 275, doi: [10.1007/s10569-016-9690-3](https://doi.org/10.1007/s10569-016-9690-3)
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Kepler Planet-Detection Mission: Introduction and First Results*, Science, 327, 977, doi: [10.1126/science.1185402](https://doi.org/10.1126/science.1185402)
- Boyle, A. W., Bouma, L. G., & Mann, A. W. 2026, *The TESS All-Sky Rotation Survey: Periods for 944,056 Stars Within 500 pc*, <https://arxiv.org/abs/2603.05586>
- Breimann, A. A., Matt, S. P., & Naylor, T. 2021, *Statistical Fitting of Evolution Models to Rotation Rates of Sun-Like Stars*, The Astrophysical Journal, 913, 75, doi: [10.3847/1538-4357/abf0a3](https://doi.org/10.3847/1538-4357/abf0a3)
- Brun, A. S., & Browning, M. K. 2017, *Magnetism, dynamo action and the solar-stellar connection*, Living Reviews in Solar Physics, 14, 1, doi: [10.1007/s41116-017-0007-8](https://doi.org/10.1007/s41116-017-0007-8)
- Brun, A. S., Strugarek, A., Varela, J., et al. 2017a, *On differential rotation and overshooting in solar-like stars*, ApJ, 836, 192, doi: [10.3847/1538-4357/aa5c40](https://doi.org/10.3847/1538-4357/aa5c40)
- . 2017b, *On differential rotation and overshooting in solar-like stars*, ApJ, 836, 192, doi: [10.3847/1538-4357/aa5c40](https://doi.org/10.3847/1538-4357/aa5c40)
- Buchner, J. 2021, *UltraNest – a robust, general purpose Bayesian inference engine*, Journal of Open Source Software, 6, 3001, doi: [10.21105/joss.03001](https://doi.org/10.21105/joss.03001)
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue*, Astronomy & Astrophysics, 564, A125, doi: [10.1051/0004-6361/201322971](https://doi.org/10.1051/0004-6361/201322971)
- Burkart, J., Quataert, E., & Arras, P. 2014, *Dynamical resonance locking in tidally interacting binary systems*, Monthly Notices of the Royal Astronomical Society, 443, 2957, doi: [10.1093/mnras/stu1366](https://doi.org/10.1093/mnras/stu1366)
- Cameron, A. C., & Jardine, M. 2018, *Hierarchical Bayesian calibration of tidal orbit decay rates among hot Jupiters*, Monthly Notices of the Royal Astronomical Society, 476, 2542, doi: [10.1093/mnras/sty292](https://doi.org/10.1093/mnras/sty292)
- Carroll, J. A. 1933, *The Spectroscopic Determination of Stellar Rotation and its Effect on Line Profiles*, Monthly Notices of the Royal Astronomical Society, 93, 478, doi: [10.1093/mnras/93.7.478](https://doi.org/10.1093/mnras/93.7.478)
- Casey, A. R., Ho, A. Y. Q., Ness, M., et al. 2019, *Tidal interactions between binary stars drives lithium production in low-mass red giants*, The Astrophysical Journal, 880, 125, doi: [10.3847/1538-4357/ab27bf](https://doi.org/10.3847/1538-4357/ab27bf)
- Charbonneau, P. 2020, *Dynamo Models of the Solar Cycle*, Living Reviews in Solar Physics, 17, 4, doi: [10.1007/s41116-020-00025-6](https://doi.org/10.1007/s41116-020-00025-6)
- Chontos, A., Huber, D., Latham, D. W., et al. 2019, *The Curious Case of KOI 4: Confirming Kepler’s First Exoplanet Detection*, The Astronomical Journal, 157, 192, doi: [10.3847/1538-3881/ab0e8e](https://doi.org/10.3847/1538-3881/ab0e8e)
- Coffaro, M., Stelzer, B., & Orlando, S. 2022, *Characterizing X-ray activity cycles of young solar-like stars with solar observations*, Astron Nachr, 343, e210066, doi: [10.1002/asna.20210066](https://doi.org/10.1002/asna.20210066)
- Cohen, N., Hamann, J., & Malhotra, A. 2026, *Bayesian optimisation for Bayesian evidence (BOBE) – a fast and efficient likelihood emulator for model selection*, <https://arxiv.org/abs/2601.11150>
- Collier Cameron, A. 2007, *Differential rotation on rapidly rotating stars*, AN, 328, 1030, doi: [10.1002/asna.200710880](https://doi.org/10.1002/asna.200710880)
- Collier Cameron, A., Donati, J.-F., & Semel, M. 2002, *Stellar differential rotation from direct star-spot tracking*, Monthly Notices of the Royal Astronomical Society, 330, 699, doi: [10.1046/j.1365-8711.2002.05120.x](https://doi.org/10.1046/j.1365-8711.2002.05120.x)
- Colombo, G., & Shapiro, I. I. 1966, *The Rotation of the Planet Mercury*, The Astrophysical Journal, 145, 296, doi: [10.1086/148762](https://doi.org/10.1086/148762)
- Correia, A. C. M., & Laskar, J. 2004, *Mercury’s capture into the 3/2 spin-orbit resonance as a result of its chaotic dynamics*, Nature, 429, 848, doi: [10.1038/nature02609](https://doi.org/10.1038/nature02609)
- Counselman, III, C. C. 1973, *Outcomes of Tidal Evolution*, \apj, 180, 307, doi: [10.1086/151964](https://doi.org/10.1086/151964)
- Cramér, H. 1946, *Mathematical Methods of Statistics* (Princeton, NJ: Princeton University Press)
- Cranmer, K., Brehmer, J., & Louppe, G. 2020, *The frontier of simulation-based inference*, Proceedings of the National Academy of Sciences, 117, 30055, doi: [10.1073/pnas.1912789117](https://doi.org/10.1073/pnas.1912789117)
- Cranmer, S. R., & Saar, S. H. 2011, *Testing a Predictive Theoretical Model for the Mass Loss Rates of Cool Stars*, \apj, 741, 54, doi: [10.1088/0004-637X/741/1/54](https://doi.org/10.1088/0004-637X/741/1/54)
- David, T. J., Hillenbrand, L. A., Cody, A. M., Carpenter, J. M., & Howard, A. W. 2015, *K2 Discovery of Young Eclipsing Binaries in Upper Scorpius: Direct Mass and Radius Determinations for the Lowest Mass Stars and Initial Characterization of an Eclipsing Brown Dwarf Binary*, The Astrophysical Journal, 816, 21, doi: [10.3847/0004-637X/816/1/21](https://doi.org/10.3847/0004-637X/816/1/21)
- David, T. J., Conroy, K. E., Hillenbrand, L. A., et al. 2016, *New Pleiades Eclipsing Binaries and a Hyades Transiting System Identified by K2*, The Astronomical Journal, 151, 112, doi: [10.3847/0004-6256/151/5/112](https://doi.org/10.3847/0004-6256/151/5/112)
- Dittmann, A. 2024, *Notes on the Practical Application of Nested Sampling: MultiNest, (Non)convergence, and Rectification*, The Open Journal of Astrophysics, 7, doi: [10.33232/001c.123872](https://doi.org/10.33232/001c.123872)
- Donahue, R. A., Saar, S. H., & Baliunas, S. L. 1996, *A Relationship between Mean Rotation Period in Lower Main-Sequence Stars and Its Observed Range*, The Astrophysical Journal, 466, 384, doi: [10.1086/177517](https://doi.org/10.1086/177517)

- Duquennoy, A., & Mayor, M. 1991, *Multiplicity among Solar Type Stars in the Solar Neighbourhood - Part Two - Distribution of the Orbital Elements in an Unbiased Sample*, *Astron. & Astrophys.*, 248, 485
- Efroimsky, M., & Williams, J. G. 2009, *Tidal torques. A critical review of some techniques*, *Celestial Mechanics and Dynamical Astronomy*, 104, 257, doi: [10.1007/s10569-009-9204-7](https://doi.org/10.1007/s10569-009-9204-7)
