

SEMI-PARAMETRIC METHODS TO AID IN THE DETECTION AND CHARACTERIZATION  
OF DISTANT WORLDS AROUND SMALL STARS

by

Ryan Cloutier

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Astronomy & Astrophysics  
University of Toronto

# Abstract

Semi-parametric methods to aid in the detection and characterization of distant worlds around small stars

Ryan Cloutier

Doctor of Philosophy

Graduate Department of Astronomy & Astrophysics

University of Toronto

2019

Statistical studies of the exoplanet population have provided key insights into their formation histories and evolution. Developing an understanding of the planet formation process requires both accurate and precise measurements of fundamental planetary parameters such as their masses, radii, and orbital characteristics. Planetary systems orbiting low mass stars represent superlative opportunities to characterize exoplanets in detail as they are plentiful within the solar neighbourhood and they are known to commonly host multi-planet systems of terrestrial to Neptune-sized planets that are more easily detectable than similar planets around Sun-like stars. One major deterrent to the characterization of such planets is stellar activity and its manifestation in radial velocity (RV) measurements that can mask and even mimic planetary signals.

In this thesis, I develop and implement a formalism aimed at enabling the detection and precise and accurate characterization of planetary systems around low mass stars. This formalism is based around semi-parametric Gaussian process (GP) regression models that are used to simultaneously model planets, in both RV and transit light curve observations, and the temporal covariance structure arising from stellar activity produced by magnetically active regions on the stellar surface. The GP formalism is applied to synthetic RV datasets emulating the upcoming planet detection survey using the near-infrared spectropolarimeter SPIRou and to synthetic optical and near-infrared measurements of the expected population of transiting planets discovered with TESS. I also apply the GP formalism to activity modelling in the K2-18 planetary system from which an accurate and precise planetary mass of its transiting temperate sub-Neptune is inferred along with the presence of an additional planet in the system using RV measurements from HARPS and CARMENES. Lastly, I extend the GP formalism to the treatment of stellar photometric variability and systematic effects in TESS light curves and uncover a number of candidate transiting planets around low mass stars in the first two TESS sectors.

Applications of GPs for the detection and characterization of exoplanets will prove to be a crucial tool

for developing a global understanding of planet formation and in revealing how common the conditions for life like our own are within our galaxy.

*To my wife and future children.*

## Acknowledgements

To each of my supervisors Kristen and René for their ongoing support over the years. Both professionally and monetarily. That support along with the wealth of opportunities that they provided for me to lead, collaborate, and present my work all over the world has been a continuous source of motivation without which I would never have reached this point. Thank you.

To Bob Abraham for his overly enthusiastic approach to teaching AST 121: the Origin and Evolution of the Universe. The energy in his lecturing and his obvious love for astronomy was enough to convince teenage Ryan to transfer into an astronomy major nearly a decade ago.

To the many astronomy grad students that I've had the pleasure of working with over the past five years, be it on research projects, shared TA duties, or through outreach activities like Astrotours. As I said, it's been a pleasure. Extra special thanks go out my past office mates for their company and (mostly) professional conduct at work. Extra extra special thanks go out the grad students with whom I shared a cohort: Alex, Dana, Epson, George, Nilou, Rob (sort of), Sasha, and Siqi. Thanks particularly for the many months of qualifying exam prep sessions. Without all of their support during those sessions I'm not entirely convinced that I would have gotten through it.

To the supporting cast of astrophysicists, geneticists, and other UofT grad students on all of the hockey, soccer, and softball teams. Thank you for helping me to balance my work and life with those late night games and playoff pushes.

To my friends outside of school: Brendon, Chris, Jesse, Lien, Sarah, and Solomon. Thank you for all of the random get-togethers for birthday bbqs, band practices, outdoor hockey games downtown, game nights, Caribbean cruises, my wedding, etc... In the years leading up to and continuing throughout this degree, these things have again helped me to maintain a certain balance that has undoubtedly helped me get to this point.

To my family especially, and for *many* more reasons than I can fit here. Specifically to my Mom for—as she says—giving me all of her physics brain cells that she didn't need back when she was in school. To my Dad for always fostering an environment in which learning and striving to be successful in school and in work was of the highest priority. And not to Chevonne.

Lastly to my wife, Alyssa. For cooking lots of my meals. For cleaning up after those meals. For taking care of me when I get sick and taking me to the hospital at three in the morning because my leg feels weird. For the breakaway passes. For continuing to put up with my brooding. For not waking me up after every dream you have that I do something wrong in. For already a year and a half of marriage! And for coming with me to Boston. I'm really looking forward to it.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Exoplanets . . . . .	1
1.2	M dwarf stars . . . . .	1
1.3	Methods of Detecting Exoplanets . . . . .	2
1.4	Exoplanet Detection via Stellar Radial Velocity . . . . .	5
1.4.1	Radial velocity curves . . . . .	5
1.4.2	Fundamental planet parameters from the radial velocity method . . . . .	7
1.4.3	Caveats to Eq. 1.11 . . . . .	8
1.5	Measuring Radial Velocities . . . . .	8
1.5.1	Stellar Spectroscopy . . . . .	8
1.5.2	Radial Velocity Error Budget Overview . . . . .	9
1.6	Point-form Thesis: Introduction . . . . .	20
<b>2</b>	<b>Stellar Activity Modelling in Radial Velocity Time Series</b>	<b>22</b>
2.1	An Overview of Techniques for Stellar Activity Mitigation . . . . .	22
2.1.1	Stellar Activity as a Scalar Parameter . . . . .	22
2.1.2	Correlations with Activity Indicators . . . . .	22
2.1.3	Photometric Modelling: the $\mathbf{FF}'$ Method . . . . .	25
2.1.4	Pre-whitening . . . . .	26
2.1.5	Deterministic Model Fitting . . . . .	26
2.1.6	Physical Models of Active Regions . . . . .	27
2.1.7	Line by line radial velocities . . . . .	27
2.2	Gaussian Process Regression for Semi-Parametric Activity Modelling . . . . .	28
2.2.1	The one-dimensional Gaussian distribution . . . . .	29
2.2.2	The two-dimensional Gaussian distribution . . . . .	29
2.2.3	The $k$ -dimensional Gaussian distribution . . . . .	33
2.2.4	Gaussian process regression modelling . . . . .	35
2.2.5	Demonstrating the effectiveness of a GP activity model: a toy model . . . . .	42
2.2.6	Gaussian process regression packages . . . . .	44
2.3	Point-form Thesis: Stellar Activity Modelling in Radial Velocity Time Series . . . . .	44
<b>3</b>	<b>Simulating the SPIRou Legacy Survey-Planet Search</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Un Spectro-Polarimètre Infra-Rouge . . . . .	47

3.3	Stellar Input Catalog . . . . .	47
3.3.1	Stellar Sample . . . . .	47
3.3.2	Physical Models of Stellar RV Activity . . . . .	49
3.4	Simulated Survey . . . . .	55
3.4.1	Window Functions . . . . .	55
3.4.2	Radial Velocity Time Series Construction . . . . .	57
3.5	Planet Sample . . . . .	59
3.5.1	Planet Sample from known Occurrence rates . . . . .	59
3.5.2	Modifications to the Planet Sample . . . . .	60
3.6	Activity Mitigation . . . . .	63
3.6.1	Overview of the Gaussian Process Formalism . . . . .	63
3.6.2	Modelling RV Activity with GP Regression . . . . .	66
3.6.3	GP Activity Model Performance . . . . .	67
3.7	Automated Planet Detection . . . . .	70
3.7.1	Establishing Putative Planetary Detections . . . . .	70
3.7.2	Model Selection . . . . .	73
3.7.3	Vetting of Planet Detections . . . . .	73
3.7.4	Summary of the Automated Planet Detection Algorithm . . . . .	74
3.8	SLS-PS Sensitivity . . . . .	75
3.8.1	Detection Sensitivity to HZ Planets . . . . .	77
3.9	SLS-PS Predicted Yield . . . . .	78
3.9.1	Giant Planet Detections . . . . .	80
3.10	The effect of an increased planet frequency around late M dwarfs . . . . .	80
3.11	Measuring the RV Planet Frequency . . . . .	83
3.11.1	Recovering planet frequency . . . . .	83
3.11.2	Measuring $\eta_{\oplus}$ . . . . .	85
3.12	Direct Imaging of Nearby Planetary Systems . . . . .	86
3.13	Comparison of Different Versions of the SLS-PS . . . . .	88
3.13.1	Optimized: the optimal survey strategy . . . . .	89
3.13.2	Closest: the closest M dwarfs . . . . .	90
3.13.3	Large: lots of stars . . . . .	90
3.13.4	Short: the effect of fewer observations . . . . .	90
3.13.5	Degraded: a degraded RV measurement precision . . . . .	90
3.13.6	Dark: no window function restrictions . . . . .	91
3.14	Appendix . . . . .	91
3.14.1	Computing $\sigma_K$ from the Fisher Information Matrix . . . . .	91
4	<b>Radial velocity follow-up of TESS discoveries</b>	95
4.1	Introduction . . . . .	95
4.2	Modelling the total observing time required to measure a transiting planet's mass . . . . .	96
4.2.1	Calculating $\sigma_K$ from the Fisher Information . . . . .	96
4.2.2	Calculating $\sigma_{RV}$ for TOIs in the photon-noise limit . . . . .	99
4.2.3	Additional sources of RV noise . . . . .	101
4.2.4	Estimating RV noise due to stellar activity . . . . .	102

4.2.5	Estimating RV noise due to unseen planets . . . . .	104
4.2.6	Exposure time calculator . . . . .	105
4.3	Model Comparison to Observations . . . . .	105
4.4	Overview of computing $N_{\text{RV}}$ for the expected TESS planet population . . . . .	108
4.5	Results for the TESS sample . . . . .	111
4.5.1	Detecting all TESS planet masses at $3\sigma$ . . . . .	111
4.5.2	Science Case 1: mass characterization of 50 TESS planets with $r_p < 4 R_{\oplus}$ . . . . .	115
4.5.3	Science Case 2: informing the mass-radius relation of planets across the radius valley	120
4.5.4	Science Case 3: characterization of temperate Earths & super-Earths . . . . .	120
4.5.5	Science Case 4: characterization of favorable JWST follow-up targets . . . . .	123
4.6	Discussion and Conclusions . . . . .	124
4.7	Appendix . . . . .	128
4.7.1	Radial Velocity Follow-up Calculator . . . . .	128
4.7.2	Fisher information with a quasi-periodic Gaussian process regression model . . . . .	130
<b>5</b>	<b>RV characterization of the K2-18 planetary system</b> . . . . .	<b>134</b>
5.1	Introduction . . . . .	134
5.2	Observations . . . . .	135
5.2.1	HARPS spectra . . . . .	135
5.2.2	K2 photometry . . . . .	136
5.3	Periodogram analysis . . . . .	137
5.4	Joint modelling of planets and correlated RV activity . . . . .	139
5.4.1	Training the GP activity model on ancillary time series . . . . .	139
5.4.2	Joint modelling of RVs . . . . .	142
5.5	Results . . . . .	142
5.5.1	Results from RV Data Analysis . . . . .	142
5.5.2	RV model comparison . . . . .	144
5.6	Searching for transits of K2-18c . . . . .	146
5.7	Dynamical stability and eccentricity restrictions . . . . .	149
5.8	Discussion . . . . .	152
<b>6</b>	<b>Reconciling the planetary interpretation of K2-18c</b> . . . . .	<b>157</b>
6.1	Introduction . . . . .	157
6.2	Sub-optimal window functions . . . . .	158
6.2.1	Detecting the 9-day signal in synthetic RV timeseries . . . . .	159
6.2.2	Identifying anomalous CARMENES observations . . . . .	162
6.3	Chromatic dependence of the 9-day signal with HARPS . . . . .	164
6.4	Temporal dependence of the 9-day signal with HARPS . . . . .	165
6.5	Simultaneous RV modelling of planets and correlated noise . . . . .	167
6.5.1	CARMENES RVs . . . . .	170
6.5.2	All HARPS RVs . . . . .	170
6.5.3	Joint HARPS+CARMENES RVs . . . . .	171
6.5.4	Overlapping HARPS and CARMENES window functions . . . . .	172
6.5.5	Model comparison . . . . .	172

6.6	Discussion and conclusions . . . . .	174
6.6.1	Improved stellar parameters based on Gaia DR2 . . . . .	175
6.6.2	Precise planetary parameters . . . . .	177
6.7	Appendix . . . . .	178
6.7.1	Iterative radial velocity timeseries and GLSP figures from Sect. 6.5 . . . . .	178
<b>7</b>	<b>The search for transiting M dwarf planets with TESS</b>	<b>185</b>
7.1	Introduction . . . . .	185
7.2	Stellar Sample . . . . .	187
7.2.1	Initial stellar sample . . . . .	187
7.2.2	Refined stellar sample based on Gaia DR2 . . . . .	187
7.3	Overview of the ORION transit detection pipeline . . . . .	192
7.3.1	TESS light curve acquisition . . . . .	193
7.3.2	Initial light curve de-trending . . . . .	193
7.3.3	Linear transit search . . . . .	196
7.3.4	Periodic transit search . . . . .	202
7.3.5	Automated planet vetting . . . . .	206
7.3.6	Joint systematic plus transiting planet modelling . . . . .	209
7.4	ORION planet search around low mass stars in TESS sectors 1 & 2 . . . . .	210
7.4.1	Manual vetting of ORION planet candidates . . . . .	212
7.4.2	Statistical validation of transiting planet candidates . . . . .	213
7.4.3	Querying Gaia sources to supplement statistical validation calculations . . . . .	216
7.4.4	Population of planet candidates . . . . .	218
7.5	Discussion . . . . .	224
7.5.1	Outstanding vetting criteria . . . . .	224
7.5.2	Discussions of individual systems with planet candidates . . . . .	224
7.5.3	TOIs undetected by ORION . . . . .	226
7.5.4	Prospects for follow-up observations . . . . .	227
7.5.5	Comparison to yield simulations . . . . .	231
<b>8</b>	<b>Conclusions and Future Work</b>	<b>233</b>
8.1	On the Accuracy of the SLS-PS Planet Yield Predictions . . . . .	233
8.2	Foreseen Improvements to the RVFC . . . . .	234
8.3	Improving our Physical Understanding of M dwarf Stellar Activity . . . . .	235
8.4	Detailing the Accuracy of RV Planet Models in the Presence of a GP . . . . .	236
8.5	Characterizing the Performance of ORION . . . . .	236
<b>9</b>	<b>Appendix</b>	<b>238</b>
9.1	Prospects for Detecting the Rossiter-McLaughlin Effect of Earth-like Planets: the test case of TRAPPIST-1b and c ( <a href="#">Cloutier &amp; Triaud, 2016</a> ) . . . . .	238
9.2	On the Radial Velocity Detection of Additional Planets in Transiting, Slowly Rotating M dwarf Systems: The Case of GJ 1132 ( <a href="#">Cloutier et al., 2017a</a> ) . . . . .	239
9.3	Near-InfraRed Planet Searcher to Join HARPS on the ESO 3.6-metre Telescope ( <a href="#">Bouchy et al., 2017</a> ) . . . . .	239

9.4	Radial velocity planet evidence . . . . .	240
9.5	Radial Velocity Follow-up of GJ 1132 with HARPS: a Precise Mass for Planet ‘b’ and the Discovery of a Second Planet ( <a href="#">Bonfils et al., 2018</a> ) . . . . .	241
9.6	A Second Terrestrial Planet Orbiting the Nearby M Dwarf LHS 1140 ( <a href="#">Ment et al., 2019</a> ) . . . . .	241
9.7	Three terrestrial planets around L 98-59 . . . . .	241
	<b>Bibliography</b>	<b>242</b>

# List of Tables

1.1	Summary of Radial Velocity Stellar Activity Sources . . . . .	14
3.1	Gaussian Process Hyperparameter Priors Used in Training . . . . .	67
3.2	Overview of SLS-PS Versions . . . . .	94
4.1	Adopted Spectral Bands . . . . .	101
4.2	Summary of RV Observations for Known Transiting Planets with White RV noise . . . . .	107
4.3	Summary of RV Observations for Known Transiting Planets with Red RV noise . . . . .	107
4.4	Fiducial Spectrograph Specifications . . . . .	110
4.5	Descriptions of RVFC Input Parameters . . . . .	129
4.6	Stellar parameters from the <a href="#">S15</a> synthetic catalog . . . . .	132
4.7	Median radial velocity noise sources and follow-up calculations for $3\sigma$ planet mass detections of the <a href="#">S15</a> synthetic catalog . . . . .	133
5.1	Summary of RV models and adopted priors . . . . .	143
5.2	Model parameters of the K2-18 planetary system . . . . .	155
5.3	HARPS time series . . . . .	156
6.1	Summary of the RV model parameter priors used for all models throughout this study .	169
6.2	Marginal likelihood estimations and Bayes factors for various RV datasets and models .	174
6.3	K2-18 model parameters from the HARPS+CARMENES joint RV analysis . . . . .	183
6.4	Full HARPS time series from <a href="#">C17a</a> and this work . . . . .	184
7.1	Model parameter priors . . . . .	195
7.2	Descriptions of the free parameters controlling the performance of <b>ORION</b> . . . . .	199
7.3	<b>vespa</b> input parameters and inferred probabilities of transiting planet and astrophysical false positive models for our 24 objects of interest . . . . .	215
7.4	Planetary parameters for our 16 vetted candidates . . . . .	220
7.5	Metric values indicating the feasibility of a variety of follow-up programs for our 16 vetted candidates . . . . .	228

# List of Figures

1.1	Methods of exoplanet detection and characterization.	3
1.2	Animation of observed radial velocities for different orbital configurations.	6
1.3	Near-IR spectra of early-to-late M dwarfs.	10
1.4	Image of spectral orders from the HARPS echelle spectrograph.	11
1.5	Doppler information content over spectral band passes and stellar properties.	13
1.6	Atmospheric transmission from 0.3-5 $\mu\text{m}$ .	14
1.7	Close-in view of the convection cells on the solar surface.	16
1.8	Illustration of the effect of active regions on the integrated stellar line profile.	17
1.9	M dwarf activity fraction versus spectral type and stellar rotation.	18
1.10	Empirical distribution of M dwarf photometric rotation periods.	19
2.1	Correlation between stellar RVs and the $V_{\text{span}}$ activity indicator.	23
2.2	Illustrations of CCF shape parameters.	24
2.3	Illustration of the $FF'$ method.	26
2.4	Principal component analysis of idealized spectra containing either spots, faculae, or planets.	28
2.5	Gaussian random variable in one dimension.	30
2.6	Uncorrelated Gaussian random variables in two dimensions.	31
2.7	Correlated Gaussian random variables in two dimensions.	32
2.8	Linear representation of Gaussian random variables in two dimensions.	34
2.9	Linear representation of Gaussian random variables in five dimensions.	34
2.10	Example of covariance matrices.	36
2.11	Behaviour of GPs with different covariance kernels and hyperparameters.	38
2.12	Samples from a quasi-periodic Gaussian process prior distribution.	40
2.13	Samples from a quasi-periodic Gaussian process predictive distribution.	41
2.14	Synthetic RV time series used to demonstrate the effectiveness of GP activity modelling.	43
2.15	Comparison of injected to recovered RV signals.	44
3.1	Stellar parameters of sources targetted in the simulated <i>SPIRou Legacy Survey-Planet Search</i> .	48
3.2	M dwarf rotation period distributions.	50
3.3	RV activity rms versus stellar rotation period.	54
3.4	Lomb-Scargle periodogram of an expected SPIRou window function during the SLS-PS.	57
3.5	RV measurement uncertainty versus $J$ -band magnitude and $v \sin i_s$ .	59
3.6	Masses and radii of synthetic planets in the simulated SLS-PS.	61

3.7	Input planet occurrence rates in the simulated SLS-PS. . . . .	64
3.8	Histogram of planetary system multiplicities in the simulated SLS-PS. . . . .	65
3.9	Examples of GP activity models in simulated SPIRou timeseries. . . . .	68
3.10	GP activity model performance in simulated SPIRou RV timeseries. . . . .	69
3.11	Illustrating the iterative planet search with Lomb-Scargle periodograms. . . . .	71
3.12	SPIRou planet yields versus $m_p \sin i$ throughout the vetting process. . . . .	75
3.13	Detection sensitivity maps in the SLS-PS. . . . .	77
3.14	Planet yield maps from the SLS-PS. . . . .	79
3.15	One realization of targetted and detected planets in the SLS-PS versus insolation and $m_p \sin i$ . . . . .	80
3.16	Planet yield results in the SLS-PS with modified input occurrence rates. . . . .	82
3.17	Planet occurrence rate maps derived from the resuls of the SLS-PS. . . . .	85
3.18	Expected measurement of $\eta_{\oplus}$ from the SLS-PS. . . . .	86
3.19	One realization of targetted and detected planets in the SLS-PS versus angular separation and reflected light contrast. . . . .	88
4.1	Observed $N_{RV}$ versus expected $N_{RV}$ as calculated using our formalism. . . . .	108
4.2	Histograms of the expected photon-noise limited RV precision for TOIs. . . . .	109
4.3	Comparing the optical and near-IR RV photon-noise limits for TOIs. . . . .	112
4.4	Cumulative mean observing times to measure RV masses of TOI planets versus spectral type. . . . .	114
4.5	Cumulative mean observing times to measure RV masses of TOI planets versus planet type. . . . .	115
4.6	Cumulative mean observing time required to complete the TESS level one science requirement. . . . .	117
4.7	Identification of the ‘best’ RV follow-up targets from TESS. . . . .	119
4.8	Cumulative mean observing times to measure RV masses of TOI planets spanning the radius valley. . . . .	121
4.9	Cumulative mean observing times to measure RV masses of potentially habitable planets transiting TOIs. . . . .	122
4.10	Cumulative mean observing times to measure RV masses of TOI planets of interest for transmission spectroscopy follow-up with JWST. . . . .	125
5.1	K2 photometric light curve of K2-18. . . . .	137
5.2	Lomb-Scargle periodograms of the HARPS timeseries and K2 photometry. . . . .	138
5.3	Posterior PDF of the GP hyperparameters of the K2 photometry. . . . .	141
5.4	Posterior PDF of the model parameters of the HARPS RVs. . . . .	145
5.5	Various HARPS RV components modelled for K2-18. . . . .	147
5.6	Phase-folded HARPS RVs depicting the planetary signals from K2-18b and c. . . . .	148
5.7	Search for transits of the RV planet K2-18c in the K2 photometry. . . . .	149
5.8	Stability fraction versus the injected orbital eccentricities of K2-18b and c. . . . .	151
5.9	Planetary mass-radius space for sub-Neptune-sized planets including K2-18b. . . . .	153
6.1	The effect of sampling on the probability of the periodic signal at $P_c = 8.962$ days in synthetic RV timeseries. . . . .	161

6.2	The sensitivity of the probability of the periodic signal $P_c = 8.962$ days with different weighting schemes in HARPS and CARMENES.	162
6.3	Leave-one-out cross-validation on the CARMENES RV timeseries.	163
6.4	Bayesian Lomb-Scargle periodograms of the chromatic HARPS RVs.	166
6.5	Temporal variations in the probability of the periodic signal at $P_c = 8.962$ days and HARPS Na D timeseries over three observing seasons.	167
6.6	Marginalized posterior PDFs of $K_b$ and $K_c$ with HARPS and CARMENES.	176
6.7	RV timeseries and models from analysis of the K2-18 CARMENES RVs.	179
6.8	RV timeseries and models from analysis of the K2-18 HARPS RVs.	180
6.9	RV timeseries and models from analysis of the joint K2-18 RVs from HARPS and CARMENES.	181
6.10	RV timeseries and models from analysis of the joint K2-18 RVs from HARPS and CARMENES during their overlap between February and June 2017.	182
7.1	Stellar parameters for M dwarf TICs in TESS sectors 1 and 2.	188
7.2	Comparison of stellar radii derived from Gaia parallaxes to values in the TIC-7.	191
7.3	Examples of GP detrending of TESS light curves.	197
7.4	Results of the linear search for repeating transit-like events around TIC 234994474.	203
7.5	Results of the periodic search for repeating transit-like events around TIC 234994474.	204
7.6	Examples of flares detected in the TIC 25200252 light curve.	208
7.7	Summary output for TIC 234994474 from ORION.	211
7.8	Gaia sources around TICs with ORION planet candidates.	217
7.9	Example of scattered light within the common aperture of nearby TICs.	219
7.10	Phase-folded transit light curves of the ORION planet candidate sample.	222
7.11	Planetary parameters for the ORION planet candidate sample.	223
7.12	Stellar parameters for the ORION planet candidate sample host stars.	223
7.13	Follow-up diagnostics for the ORION planet candidate sample.	230
9.1	Measuring the sky-projected spin-orbit angle using the RM effect.	239

# Chapter 1

## Introduction: radial velocity studies of planets around small stars

### 1.1 Exoplanets

Since the first detection of an exo-solar planet around a main sequence star (Mayor & Queloz, 1995) the quest to find and study additional planets outside of our own solar system has been met with much enthusiasm and success. In recent years, numerous purpose-built instruments have been designed to detect exoplanets using a variety of methods each with their own unique sensitivities and biases. Combining the results of multiple observational methods provides a nearly global overview of the exoplanet population and thus provides insight into the dominant pathways of planet formation and towards an understanding of how common the conditions for life like our own are within our galaxy.

The results from many such dedicated exoplanet missions have been astounding. At the time of writing there were 3926 confirmed exoplanets<sup>1</sup> with many more thousands of planet candidates. Furthermore, there exists an extensive diversity among planetary systems including the existence of certain exotic subsets of exoplanets for which analogs do not exist within our own solar system. Some of these extreme systems include i) the so-called hot Jupiters (e.g. 51 Pegasi b; Mayor & Queloz 1995, HD 189733b; Bouchy et al. 2005, & HD 209458b; Mazeh et al. 1999; Charbonneau et al. 2000): gas giant planets with masses comparable to or exceeding the mass of Jupiter but with orbital periods of  $\lesssim 10$  days, ii) circumbinary planets (e.g. Kepler-16b; Doyle et al. 2011): planets that orbit stellar binaries, and iii) the most common planet types in the local universe (Petigura et al., 2013a), namely super-Earths and sub-Neptunes whose radii span  $\sim 1.5 - 4 R_{\oplus}$  and over a wide range of orbital separations (e.g. GJ 1214b; Charbonneau et al. 2009, Kepler 440b; Torres et al. 2015, & K2-18b; Foreman-Mackey et al. 2015; Montet et al. 2015).

---

<sup>1</sup><https://exoplanets.nasa.gov/>; accessed on April 1, 2019.

## 1.2 M dwarf stars

The majority of the work presented throughout this thesis focuses on planetary systems around M dwarf stars. M dwarfs embody the low mass end of the main sequence phase of stellar evolution and dominate the stellar initial mass function (Kroupa, 2001) which, when coupled with their extended main sequence lifetimes, causes M dwarfs to dominate the spectral type distribution throughout the galaxy while outnumbering Sun-like stars 4:1 in the solar neighbourhood (Winters et al., 2015). In addition to being plentiful within the nearby Universe, M dwarf stars are known to frequently host multi-planet systems of Earth to Neptune-sized planets (Dressing & Charbonneau, 2015; Gaidos et al., 2016) that are more easily detectable in radial velocity around M dwarfs than around Sun-like stars due to the their lower masses (see Sect. 1.4.2). What’s more is that the habitable zone around M dwarfs (Kasting et al., 1993; Kopparapu et al., 2013), where the probability of hosting potentially habitable planets is maximized due to the temperate planetary equilibrium temperatures there, exists at much smaller orbital separations than around Sun-like stars thus making the characterization of temperate M dwarf planets a much more tractable problem observationally.

A critical caveat to the aforementioned benefits of focusing on planet searches around M dwarfs is that many of these stars tend to be active (see Sect. 1.5.2). This fact not only deters the prospect of finding signs of life like us on M dwarf planets (e.g. Owen & Mohanty, 2016; O’Malley-James & Kaltenegger, 2017) but it also inhibits our ability to obtain accurate and precise empirical parameters of their planetary systems. Overcoming this problem is one of the primary topics of this thesis.

M dwarfs effective temperatures span a wide range of values:  $T_{\text{eff}} \in [2700, 3800]$  K. Their relatively cool temperatures compared to that of the Sun (i.e.  $T_{\text{eff},\odot} = 5777$  K) cause their spectra to peak at near-IR wavelengths between  $\sim 0.8 - 1.1 \mu\text{m}$  rather than in the optical. As such, many of the next generation of radial velocity spectrographs are being adapted from optical instruments to the development of near-IR spectrographs (e.g. SPIRou; Delfosse et al. 2013b; Artigau et al. 2014, NIRPS; Bouchy et al. 2017, CARMENES; Quirrenbach et al. 2014, HPF; Mahadevan et al. 2012, IRD; Tamura et al. 2012). Spectrographs in the near-IR wavelength domain are optimized to observe cool M dwarfs by maximizing the flux received from those stars. However, as will be discussed throughout this thesis, a number of issues and benefits accompany the use of near-IR spectrographs including the increased contamination of stellar spectra by telluric spectral features and the reduced effects of radial velocity stellar activity from active regions (Sect. 1.5.2).

## 1.3 Methods of Detecting Exoplanets

Numerous methods for the direct or indirect detection of exoplanets have been successfully demonstrated or proposed. All such methods provide complementary access to the exoplanetary parameter space thus enabling a unique view of the exoplanet population. A non-exhaustive list of prominent exoplanet discovery methods follows and are visualized in Fig. 1.1.

*Radial velocities:* the measurement of variations in the Doppler shift of star around the barycenter of its planetary system due to the gravitational influence of planetary-mass companions. This method

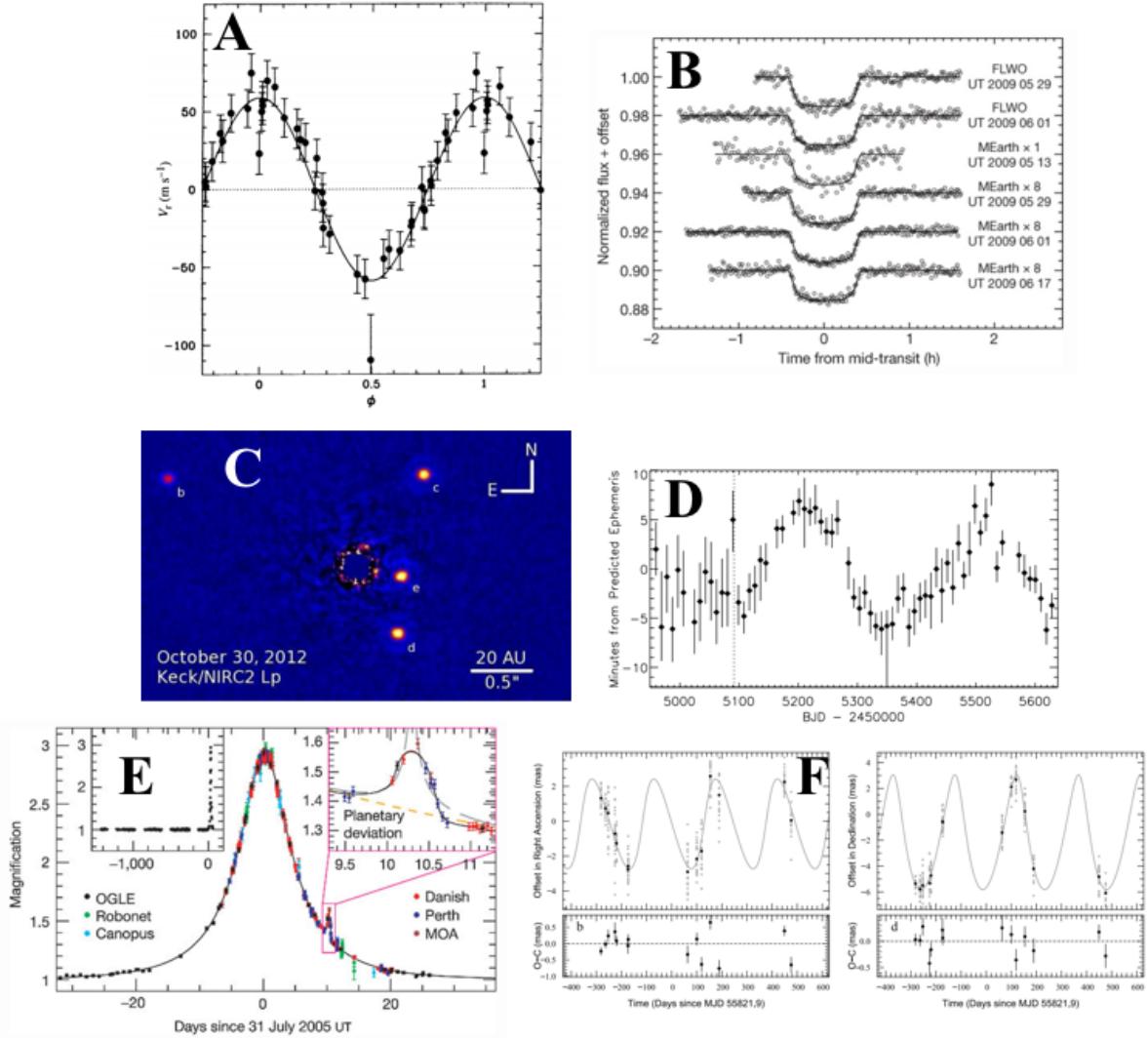


Figure 1.1: Examples of observables for exoplanet detections using A) the radial velocity method ([Mayor & Queloz, 1995](#)), B) planetary transits ([Charbonneau et al., 2009](#)), C) direct imaging ([Currie et al., 2014](#)), D) transit timing variations ([Ballard et al., 2011](#)), E) gravitational microlensing ([Beaulieu et al., 2006](#)), and F) astrometry ([Sahlmann et al., 2013](#)).

is used to measure the planet-star orbital separation and places a lower limit on the planet's mass. The radial velocity method favours the detection of massive, short-period planets around inactive stars (see Sect. 7.5.4 for more details).

*Photometric transits:* monitoring the brightness of a star over time and searching for periodically reoccurring brightness dips due to the occultation of the star by an orbiting planet. This method is highly complementary to the radial velocity method as it measures the planet's radius and orbital inclination where the latter can be used to break the mass-inclination degeneracy suffered by radial velocity planet detections. The recovery of radial velocity planet masses and planet radii from transit observations together provide views of planetary bulk densities that can be used to distinguish predominantly rocky planets from planets with significant fractions of their sizes in gaseous envelopes. The transit method is the most successful planet detection method as a single wide-field camera can be used to monitor many stars simultaneously and continuously. However, this method requires planets to have nearly edge-on orbits relative to the observer and thus favours the detection of close-in giant planets orbiting small stars.

*Direct imaging:* spatially resolving planets from their bright host stars and thus the direct detection of photons originating from the planet. The resulting planetary luminosity measurements can be compared to forward models of planetary evolution to infer planet masses and possible formation pathways (Marley et al., 2007). Multi-epoch observations can also reveal the nature of planetary orbits (Wang et al., 2018). Due to the extremely small planet-star contrasts that are required to directly detect planetary companions, direct imaging favours the detection of young massive planets on wide orbits in the thermal infrared where the planet-star contrast is maximized.

*Transit timing variations:* or TTVs, represent deviations in the expected transit times of a transiting planet under the assumption of a linear ephemeris. Such deviations may be attributed to an unseen planetary companion whose gravitational influence on the transiting planet is observable either through close encounters or mean motion resonances. The amplitude and periodicity of a transiting planet's TTVs can be used to constrain the mass and orbital period of the perturbing planet although a mass-eccentricity degeneracy exists for planet pairs close to mean motion resonances (Lithwick et al., 2012).

*Gravitational microlensing:* offers a unique method for probing the planet population out to large distances within the Milky Way as the detected photons do not originate from the planetary system itself but rather from a distant bright source. As the gravitational field of a foreground star (i.e. the lens) occults a distant bright source, the source appears to brighten as a result of the gravitational lensing effect. In the case for which the lens' gravitational potential deviates significantly from radial symmetry because of a gravitational anomaly induced by a planetary companion, that inhomogeneity in the gravitational field of the lens will disrupt the smoothness of the field and cause an anomalous brightening that reveals the presence of a co-moving companion with a measurable mass. One drawback of this method however is that observations of the same planetary system are not repeatable due to the very precise alignment of the source and lens that is required.

*Astrometry:* is similar to the radial velocity method wherein a planetary-mass companion causes its host star to periodically wobble. The astrometry method is unique to the radial velocity method as the

observed variations occur in the plane of the sky in position space, rather than in velocity space. This method is used to measure the orbital separation of a star and planet and measure the planet's mass. Because the variation in the star's position needs to be resolved, the astrometry method favours nearby planetary systems with massive planets on wide-orbits.

## 1.4 Exoplanet Detection via Stellar Radial Velocity

### 1.4.1 Radial velocity curves

The radial velocity (RV) method of detecting and characterizing exoplanets is largely the focus of the work presented herein and so here I provide an overview of its formalism.

The RV method represents a simple application of Newton's third law: “*for every action, there is an equal and opposite reaction*”. In this case, the presence of a planetary companion around a host star displaces the barycenter of that two-body system (i.e. the centre-of-mass) from the centre of the star itself. Because of this, a star with a sole companion will execute a Keplerian orbit about the system's barycenter. The radial velocity of such a star is

$$V_r(t) = \gamma_0 + K[\cos(\nu(t, P, T_0) + \omega) + e \cos \omega], \quad (1.1)$$

where  $t$  is the independent time coordinate,  $\gamma_0$  is the systemic velocity of the star relative to the observer,  $K$  is the semi-amplitude of the RV variation,  $P$  is the orbital period of the star and planet,  $T_0$  is the reference epoch or time of inferior conjunction, and the orbital elements  $e$  and  $\omega$  are the orbital eccentricity and argument of periastron respectively. Note that  $\omega$  is undefined for circular orbits with  $e = 0$  and is assigned the conventional value of  $\pi/2$  in such a scenario. Eq. 1.1 is known as the Keplerian orbital solution. The true anomaly

$$\nu(t, P, T_0) = 2 \arctan \left( \sqrt{\frac{1+e}{1-e}} \tan \frac{E(t)}{2} \right) \quad (1.2)$$

defines the position of the star on its orbit relative to the argument of periastron and in the direction of motion. It is derived from another angular parameter, the eccentric anomaly  $E(t)$ . The eccentric anomaly can be solved for using the mean anomaly  $M(t)$  according to Kepler's equation

$$M(t) = E(t) - e \sin E(t), \quad (1.3)$$

where the mean anomaly is

$$M(t) = \frac{2\pi}{P}(t - \tau_{\text{peri}}), \quad (1.4)$$

and  $\tau_{\text{peri}}$  is the epoch of pericenter passage. Because Eq. 1.3 does not have a closed-form solution, in practice it is solved using iterative numerical techniques. Examples of Keplerian orbits using Eq. 1.1 are shown in Fig. 1.2 for two 5 Earth mass planets each with an orbital period of 2 days around a  $0.2 M_\odot$  star. The first planet is on a perfectly circular orbit and therefore induces a purely sinusoidal RV variation on the host star while the second planet's eccentricity is 0.3 and exhibits a RV variation that

clearly differs from sinusoidal with the addition of the second cosine term in Eq. 1.1.

The Keplerian orbital solution is valid in the absence of dynamical perturbations to the two-body system. For Keplerian orbits, only the true anomaly in Eq. 1.1 varies with time. However, if additional bodies are present in the system such as in circumbinary systems or in high multiplicity planetary systems, then the parameters in Eq. 1.1 need not be fixed and we instead refer to the measured orbit as the osculating orbit that may vary between orbital cycles. Note that if the timescale for perturbations to the stellar orbital solution is comparable to or less than the time interval over which RV observations are obtained, then the Keplerian orbital solution is an insufficient description of the observations and one must consider a more complicated treatment of the dynamics typically in the form of N-body simulations.