- Eiff, M. A.-v., & Reiners, A. 2012, *New measurements of rotation and differential rotation in A-F stars: Are there two populations of differentially rotating stars?*, *A&A*, 542, A116, doi: [10.1051/0004-6361/201118724](https://doi.org/10.1051/0004-6361/201118724)
- El-Badry, K., Ting, Y.-S., Rix, H.-W., et al. 2018, *Discovery and characterization of 3000+ main-sequence binaries from APOGEE spectra*, *MNRAS*, 476, 528, doi: [10.1093/mnras/sty240](https://doi.org/10.1093/mnras/sty240)
- El Gammal, J., Schöneberg, N., Torrado, J., & Fidler, C. 2023, *Fast and robust Bayesian inference using Gaussian processes with GPry*, *Journal of Cosmology and Astroparticle Physics*, 2023, 021, doi: [10.1088/1475-7516/2023/10/021](https://doi.org/10.1088/1475-7516/2023/10/021)
- Estrela, R., & Valio, A. 2016, *STELLAR MAGNETIC CYCLES IN THE SOLAR-LIKE STARS KEPLER-17 AND KEPLER-63*, *ApJ*, 831, 57, doi: [10.3847/0004-637X/831/1/57](https://doi.org/10.3847/0004-637X/831/1/57)
- Fabrycky, D., & Tremaine, S. 2007, *Shrinking Binary and Planetary Orbits by Kozai Cycles with Tidal Friction**, *The Astrophysical Journal*, 669, 1298, doi: [10.1086/521702](https://doi.org/10.1086/521702)
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics*, *Monthly Notices of the Royal Astronomical Society*, 398, 1601, doi: [10.1111/j.1365-2966.2009.14548.x](https://doi.org/10.1111/j.1365-2966.2009.14548.x)
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, *Importance Nested Sampling and the MultiNest Algorithm*, *The Open Journal of Astrophysics*, 2, 10.21105/astro.1306.2144, doi: [10.21105/astro.1306.2144](https://doi.org/10.21105/astro.1306.2144)
- Ferraz-Mello, S., Rodríguez, A., & Hussmann, H. 2008, *Tidal friction in close-in satellites and exoplanets: The Darwin theory re-visited*, *Celestial Mechanics and Dynamical Astronomy*, 101, 171, doi: [10.1007/s10569-008-9133-x](https://doi.org/10.1007/s10569-008-9133-x)
- Fleming, D. P., Barnes, R., Davenport, J. R. A., & Luger, R. 2019, *Rotation Period Evolution in Low-mass Binary Stars: The Impact of Tidal Torques and Magnetic Braking*, *The Astrophysical Journal*, 881, 88, doi: [10.3847/1538-4357/ab2ed2](https://doi.org/10.3847/1538-4357/ab2ed2)
- Fleming, D. P., Barnes, R., Graham, D. E., Luger, R., & Quinn, T. R. 2018, *On the Lack of Circumbinary Planets Orbiting Isolated Binary Stars*, *The Astrophysical Journal*, 858, 86, doi: [10.3847/1538-4357/aabd38](https://doi.org/10.3847/1538-4357/aabd38)
- Fleming, D. P., Barnes, R., Luger, R., & VanderPlas, J. T. 2020, *On the XUV Luminosity Evolution of TRAPPIST-1*, *The Astrophysical Journal*, 891, 155, doi: [10.3847/1538-4357/ab77ad](https://doi.org/10.3847/1538-4357/ab77ad)
- Fleming, D. P., & VanderPlas, J. 2018, *approxposterior: Approximate Posterior Distributions in Python*, *Journal of Open Source Software*, 3, 781, doi: [10.21105/joss.00781](https://doi.org/10.21105/joss.00781)
- Foreman-Mackey, D. 2016, *corner.py: Scatterplot matrices in Python*, *The Journal of Open Source Software*, 1, doi: [10.21105/joss.00024](https://doi.org/10.21105/joss.00024)
- . 2024, *tinygp: The tiniest of Gaussian Process libraries*, doi: [10.5281/zenodo.10463641](https://doi.org/10.5281/zenodo.10463641)
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *Fast and Scalable Gaussian Process Modeling with Applications to Astronomical Time Series*, *The Astronomical Journal*, 154, 220, doi: [10.3847/1538-3881/aa9332](https://doi.org/10.3847/1538-3881/aa9332)
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *Fast and scalable Gaussian process modeling with applications to astronomical time series*, *The Astronomical Journal*, 154, 220, doi: [10.3847/1538-3881/aa9332](https://doi.org/10.3847/1538-3881/aa9332)
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2023, *celerite2: Fast and scalable Gaussian Processes in one dimension*, *Astrophysics Source Code Library*, record ascl:2310.001. <http://ascl.net/2310.001>
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *emcee: The MCMC Hammer*, *Publications of the Astronomical Society of the Pacific*, 125, 306, doi: [10.1086/670067](https://doi.org/10.1086/670067)
- Forgács-Dajka, E., Dobos, L., & Ballai, I. 2021, *Time-dependent properties of sunspot groups: I. Lifetime and asymmetric evolution*, *Astronomy & Astrophysics*, 653, A50, doi: [10.1051/0004-6361/202140731](https://doi.org/10.1051/0004-6361/202140731)
- Fuller, J., & Lai, D. 2012, *Dynamical tides in compact white dwarf binaries: tidal synchronization and dissipation*, *Mon. Not. R. Astron. Soc.*, 421, 426, doi: [10.1111/j.1365-2966.2011.20320.x](https://doi.org/10.1111/j.1365-2966.2011.20320.x)
- Gastine, T., Yadav, R. K., Morin, J., Reiners, A., & Wicht, J. 2013, *From Solar-like to Antisolar Differential Rotation in Cool Stars*, *Monthly Notices of the Royal Astronomical Society: Letters*, 438, L76, doi: [10.1093/mnrasl/slt162](https://doi.org/10.1093/mnrasl/slt162)
- Gelman, A., Carlin, J. B., Stern, H. S., et al. 2013, *Bayesian Data Analysis*, Third Edition (CRC Press)
- Gilbert-Janizek, S., Barnes, R. K., Driscoll, P. E., et al. 2026, *A whole-planet model of the Earth without life for terrestrial exoplanet studies*, *arXiv*, doi: [10.48550/arXiv.2602.02267](https://doi.org/10.48550/arXiv.2602.02267)
- Giles, H. A. C., Cameron, A. C., & Haywood, R. D. 2017, *A Kepler Study of Starspot Lifetimes with Respect to Light Curve Amplitude and Spectral Type*, *Monthly Notices of the Royal Astronomical Society*, 472, 1618, doi: [10.1093/mnras/stx1931](https://doi.org/10.1093/mnras/stx1931)
- Gillen, E., Hillenbrand, L. A., David, T. J., et al. 2017, *New low-mass eclipsing binary systems in Praesepe discovered by K2*, *The Astrophysical Journal*, 849, 11, doi: [10.3847/1538-4357/aa84b3](https://doi.org/10.3847/1538-4357/aa84b3)
- Ginat, Y. B., & Perets, H. B. 2021, *An Analytical, Statistical Approximate Solution for Dissipative and non-Dissipative Binary-Single Stellar Encounters*, *Physical Review X*, 11, 031020, doi: [10.1103/PhysRevX.11.031020](https://doi.org/10.1103/PhysRevX.11.031020)
- Gizon, L., & Solanki, S. K. 2003, *Determining the Inclination of the Rotation Axis of a Sun-like Star*, *The Astrophysical Journal*, 589, 1009, doi: [10.1086/374715](https://doi.org/10.1086/374715)
- Goldreich, P., & Peale, S. 1966, *Spin-orbit coupling in the solar system*, *The Astronomical Journal*, 71, 425, doi: [10.1086/109947](https://doi.org/10.1086/109947)