### 1.4.2 Fundamental planet parameters from the radial velocity method

If we consider the simplest planetary system containing a star and a planet, each body will orbit their common centre-of-mass (COM) whose position vector is

$$\mathbf{r}_{\text{COM}} = \frac{M_s \mathbf{r}_s + m_p \mathbf{r}_p}{M_s + m_p}, \quad (1.5)$$

where  $M_s$  and  $m_p$  are the stellar and planetary masses respectively and  $\mathbf{r}_s$  and  $\mathbf{r}_p$  are their position vectors. Let us now take a barycentric view of this two-body system such that the COM coordinate is located at the origin ( $|\mathbf{r}_{\text{COM}}| = 0$ ). Further assuming circular orbits (i.e.  $e = 0$ ) and working in absolute scalar terms, the orbital distances of the star and planet are related via

$$a_s M_s = a_p m_p, \quad (1.6)$$

where we have written  $|\mathbf{r}_s| = a_s$  and  $|\mathbf{r}_p| = a_p$ . For a circular orbit, the stellar velocity around the COM has the constant value

$$v_s = \frac{2\pi a_s}{P}. \quad (1.7)$$

As the star orbits the COM the maximum observed radial velocity of the star is realized when the magnitude of the star's radial velocity vector equals its velocity  $v_s$ . At these points in the orbit,  $V_r(t) = \max(V_r(t)) = v_s = K$ . Because the RV curve, and hence  $K$ , is observable, we can relate  $K$  to the masses of the star and planet by combining the velocity equation (Eq. 1.7) with the COM equation (Eq. 1.6) and Kepler's third law:

$$\max(V_r(t)) = K = \frac{2\pi a_s}{P}, \quad (1.8)$$

$$= \frac{2\pi}{P} \left( \frac{m_p}{M_s} \right) a_p, \quad (1.9)$$

$$= \frac{2\pi}{P} \left( \frac{m_p}{M_s} \right) \left( \frac{GM_s}{4\pi^2} \right)^{1/3} P^{2/3}, \quad (1.10)$$

$$= \left( \frac{2\pi G}{M_s^2 P} \right)^{1/3} m_p. \quad (1.11)$$

Figure 1.2: An animation of the orbits of two (non-interacting) planets of equal mass and on 2 day orbits around a 0.2 solar mass star. The first planet has a circular Keplerian orbit ( $e = 0$ ) while the orbit of the second exhibits some ellipticity with  $e = 0.3$ . The lower panel depicts the corresponding RV variation induced on the host star by each planet independently.

Eq. 1.11 contains two quantities that are directly observable from the RV time series. Namely, the orbital period  $P$  and semi-amplitude  $K$ . Observers will often seek an independent constraint on the stellar mass  $M_s$  from say the comparison of the stellar spectrum to spectral templates of known stellar mass from stellar evolution models (e.g. Muirhead et al., 2012a) or from empirical relations between optical or near-IR stellar luminosities and dynamically measured stellar masses (e.g. Benedict et al., 2016; Mann et al., 2019). Prior knowledge of the stellar mass enables the mass of the planetary companion to be inferred from RV time series using Eq. 1.11. The uncertainties in the resulting planet mass measurement are typically dominated by the noise in the dataset which directly affects the measurement precision of  $P$  and  $K$ . Additional uncertainties in  $m_p$  are also introduced by the uncertainty in  $M_s$  which motivates the need to develop a detailed understanding of fundamental stellar parameters to enable precise exoplanet measurements.

Looking more closely at Eq. 1.11 we can see that the RV semi-amplitude scales linearly with planet mass but has a negative scaling with stellar mass and orbital period. This makes sense because the RV effect is a manifestation of a gravitational effect. An increased perturbing mass (i.e.  $m_p$ ) will result in an increased perturbation to the central body which itself becomes more difficult to perturb if more massive. Also, the effect weakens as the distance between the two bodies is increased (i.e. larger  $P$ ). The RV method therefore has a natural observational bias towards massive, close-in planets around low mass stars as the combination of these parameters will maximize the RV signal and increase the chances of that signal being detected above the noise (see Sect. 1.5.2). This explains why the first planets to be detected with the RV method were hot Jupiters (e.g. Mayor & Queloz, 1995).

### 1.4.3 Caveats to Eq. 1.11

In the derivation of the RV semi-amplitude (Eq. 1.11) we have neglected two important geometric effects. The first is that of eccentricity which, if non-zero, causes the stellar velocity to vary throughout its orbit such that  $dv_s/dt \neq 0$ . Instead  $\max(V_r(t)) > 2\pi a_s/P$  because the conservation of angular momentum requires that  $v_s(\tau_{\text{peri}})$  be greater than elsewhere on the orbit. The corresponding correction factor to Eq. 1.11 is  $(1 - e^2)^{-1/2}$ .

The second effect is that of inclination. When attempting to measure the planetary mass from RVs the orbital inclination of the planetary orbit<sup>2</sup> is degenerate with  $m_p$  and is entirely unknown unless the planet has an a-priori measurement of its orbital inclination from either transit or astrometric observations. But the majority of planets are neither transiting nor induce a resolvable astrometric variation and thus have unknown orbital inclinations. For these planets, RV measurements are only sensitive to a mass lower limit  $m_p \sin i$  rather than to the absolute planet mass. The  $\sin i$  correction factor also illustrates that an RV signal is only non-zero when the observer's orientation is not exactly orthogonal to the normal of the star's orbital plane; i.e. face-on orientations with  $i = 0$ .

The inclusion of these correction factors gives a more complete description of the RV semi-amplitude

---

<sup>2</sup>Orbital inclination is conventionally measured relative to the plane of the sky; i.e. an edge-on orbit has  $i = 90^\circ$ .

compared to Eq. 1.11:

$$K = \left( \frac{2\pi G}{M_s^2 P} \right)^{1/3} \frac{m_p \sin i}{\sqrt{1 - e^2}}. \quad (1.12)$$

## 1.5 Measuring Radial Velocities

### 1.5.1 Stellar Spectroscopy

Radial velocity variations induced by planetary companions cause the star to become periodically Doppler shifted. These wavelength shifts can be detected using high resolution spectroscopy as spectral features originating from the stellar photosphere are shifted from their rest wavelength  $\lambda_{\text{rest}}$  as the star's velocity changes along the line-of-sight according to

$$\frac{\Delta\lambda}{\lambda_{\text{rest}}} = \frac{V_r}{c}, \quad (1.13)$$

where  $\Delta\lambda = \lambda_{\text{obs}} - \lambda_{\text{rest}}$  and  $\lambda_{\text{obs}}$  is the observed wavelength of the spectral features. In practice, RV shifts are often measured from the cross-correlation of the observed stellar spectrum with a master empirical spectrum or template mask that is shifted in velocity space to map their correlation over relative velocities (Astudillo-Defru et al., 2015). The resulting RV precision is enhanced when the individual velocities of many lines can be averaged. This makes precision RV measurements especially tractable for M dwarfs as they exhibit a plethora of spectral lines that become more abundant towards later type stars. A sample of infrared spectra of early-to-late M dwarfs are shown in Fig. 1.3 to demonstrate their wealth of absorption features (Rayner et al., 2009).

The aforementioned spectroscopic measurements are made using a stabilized spectrograph at high spectral resolution ( $R = \lambda/\Delta\lambda$ ) that contains a grating to diffract or spatially disperse the incoming source light into its various colour components. Echelle spectrographs in particular include an additional high order grating that separates the light into spectral orders that are subsequently cross-dispersed spatially onto the detector. The dispersed spectral orders create the observed stellar spectrum (Fig. 1.4). After 1D spectral extraction, the spectrum is cross-correlated with the aforementioned master spectrum or template mask wherein prominent lines are averaged to create the cross-correlation function (CCF) of the star at the epoch of observation. Traditionally, the CCF is typically fit with a Gaussian line profile whose best-fit mean is equal to the measured radial velocity (Pepe et al., 2002). It should be noted however that the net convective blueshift in stellar photospheres causes the native CCF to be asymmetric rather than purely Gaussian even in the absence of surface inhomogeneities (Gray, 1989). Active regions on the stellar surface arising from the presence of magnetic fields can have additional effects on the observed CCF causing the line profile to exhibit additional and localized non-Gaussianities. These distortions can in principle be used to model stellar activity in cases for which the distortions can be detected at high signal-to-noise and their origin dominates the astrophysical RV budget (as opposed to being due to systematic effects).

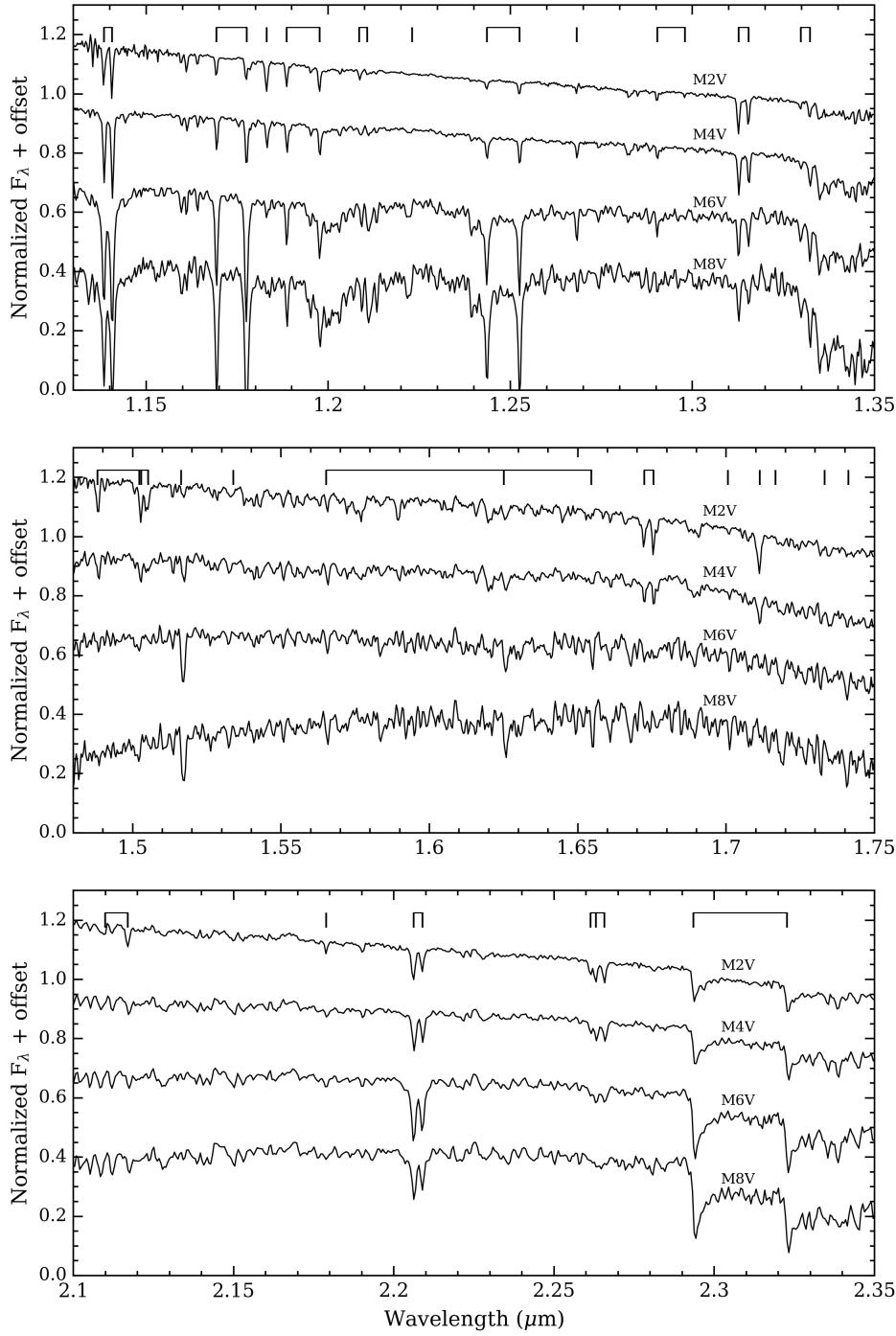


Figure 1.3: Medium resolution spectra ( $R = \lambda/\Delta\lambda \sim 2000$ ) for four early-to-late M dwarfs in the J band (top), H band (middle), and K<sub>S</sub> band (bottom) from the SpeX spectrograph at NASA's Infrared Telescope Facility.

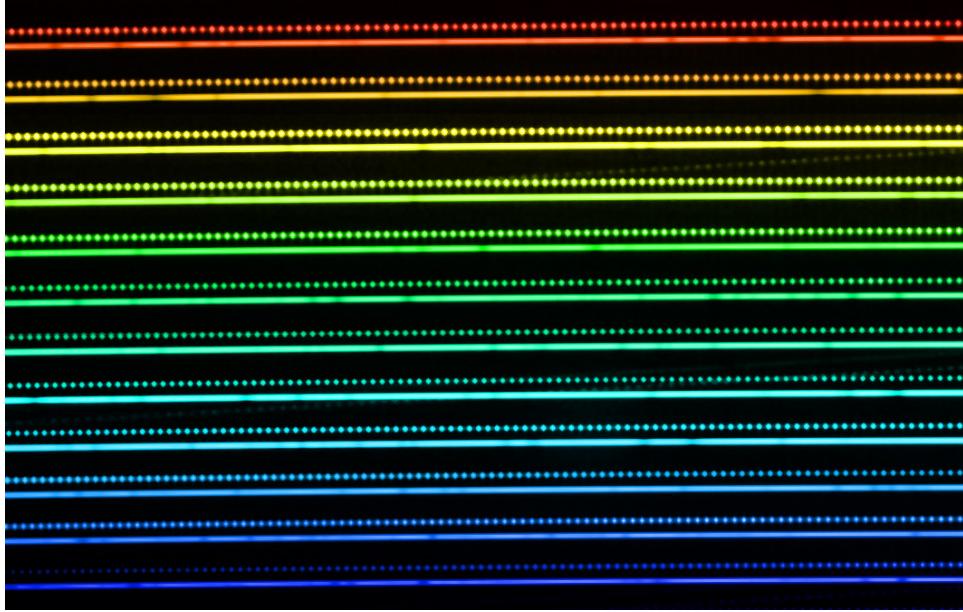


Figure 1.4: A colourized image of the HARPS detector showing a subset of its 72 spectral orders. Each row pair contains the stellar spectrum (solid spectrum) over a portion of the HARPS wavelength domain and the corresponding wavelength reference from its laser frequency comb (dotted spectrum). Accurate wavelength calibration enables monitoring of stellar absorption features over time that may reveal periodic Doppler shifts possibly indicative of a planetary companion. Image credit: *European Southern Observatory*.

### 1.5.2 Radial Velocity Error Budget Overview

The complete error budget when conducting precision RV measurements features many contributing factors. Among these are instrument/facility limitations, the fundamental photon noise limit, imperfect corrections of telluric contamination, and astrophysical noise. Here we review the RV error budget with a particular focus on stellar activity as it pertains to the work presented in this thesis.

#### Instrumentation Challenges

RV spectrographs have an inherent RV precision that is set by a multitude of individual error sources that hinder a user's ability to measure infinitesimally precise RVs (Podgorski et al., 2014; Halverson et al., 2016). These issues are intrinsic to the experimental hardware and their relative contributions vary significantly between instruments with often not a single source of error dominating the instrumental error budget.

A non-exhaustive list of prominent issues with RV spectrographs includes

- thermal stability,
- pressure stability,
- wavelength calibration,
- vibration control,
- drift control,
- detector imperfections (i.e. readout noise, intrapixel quantum efficiencies, etc),
- internal scattered light,

- fibre modal noise,
- tracking errors,
- focus errors,
- etc.

The aforementioned RV error sources arise from the design and operation of the instrument and telescope facility. As such, their mitigation is set by the hardware specifications and cannot be significantly improved by post processing. These challenges associated with building state-of-the-art RV instrumentation are not the focus of this thesis but it is important to note their contributions to the overall RV error budget as together they set a limit to the highest possible RV precision attainable with a given instrument.

### Photon noise limit

The photon noise limit represents the fundamental limit to the RV measurement precision that can be obtained from a stellar spectrum  $A_i$  where  $i$  is the index over wavelength elements. The photon noise limited RV precision from [Bouchy et al. \(2001\)](#) is

$$\sigma_\gamma = \frac{c}{S/N \cdot Q} \quad (1.14)$$

where  $c$  is the speed of light, S/N is the signal-to-noise of the spectrum over its complete spectral domain, and

$$Q = \frac{\sqrt{\sum_i W_i}}{\sqrt{\sum_i A_i}} \quad (1.15)$$

is known as the quality factor and represents the density of RV information content in the spectrum  $A_i$  measured in photoelectrons. The weighting function is given by

$$W_i = \left( \frac{\lambda_i^2}{A_i} \right) \left( \frac{\partial A_i}{\partial \lambda_i} \right)^2. \quad (1.16)$$

By the dependence of  $\sigma_\gamma$  on the first spectral derivative of the observed spectrum (Eq. 1.16), it is clear that high S/N spectra with a high density of sharp features provide the best possible RV precision. Because the sharpness of the lines is important, obtaining high resolution spectra has a direct benefit on minimizing  $\sigma_\gamma$ . Similarly, stars with low levels of collisional and rotational broadening (i.e. low  $\log g$  and  $v \sin i_s$ ) are favourable targets for minimizing  $\sigma_\gamma$ . Also because line density is important, cooler stars or stars with a high metallicity are favourable as they exhibit more molecular and metal features over which counting errors can be minimized. Fig. 1.5 demonstrates these dependencies over the spectral bands  $UBVRIYJHK_S$ .

### Telluric contamination

In addition to error floors arising from the instrumentation and the photon noise limit, other physical sources of RV errors exist that need to be accounted for in the post-processing software. One example is the effect of telluric contamination of the stellar spectrum by spectral features originating from Earth's atmosphere. Because all high precision RV experiments to date are executed from the ground the incoming starlight must traverse much of the Earth's atmosphere at which time telluric absorption and emission features become imprinted on the incident spectrum. This effect is much more prominent in

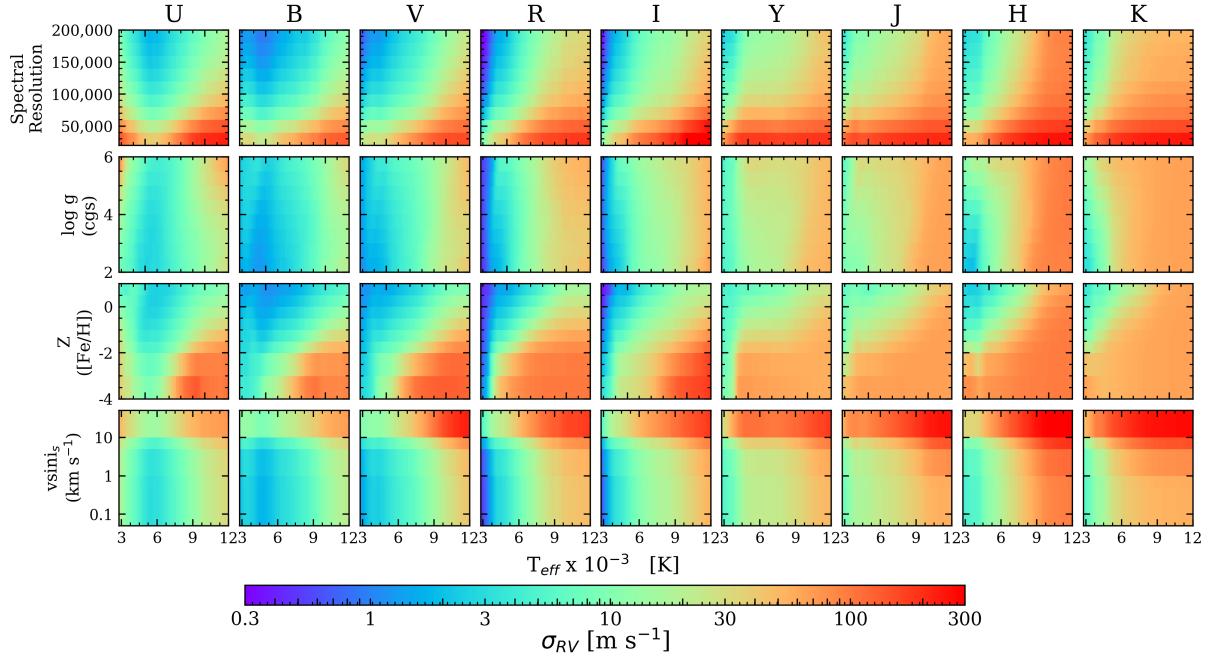


Figure 1.5: Distributions of the photon noise limited RV precision in the  $UBVR\text{I}YJHK_{\text{S}}$  bands and as a function of spectral resolution,  $\log g$ , metallicity,  $v \sin i_s$ , and stellar effective temperature. The distribution shown in each subpanel is marginalized over all parameters not depicted on either of the panel’s axes and is therefore not representative of true  $\sigma_{\gamma}$  values particularly in the near-IR bands in which telluric contamination is not properly treated.

the infrared where the atmosphere becomes more opaque over wider spectral bands than in the visible due in large part to absorption bands by various greenhouse gases such as  $\text{H}_2\text{O}$  (Fig. 1.6). In the visible, small regions in wavelength space that are known to be contaminated by tellurics can be simply masked without significant loss of RV content. Whereas this method if applied to near-IR spectrographs eliminates such a large fraction of the available RV information that it effectively negates the benefits of developing near-IR RV spectrographs at all. Instead, near-IR RVs must rely on more sophisticated methods of telluric spectrum extraction and removal based on atmospheric transmission models (e.g. Vacca et al., 2003; Bertaux et al., 2014) and/or data-driven methods (e.g. Artigau et al., 2014; Bedell et al., 2019).