- Goldreich, P., & Soter, S. 1966, *Q in the solar system*, *Icarus*, 5, 375, doi: [10.1016/0019-1035\(66\)90051-0](https://doi.org/10.1016/0019-1035(66)90051-0)
- Golub, G. H., & Van Loan, C. F. 2013, *Matrix Computations*, 4th edn. (Baltimore: Johns Hopkins University Press)
- Gordon, T. A., Davenport, J. R. A., Angus, R., et al. 2021, *Stellar Rotation in the K2 Sample: Evidence for Modified Spin-down*, *The Astrophysical Journal*, 913, 70, doi: [10.3847/1538-4357/abf63e](https://doi.org/10.3847/1538-4357/abf63e)
- Gould, N. I. M., Lucidi, S., Roma, M., & Toint, P. L. 1999, *Solving the Trust-Region Subproblem using the Lanczos Method*, *SIAM Journal on Optimization*, 9, 504, doi: [10.1137/S1052623497322735](https://doi.org/10.1137/S1052623497322735)
- Gray, D. F. 1976, *The observation and analysis of stellar photospheres*. <https://ui.adsabs.harvard.edu/abs/1976oasp.book.....G>

- . 1977, *A quest for differential stellar rotation in A stars*, The Astrophysical Journal, 211, 198, doi: [10.1086/154918](https://doi.org/10.1086/154918)
- Greenberg, R.** 1974, *Outcomes of tidal evolution for orbits with arbitrary inclination*, *icarus*, 23, 51, doi: [10.1016/0019-1035\(74\)90103-1](https://doi.org/10.1016/0019-1035(74)90103-1)
- . 2009, *FREQUENCY DEPENDENCE OF TIDAL Q*, The Astrophysical Journal, 698, L42, doi: [10.1088/0004-637X/698/1/L42](https://doi.org/10.1088/0004-637X/698/1/L42)
- Guerrero, G., Smolarkiewicz, P. K., Kosovichev, A. G., & Mansour, N. N.** 2013a, *Differential rotation in solar-like stars from global simulations*, The Astrophysical Journal, 779, 176, doi: [10.1088/0004-637X/779/2/176](https://doi.org/10.1088/0004-637X/779/2/176)
- . 2013b, *Differential rotation in solar-like stars from global simulations*, The Astrophysical Journal, 779, 176, doi: [10.1088/0004-637X/779/2/176](https://doi.org/10.1088/0004-637X/779/2/176)
- Hall, D. S.** 1972, *A T Tauri-like Star in the Eclipsing Binary RS Canum Venaticorum*, Publications of the Astronomical Society of the Pacific, 84, 323, doi: [10.1086/129291](https://doi.org/10.1086/129291)
- . 1991, in *Lecture Notes in Physics*, Vol. 380, The Sun and Cool Stars: activity, magnetism, dynamos, ed. I. Tuominen, D. Moss, & G. Rüdiger (Berlin, Heidelberg: Springer), 353–370, doi: [10.1007/3-540-53955-7_156](https://doi.org/10.1007/3-540-53955-7_156)
- Hamer, J. H., & Schlaufman, K. C.** 2019, *Hot Jupiters Are Destroyed by Tides While Their Host Stars Are on the Main Sequence*, The Astronomical Journal, 158, 190, doi: [10.3847/1538-3881/ab3c56](https://doi.org/10.3847/1538-3881/ab3c56)
- Hamers, A. S., Rantala, A., Neunteufel, P., Preece, H., & Vynatheya, P.** 2021, *Multiple Stellar Evolution: a population synthesis algorithm to model the stellar, binary, and dynamical evolution of multiple-star systems*, Monthly Notices of the Royal Astronomical Society, 502, 4479, doi: [10.1093/mnras/stab287](https://doi.org/10.1093/mnras/stab287)
- Hansen, B. M. S.** 2010, *CALIBRATION OF EQUILIBRIUM TIDE THEORY FOR EXTRASOLAR PLANET SYSTEMS*, The Astrophysical Journal, 723, 285, doi: [10.1088/0004-637X/723/1/285](https://doi.org/10.1088/0004-637X/723/1/285)
- Hathaway, D. H.** 2015, *The Solar Cycle*, Living Reviews in Solar Physics, 12, 1, doi: [10.1007/lrsp-2015-4](https://doi.org/10.1007/lrsp-2015-4)
- Hatzes, A. P.** 2019, in *The Doppler Method for the Detection of Exoplanets*, 2514-3433 (IOP Publishing), 13–1 to 13–14, doi: [10.1088/2514-3433/ab46a3ch13](https://doi.org/10.1088/2514-3433/ab46a3ch13)
- Head, T., Kumar, M., Nahrstaedt, H., Louppe, G., & Shcherbatyi, I.** 2020, *scikit-optimize/scikit-optimize*, Zenodo, doi: [10.5281/zenodo.4014775](https://doi.org/10.5281/zenodo.4014775)
- Heller, R., Leconte, J., & Barnes, R.** 2011, *Tidal obliquity evolution of potentially habitable planets*, *\aap*, 528, A27, doi: [10.1051/0004-6361/201015809](https://doi.org/10.1051/0004-6361/201015809)
- Herman, J., & Usher, W.** 2017, *SALib: An open-source Python library for Sensitivity Analysis*, Journal of Open Source Software, 2, 97, doi: [10.21105/joss.00097](https://doi.org/10.21105/joss.00097)
- Hobson-Ritz, M., Birky, J., Peterson, L., et al.** 2025, *Tidal Synchronization of TESS Eclipsing Binaries*, arXiv e-prints, arXiv:2501.04082, doi: [10.48550/arXiv.2501.04082](https://doi.org/10.48550/arXiv.2501.04082)
- Hoffman, M. D., & Gelman, A.** 2014, *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*, Journal of Machine Learning Research, 15, 1593. <http://jmlr.org/papers/v15/hoffman14a.html>
- Hogg, D. W., Bovy, J., & Lang, D.** 2010, *Data analysis recipes: Fitting a model to data*, arXiv:1008.4686 [astro-ph, physics:physics]. [http://arxiv.org/abs/1008.4686](https://arxiv.org/abs/1008.4686)
- Howell, S. B., Sobeck, C., Haas, M., et al.** 2014, *The K2 Mission: Characterization and Early Results*, *\pasp*, 126, 398, doi: [10.1086/676406](https://doi.org/10.1086/676406)
- Hunter, J. D.** 2007, *Matplotlib: A 2D Graphics Environment*, Computing in Science and Engineering, 9, 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Hut, P.** 1980, *Stability of tidal equilibrium*, Astronomy and Astrophysics, 92, 167. <https://ui.adsabs.harvard.edu/abs/1980A&A...92..167H/abstract>
- . 1981, *Tidal evolution in close binary systems*, Astronomy and Astrophysics, 99, 126. <http://adsabs.harvard.edu/abs/1981A%26A...99..126H>
- Hvarfner, C., Hellsten, E. O., & Nardi, L.** 2024, *Vanilla Bayesian Optimization Performs Great in High Dimensions*. <https://arxiv.org/abs/2402.02229>
- Ikuta, K., Machara, H., Notsu, Y., et al.** 2020, *Starspot Mapping with Adaptive Parallel Tempering. I. Implementation of Computational Code*, The Astrophysical Journal, 902, 73, doi: [10.3847/1538-4357/abb5f1](https://doi.org/10.3847/1538-4357/abb5f1)
- Ireland, L. G., Matt, S. P., Davey, C. R., et al.** 2022, *Effect of Differential Rotation on the Magnetic Braking of Low-mass and Solar-like Stars: A Proof-of-concept Study*, The Astrophysical Journal, 925, 100, doi: [10.3847/1538-4357/ac3e69](https://doi.org/10.3847/1538-4357/ac3e69)
- Jackson, B., Barnes, R., & Greenberg, R.** 2009, *Observational Evidence for Tidal Destruction of Exoplanets*, *\apj*, 698, 1357, doi: [10.1088/0004-637X/698/2/1357](https://doi.org/10.1088/0004-637X/698/2/1357)
- Jackson, B., Greenberg, R., & Barnes, R.** 2008, *Tidal Evolution of Close-in Extrasolar Planets*, *\apj*, 678, 1396, doi: [10.1086/529187](https://doi.org/10.1086/529187)