### Stellar activity

There exists a variety of physical processes in the photospheres and chromospheres of stars that can lead to temporally correlated RV signals that are collectively referred to as *stellar activity*. Alternatively, activity sources that act on short timescales (i.e. seconds to minutes) and are unrelated to magnetic processes are sometimes referred to as stellar jitter. Particularly in RV planet studies, stellar activity signals are the bane of observers searching for sub-Neptune-sized planets because activity signals can mask and in certain instances even mimic the typically small planetary signals of interest.

Depending on the exact physical process these signals can have widely differing timescales and manifestations in spectroscopic, photometric, and polarimetry observations. Activity signals from active

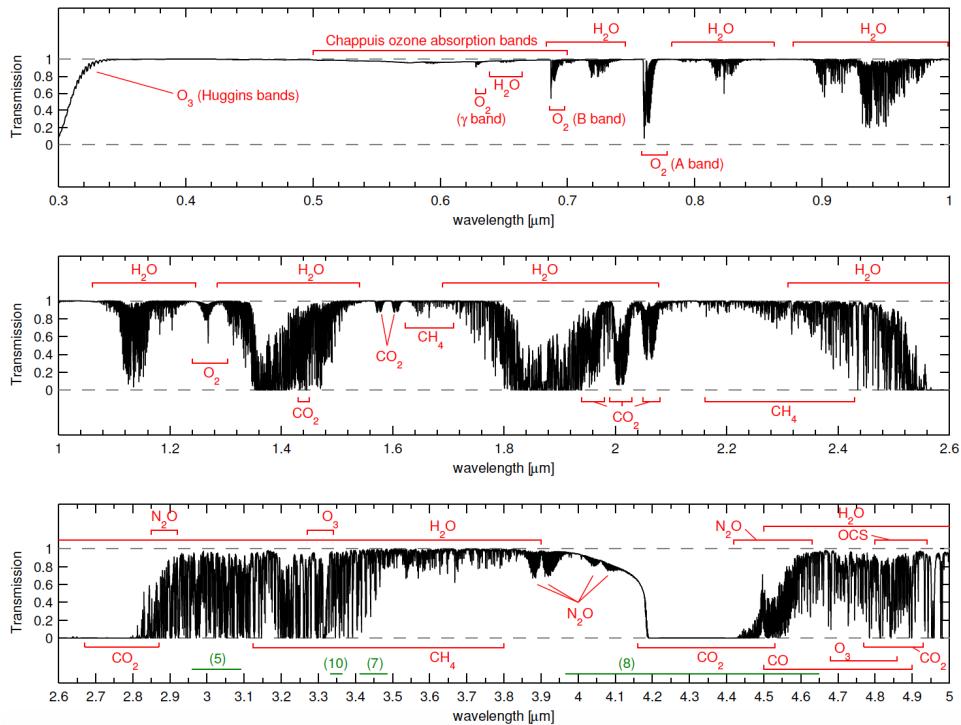


Figure 1.6: A model of telluric absorption from  $0.3\text{--}5 \mu\text{m}$  calculated using the LBLRTM radiative transfer code (Clough et al., 2005). The atmosphere becomes increasingly opaque towards infrared wavelengths with narrow observing windows at the  $YJHK_{\text{S}}LL'$  bands. Telluric absorption in the optical is dominated by O<sub>2</sub>, O<sub>3</sub>, and H<sub>2</sub>O while the near-IR wavelengths feature strong H<sub>2</sub>O absorption bands in addition to other strong absorbers such as CO, CO<sub>2</sub>, CH<sub>4</sub>, OCS, and N<sub>2</sub>O. (Image credit: Smette et al., 2015)

Table 1.1: Summary of Radial Velocity Stellar Activity Sources

Activity Source	Typical Timescale	Typical Signal Amplitude	Notes
Oscillations	minutes	few m s <sup>-1</sup>	Can be mitigated with sufficiently long exposure times or multiple observations per night <sup>1</sup> .
Granulation	minutes	few m s <sup>-1</sup>	See oscillations.
Flares	minutes-hours	few m s <sup>-1</sup>	Has distinct photometric and spectroscopic signatures. Observations during a flaring event should be excluded from planet searches <sup>2</sup> .
Active Regions	a few stellar rotation periods	sub-m s <sup>-1</sup> → tens of m s <sup>-1</sup>	Timescale and amplitude depend heavily on the active region sizes and on stellar rotation <sup>3,4</sup> . See Fig. 1.10 for the distribution of M dwarf $P_{\text{rot}}$ .
Gravitational redshift	days-years	cm s <sup>-1</sup>	Timescale depends on the mechanism for causing significant changes in the stellar radius <sup>5</sup> .
Magnetic Cycles	$\gtrsim$ 2 years	few m s <sup>-1</sup>	Only important when searching for long period planets or planets in time series with long baselines <sup>6,7</sup> .

**Notes.** (1) Dumusque et al. (2011b), (2) Reiners (2009), (3) Dumusque et al. (2011a), (4) Giles et al. (2017), (5) Cegla et al. (2012), (6) Santos et al. (2010), (7) Robertson et al. (2014).

regions such as cool star spots and chromospheric plages are also wavelength dependent unlike Doppler variations induced by planetary companions (Tal-Or et al., 2018). Ancillary time series to RV measurements derived from any of the aforementioned observations may therefore be used to disentangle stellar activity from achromatic Doppler shifts induced by a companion. Such observables include spectroscopic activity indicators such as the  $H\alpha$  and Ca II H & K indices that are magnetically sensitive features originating from the hot stellar chromosphere, line deformation diagnostics (e.g. CCF shape parameters; Boisse et al. 2011, line-by-line analyses Davis et al. 2017; Dumusque 2018), plus Doppler imaging of the stellar surface (Hébrard et al., 2016) (see Sect. 2.1.2 for details).

Examples of sources of stellar activity are summarized below while a brief overview is given in Table 1.1.

*Flares:* magnetically active stars can undergo energetic flares or coronal mass ejections originating from the stellar atmosphere. The exact flare physics in M dwarfs remains uncertain but clearly requires strong magnetic fields that can be sustained by turbulent convective motions and rotation (Browning, 2008). Flare events are easily identifiable as they exhibit a characteristically rapid increase in brightness (or in the intensity of certain emission lines; e.g. Ca II, He I,  $H\alpha$ , etc. Schmidt et al. 2012) followed by an exponential decay over hours. Because RV measurements affected by flares can be easily identified by the characteristic line intensity spike over short timescales, such measurements are commonly removed from planet searches (Reiners, 2009).

*Stellar Oscillations:* small scale mechanical perturbations to the internal stellar structure can give rise to oscillation modes. If not damped, these oscillations can propagate through the stellar interior before rebounding at the surface of the star thus creating observable pulses in brightness and in the RVs at a few m s<sup>-1</sup> and on timescales of a few to tens of minutes (Bedding et al., 2001). This phenomenon is predominantly observed in Sun-like and post main sequence stars. Due to the short timescale of pulsation variations in RV measurements of pulsating stars, the corresponding jitter is often mitigated with sufficiently long exposure times (Lovis et al., 2005; Dumusque et al., 2011b).

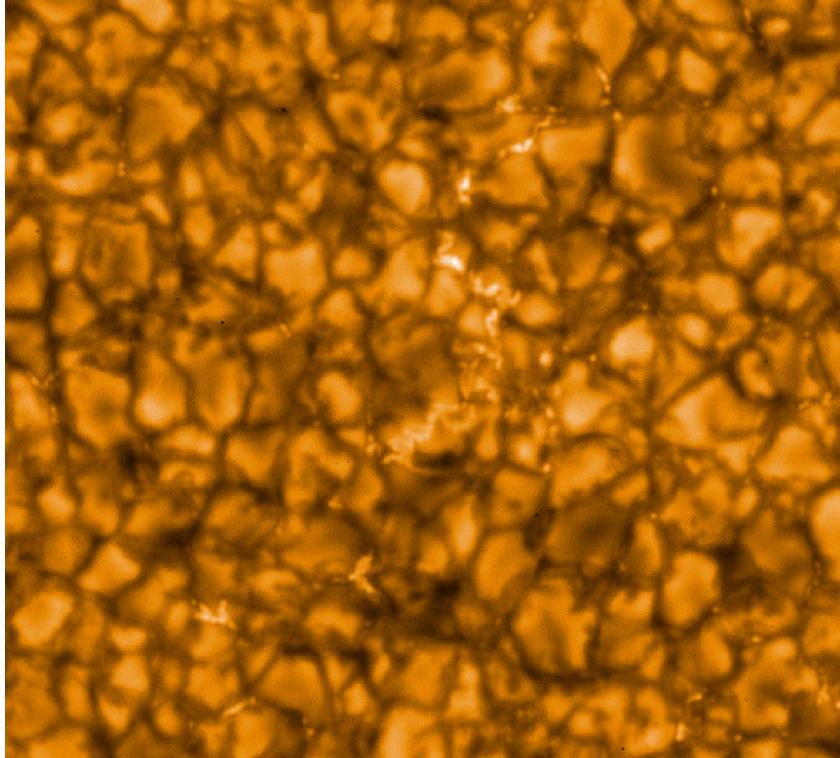


Figure 1.7: An image of a patch of the solar surface taken with NASA’s Hinode Optical Telescope. The bright regions are the result of rising hot fluid parcels whereas the cooler, ‘intergranular lanes’ reveal where the gas has cooled and descends back into the solar interior. (Image credit: Hinode JAXA/NASA/PPARC)

*Granulation:* stars with outer convection zones like the Sun and M dwarfs exhibit granulation patterns at their surfaces. The granulation pattern is the result of convection cells penetrating the surface; hot fluid parcels rising to the surface before cooling and descending back into the stellar interior (see Fig. 1.7). These patches of hot rising fluid are therefore brighter than the surrounding regions of cold descending fluid. Furthermore, the relative fractional coverage of a star’s surface by rising fluid parcels is in general greater than that of descending ones as evidenced in Fig. 1.7. Along the line-of-sight, the RV component of photons emitted by the rising parcels at the photospheric boundary will be blueshifted whereas the receding fluid within the ‘intergranular lanes’ will be redshifted. The domination of the stellar surface area by hot parcels results in a granulated star having a net *convective blueshift* that in turn creates an asymmetry in the disk-integrated spectral line profile with additional power at blueshifted velocities. The effect of granulation is also empirically known to decrease towards cool M dwarf stars (Dumusque et al., 2011b; Meunier et al., 2017).

The short lifetimes of typical granules ( $\sim 10$  minutes; Hall, 2008; Gilliland et al., 2011) implies that brightness variability of the surface varies on a similar timescale. However, ‘supergranules’ may also persist over longer timescales albeit with smaller RV contributions ( $\lesssim 1 \text{ m s}^{-1}$ ; Rincon et al., 2017) The corresponding RV jitter is dependent on the time variability of the relative velocity of convective cells across the stellar surface. Noting that the relative sizes of granules varies over time, so too does the

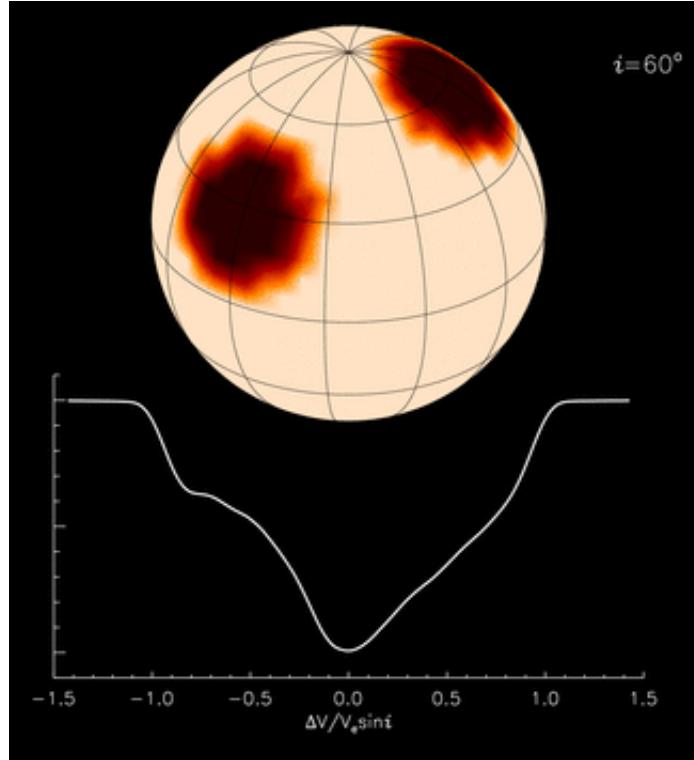


Figure 1.8: An snapshot example of a spotted stellar photosphere and its effect on the disk-integrated spectral line profile. The cool spots depicted here disrupt the symmetry of the line profile according to the temperature difference between the spot and unspotted photosphere at its position on the stellar disk. (Image credit: Kochukhov, 2016)

resulting RV jitter whose net effect (hot and cold regions partially cancel each other) is typically at the level of a few  $\text{m s}^{-1}$  (Lindegren & Dravins, 2003).

*Active Regions:* like the Sun, the photospheres of active M dwarfs are littered with active regions such as dark photospheric starspots, hot faculae, or plages in the chromosphere. Indeed, the activity signals from active regions appear to dominate the RV activity budget on M dwarfs (Lindegren & Dravins, 2003). These localized regions of either hot or cold gas relative to the surrounding photosphere become trapped by magnetic field loops reconnecting with the stellar surface. Active regions create a temporally evolving RV signal and spectral line profile distortion due to their distinct emitting temperatures and their orientation as they rotate in and out of view (see Fig. 1.8). In the Sun, these structures have long been known to be associated with strong local magnetic fields as evidenced by the observed Zeeman splitting of lines emitted from these active regions (Hale, 1908).

The activity timescale from active regions is modulated by the stellar rotation. The stellar rotation period  $P_{\text{rot}}$  can often be measured for active stars with non-zero inclination from the quasi-periodicity of their photometric variability. The spectroscopic method of measuring a star's rotation state is done by measuring the projected stellar rotation velocity  $v \sin i_s$ <sup>3</sup> via the rotational broadening of its spectral

---

<sup>3</sup>Slowly rotating stars with correspondingly small  $v \sin i_s$  values may not be detectable with insufficient spectral resolution. In such cases, only an upper limit on  $v \sin i_s$  can be measured.

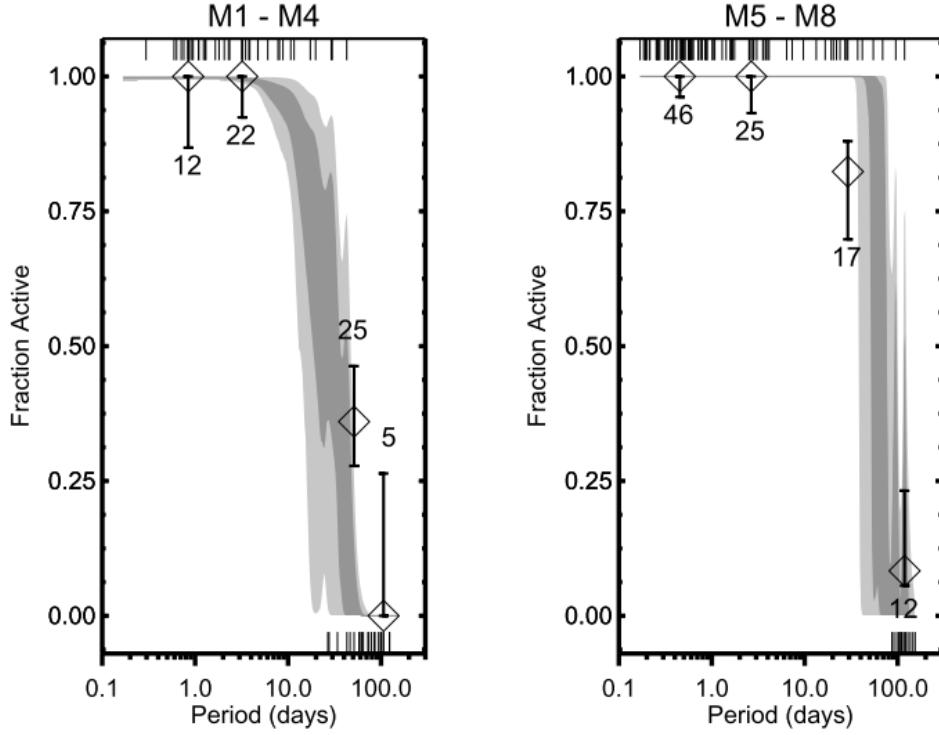


Figure 1.9: The fraction of active M dwarfs as a function of stellar rotation period for early-to-mid (*left panel*) and mid-to-late (*right panel*) M dwarfs. An individual star's activity flag is based on the ratio of its H $\alpha$  emission luminosity to its bolometric luminosity. The activity fraction of stars is seen to decrease as rotation slows although mid-to-late M dwarfs can maintain a higher activity fraction than early-to-mid M dwarfs within a given  $P_{\text{rot}}$  bin. (Image credit: West et al., 2015)

features (Gray, 2008). Because active regions are associated with magnetic activity and stellar magnetic fields require rotation in order to be sustained over long timescales, rapidly rotating late-type stars tend to be more magnetically active. Fig. 1.9 demonstrates that the fraction of active M dwarfs is sustained at slower rotation states in later M dwarfs than in early M dwarfs and that the activity fraction in general decreases as the stars spin-down via magnetized braking from stellar winds (Skumanich, 1972). A measure of stellar rotation can therefore be used as a first-order indicator of a star's activity level.

Fig. 1.10 depicts the rotation period distribution of M dwarfs in the solar neighbourhood (Newton et al., 2016) and in the Kepler field up to  $P_{\text{rot}} < 70$  days (McQuillan et al., 2013). Two distinct populations exist, namely a rapidly rotating population ( $P_{\text{rot}} \lesssim 3$  days) and a slowly rotating population ( $P_{\text{rot}} \gtrsim 50$  days). It has been proposed that these two populations have a distinct range of ages with the slow rotators belonging to an older stellar population based on their galactic kinematics (Irwin et al., 2011) and resulting from magnetic-braking over time. Indeed, a robust relation exists for GK and early M dwarf stars between stellar mass, age, and rotation (gyrochronology; Barnes, 2003).