- Jermyn, A. S., Tayar, J., & Fuller, J.** 2020, *Differential rotation in convective envelopes: constraints from eclipsing binaries*, Monthly Notices of the Royal Astronomical Society, 491, 690, doi: [10.1093/mnras/stz2983](https://doi.org/10.1093/mnras/stz2983)
- Jones, D. R., Schonlau, M., & Welch, W. J.** 1998, *Efficient Global Optimization of Expensive Black-Box Functions*, Journal of Global Optimization, 13, 455, doi: [10.1023/A:1008306431147](https://doi.org/10.1023/A:1008306431147)
- Jones, E., Oliphant, T., Peterson, P., & others.** 2001, *SciPy: Open source scientific tools for Python*. <http://www.scipy.org/>
- Jones, G. L., & Hobert, J. P.** 2001, *Honest Exploration of Intractable Probability Distributions via Markov Chain Monte Carlo*, Statistical Science, 16, 312, doi: [10.1214/ss/1015346317](https://doi.org/10.1214/ss/1015346317)
- Kandasamy, K., Schneider, J., & Póczos, B.** 2017, *Query efficient posterior estimation in scientific experiments via Bayesian active learning*, Artificial Intelligence, 243, 45, doi: <https://doi.org/10.1016/j.artint.2016.11.002>
- Käpylä, P. J., Browning, M. K., Brun, A. S., Guerrero, G., & Warnecke, J.** 2023, *Simulations of Solar and Stellar Dynamos and Their Theoretical Interpretation*, Space Science Reviews, 219, 58, doi: [10.1007/s11214-023-01005-6](https://doi.org/10.1007/s11214-023-01005-6)
- Kaula, W. M.** 1966, *Theory of satellite geodesy. Applications of satellites to geodesy*
- Kawaler, S. D.** 1988, *Angular momentum loss in low-mass stars*, *\textbackslashtextbackslashashapj*, 333, 236, doi: [10.1086/166740](https://doi.org/10.1086/166740)

- Kipping, D. M. 2012, *An analytic model for rotational modulations in the photometry of spotted stars: An analytic model for rotational modulations*, Monthly Notices of the Royal Astronomical Society, 427, 2487, doi: [10.1111/j.1365-2966.2012.22124.x](https://doi.org/10.1111/j.1365-2966.2012.22124.x)
- Klein, B., Aigrain, S., Cretignier, M., et al. 2025, *Using Doppler Imaging to model stellar activity and search for planets around Sun-like stars*, Monthly Notices of the Royal Astronomical Society, 542, 2714, doi: [10.1093/mnras/staf1337](https://doi.org/10.1093/mnras/staf1337)
- Kohavi, R. 1995, in Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI) (Morgan Kaufmann), 1137–1143
- Kounkel, M., Covey, K. R., Stassun, K. G., et al. 2021, *Double-lined spectroscopic binaries in the APOGEE DR16 and DR17 data*, The Astronomical Journal, 162, 184, doi: [10.3847/1538-3881/ac1798](https://doi.org/10.3847/1538-3881/ac1798)
- Küker, M., & Rüdiger, G. 2011, *Differential rotation and meridional flow on the lower zero-age main sequence*, Astronomische Nachrichten, 332, 933, doi: [10.1002/asna.201111628](https://doi.org/10.1002/asna.201111628)
- Kővári, Z., Oláh, K., Kriskovics, L., et al. 2017, *Rotation–differential rotation relationships for late-type single and binary stars from Doppler imaging*, Astron Nachr, 338, 903, doi: [10.1002/asna.201713400](https://doi.org/10.1002/asna.201713400)
- Lai, D. 1997, *Dynamical Tides in Rotating Binary Stars*, The Astrophysical Journal, 490, 847, doi: [10.1086/304899](https://doi.org/10.1086/304899)
- Lao, J., Suter, C., Langmore, I., et al. 2020, *tfp.mcmc: Modern Markov Chain Monte Carlo Tools Built for Modern Hardware*. <https://arxiv.org/abs/2002.01184>
- Leconte, J., Chabrier, G., Baraffe, I., & Levrard, B. 2010, *Is tidal heating sufficient to explain bloated exoplanets? Consistent calculations accounting for finite initial eccentricity*, Astronomy and Astrophysics, 516, A64, doi: [10.1051/0004-6361/201014337](https://doi.org/10.1051/0004-6361/201014337)
- Levrard, B., Correia, A. C. M., Chabrier, G., et al. 2007, *Tidal dissipation within hot Jupiters: a new appraisal*, Astronomy & Astrophysics, 462, L5, doi: [10.1051/0004-6361:20066487](https://doi.org/10.1051/0004-6361:20066487)
- Lu, Y. L., Angus, R., Agüeros, M. A., et al. 2024, *In This Day and Age: An Empirical Gyrochronology Relation for Partially and Fully Convective Single Field Stars*, The Astronomical Journal, 167, 159, doi: [10.3847/1538-3881/ad28b9](https://doi.org/10.3847/1538-3881/ad28b9)
- Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019, *STARRY: Analytic Occultation Light Curves*, The Astronomical Journal, 157, 64, doi: [10.3847/1538-3881/aae8e5](https://doi.org/10.3847/1538-3881/aae8e5)
- Luger, R., Bedell, M., Foreman-Mackey, D., et al. 2021a, *Mapping stellar surfaces III: An Efficient, Scalable, and Open-Source Doppler Imaging Model*, arXiv, doi: [10.48550/arXiv.2110.06271](https://doi.org/10.48550/arXiv.2110.06271)
- Luger, R., Foreman-Mackey, D., & Hedges, C. 2021b, *Mapping Stellar Surfaces. II. An Interpretable Gaussian Process Model for Light Curves*, The Astronomical Journal, 162, 124, doi: [10.3847/1538-3881/abfdb9](https://doi.org/10.3847/1538-3881/abfdb9)
- Luger, R., Foreman-Mackey, D., Hedges, C., & Hogg, D. W. 2021c, *Mapping Stellar Surfaces. I. Degeneracies in the Rotational Light-curve Problem*, The Astronomical Journal, 162, 123, doi: [10.3847/1538-3881/abfdb8](https://doi.org/10.3847/1538-3881/abfdb8)
- Lund, M. N., Miesch, M. S., & Christensen-Dalsgaard, J. 2014, *DIFFERENTIAL ROTATION IN MAIN-SEQUENCE SOLAR-LIKE STARS: QUALITATIVE INFERENCE FROM ASTEROSEISMIC DATA*, ApJ, 790, 121, doi: [10.1088/0004-637X/790/2/121](https://doi.org/10.1088/0004-637X/790/2/121)
- Lurie, J. C., Vyhmeister, K., Hawley, S. L., et al. 2017, *Tidal Synchronization and Differential Rotation of KeplerEclipsing Binaries*, The Astronomical Journal, 154, 250, doi: [10.3847/1538-3881/aa974d](https://doi.org/10.3847/1538-3881/aa974d)
- MacDonald, G. J. F. 1964, *Tidal friction*, Reviews of Geophysics, 2, 467, doi: [10.1029/RG002i003p00467](https://doi.org/10.1029/RG002i003p00467)
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2015, *The Apache Point Observatory Galactic Evolution Experiment (APOGEE)*, \aj, 154, 94, doi: [10.3847/1538-3881/aa784d](https://doi.org/10.3847/1538-3881/aa784d)
- Mardling, R. A., & Aarseth, S. J. 2001, *Tidal interactions in star cluster simulations*, Monthly Notices of the Royal Astronomical Society, 321, 398, doi: [10.1046/j.1365-8711.2001.03974.x](https://doi.org/10.1046/j.1365-8711.2001.03974.x)
- Marilli, E., Frasca, A., Covino, E., et al. 2007, *Rotational periods of solar-mass young stars in Orion*, \aap, 463, 1081, doi: [10.1051/0004-6361:20066458](https://doi.org/10.1051/0004-6361:20066458)
- Mathieu, R. D., & Mazeh, T. 1988, *The Circularized Binaries in Open Clusters: A New Clock for Age Determination*, Astrophys. J., 326, 256, doi: [10.1086/166087](https://doi.org/10.1086/166087)