Another important point regarding active regions is that their lifetimes are often one to a few stellar rotation timescales with those timescales being extended for later type stars (Giles et al., 2017). That is that active regions on M dwarfs appear to be long-lived compared to around Sun-like stars. The evolu-

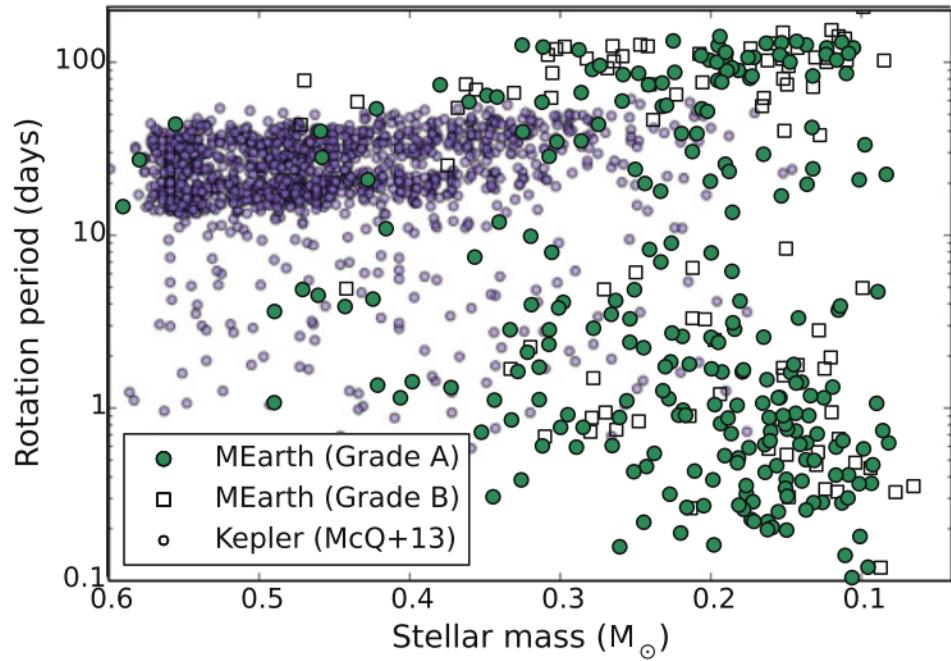


Figure 1.10: Rotation period versus stellar mass for solar neighbourhood and Kepler field M dwarfs. Periods of the Kepler stars are intentionally capped at  $P_{\text{rot}} = 70$  days. The distribution of  $P_{\text{rot}}$  appears to be bi-modal featuring a population of fast ( $P_{\text{rot}} \lesssim 3$  days) and slowly ( $P_{\text{rot}} \gtrsim 50$  days) rotating M dwarfs with a dearth of stars at intermediate  $P_{\text{rot}}$ . These features are thought to be indicative of two distinctly aged populations. (Image credit: [Newton et al., 2016](#))

tion of active regions over rotation cycles implies that the associated stellar activity can vary on similar timescales. M dwarf active region lifetimes can range from days to months and given that those lifetimes are primarily determined by the active region sizes as they diffuse into the photosphere (Berdyugina, 2005), larger volume/surface ratios results in longer lifetimes (Petrovay & van Driel-Gesztelyi, 1997; Giles et al., 2017). The lifetimes of M dwarf active regions will have important implications when attempting to model stellar activity in RV time series in subsequent chapters.

*Gravitational redshift:* the contraction and expansion of stars throughout their evolution will introduce an RV activity signal. Even small variations at the level of hundredths of a percent will have manifestations at the  $10 \text{ cm s}^{-1}$  level which is sufficient to mask an Earth-like planet around a Sun-like star (Lindegren & Dravins, 2003). In low mass stars, the suppression of surface convection by strong magnetic fields may result in stellar radius fluctuations and consequently an activity signal due to gravitational redshift that varies on the timescale of variations in the star's magnetic field strength of days to years (Cegla et al., 2012).

*Magnetic Cycles:* the Sun undergoes a long-term magnetic cycle with an 11 year period as measured by the fractional coverage or number of sunspots and faculae visible on its surface (Hathaway, 2010). At solar maximum, the peak in magnetic activity over the solar cycle, the Sun i) experiences strong prominences and coronal mass ejections and ii) contains a large number of active regions that contribute to a higher level of activity in RV time series as well as in other observables (Maunder, 1904). Long-term magnetic activity cycles have been observed in long-baseline spectroscopic monitoring programs (Santos et al., 2010). From solar observations, the flux in the Ca II H & K lines are known to correlate with the number of sunspots. Numerous results from similar surveys have demonstrated that the majority of Sun-like stars do exhibit magnetic activity cycles with a range of periods from  $\sim 7 - 30$  years (Duncan et al., 1991; Lockwood et al., 1997; Baliunas et al., 1998).

As it pertains to RV planet searches, activity cycles have been observed and characterized in Sun-like stars (Baliunas et al., 1995). Even when complete cycles remain unresolved, subsections of the cyclic trends have been seen in certain mature RV surveys to be correlated with the  $\log R'_{\text{HK}}$  spectroscopic activity indicator (e.g. Lovis et al., 2011). A common tactic to correct for such long-term activity trends in RV time series is to use a linear relation between the RVs and the contemporaneous  $\log R'_{\text{HK}}$  time series although residual jitter often persists (Dumusque et al., 2012). The corresponding RV signal from magnetic cycles can reach tens of  $\text{m s}^{-1}$ . Luckily, when searching for planets whose orbital timescales are much less than the period of the star's magnetic cycle, the corresponding activity is not often a major concern.

## 1.6 Point-form Thesis: Introduction

- 1.1 **Exoplanets:** exoplanets in the local universe are extremely abundant and quite diverse compared to the planets in our own solar system.
- 1.2 **M dwarf stars:** M dwarfs represent superlative opportunities to characterize planetary systems in detail if we can mitigate their prominent stellar activity signatures.

- **1.3 Methods of Detecting Exoplanets:** there exists a number of methods of detecting exoplanets, each with its own sensitivities and biases.
- **7.5.4 Exoplanet Detection via Stellar Radial Velocity:** the gravitational influence of planetary companions on their host star induces a periodically varying Doppler shift that may be detectable and used to measure an exoplanet’s minimum mass and orbital properties.
- **1.5 Measuring Radial Velocities:** hi-resolution spectrographs are used to measure the Doppler shift of stellar absorption features. The precision of such measurements is often impeded by sources of error from the instrumentation, counting statistics, telluric contamination, and stellar activity.

# Chapter 2

# Stellar Activity Modelling in Radial Velocity Time Series

As discussed in Sect. 1.5.2, there exists a multitude of physical processes ongoing within the photospheres and chromospheres of active stars that produce observable signatures with a variety of amplitudes and timescales (see Table 1.1). The subsequent sections discuss a variety of techniques that have been used to model and consequently mitigate the effects of stellar activity in RV time series.

## 2.1 An Overview of Techniques for Stellar Activity Mitigation

### 2.1.1 Stellar Activity as a Scalar Parameter

Back when the first giant exoplanets were being discovered with RVs, typically a treatment of stellar activity was not implemented. The reason being that the measurement precisions of those datasets were insufficient to resolve the temporal structure of RV activity for any but the most active stars. Many observers however did report the root-mean-square (rms) of their RV residuals following the removal of their best-fit planet model (e.g. [Mayor & Queloz, 1995](#); [Butler & Marcy, 1996](#)). In many cases it was revealed that the residual rms exceeded the characteristic RV measurement uncertainty and thus alluded to the presence of additional jitter signals that may or may not vary significantly with time.

In many of the subsequent RV planet searches the apparent jitter was characterized by an additive scalar jitter  $s$ . The free parameter  $s$  was used to characterize the level of RV jitter as it was added in quadrature to the RV measurement uncertainties when evaluating the objective function during any analysis equivalent or analogous to a  $\chi^2$ -minimization routine. The nature of this method assumes a fixed level of dispersion due to jitter that is constant in time. As such, the resulting measurement of  $s$  does little more than inform us of the jitter amplitude as any temporal evolution of activity due to finite active region lifetimes or magnetic cycles remained unresolved.

### 2.1.2 Correlations with Activity Indicators

Stellar RV observations are known to be affected by both planetary companions as well as by stellar activity. Disentangling those signals in RV time series therefore benefits significantly from ancillary

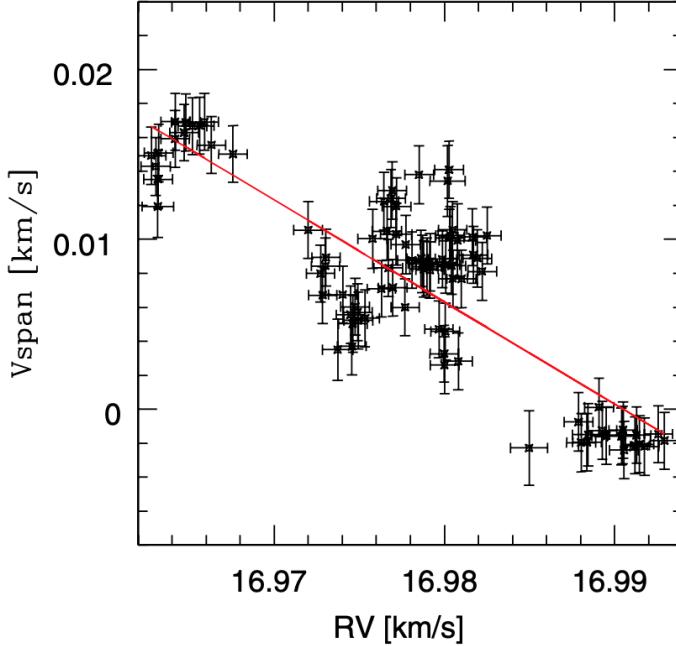


Figure 2.1: The correlation of the  $V_{\text{span}}$  activity indicator with the RVs for the active Sun-like star HD 17051 from HARPS spectra. The solid line depicts the best-fit linear relation from least squares fitting. The fitted relation is used to de-correlate the RVs for the effects of stellar activity as probed by the  $V_{\text{span}}$  time series. (Image credit: [Boisse et al., 2011](#)).

time series that are sensitive to stellar activity only ([Boisse et al., 2009](#)). The classical implementation of de-correlation by an activity indicator is to derive time series of one or many spectroscopic activity indicators whose sampling is simultaneous with the RVs and then fitting an often linear relation between those datasets (Fig. 2.1) to account for the temporal evolution of activity over the observational baseline. The relation, when fitted simultaneously with planetary solutions, allows the RVs to be de-correlated jointly with the measurement of the planetary parameters. This technique has been shown to be reasonably effective when the stellar rotation period  $P_{\text{rot}}$  is well constrained, the planetary orbital period is distinct from  $P_{\text{rot}}$ , the amplitude of the planetary signal exceeds that of the activity signal by  $\gtrsim 30\%$ , and the stellar rotation period is well sampled over multiple cycles ([Boisse et al., 2011](#)).

There exists a number of activity indicators that can be derived from the stellar spectra. Their definitions and physical motivations are summarized below.

$\log R'_{HK}$ : the Ca II H & K resonance lines are excited by non-thermal heating and act as a sensitive indicator of chromospheric structure and particularly of the presence of bright chromospheric plages. For optical spectra with access to the Ca II H & K lines centered on 3968.47 Å and 3933.66 Å, the Mt. Wilson S-index is

$$\text{S-index} \propto \frac{\Psi_H + \Psi_K}{\Psi_B + \Psi_V} \quad (2.1)$$

where  $\Psi_H$  and  $\Psi_K$  represent the narrowband ( $\sim 1.1$  Å wide) fluxes in the cores of the H & K lines of the

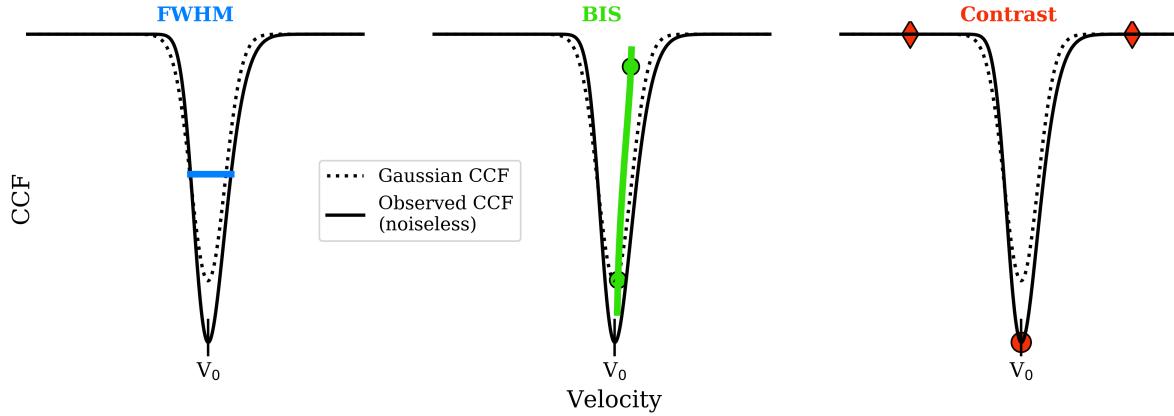


Figure 2.2: Illustrations of measurements of the CCF width, asymmetry, and contrast via the full width half maximum (FWHM), bisector (BIS), and contrast shape parameters. The FWHM characterizes broadening of the CCF due to active regions at unique velocities on the stellar disk. The BIS characterizes the asymmetry of the line profile by computing the difference in average velocity in the upper and lower portions of the CCF (green circles in the middle panel). The contrast characterizes the depth of the CCF by taking the average of the CCF in the continuum (diamonds) and its difference with the CCF flux at the stellar velocity  $V_0$ .

Ca II doublet. The S-index is normalized by total flux in the  $B$  and  $V$  continuum bands which are 20 Å wide broad bands centered on 3900 and 4000 Å respectively. From the S-index, the  $\log R'_{\text{HK}}$  indicator is derived using a variety of formulations that are all proportional to the S-index and attempt to isolate the chromospheric component of  $\Psi_H$  and  $\Psi_K$  from the photospheric component using normalization factors that are dependent on  $T_{\text{eff}}$  and level of stellar activity (Lovis et al., 2011). The  $\log R'_{\text{HK}}$  indicator has been well characterized as a sensitive probe of activity on both Sun-like stars and M dwarfs (e.g. Wright et al., 2004; Lovis et al., 2011; Astudillo-Defru et al., 2017a).

$H\alpha$ : similarly to the  $\log R'_{\text{HK}}$  index,  $H\alpha$  photons ( $\lambda = 6562.81$  Å) are emitted by hot plages and thus act as a tracer of stellar activity originating from the chromosphere.  $H\alpha$  may act as a more suitable spectroscopic activity indicator for cool M dwarfs which lack significant flux in the Ca II H & K lines (Robertson et al., 2014). In practical terms, similar spectral features tracing a star's chromospheric structure may alternatively be used as activity indicators depending on the accessible wavelengths of the employed spectrograph (e.g. He I, Na I, etc).

*CCF shape parameters:* the cross-correlation function (CCF) of a stellar spectrum represents its average line profile at typically a high S/N. As was seen in Fig. 1.8, the presence of active regions distorts the Gaussianity of the CCF thus alluding to the presence and nature of those active regions. A number of shape parameters to the fitted CCF may be used to characterize stellar activity using the same CCF from which the stellar RVs themselves are derived. However, these spectroscopic diagnostics are only useful when they are robustly derived from high S/N spectra of active stars where their effects on the CCF are clearly discernible (Desort et al., 2007). Examples of three shape parameters are visualized in Fig. 2.2 and described below.

*FWHM:* the full width half maximum (FWHM) of the CCF is commonly used as a measure of CCF's

width. As an active region transverses across the differentially Doppler-shifted limbs of a rotating star its temperature contrast with the stellar photosphere adds additional power to the CCF in the velocity direction opposite the occulted limb. This results in a broadening of the CCF on one side of its mean velocity  $V_0$ . As illustrated in the first panel in Fig. 2.2, the FWHM can be measured from the width of the CCF at a specified location.

*BIS:* recall that the broadening of the CCF due to a single active region is asymmetric about  $V_0$ . A common shape parameter to characterize the resulting asymmetry is the bisector (BIS; Queloz et al., 2001). Similar parameters have also been proposed as measures of CCF asymmetry such as the curvature of the bisector line (Hatzes, 1996), the velocity span (Boisse et al., 2011), and the BIS inverse slope (Queloz et al., 2001). Their definitions are closely related and rely on computing the weighted velocity in the upper and lower portions of the CCF (see middle panel of Fig. 2.2) and calculating their difference. A symmetric line profile will therefore have a BIS= 0 whereas asymmetric profiles affected by active regions will be non-zero and whose exact value and units will depend on the definition of the asymmetry parameter used.

*Contrast:* (Boisse et al., 2009) noted that increasingly active stars, based on their fractional coverage by dark star spots, tend to have shallower line profiles. This is the direct result of such stars featuring more significant contributions to their average line profiles from cool and therefore dimmer regions. This effect is often characterized by the CCF contrast which is equal to the difference between the average baseline of the CCF in the continuum (located many FWHMs away from  $V_0$  in the CCF wings) and the flux at the stellar velocity  $V_0$ .

### 2.1.3 Photometric Modelling: the $FF'$ Method

An alternative to using simultaneous spectroscopic activity indicators to temporally de-correlate RV measurements for the effects of stellar activity is to use contemporaneous photometry. Aigrain et al. (2012) present the  $FF'$  formalism that uses photometry to model two distinct RV activity components resulting from the active regions that give rise to photometric variability. In this way the photometry is used to infer a quantity that is not directly observable (i.e. the fractional spot coverage) from which various RV activity components are derived using the  $FF'$  formalism (Rajpaul et al., 2015).

Aigrain et al. (2012) argue that in the limit of small active regions with non-complex configurations that the fractional coverage of the stellar disk by active regions  $F(t)$  is related to the photometric flux  $\Psi(t)$  via

$$\Psi(t) = 1 - F(t). \quad (2.2)$$

Eq. 2.2 reads that the fractional spot coverage is just one minus the observed stellar flux in normalized units such that the star's brightness equals unity in the absence of active regions. The corresponding RV signals from active regions, that also give rise to the photometric variations, are known as the rotation effect  $\Delta RV_{\text{rot}}$  and the convective blueshift effect  $\Delta RV_{\text{conv}}$ . The former arises from active regions occulting the differentially Doppler-shifted stellar limbs and thereby suppressing the flux of photons with a particular Doppler-shift. Because the rotation effect affects the RVs in proportion to the spot coverage

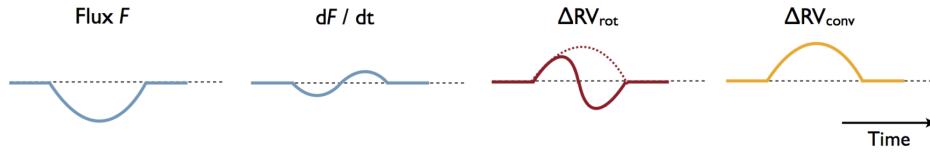


Figure 2.3: The  $FF'$  method for deriving the rotation and convective blueshift components of stellar RV activity from contemporaneous photometry. Here the photometric stellar flux is used to infer the fractional spot coverage  $F(t)$ . The rotation effect scales with  $F(t)\dot{F}(t)$  and thus varies with half the rotation period while the convective blueshift effect scales as  $F^2(t)$ . (Image credit: [Haywood, 2015](#)).

and exhibits a sign change as the spot crosses from the blueshifted limb to the redshifted limb over a single rotation cycle, the rotation effect is

$$\Delta RV_{\text{rot}}(t) \propto F(t)\dot{F}(t), \quad (2.3)$$

where the first time derivative of the fractional active region coverage is  $\dot{F}(t) = -\dot{\Psi}(t)$  from Eq. 2.2. As such, the rotation effect tends to dominate on rapidly rotating stars ( $P_{\text{rot}} \lesssim$  a few days) wherein  $\dot{F}(t)$  is large. Conversely, the convective blueshift effect operates by active regions disrupting the homogeneity of the stellar disk on which convective blueshift is ongoing everywhere. This disruption scales with the fractional spot coverage and the angle between the spot normal and the line-of-sight which results in

$$\Delta RV_{\text{conv}}(t) \propto F^2(t). \quad (2.4)$$

The convective blueshift effect tends to dominate more slowly rotating spotted stars such as the Sun ([Haywood et al., 2016](#)).

Fig. 2.3 demonstrates both the rotation and convective blueshift effects on the RVs and how they are related to the photometric detriment incurred by a single spot over a portion of one stellar rotation cycle.

#### 2.1.4 Pre-whitening

This method aims at identifying strong periodicities in an RV time series by modelling said periodic signals with sinusoids and iteratively removing them until only residual white noise remains. This method is known as pre-whitening and is based on analyzes of multi-periodic stellar oscillation modes. Pre-whitening operates by performing a Fourier transform of the data and removing sinusoidal functions from the raw data whose periods are identified by the frequency analysis ([Queloz et al., 2009](#)). Pre-whitening however assumes that all signals are long-lived and strictly periodic which is not always applicable to the stellar activity signals that this method aims to mitigate.

#### 2.1.5 Deterministic Model Fitting

RV time series exhibiting strong coherent signals at  $P_{\text{rot}}$  or one of its low order harmonics<sup>1</sup> may be well modelled by a deterministic function in the form of a sinusoid at  $P_{\text{rot}}$ ,  $P_{\text{rot}}/2$ ,  $P_{\text{rot}}/3$ , etc ([Boisse](#)

<sup>1</sup>For example the first harmonic of  $P_{\text{rot}}$  can dominate for rapidly rotating spotted stars whose RV activity is dominated by the rotation effect (see Sect. 2.1.3).

(et al., 2011). In general, such models are incomplete because they neglect any temporal evolution of active regions that are known to vary in their sizes, temperatures, and spatial distributions over adjacent rotation cycles (Giles et al., 2017). These nuances are not captured in deterministic sinusoidal models although this methodology may still be applicable for time series with short observational baselines, compared to  $P_{\text{rot}}$ , of stars with long-lived active region groups.

### 2.1.6 Physical Models of Active Regions

Another seemingly logical approach to modelling stellar activity is to parameterize a physical model of active regions and sample those parameter posteriors jointly with planet models. This direct spot modelling has been considered in Giguere et al. (2016) who fit a multi-spot model to contemporaneous photometry and spectroscopic time series to determine the physical nature of spots on the early K dwarf  $\epsilon$  Eridani. This method however suffers from a number of issues with attempting to sample such a large parameter space such as the large computational requirements and inherent degeneracies in spot parameters (Giguere et al., 2016). Other prominent packages exist for generating forward models of observables from physical spot models (SOAP; Boisse et al. 2012, SOAP 2.0; Dumusque et al. 2014a, StarSim; Herrero et al. 2016) although I am unaware of their usage for the purpose of directly modelling spots in RV time series.

### 2.1.7 Line by line radial velocities

Different spectral lines are known to be uniquely affected by stellar activity due to their unique sensitivities to pressure, temperature, embedded magnetic field strength, and local fluid velocities, all of which are influenced by stellar activity (Davis et al., 2017; Wise et al., 2018). These effects from stellar activity on spectral line shapes are distinct from Doppler-induced line shifts that are identical to all lines. Dumusque (2018) investigated the prospect of deriving stellar RVs line by line. The resulting RVs derived from individual lines were shown to be consistent with the archival RVs derived from the HARPS data reduction software. The derivation of RVs in this way is postulated to mitigate anomalous RV signals from stellar activity if spectral features that are minimally affected by activity are utilized. Intelligent identification of lines that are minimally sensitive to activity remains an active area of research.

Davis et al. (2017) also demonstrated the power of analyzing individual spectral lines to identify the effect of stellar activity on measured RVs. Principal component analysis was run on sets of idealized spectra from stars hosting either an equatorial spot, an equatorial facula, or a planetary companion on a circular orbit. As shown in Fig. 2.4, the maximum variance contained in the first principal component of the Doppler-shifted spectrum (i.e. with no activity) is maximized where the slope of spectral features is greatest and is identical to all spectral features as they are each equivalently affected by Doppler shifts. Lines in the remaining synthetic spectra that are affected by either a spot or facula are not identically affected due to the differential sensitivity of various lines to their physical environment that in turn is influenced by the presence of active regions. Fig. 2.4 reveals examples of strongly varying lines from Ti I and Ni I at 5009.5 Å and 5010.8 Å respectively. Davis et al. (2017) also quantify the wealth of information contained in the spectra as a function of S/N and spectral resolution and argue for the need to obtain high resolution spectra to finely resolve line profiles and exploit their inherent activity information content.

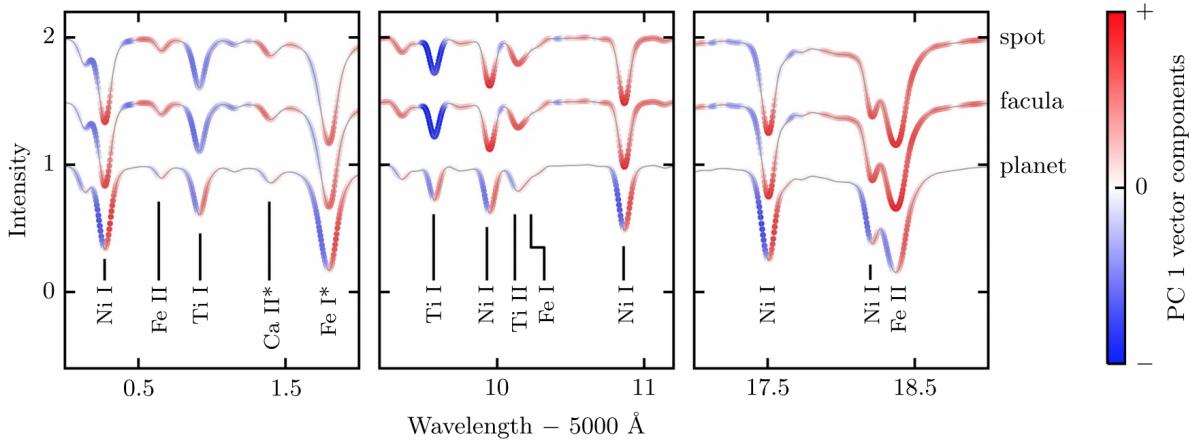


Figure 2.4: The values of the first principal component for three spectra of a star with either an equatorial spot, an equatorial facula, or a circular Keplerian orbit. The maximum variance (with either sign) in the Doppler-shifted spectrum is dominated by regions with the greatest slope and is the same for all spectral features. Variance in the active spectra and not equal within each line as not all lines respond identically to in the presence of active regions. (Image credit: [Davis et al., 2017](#))

## 2.2 Gaussian Process Regression for Semi-Parametric Activity Modelling

Many of the shortcomings of the activity mitigation techniques discussed in the previous sections are based on their incompleteness and their inability to self-consistently characterize model uncertainties. For example, correlations with activity-sensitive time series are often incomplete because the chosen indicator is not sensitive to all activity sources present in the RVs and often result in large RV residuals after de-correlation. Similarly, deterministic activity models lack sensitivity to stochastic changes in activity levels due to the evolution of active regions between adjacent rotation cycles or over significant subsets of long-term magnetic cycles.

Gaussian process regression models represent one such tool that aims to overcome many of these shortcomings by jointly modelling planets and activity in a self-consistent manner. This technique, originally pioneered on active Sun-like stars (e.g. [Haywood et al., 2014](#); [Grunblatt et al., 2015](#); [Faria et al., 2016](#); [López-Morales et al., 2016](#)), has been shown to reconcile disparate solutions between observations from multiple spectrographs ([Rajpaul et al., 2017](#); [Cloutier et al., 2019b](#)), to reconcile RV solutions for planetary systems whose favoured models are ambiguous ([Rajpaul et al., 2017](#); [Cloutier et al., 2019b](#)), and to disentangle neighbouring periodic signals from planets, stellar rotation, and/or from window functions ([Rajpaul et al., 2016](#); [Cloutier et al., 2017b](#)). In the following sections I will give an overview of what Gaussian processes are and how I implement their formalism in the context of modelling stellar activity in RV time series.

Before proceeding it is important to note that Gaussian process regression models of activity are not the be all end all when it comes to activity modelling in RV time series. As will be argued throughout this thesis though, Gaussian process activity modelling does provide a tractable framework with many

desirable properties for deriving models of stochastic activity signals in RVs.

### 2.2.1 The one-dimensional Gaussian distribution

A Gaussian process (GP) is defined as *a collection of random variables, any finite subset of which has a joint Gaussian distribution*. In other words, any random process<sup>2</sup> for which all finite subsets have a multivariate Gaussian distribution, is a Gaussian process (Rasmussen & Williams, 2005). The following discussion is intended to help the reader develop an intuition of what a Gaussian process is. Much of the information provided and the pedagogical approach taken throughout was acquired from a video lecture by Prof. David MacKay<sup>3</sup>.

In order to develop a visual intuition of the definition of a Gaussian process let us first consider a Gaussian random variable in one dimension. So, imagine some random process that draws a single value of a random variable  $X$  for each realization of an experiment. As an astronomer with a life outside of astronomy, and furthermore as an astronomer who spends many of his off-work hours being pelted with disks of vulcanized rubber, my favourite Gaussian random variable has to be the ‘goals against average’ (GAA) statistic:

$$X = GAA \equiv \frac{\text{total number of goals against}}{\text{total number of games played}}. \quad (2.5)$$

As a Gaussian random variable, the results of repeated experiments (i.e. repeated draws from the empirical *GAA* distribution for various goaltenders) can be expressed by the Gaussian probability density function (PDF) of the form

$$\mathcal{N}(X|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right), \quad (2.6)$$

that is parameterized by  $\mu = \text{mean}(X)$  and  $\sigma = \text{std}(X)$ ; the mean and standard deviation of  $X$ . As shown in Fig. 2.5, the empirical distribution of the *GAA* for goaltenders in the National Hockey League (NHL) over the past five seasons closely resembles a Gaussian distribution with a mean of 2.64 and a standard deviation of 0.40 goals against per game. Hence, my favourite Gaussian random variable.

Before a new season begins we can write down a prior distribution on the *GAA* based on the prior knowledge we have obtained from previous NHL seasons; i.e.  $\mathcal{N}(GAA|2.64, 0.40)$ . This distribution has a unique expectation value and variance describing what we expect a goaltender’s *GAA* to be in the up-coming season and, as crucially, how precisely that value is constrained.

### 2.2.2 The two-dimensional Gaussian distribution

Of course we can introduce a second Gaussian random variable and extend the concept of the one-dimensional Gaussian distribution to two dimensions. This is done by recasting the expression for the Gaussian PDF in  $k > 1$  dimensions where  $k = 2$  in this scenario. Rewriting Eq. 2.6 in  $k$  dimensions is achieved by replacing the scalar mean and standard deviation with a mean  $k$ -vector  $\mu$  and a  $k \times$

---

<sup>2</sup>A random process is any collection of random variables indexed by an independent variable such as time.

<sup>3</sup><https://www.youtube.com/watch?v=NegVuuHwa8Q>

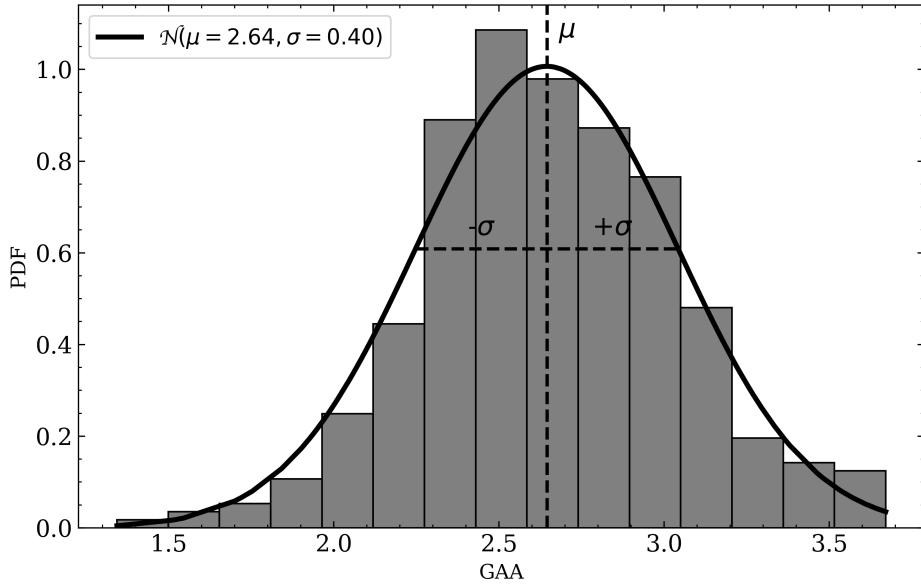


Figure 2.5: The distribution of  $GAA$  for NHL goaltenders over the past five seasons as an example of a Gaussian random variable. The solid line depicts the Gaussian probability density function with mean and standard deviation 2.64 and 0.40 goals against per game.

$k$  covariance matrix  $\mathbf{K}$  respectively. The updated expression for the so-called multivariate Gaussian distribution in  $k$  dimensions is

$$\mathcal{N}(\mathbf{X}|\boldsymbol{\mu}, \mathbf{K}) = \frac{1}{\sqrt{(2\pi)^k |\mathbf{K}|}} \exp\left(-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu})^T \mathbf{K}^{-1} (\mathbf{X} - \boldsymbol{\mu})\right). \quad (2.7)$$

The  $k$ -vector  $\boldsymbol{\mu}$  represents the mean value of each random variable in  $\mathbf{X} = (X_1, \dots, X_k)^T$ . The elements of the covariance matrix  $\mathbf{K}$  represent the covariances between each pair of random variables. The diagonal elements of  $\mathbf{K}$  therefore represent the covariance of each random variable with itself, or equivalently the variance of that random variable  $\sigma^2$ . The covariance between two random variables  $X_1$  and  $X_2$  can be computed explicitly in terms of their expectation values  $E[X_1]$  and  $E[X_2]$  via

$$\text{cov}(X_1 X_2) = E[(X_1 - E[X_1])(X_2 - E[X_2])] \quad (2.8)$$

$$= E[X_1 X_2] - E[X_1]E[X_2]. \quad (2.9)$$

Note that the expectation value of a Gaussian random variable is simply its mean.

### Uncorrelated variables

Consider the special case of when the two Gaussian random variables under consideration  $X_1$  and  $X_2$  are uncorrelated; i.e.  $\text{cov}(X_1 X_2) = 0$ . A quick example of this is to let  $X_1 = GAA$  and sample  $X_2$  from an independent Gaussian distribution with zero mean and unit variance. The nature of sampling  $X_2$  in

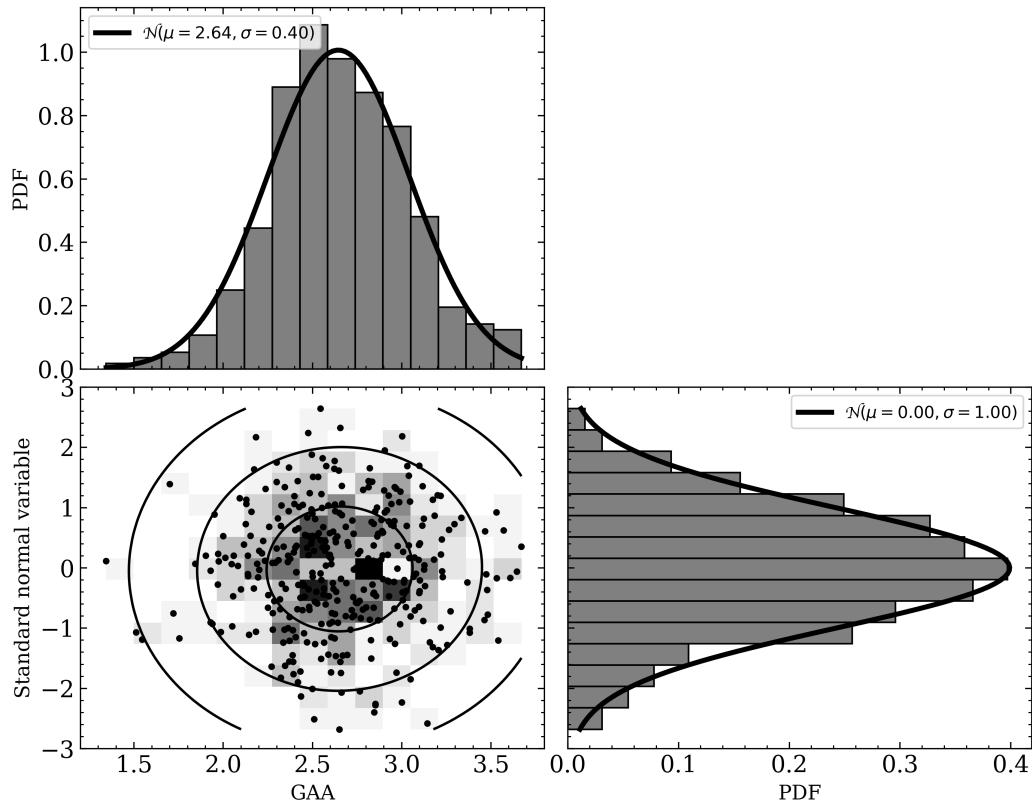


Figure 2.6: The marginalized and joint distributions of the Gaussian random variables  $GAA$  and draws from the standard normal distribution (i.e.  $\mathcal{N}(0, 1)$ ). The solid lines overlaid on their joint distribution represent the 1, 2, and  $3\sigma$  contours. The covariance matrix (Eq. 2.10) has off-diagonal elements equal to zero indicating that the two variables are uncorrelated; a property that is largely discernible from the lack of correlation in their joint distribution.

this way ensures that  $X_1$  and  $X_2$  are indeed uncorrelated. The corresponding covariance matrix

$$\mathbf{K}(GAA, X_2) = \begin{bmatrix} 0.15 & 0.00 \\ 0.00 & 0.98 \end{bmatrix} \quad (2.10)$$

is calculated using Eq. 2.9, is diagonal, and confirms that  $X_1$  and  $X_2$  are uncorrelated as the off-diagonal elements of  $\mathbf{K}$  are zero. This fact can also be visually discerned in the joint  $X_1 X_2$  distribution in Fig. 2.6. Because the variables  $X_1$  and  $X_2$  are uncorrelated in this example, the measurement of a value of  $X_1$  does nothing to inform us of its corresponding  $X_2$  value. That is that the prior distribution  $\mathcal{N}(\mathbf{X}|\boldsymbol{\mu}, \mathbf{K})$  is uninformative with regards to  $X_2$  in a sample for which the corresponding value of  $X_1$  is known.