- Mathieu, R. D., Meibom, S., & Dolan, C. J. 2004, *The Tidal Circularization Cutoff Period of the Old Open Cluster NGC188*, The Astrophysical Journal, 602, L121, doi: [10.1086/382686](https://doi.org/10.1086/382686)
- Mathur, S., Huber, D., Batalha, N. M., et al. 2017, *Revised Stellar Properties of Kepler Targets for the Q1-17 (DR25) Transit Detection Run*, The Astrophysical Journal Supplement Series, 229, 30, doi: [10.3847/1538-4365/229/2/30](https://doi.org/10.3847/1538-4365/229/2/30)
- Matson, R. A., Gies, D. R., Guo, Z., & Orosz, J. A. 2016, *FUNDAMENTAL PARAMETERS OF KEPLER ECLIPSING BINARIES. I. KIC 5738698*, The Astronomical Journal, 151, 139, doi: [10.3847/0004-6256/151/6/139](https://doi.org/10.3847/0004-6256/151/6/139)
- Matt, S. P., Brun, A. S., Baraffe, I., Bouvier, J., & Chabrier, G. 2015, *The Mass-dependence of Angular Momentum Evolution in Sun-like Stars*, \apjl, 799, L23, doi: [10.1088/2041-8205/799/2/L23](https://doi.org/10.1088/2041-8205/799/2/L23)
- . 2019, *Erratum: “The Mass-dependence of Angular Momentum Evolution in Sun-like Stars”* (2015, ApJL, 799, L23), \apjl, 870, L27, doi: [10.3847/2041-8213/aaf1b](https://doi.org/10.3847/2041-8213/aaf1b)
- Matt, S. P., MacGregor, K. B., Pinsonneault, M. H., & Greene, T. P. 2012, *MAGNETIC BRAKING FORMULATION FOR SUN-LIKE STARS: DEPENDENCE ON DIPOLE FIELD STRENGTH AND ROTATION RATE*, The Astrophysical Journal Letters, 754, L26, doi: [10.1088/2041-8205/754/2/L26](https://doi.org/10.1088/2041-8205/754/2/L26)
- Mayor, M., & Mermilliod, J. C. 1984, in *Observational Tests of the Stellar Evolution Theory*, 411
- Mazeh, T. 2008, *Observational Evidence for Tidal Interaction in Close Binary Systems*, EAS Publications Series, 29, 1, doi: [10.1051/eas:0829001](https://doi.org/10.1051/eas:0829001)
- McQuillan, A., Mazeh, T., & Aigrain, S. 2014, *Rotation Periods of 34,030 Kepler Main-sequence Stars: The Full Autocorrelation Sample*, \apjs, 211, 24, doi: [10.1088/0067-0049/211/2/24](https://doi.org/10.1088/0067-0049/211/2/24)
- Meibom, S., Barnes, S. A., Platais, I., et al. 2015, *A spin-down clock for cool stars from observations of a 2.5-billion-year-old cluster*, \nat, 517, 589, doi: [10.1038/nature14118](https://doi.org/10.1038/nature14118)
- Meibom, S., & Mathieu, R. D. 2005, *A Robust Measure of Tidal Circularization in Coeval Binary Populations: The Solar-Type Spectroscopic Binary Population in the Open Cluster M35*, \apj, 620, 970, doi: [10.1086/427082](https://doi.org/10.1086/427082)

- Meibom, S., Mathieu, R. D., & Stassun, K. G. 2006, *An Observational Study of Tidal Synchronization in Solar-Type Binary Stars in the Open Clusters M35 and M34*, *\apj*, 653, 621, doi: [10.1086/508252](https://doi.org/10.1086/508252)
- Menezes, F., Araújo, A., & Valio, A. 2024, *Probing the magnetic fields of starspots with transit mapping*, *A&A*, 691, L12, doi: [10.1051/0004-6361/202452071](https://doi.org/10.1051/0004-6361/202452071)
- Messina, S., & Guinan, E. F. 1999, *Starspot Activity, Differential Rotation, and Activity Cycles on Young Solar-Type Stars*, *Astronomy and Astrophysics*, 347, 249
- Miesch, M. S., Brun, A. S., DeRosa, M. L., & Toomre, J. 2008, *Structure and Evolution of Giant Cells in Global Models of Solar Convection*, *The Astrophysical Journal*, 673, 557, doi: [10.1086/523838](https://doi.org/10.1086/523838)
- Miesch, M. S., Elliott, J. R., Toomre, J., et al. 2000, *Three-Dimensional Spherical Simulations of Solar Convection.I. Differential Rotation and Pattern Evolution Achieved with Laminar and Turbulent States*, *The Astrophysical Journal*, 532, 593, doi: [10.1086/308555](https://doi.org/10.1086/308555)
- Mirouh, G. M., Hendriks, D. D., Dykes, S., Moe, M., & Izzard, R. G. 2023, *Detailed equilibrium and dynamical tides: impact on circularization and synchronization in open clusters*, *Monthly Notices of the Royal Astronomical Society*, 524, 3978, doi: [10.1093/mnras/stad2048](https://doi.org/10.1093/mnras/stad2048)
- Moe, M., & Kratter, K. M. 2018, *Dynamical Formation of Close Binaries during the Pre-main-sequence Phase*, *The Astrophysical Journal*, 854, 44, doi: [10.3847/1538-4357/aaa6d2](https://doi.org/10.3847/1538-4357/aaa6d2)
- Morris, B. M., Curtis, J. L., Sakari, C., Hawley, S. L., & Agol, E. 2019a, *Stellar Properties of Active G and K Stars: Exploring the Connection between Starspots and Chromospheric Activity*, *AJ*, 158, 101, doi: [10.3847/1538-3881/ab2e04](https://doi.org/10.3847/1538-3881/ab2e04)
- Morris, B. M., Davenport, J. R. A., Giles, H. A. C., et al. 2019b, *The solar benchmark: rotational modulation of the Sun reconstructed from archival sunspot records*, *Monthly Notices of the Royal Astronomical Society*, 484, 3244, doi: [10.1093/mnras/stz199](https://doi.org/10.1093/mnras/stz199)
- Morris, B. M., Hebb, L., Davenport, J. R. A., Rohn, G., & Hawley, S. L. 2017, *The Starspots of HAT-P-11: Evidence for a Solar-like Dynamo*, *ApJ*, 846, 99, doi: [10.3847/1538-4357/aa8555](https://doi.org/10.3847/1538-4357/aa8555)
- Morton, T. D., & Winn, J. N. 2014, *OBLIQUITIES OF KEPLER STARS: COMPARISON OF SINGLE- AND MULTIPLE-TRANSIT SYSTEMS*, *The Astrophysical Journal*, 796, 47, doi: [10.1088/0004-637X/796/1/47](https://doi.org/10.1088/0004-637X/796/1/47)
- Murphy, K. P. 2012, *Machine learning: a probabilistic perspective*, *Adaptive computation and machine learning series* (Cambridge, MA: MIT Press)
- Naoz, S. 2016, *The Eccentric Kozai-Lidov Effect and Its Applications*, *Annual Review of Astronomy and Astrophysics*, 54, 441, doi: [10.1146/annurev-astro-081915-023315](https://doi.org/10.1146/annurev-astro-081915-023315)
- Nelder, J. A., & Mead, R. 1965, *A Simplex Method for Function Minimization*, *The Computer Journal*, 7, 308, doi: [10.1093/comjnl/7.4.308](https://doi.org/10.1093/comjnl/7.4.308)
- Netto, Y., & Valio, A. 2020, *Stellar magnetic activity and the butterfly diagram of Kepler-63*, *A&A*, 635, A78, doi: [10.1051/0004-6361/201936219](https://doi.org/10.1051/0004-6361/201936219)
- Nicholson, B. A., & Aigrain, S. 2022, *Quasi-periodic Gaussian processes for stellar activity: From physical to kernel parameters*, *Mon Not R Astron Soc*, 515, 5251, doi: [10.1093/mnras/stac2097](https://doi.org/10.1093/mnras/stac2097)
- Nielsen, M. B., Schunker, H., Gizon, L., & Ball, W. H. 2015, *Constraining differential rotation of Sun-like stars from asteroseismic and starspot rotation periods*, *A&A*, 582, A10, doi: [10.1051/0004-6361/201526615](https://doi.org/10.1051/0004-6361/201526615)
- Nogueira, F. 2014, *Bayesian Optimization: Open source constrained global optimization tool for Python*. <https://github.com/bayesian-optimization/BayesianOptimization>
- Noraz, Q., Breton, S. N., Brun, A. S., et al. 2022a, *Hunting for Anti-Solar Differentially Rotating Stars Using the Rossby Number - An Application to the Kepler Field*, *Astronomy & Astrophysics*, 667, A50, doi: [10.1051/0004-6361/202243890](https://doi.org/10.1051/0004-6361/202243890)
- Noraz, Q., Brun, A. S., Strugarek, A., & Depambour, G. 2022b, *Impact of anti-solar differential rotation in mean-field solar-type dynamos. Exploring possible magnetic cycles in slowly rotating stars*, *Astronomy and Astrophysics*, 658, A144, doi: [10.1051/0004-6361/202141957](https://doi.org/10.1051/0004-6361/202141957)
- . 2022c, *Impact of Anti-Solar Differential Rotation in Mean-Field Solar-Type Dynamos: Exploring Possible Magnetic Cycles in Slowly Rotating Stars*, *Astronomy & Astrophysics*, 658, A144, doi: [10.1051/0004-6361/202141946](https://doi.org/10.1051/0004-6361/202141946)
- Ogilvie, G. I. 2013, *Tides in rotating barotropic fluid bodies: the contribution of inertial waves and the role of internal structure*, *Monthly Notices of the Royal Astronomical Society*, 429, 613, doi: [10.1093/mnras/sts362](https://doi.org/10.1093/mnras/sts362)
- Oleskiewicz, P., & Baugh, C. M. 2019, *Sensitivity analysis of a galaxy formation model*, *arXiv:1910.01745 [astro-ph]*, doi: [10.1093/mnras/stz3560](https://doi.org/10.1093/mnras/stz3560)
- Parzen, E. 1962, *On Estimation of a Probability Density Function and Mode*, *The Annals of Mathematical Statistics*, 33, 1065
- Patel, R., & Penev, K. 2022, *Constraining tidal quality factor using spin period in eclipsing binaries*, *Mon. Not. R. Astron. Soc.*, 512, 3651, doi: [10.1093/mnras/stac203](https://doi.org/10.1093/mnras/stac203)
- Patel, R., & Penev, K. 2022, *Constraining Tidal Quality Factor using Spin Period in Eclipsing Binaries*, *Monthly Notices of the Royal Astronomical Society*, stac203, doi: [10.1093/mnras/stac203](https://doi.org/10.1093/mnras/stac203)
- Patra, K. C., Winn, J. N., Holman, M. J., et al. 2017, *The Apparently Decaying Orbit of WASP-12b*, *\aj*, 154, 4, doi: [10.3847/1538-3881/aa6d75](https://doi.org/10.3847/1538-3881/aa6d75)
- . 2020, *The Continuing Search for Evidence of Tidal Orbital Decay of Hot Jupiters*, *The Astronomical Journal*, 159, 150, doi: [10.3847/1538-3881/ab7374](https://doi.org/10.3847/1538-3881/ab7374)
- Paxton, B., Bildsten, L., Dotter, A., et al. 2010, *MODULES FOR EXPERIMENTS IN STELLAR ASTROPHYSICS (MESA)*, *The Astrophysical Journal Supplement Series*, 192, 3, doi: [10.1088/0067-0049/192/1/3](https://doi.org/10.1088/0067-0049/192/1/3)
- Peale, S. J. 1969, *Generalized Cassini's Laws*, *The Astronomical Journal*, 74, 483, doi: [10.1086/110825](https://doi.org/10.1086/110825)
- Penev, K., Bouma, L. G., Winn, J. N., & Hartman, J. D. 2018, *Empirical Tidal Dissipation in Exoplanet Hosts From Tidal Spin-up*, *The Astronomical Journal*, 155, 165, doi: [10.3847/1538-3881/aaaf71](https://doi.org/10.3847/1538-3881/aaaf71)
- Penev, K., Zhang, M., & Jackson, B. 2014, *POET: A Model for Planetary Orbital Evolution Due to Tides on Evolving Stars*, *Publications of the Astronomical Society of the Pacific*, 126, 553, doi: [10.1086/677042](https://doi.org/10.1086/677042)

- Penev, K. M., & Schussler, J. A. 2022, *Comprehensive Bayesian Modeling of Tidal Circularization in Open Cluster Binaries part I: M 35, NGC 6819, NGC 188*, Monthly Notices of the Royal Astronomical Society, stac2618, doi: [10.1093/mnras/stac2618](#)
- Perez-Cruz, F. 2008, in 2008 IEEE International Symposium on Information Theory, 1666–1670, doi: [10.1109/ISIT.2008.4595271](#)
- Piskunov, N. E., & Rice, J. B. 1993, *Spectral Line Profiles from the Surface of a Rotating Star*, Publications of the Astronomical Society of the Pacific, 105, 1415, doi: [10.1086/133330](#)
- Price-Whelan, A. M., & Goodman, J. 2018, *Binary companions of evolved stars in APOGEE DR14: Orbital circularization*, The Astrophysical Journal, 867, 5, doi: [10.3847/1538-4357/aae264](#)
- Price-Whelan, A. M., Hogg, D. W., Rix, H.-W., et al. 2020, *Close Binary Companions to APOGEE DR16 Stars: 20,000 Binary-star Systems Across the Color–Magnitude Diagram*, The Astrophysical Journal, 895, 2, doi: [10.3847/1538-4357/ab8acc](#)
- Rainer, M., Desidera, S., Borsa, F., et al. 2023, *The GAPS programme at TNG - XLIV. Projected rotational velocities of 273 exoplanet-host stars observed with HARPS-N*, A&A, 676, A90, doi: [10.1051/0004-6361/202245564](#)
- Rajaratnam, B., & Sparks, D. 2015, *MCMC-Based Inference in the Era of Big Data: A Fundamental Analysis of the Convergence Complexity of High-Dimensional Chains*. <https://arxiv.org/abs/1508.00947>
- Rao, C. R. 1945, *Information and the Accuracy Attainable in the Estimation of Statistical Parameters*, Bulletin of the Calcutta Mathematical Society, 37, 81
- Rasmussen, C. E., & Williams, C. K. I. 2006, *Gaussian Processes for Machine Learning*
- Rebull, L. M., Stauffer, J. R., Megeath, S. T., Hora, J. L., & Hartmann, L. 2006, *A Correlation between Pre-Main-Sequence Stellar Rotation Rates and IRAC Excesses in Orion*, *ApJ*, 646, 297, doi: [10.1086/504865](#)
- Reiners, A., & Mohanty, S. 2012, *Radius-dependent Angular Momentum Evolution in Low-mass Stars. I.*, *ApJ*, 746, 43, doi: [10.1088/0004-637X/746/1/43](#)
- Reiners, A., & Schmitt, J. H. M. M. 2002a, *On the feasibility of the detection of differential rotation in stellar absorption profiles*, *Astronomy and Astrophysics*, 384, 155, doi: [10.1051/0004-6361:20011801](#)
- . 2002b, *On the feasibility of the detection of differential rotation in stellar absorption profiles*, *Astronomy and Astrophysics*, 384, 155, doi: [10.1051/0004-6361:20011801](#)
- . 2003, *Rotation and differential rotation in field F- and G-type stars*, *Astronomy and Astrophysics*, 398, 647, doi: [10.1051/0004-6361:20021642](#)
- Reinhold, T., & Gizon, L. 2015, *Rotation, differential rotation, and gyrochronology of active Kepler stars*, *Astronomy and Astrophysics*, 583, A65, doi: [10.1051/0004-6361/201526216](#)
- Repetto, S., & Nelemans, G. 2014, *The coupled effect of tides and stellar winds on the evolution of compact binaries*, Monthly Notices of the Royal Astronomical Society, 444, 542, doi: [10.1093/mnras/stu1454](#)
- Rice, J. B. 2002, *Doppler imaging of stellar surfaces – techniques and issues*, *Astronomische Nachrichten*, 323, 220, doi: [10.1002/1521-3994\(200208\)323:3/4<220::AID-ASNA220>3.0.CO;2-P](#)