### Correlated variables

But what if next we consider an alternate Gaussian random variable. Namely  $X_2 = SV\%$  where

$$SV\% \equiv \frac{\text{total number of saves made}}{\text{total number of shots faced}} \quad (2.11)$$

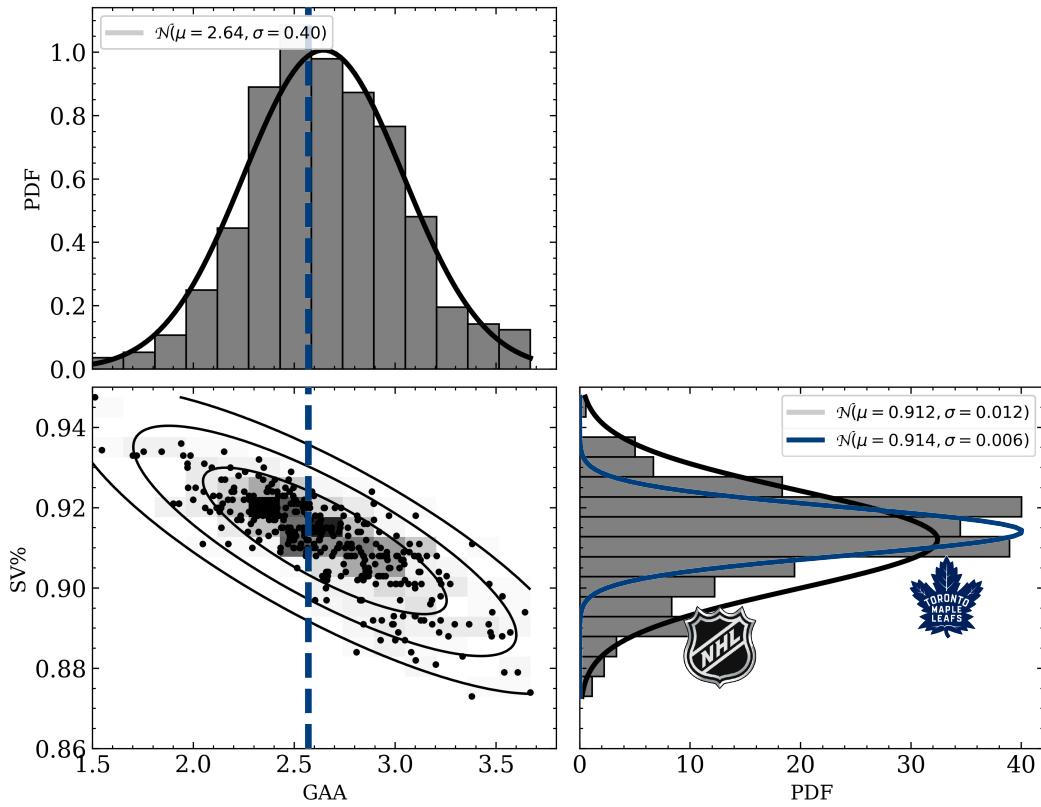


Figure 2.7: The marginalized and joint distributions of the Gaussian random variables  $GAA$  and  $SV\%$ . The covariance matrix (Eq. 2.12) has negative off-diagonal elements describing the degree of anti-correlation between the two variables; a property that is largely discernible from their joint distribution as high values of the  $SV\%$  tend to correspond to a lower  $GAA$ . The correlation can be used to inform the value of one variable given a measurement of the other. This is demonstrated as the measured  $GAA$  for the Toronto Maple Leafs' goaltender (vertical dashed line) provides some additional information on the corresponding  $SV\%$  whose posterior given the measured  $GAA$  is more tightly constrained than the  $SV\%$  distribution for the entire NHL.

is a measure of a goaltender's save percentage. In this scenario one might expect the variables  $GAA$  and  $SV\%$  to have some degree of correlation as a "good" goaltender who boasts a low  $GAA$  probably does so because of their high  $SV\%$ . Indeed the covariance matrix of the  $GAA$  and  $SV\%$  for NHL goaltenders over the past five seasons is

$$\mathbf{K}(GAA, SV\%) = \begin{bmatrix} 0.1496 & -0.0037 \\ -0.0037 & 0.0001 \end{bmatrix} \quad (2.12)$$

and has non-zero off-diagonal elements. Furthermore, the off-diagonal elements are negative which is indicative of the anti-correlation between the  $GAA$  and the  $SV\%$  that was just postulated. This strong correlation is easily visualized in their joint distribution in Fig. 2.7.

The fact that the two variables are correlated implies that knowledge of one variable provides some additional constraining power on the value of the second. This demonstrated in Fig. 2.7 wherein we measure the value of the  $GAA$  for the Toronto Maple Leafs' goaltender Freddie Andersen to be 2.57

goals per game. Not bad. Not great, but not bad. By measuring this value we can establish a posterior distribution on Freddie's corresponding  $SV\%$  given his  $GAA$ :  $p(SV\%|GAA = 2.57)$ . Because the variables are dependent on each other according to Eq. 2.12, this posterior distribution is narrower (i.e. more precise) than the full  $SV\%$  distribution. Indeed this is evidenced in the lower right panel of Fig. 2.7 that compares the two  $SV\%$  distributions and reveals that the dispersion in  $p(SV\%|GAA = 2.57)$  has approximately half of the dispersion as does the full  $SV\%$  distribution. The degree of improvement in the predictive distribution of  $X_2$  given  $X_1$  is dependent on how strongly correlated the two variables are.

### 2.2.3 The $k$ -dimensional Gaussian distribution

So far we have considered a pair of correlated Gaussian random variables whose PDF is given by Eq. 2.7 for  $k = 2$  and whose covariance matrix is Eq. 2.12. Next we want to extend the multivariate Gaussian distribution to arbitrarily large  $k$  and visualize that distribution which cannot be done with the conventional contour-based visualization (Figs. 2.6 & 2.7) beyond two dimensions. To resolve this restriction we now adopt an alternative visualization of a multivariate Gaussian distribution that I refer to as its *linear representation*.

This new visualization is depicted in Fig. 2.8 for the  $GAA$  and the  $SV\%$  alongside the conventional representation. Specifically, Fig. 2.8 depicts a finite number of samples from the two-dimensional Gaussian distribution shown in Fig. 2.7. Each identically coloured line in the linear representation depicts the values of the scaled  $GAA$  and the  $SV\%$ <sup>4</sup> for a single sample from their joint distribution. Furthermore, the correlation between the two scaled variables, that we now insist be a positive correlation, is qualitatively apparent as a low value of the scaled  $GAA$  tends to correspond to a low value of the scaled  $SV\%$ . If the two variables were weakly correlated then we would expect more crossing of curves in the linear representation than is seen in Fig. 2.8. What's more is that the distribution of values in the vertical direction represent samples from each variable's one-dimensional marginalized distribution such that the samples are indicative of the variable's mean and standard deviation.

The benefit of adopting the new linear representation is that it can be extended to arbitrarily large dimensions. Fig. 2.9 depicts the linear representation of a  $k = 5$  dimensional space that includes three new Gaussian random variables that are defined to be positively correlated with the scaled  $GAA$  and  $SV\%$ . The left panel in Fig. 2.9 depicts samples from the five-dimensional Gaussian distribution as well as each variable's mean and  $1\sigma$  dispersion. The right panel of Fig. 2.9 illustrates the effect of how the measurement of one of the variables informs the posteriors of the other variables as a result of their correlation. That is that when the third variable in Fig. 2.9 is measured and its uncertainty goes to zero, the posterior distributions on the remaining variables are narrowed. As more and more variables are measured, constraints on the unmeasured variables would similarly increase in a way that is quantified by the correlations between the measured and unmeasured variables. The notion that the value of an unseen variable can be informed by the measurement of another, if those variables are correlated, has major implications for predictive models as we shall see in Sect. 2.2.4.

---

<sup>4</sup>Note that the  $GAA$  and  $SV\%$  are scaled to a common unit such that they span roughly the same region of the y-axis in the Fig. 2.8 linear representation. The scaling is also such that the correlation between the scaled  $GAA$  and  $SV\%$  is now positive (i.e. not anti-correlated).

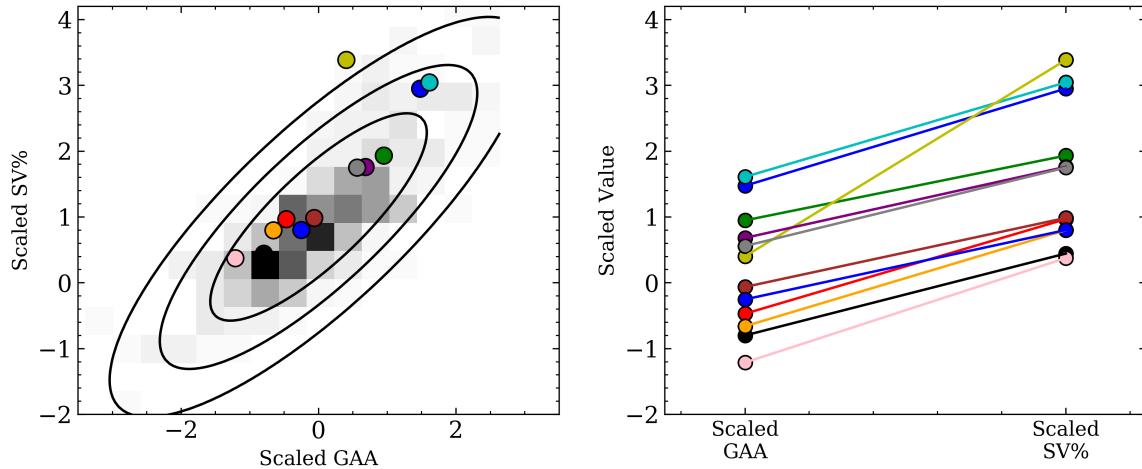


Figure 2.8: Visual comparison of samples from the two-dimensional Gaussian distribution depicted in the conventional way (*left panel*) and in the new linear representation (*right panel*). The  $GAA$  and  $SV\%$  variables are scaled to a common unit with a positive correlation and offset for clarity.

## 2.2.4 Gaussian process regression modelling

Upon inspection of Fig. 2.9, one might imagine changing the “variables” axis to an independent variable such as time and the “scaled values” axis to an observable such as stellar RVs. In this way we would be looking at a time series and particularly at the correlation between adjacent measurements. Because RV signals arising from active stars are temporally correlated as active regions rotate in and out of view of the observer, it is clear that modelling the correlation between adjacent RV observations could be used to inform our understanding of that signal in a way that is independent of attempting to model the physical nature of the signal.

Indeed one-dimensional GP regression models are a convenient way to model correlations between observations in a semi-parametric way. In particular, GPs provide a flexible framework to perform Bayesian inference on functions. Their flexibility stems from their semi-parametric nature which makes them well-suited to modelling stochastic processes like stellar activity that lack a deterministic functional form. GP regression works to model the covariance between data points simultaneously with deterministic components such as planets on Keplerian orbits. Given a parameterization of the data covariance structure, the GP is by definition a multivariate Gaussian distribution with a well-defined mean function and variance that represent our model of the correlated observables and its uncertainty (Rasmussen & Williams, 2005).

### Covariance kernel functions

Modelling of temporally correlated stellar activity with a GP regression model first requires a parameterization of the data covariance. This parameterization of an analytical covariance kernel through a

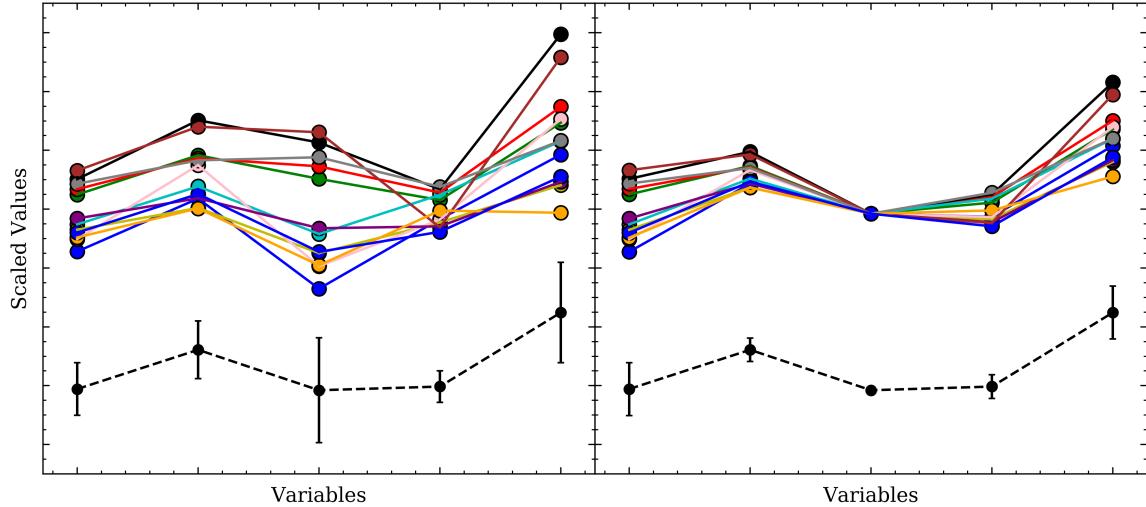


Figure 2.9: The linear representation of samples from a five-dimensional Gaussian distribution with each sample depicted with a particular colour in each panel. The mean and  $1\sigma$  dispersion of each variable is depicted by the offset black markers and error bars. The *right panel* illustrates how the measurement of the third variable constrains the values of the remaining variables due to their correlations.

small set of hyperparameters<sup>5</sup> is why GP regression models are often referred to as semi-parametric. The covariance matrix of the GP is written in terms of the covariance kernel function  $k(t, t')$  as

$$\mathbf{K}_{ij} = k(t_i, t_j) + \delta_{ij}(\sigma_{\text{RV}}^2(t_i) + s^2) \quad (2.13)$$

where the indices  $i, j$  run up to the number of RV measurements (i.e.  $i, j = 1, \dots, N_{\text{RV}}$ ). Because of the Dirac delta function  $\delta_{ij}$  in Eq. 2.13, the diagonal of the covariance matrix  $\mathbf{K}$  includes direct contributions from the RV measurement uncertainties  $\sigma_{\text{RV}}$  plus an additive scalar jitter  $s$  designed to absorb excess white noise from systematics and/or from other temporally independent noise sources.

A number of covariance kernels have been proposed for the treatment of stellar RV activity and are summarized below (Grunblatt et al., 2015). Their corresponding covariance matrices are also shown in Fig. 2.10.

*Squared exponential:*

$$k_{\text{SE}}(t_i, t_j) = a^2 \exp \left[ -\frac{(t_i - t_j)^2}{2\lambda^2} \right], \quad (2.14)$$

has two hyperparameters  $\Theta = \{a, \lambda\}$ . Namely, the covariance amplitude  $a$  in units of the observable, and an exponential timescale  $\lambda$  describing how much time is required for measurements to effectively forget previous values. With this covariance kernel, measurements made closely in time are more highly correlated than well separated measurements and the degree of correlation for a given time separation will be larger for smaller  $\lambda$ .

<sup>5</sup>“Small” relative to the number of RV measurements.

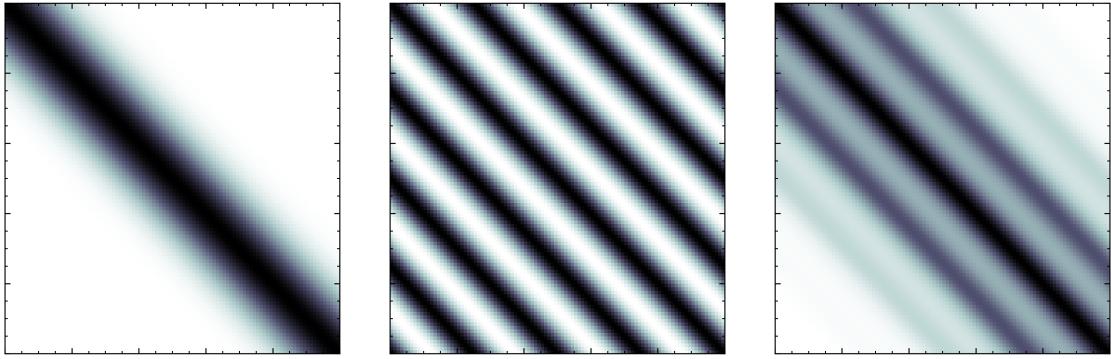


Figure 2.10: Examples of covariance matrices with a squared exponential covariance kernel (*left*), a periodic covariance kernel (*middle*), and a quasi-periodic covariance kernel (*right*). In the latter example the exponential timescale exceeds the periodic timescale such that the nearly periodic nature of the covariance is still discernible for points closely separated in time (i.e. near the matrix diagonal).

*Periodic:*

$$k_P(t_i, t_j) = a^2 \exp \left[ -\Gamma^2 \sin^2 \left( \frac{\pi|t_i - t_j|}{P} \right) \right], \quad (2.15)$$

has three hyperparameters  $\Theta = \{a, \Gamma, P\}$  that includes the familiar covariance amplitude  $a$ . The periodic kernel also contains a coherence parameter  $\Gamma$  and a periodic timescale  $P$  within the sinusoidal term. With this covariance kernel, measurements that are separated by an integer multiple of  $P$  will be highly correlated.

*Quasi-periodic:*

$$k_{QP}(t_i, t_j) = a^2 \exp \left[ -\frac{(t_i - t_j)^2}{2\lambda^2} - \Gamma^2 \sin^2 \left( \frac{\pi|t_i - t_j|}{P} \right) \right], \quad (2.16)$$

is the product of the squared exponential and periodic covariance kernels and therefore contains all four unique hyperparameters among the two kernels  $\Theta = \{a, \lambda, \Gamma, P\}$ .

Stellar activity on M dwarfs is dominated by active regions (Lindegren & Dravins, 2003)<sup>6</sup> and the resulting RV signals evolve in a manner that is easy to understand qualitatively. Namely, we can expect a periodic component to the RV correlations because active regions rotate in and out of view at the stellar rotation period although the periodicity of the correlations need not be at  $P_{\text{rot}}$  and may instead appear at a low order harmonic of  $P_{\text{rot}}$  such as in the case of rapidly rotating spotted stars (see Sect. 2.1.3). It is worth noting that a precise value of  $P_{\text{rot}}$  may not exist for stars with significant differential rotation although that effect is known to be small for cool M dwarf stars (Barnes et al., 2005). But in addition to these periodic variations we can also expect activity arising from active regions to not be strictly periodic because of active region evolution through which their sizes, temperatures, and spatial distributions will vary. Thus we have a physical justification for the adaption of a quasi-periodic covariance kernel as it includes the first-order periodic structure of the activity signal as well as the squared exponential term

<sup>6</sup> And sometimes flares but those occur on short timescales such that they rarely affect low cadence RV observations for almost all but the most active flare stars.

whose timescale can be thought of as characterizing the lifetimes of the active regions.

The optimization of the quasi-periodic covariance kernel over competing kernel functions was demonstrated by [Grunblatt et al. \(2015\)](#) on the active Sun-like star Kepler-78. However, it should be noted that due to the semi-parametric nature of GP regression models, the exact choice of kernel and/or the exact values of its hyperparameters often do not matter as much as one might think. Fig. 2.11 illustrates this notion by modelling a set of synthetic data points with each of the three covariance functions defined above. The synthetic data are generated by a sinusoidal function plus a decreasing linear trend. Each GP does a qualitatively similar job at modelling the data and with comparable  $1\sigma$  uncertainties. The optimal kernel choice often does not matter unless the intended use of the resulting GP predictive distribution is for predictions at previously unseen epochs in which case one should consider employing cross-validation practices to minimize over-fitting.

Also depicted in Fig. 2.11 is the effect of varying the hyperparameters of a quasi-periodic GP regression model to the same synthetic data. The first curve depicts the resulting GP model using the approximate optimized hyperparameters with a periodic timescale approximately equal to the injected timescale of 50 (arbitrary units). The second GP model has the same periodic timescale but whose exponential timescale is greater by nearly three orders of magnitude relative to the first curve. Despite this, the resulting mean GP appears to continue to be a good fit to the data. The final GP model reverts back to the optimized exponential timescale but decreases the periodic timescale from 50 to 20. The effect of a shorter periodic timescale is only apparent in the regions between data points as the resulting GP again appears to do a qualitatively similar job at modelling the data despite having widely varying hyperparameters.

These exercises were intended to demonstrate that for some datasets the exact form of the covariance kernel or the exact values of the hyperparameters do not always have a drastic effect on the model prediction at the epochs of observations. Although it is clear from Fig. 2.11 that predictions from each unique GP can vary widely at times far from the epochs of observation. Hence why the identification of over-fitting GPs is necessary if the GP regression models are being used for prediction as is commonly done in the world of machine learning.