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in *Textbackslashprocspie*, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, 914320, doi: [10.1117/12.2063489](#)
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *Transiting Exoplanet Survey Satellite (TESS)*, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 014003, doi: [10.1117/1.JATIS.1.1.014003](#)
- Rosenblatt, M. 1956, *Remarks on Some Nonparametric Estimates of a Density Function*, *The Annals of Mathematical Statistics*, 27, 832
- Sagynbayeva, S., & Farr, W. M. 2026, *Polka-dotted Stars. II. Starspots and Obliquities of Kepler-17 and Kepler-63*, *ApJ*, 997, 297, doi: [10.3847/1538-4357/ae2bdb](#)
- Sagynbayeva, S., Farr, W. M., Morris, B., & Luger, R. 2025, *Polka-dotted Stars: a Hierarchical Model for Mapping Stellar Surfaces Using Occultation Light Curves and the Case of TOI-3884*, arXiv, doi: [10.48550/arXiv.2504.21852](#)
- Saltelli, A., Aleksankina, K., Becker, W., et al. 2017, *Why So Many Published Sensitivity Analyses Are False. A Systematic Review of Sensitivity Analysis Practices*, Tech. Rep. arXiv:1711.11359, arXiv, doi: [10.48550/arXiv.1711.11359](#)
- Saltelli, A., Annoni, P., Azzini, I., et al. 2010, *Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index*, *Computer Physics Communications*, 181, 259, doi: [10.1016/j.cpc.2009.09.018](#)
- Sanchis-Ojeda, R., Winn, J. N., Marcy, G. W., et al. 2013a, *KEPLER-63b: A GIANT PLANET IN A POLAR ORBIT AROUND A YOUNG SUN-LIKE STAR*, *ApJ*, 775, 54, doi: [10.1088/0004-637X/775/1/54](#)
- . 2013b, *Kepler-63b: A Giant Planet in a Polar Orbit around a Young Sun-like Star*, *The Astrophysical Journal*, 775, 54, doi: [10.1088/0004-637X/775/1/54](#)
- Santos, A. R. G., Cunha, M. S., Avelino, P. P., García, R. A., & Mathur, S. 2017, *Starspot signature on the light curve - Learning about the latitudinal distribution of spots*, *Astronomy & Astrophysics*, 599, A1, doi: [10.1051/0004-6361/201629923](#)
- Saunders, N., van Saders, J. L., Lyttle, A. J., et al. 2024, *Stellar Cruise Control: Weakened Magnetic Braking Leads to Sustained Rapid Rotation of Old Stars*, *The Astrophysical Journal*, 962, 138, doi: [10.3847/1538-4357/ad1516](#)
- Schou, J., Antia, H. M., Basu, S., et al. 1998, *Helioseismic Studies of Differential Rotation in the Solar Envelope by the Solar Oscillations Investigation Using the Michelson Doppler Imager*, *The Astrophysical Journal*, 505, 390, doi: [10.1086/306146](#)
- Schunker, H., Schou, J., Ball, W., Nielsen, M. B., & Gizon, L. 2016, *Asteroseismic inversions for radial differential rotation of Sun-like stars: ensemble fits*, *A&A*, 586, A79, doi: [10.1051/0004-6361/201527485](#)
- Schwabe, H., & Schwabe, H. 1844, *Sonnen — Beobachtungen im Jahre 1843*, *Astronomische Nachrichten*, doi: [10.1002/asna.18440211505](#)
- Sills, A., Pinsonneault, M. H., & Terndrup, D. M. 2000, *The Angular Momentum Evolution of Very Low Mass Stars*, *The Astrophysical Journal*, 534, 335, doi: [10.1086/308739](#)

- Skilling, J. 2004, in AIP Conference Proceedings, Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop, 395–405, doi: [10.1063/1.1835238](https://doi.org/10.1063/1.1835238)
- Skumanich, A. 1972, *Time Scales for CA II Emission Decay, Rotational Braking, and Lithium Depletion*, *\apj*, 171, 565, doi: [10.1086/151310](https://doi.org/10.1086/151310)
- Sobol, I. M. 2001, *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*, *Mathematics and Computers in Simulation*, 55, 271, doi: [10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6)
- Soderblom, D. R. 1983, *Rotational studies of late-type stars. II. Ages of solar-type stars and the rotational history of the Sun.*, *Astrophys. J. Supp.*, 53, 1, doi: [10.1086/190880](https://doi.org/10.1086/190880)
- Soderblom, D. R. 2010, *The Ages of Stars*, *Annual Review of Astronomy and Astrophysics*, 48, 581, doi: [10.1146/annurev-astro-081309-130806](https://doi.org/10.1146/annurev-astro-081309-130806)
- Song, H. F., Meynet, G., Maeder, A., et al. 2018, *Close binary evolution. III. Impact of tides, wind magnetic braking, and internal angular momentum transport*, *Astronomy & Astrophysics*, 609, A3, doi: [10.1051/0004-6361/201731073](https://doi.org/10.1051/0004-6361/201731073)
- Southworth, J., & Clausen, J. V. 2006, *Eclipsing Binary Stars in Open Clusters*, arXiv:astro-ph/0608016. <http://arxiv.org/abs/astro-ph/0608016>
- Speagle, J. S. 2020a, *A Conceptual Introduction to Markov Chain Monte Carlo Methods*, arXiv. <http://arxiv.org/abs/1909.12313>
- . 2020b, *dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences*, *Monthly Notices of the Royal Astronomical Society*, 493, 3132, doi: [10.1093/mnras/staa278](https://doi.org/10.1093/mnras/staa278)
- Stassun, K. G., Mathieu, R. D., Mazeh, T., & Vrba, F. J. 1999, *The Rotation Period Distribution of Pre-Main-Sequence Stars in and around the Orion Nebula*, *\aj*, 117, 2941, doi: [10.1086/300881](https://doi.org/10.1086/300881)
- Sterzik, M. F., Durisen, R. H., & Zinnecker, H. 2003, *How do binary separations depend on cloud initial conditions?*, *Astronomy & Astrophysics*, 411, 91, doi: [10.1051/0004-6361:20034219](https://doi.org/10.1051/0004-6361:20034219)
- Stone, M. 1974, *Cross-Validatory Choice and Assessment of Statistical Predictions*, *Journal of the Royal Statistical Society: Series B (Methodological)*, 36, 111, doi: [10.1111/j.2517-6161.1974.tb00994.x](https://doi.org/10.1111/j.2517-6161.1974.tb00994.x)
- Strassmeier, K. G. 2009a, *Starspots*, *The Astronomy and Astrophysics Review*, 17, 251, doi: [10.1007/s00159-009-0020-6](https://doi.org/10.1007/s00159-009-0020-6)
- . 2009b, *Starspots*, *The Astronomy and Astrophysics Review*, 17, 251, doi: [10.1007/s00159-009-0020-6](https://doi.org/10.1007/s00159-009-0020-6)
- Strassmeier, K. G., & Bopp, B. W. 1992, *Time-series photometric spot modelling. I. Parametric fits to observed light curves*, *Astronomy and Astrophysics*, 259, 183
- Swersky, K. 2017. <https://api.semanticscholar.org/CorpusID:64999289>
- Terquem, C., Papaloizou, J. C. B., Nelson, R. P., & Lin, D. N. C. 1998, *On the Tidal Interaction of a Solar-Type Star with an Orbiting Companion: Excitation of g-Mode Oscillation and Orbital Evolution*, *The Astrophysical Journal*, 502, 788, doi: [10.1086/305927](https://doi.org/10.1086/305927)
- Tokovinin, A. 2021, *Architecture of Hierarchical Stellar Systems and Their Formation*, *Universe*, 7, 352, doi: [10.3390/universe7090352](https://doi.org/10.3390/universe7090352)
- Tokovinin, A., & Moe, M. 2020, *Formation of close binaries by disc fragmentation and migration, and its statistical modelling*, *Monthly Notices of the Royal Astronomical Society*, 491, 5158, doi: [10.1093/mnras/stz3299](https://doi.org/10.1093/mnras/stz3299)
- Tokuno, T., Suzuki, T. K., & Shoda, M. 2023, *Transition of latitudinal differential rotation as a possible cause of weakened magnetic braking of solar-type stars*, *Monthly Notices of the Royal Astronomical Society*, 520, 418, doi: [10.1093/mnras/stad103](https://doi.org/10.1093/mnras/stad103)
- Toonen, S., Hamers, A., & Portegies Zwart, S. 2016, *The evolution of hierarchical triple star-systems*, *Computational Astrophysics and Cosmology*, 3, 6, doi: [10.1186/s40668-016-0019-0](https://doi.org/10.1186/s40668-016-0019-0)
- Torres, G., Curtis, J. L., Vanderburg, A., Kraus, A. L., & Rizzuto, A. 2018, *Eclipsing binaries in the open cluster Ruprecht 147. I: EPIC 219394517*, *The Astrophysical Journal*, 866, 67, doi: [10.3847/1538-4357/aadca8](https://doi.org/10.3847/1538-4357/aadca8)
- Triaud, A. H. M. J. 2018, 1375–1401, doi: [10.1007/978-3-319-55333-7_2](https://doi.org/10.1007/978-3-319-55333-7_2)
- van Saders, J. L., Ceillier, T., Metcalfe, T. S., et al. 2016, *Weakened magnetic braking as the origin of anomalously rapid rotation in old field stars*, *Nature*, 529, 181, doi: [10.1038/nature16168](https://doi.org/10.1038/nature16168)
- van Saders, J. L., & Pinsonneault, M. H. 2013, *Fast Star, Slow Star: Old Star, Young Star: Subgiant Rotation as a Population and Stellar Physics Diagnostic*, *The Astrophysical Journal*, 776, 67, doi: [10.1088/0004-637X/776/2/67](https://doi.org/10.1088/0004-637X/776/2/67)
- Vasil, G. M., Lecoanet, D., Augustson, K., et al. 2024, *The solar dynamo begins near the surface*, *Nature*, 629, 769, doi: [10.1038/s41586-024-07315-1](https://doi.org/10.1038/s41586-024-07315-1)
- Vidal, J., & Barker, A. J. 2020, *Turbulent Viscosity Acting on the Equilibrium Tidal Flow in Convective Stars*, *The Astrophysical Journal*, 888, L31, doi: [10.3847/2041-8213/ab6219](https://doi.org/10.3847/2041-8213/ab6219)
- Vissapragada, S., Chontos, A., Greklek-McKeon, M., et al. 2022, *The Possible Tidal Demise of Kepler’s First Planetary System*, *The Astrophysical Journal Letters*, 941, L31, doi: [10.3847/2041-8213/aca47e](https://doi.org/10.3847/2041-8213/aca47e)
- Vogt, S. S., & Penrod, G. D. 1983, *Doppler Imaging of Spotted Stars – Application to the RS Canum Venaticorum Star HR 1099*, *Publications of the Astronomical Society of the Pacific*, 95, 565, doi: [10.1086/131208](https://doi.org/10.1086/131208)
- Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, *Global hierarchical collapse in molecular clouds. Towards a comprehensive scenario*, *Monthly Notices of the Royal Astronomical Society*, 490, 3061, doi: [10.1093/mnras/stz2736](https://doi.org/10.1093/mnras/stz2736)
- Walker, G. A. H., Croll, B., Kuschnig, R., et al. 2007, *The Differential Rotation of κ^1 Ceti as Observed by MOST*, *The Astrophysical Journal*, 659, 1611, doi: [10.1086/511851](https://doi.org/10.1086/511851)
- Walkowicz, L. M., Basri, G., & Valenti, J. A. 2013, *The Information Content in Analytic Spot Models of Broadband Precision Light Curves*, *The Astrophysical Journal Supplement Series*, 205, 17, doi: [10.1088/0067-0049/205/2/17](https://doi.org/10.1088/0067-0049/205/2/17)
- Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, *The NumPy Array: A Structure for Efficient Numerical Computation*, *Computing in Science Engineering*, 13, 22, doi: [10.1109/MCSE.2011.37](https://doi.org/10.1109/MCSE.2011.37)
- Wang, H., & Li, J. 2018, *Adaptive Gaussian Process Approximation for Bayesian Inference with Expensive Likelihood Functions*, *Neural Computation*, 30, 3072, doi: [10.1162/neco_a_01127](https://doi.org/10.1162/neco_a_01127)

- Wang, J., Pan, Y., Fu, J., et al. 2024, *Lifetime of starspots on detached eclipsing binaries: Detecting the effects of tides on stellar activity*, *Astronomy & Astrophysics*, 690, A201, doi: [10.1051/0004-6361/202449484](https://doi.org/10.1051/0004-6361/202449484)
- Weber, E. J., & Davis, Jr., L. 1967, *The Angular Momentum of the Solar Wind*, *The Astrophysical Journal*, 148, 217, doi: [10.1086/149138](https://doi.org/10.1086/149138)
- Willems, J. C. 2000, *Modelling dynamical systems using manifest and latent variables*, *Mathematics and Computers in Simulation*, 53, 227, doi: [10.1016/S0378-4754\(00\)00209-3](https://doi.org/10.1016/S0378-4754(00)00209-3)
- Windemuth, D., Agol, E., Ali, A., & Kiefer, F. 2019, *Modeling Kepler Eclipsing Binaries: Homogeneous Inference of Orbital & Stellar Properties*, *Monthly Notices of the Royal Astronomical Society*, 489, 1644, doi: [10.1093/mnras/stz2137](https://doi.org/10.1093/mnras/stz2137)
- Witte, M. G., & Savonije, G. J. 1999, *Tidal evolution of eccentric orbits in massive binary systems. A study of resonance locking*, *Astron. & Astrophys.*, 350, 129. <https://arxiv.org/abs/astro-ph/9909073>
- Witte, M. G., & Savonije, G. J. 2002, *Orbital evolution by dynamical tides in solar type stars: Application to binary stars and planetary orbits*, *Astronomy & Astrophysics*, 386, 222, doi: [10.1051/0004-6361:20020155](https://doi.org/10.1051/0004-6361:20020155)
- Xiang, Y., Gu, S., Cameron, A. C., et al. 2020, *Doppler imaging and differential rotation of σ^2 Coronae Borealis using SONG*, *ApJ*, 893, 164, doi: [10.3847/1538-4357/ab8229](https://doi.org/10.3847/1538-4357/ab8229)
- Yee, S. W., Winn, J. N., Knutson, H. A., et al. 2019, *The Orbit of WASP-12b Is Decaying*, *The Astrophysical Journal Letters*, 888, L5, doi: [10.3847/2041-8213/ab5c16](https://doi.org/10.3847/2041-8213/ab5c16)
- Yi, Z., Wei, Y., Cheng, C. X., He, K., & Sui, Y. 2024, *Improving sample efficiency of high dimensional Bayesian optimization with MCMC*. <https://arxiv.org/abs/2401.02650>
- Zahn, J.-P. 1975, *The dynamical tide in close binaries*, *\aap*, 41, 329
- . 2008, *Tidal dissipation in binary systems*, *EAS Publications Series*, 29, 67, doi: [10.1051/eas:0829002](https://doi.org/10.1051/eas:0829002)
- Zahn, J.-P., & Bouchet, L. 1989, *Tidal evolution of close binary stars. II - Orbital circularization of late-type binaries*, *\aap*, 223, 112
- Zanazzi, J. J., & Wu, Y. 2021, *Tidal Circularization of Binaries by Resonance Locking. I. The Importance of the Pre-main Sequence*, *The Astronomical Journal*, 161, 263, doi: [10.3847/1538-3881/abf097](https://doi.org/10.3847/1538-3881/abf097)