### Hyperparameter estimation

Determination of the GP hyperparameters  $\Theta$  for a given covariance kernel is done by first specifying a mean model  $\mu$  which itself is parameterized by a set of model parameters  $\theta$ . In the case of modelling RVs, the mean model might represent our planet model whose Keplerian orbital parameters are what we are interested in measuring with the GP being used to model the correlated RV residuals arising from stellar activity. The residual vector is written as  $\mathbf{r} = \mathbf{y} - \mu$  where  $\mathbf{y}$  is the vector of  $N_{\text{RV}}$  RV measurements. To estimate all model parameters, including the GP hyperparameters and mean model parameters (i.e.  $\Psi = \Theta \cup \theta$ ), we sample their joint posterior according to Bayes theorem

$$p(\Psi|\mathbf{y}) \propto \mathcal{L}(\mathbf{y}|\Psi) \cdot \Pi(\Psi) \quad (2.17)$$

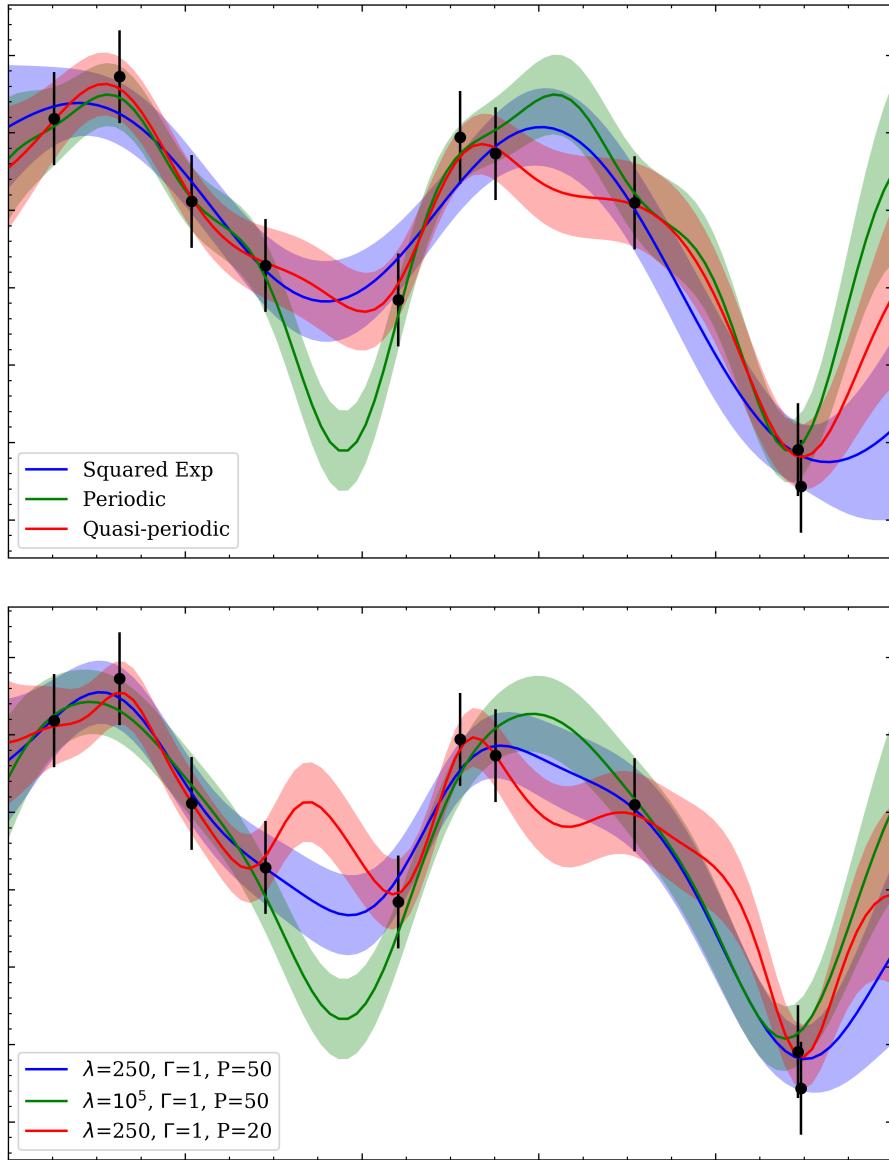


Figure 2.11: *Top panel:* the behaviour of three GP models with different covariance kernel functions on a synthetic dataset. *Bottom panel:* the behaviour of three quasi-periodic GP models with varying hyperparameters. Each GP, regardless of covariance kernel or hyperparameters, exhibits a qualitatively similar behaviour in the vicinity of the data points.

where

$$\ln \mathcal{L}(\mathbf{y}|\Psi) = -\frac{1}{2} (\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r} + \ln \det \mathbf{K} + N_{RV} \ln 2\pi) \quad (2.18)$$

is the logarithmic likelihood function given the set of model parameters  $\Psi$  and  $\Pi(\Psi)$  represents our prior on the model parameters. When estimating the GP hyperparameters in practice we often sample the logarithmic hyperparameters and adopt broad log uniform priors unless otherwise stated. Note that in Eq. 2.17 we have neglected the normalization factor that is not required for the purpose of parameter estimation but will be calculated in future chapters for the purpose of comparing competing models of a dataset.

Sampling of the  $\Psi$  posteriors is done using MCMC. In this way, our uncertainties in the Keplerian model parameters  $\theta$  are marginalized over our uncertainties in the GP hyperparameters  $\Theta$ . The exact strategies and details of our MCMC implementation are discussed in subsequent chapters as necessary. After sampling the  $\Psi$  posteriors, the values of the hyperparameters are obtained from the maximum a-posteriori point estimates of those posteriors that in turn are used to define our GP activity model.

### Model inference

From the set of maximum a-posteriori hyperparameters  $\Theta$ , we define a unique GP prior distribution. From this distribution we can sample functions or construct a predictive distribution conditioned on the data  $\mathbf{y}(\mathbf{t})$  and evaluated at some unseen epochs  $\mathbf{t}^*$ . Function prediction from such a predictive distribution is historically known as kriging. The mean function and variance of the predictive distribution are

$$\mu(\mathbf{t}^*) = \mathbf{K}(\mathbf{t}^*, \mathbf{t}) \cdot \mathbf{K}(\mathbf{t}, \mathbf{t})^{-1} \cdot \mathbf{y}(\mathbf{t}), \quad (2.19)$$

and

$$\sigma(\mathbf{t}) = \mathbf{K}(\mathbf{t}^*, \mathbf{t}^*) - \mathbf{K}(\mathbf{t}^*, \mathbf{t}) \cdot \mathbf{K}(\mathbf{t}, \mathbf{t})^{-1} \cdot \mathbf{K}(\mathbf{t}^*, \mathbf{t})^T. \quad (2.20)$$

Figs. 2.12 and 2.13 depict samples from a quasi-periodic GP prior and predictive distribution respectively. The latter is conditioned on a set of ten synthetic data points and whose mean function and standard deviation are computed using Eqs. 2.19 and 2.20. When modelling RV activity with a GP whose hyperparameters  $\Theta$  have been derived from MCMC, we take the “best-fit” of that multivariate Gaussian, namely its mean function, to be our activity model similarly to how maximum a-posteriori point estimates of Keplerian orbital parameters are used to compute a “best-fit” planet model.

### 2.2.5 Demonstrating the effectiveness of a GP activity model: a toy model

Here I will provide a quick demonstration of the effectiveness of using a semi-parametric GP to model the temporally correlated physical process of stellar RV activity. The forthcoming experiment is not intended to be representative of the performance of GP activity modelling in general but it is useful to demonstrate how a non-physical model in the form of a GP can aid in the accurate modelling of RV

Figure 2.12: Continuous samples from a quasi-periodic Gaussian process prior distribution over the continuous variable time (*x-axis*). The samples have arbitrary units. A discontinuity in the sampling occurs after 50 frames. This animation was generated using code written by João Faria (<https://joaofaria.space/blog/continuous-samples>).

Figure 2.13: Independent samples from a quasi-periodic Gaussian process predictive distribution conditioned on ten synthetic data points with arbitrary units and versus the independent variable time ( $x$ -axis). The mean function and  $1\sigma$  standard deviation are depicted as the dashed black curve and surrounding shaded region respectively.

planets and activity.

Firstly, I construct a set of synthetic RVs containing a planet, stellar activity, and noise, sampled using the actual window function of 71 HARPS observations of the M dwarf K2-18 (Cloutier et al., 2017b). The injected planetary signal takes the form of a circular Keplerian with  $P = 22$  days and  $K = 2 \text{ m s}^{-1}$ . The stellar activity signal is derived from a physical spot model from SOAP 2.0 (Dumusque et al., 2014a) that computes time series of the photometric, RV, and line shape parameters resulting from spots and/or plages on a rotating star. In this toy model I simulate a single equatorial spot with 1% fractional coverage of the visible stellar disk on a star with  $P_{\text{rot}} = 30$  days. White noise at the level of  $1 \text{ m s}^{-1}$  is also added to the RV time series. The injected time series and the synthetic RVs are shown in the top panel of Fig. 2.14. The injected activity signal is seen to have a peak-to-peak amplitude of  $\sim 8.5 \text{ m s}^{-1}$  and is dominated by the suppression of convective blueshift when the simulated starspot is present on the visible hemisphere. The RV signal due to activity is equal to zero when the lone starspot is not visible. The peak-to-peak amplitude of the injected Keplerian signal is seen to be approximately half that of the injected activity.

Next I impose an RV model that includes an activity component in the form of an untrained quasi-periodic GP plus a single Keplerian model of the planet that I treat as a transiting planet such that its linear ephemeris (i.e. period and phase) are well-constrained by its transit light curve. Note that untrained GPs can be dangerous as scenarios have been observed in which an untrained GP model of RVs absorbs some planetary signal and can thus result in misestimated planetary parameters (Ribas et al., 2018). I proceed by deriving the joint posterior of the GP hyperparameters and Keplerian parameters to compute the ‘best-fit’ GP and Keplerian models that can then be compared to their injected signal counterparts. The recovered models are shown in the bottom panel of Fig. 2.14. Here the mean GP model is depicted along with its 68% confidence interval and appears to closely resemble the injected activity signal despite being derived from a semi-parametric model of a physically induced signal.

Fig. 2.15 compares the recovered activity and planetary signals with the injected signals. The mean GP model appears to have approximately the same period and phase as the injected activity signal as it lies along the one-to-one line in Fig. 2.15. The residual RV rms between the recovered and injected activity signals is  $0.72 \text{ m s}^{-1}$  compared to the injected RV measurement uncertainties of  $1 \text{ m s}^{-1}$ . A notable exception to the inaccuracy of the GP activity model is in its failure to capture the pile up of injected RVs at  $0 \text{ m s}^{-1}$  corresponding to when the single simulated starspot is not visible. Although, the mean GP activity model at those epochs are often consistent within  $1\sigma$  of  $0 \text{ m s}^{-1}$ . Similarly the recovered Keplerian model is in phase and with the injected period owing to our treatment of that signal as one whose ephemeris is well-constrained by some unseen hypothetical data. The recovered semi-amplitude however is somewhat inaccurate at  $K = 1.83 \pm 0.55 \text{ m s}^{-1}$  which is underestimated compared to the injected semi-amplitude of  $2 \text{ m s}^{-1}$ . However it is clearly consistent within  $1\sigma$ . This  $\sim 10\%$  underestimation in the value of  $K$  causes the slight offset from unity in the slope of the line in Fig. 2.15 comparing the recovered and injected Keplerian signals.

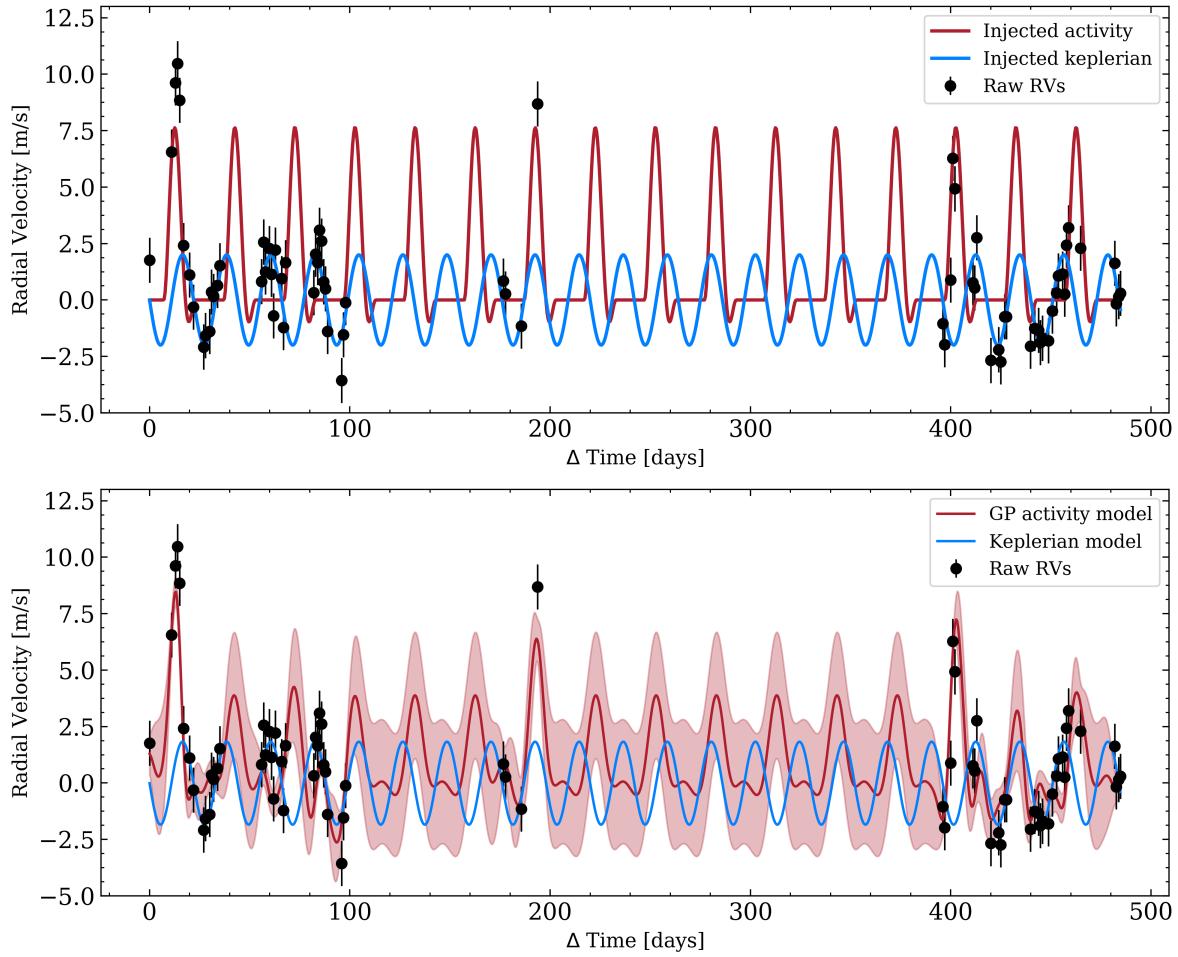


Figure 2.14: *Top:* a synthetic RV time series with injected stellar activity from a SOAP 2.0 physical model ( $P_{\text{rot}} = 30$  days; red curve), a Keplerian planetary signal ( $P = 22$  days,  $K = 2 \text{ m s}^{-1}$ ; blue curve), and noise. The RVs (black markers) are sampled with the real HARPS window function of K2-18. *Bottom:* the recovered GP activity model and Keplerian model using the maximum a-posteriori model parameters from MCMC analysis. The recovered models closely resemble their injected counterparts (see Fig. 2.15 for more details).

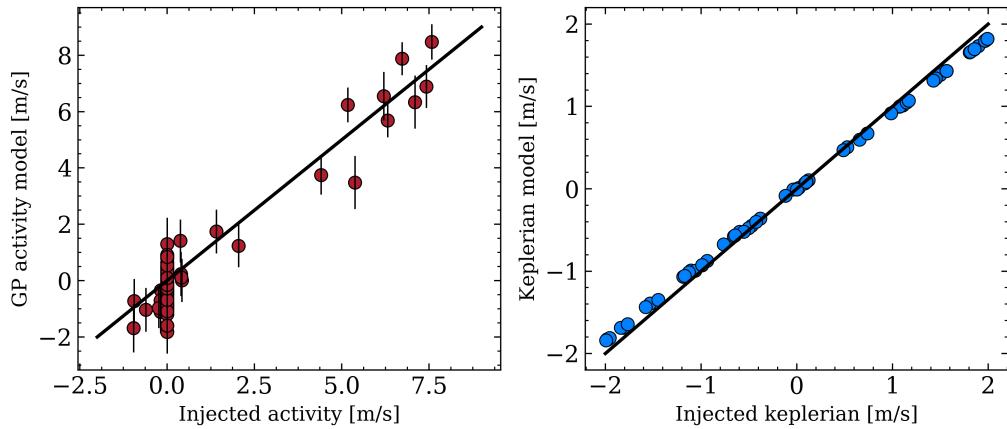


Figure 2.15: *Left panel:* comparison of the injected RV activity signal to the recovered GP activity model. The mean GP model appears to lie along the one-to-one line (*black line*) with a residual rms of  $0.72 \text{ m s}^{-1}$ . *Right panel:* comparison of the injected Keplerian signal to the recovered Keplerian model.

## 2.2.6 Gaussian process regression packages

Note that a number of packages exist for the efficient construction of GP prior distributions, function sampling, and matrix algebra required to compute the data likelihoods (see Eq. 2.18). Some such packages exist within the astronomy field as well as for implementation in broader contexts. Examples of GP packages includes `scikit-learn` (Pedregosa et al., 2011), `gpytorch` (Gardner et al., 2018; Wang et al., 2019), `george` (Ambikasaran et al., 2014; Foreman-Mackey, 2015), and `celerite` (Foreman-Mackey et al., 2017).

## 2.3 Point-form Thesis: Stellar Activity Modelling in Radial Velocity Time Series

- 2.1 An Overview of Techniques for Stellar Activity Mitigation: a multitude of techniques for the mitigation of stellar activity in RV time series exist but many of them suffer from incompleteness and/or inaccuracy.
- 2.2 Gaussian Process Regression for Semi-Parametric Activity Modelling: Gaussian process regression offers an alternative method to model stochastic activity signals within a semi-parametric Bayesian formalism that ensures consistency with modelled planetary signals.

# Chapter 3

## Predictions of planet detections with near infrared radial velocities in the up-coming\* SPIRou legacy survey-planet search

### 3.1 Introduction

The radial velocity method of detecting exoplanets is one of the most successful methods of exoplanet detection and has been widely used since the first discovery of an exoplanet around a main-sequence star over two decades ago ([Mayor & Queloz, 1995](#)). Since then numerous international teams have successfully built and used precision velocimeters to grow the population of radial velocity (RV) planets (e.g. HARPS; [Mayor et al. 2003](#), HARPS-N; [Cosentino et al. 2012a](#), HIRES; [Vogt et al. 1994](#)). Up until recently the majority of these precision velocimeters have operated in the visible wavelength regime where their sensitivity is maximized for the discovery of planets around Sun-like stars.

In recent years much interest has been generated regarding the population of exoplanets around M dwarfs. M dwarfs, with effective temperatures  $\lesssim 3800$  K and masses  $\lesssim 0.6 M_{\odot}$ , outnumber Sun-like stars in the Solar neighbourhood (within  $\sim 10$  pc) nearly 4:1 ([Henry, 2009](#)). Furthermore, M dwarfs are known to frequently host multiple small ( $r_p \leq 4 R_{\oplus}$ ) planets (e.g. [Dressing & Charbonneau, 2015](#); [Gaidos et al., 2016](#)) including a large fraction of planets within the star's habitable zone (HZ) which itself spans shorter orbital periods than around the more luminous Sun-like stars ([Kasting et al., 1993](#); [Kopparapu et al., 2013](#)). Lastly the amplitude of the radial velocity signal induced by a given planet is larger around M dwarfs than around Sun-like stars owing to their smaller masses. These favorable qualities have many astronomers committed to uncovering the M dwarf exoplanet population with purpose-built transit (e.g.

---

\*Now ongoing.

The contents of this chapter are copied verbatim from the published paper ([Cloutier et al., 2018b](#)).

MEarth; [Irwin et al. 2015](#), ExTrA [Bonfils et al. 2015](#), TRAPPIST; [Gillon et al. 2011](#), SPECULOOS; [Gillon et al. 2013](#)) and radial velocity instrumentation (e.g. SPIRou; [Delfosse et al. 2013b](#); [Artigau et al. 2014](#), NIRPS; [Bouchy et al. 2017](#), CARMENES; [Quirrenbach et al. 2014](#), HPF; [Mahadevan et al. 2012](#), IRD; [Tamura et al. 2012](#)).

One particular precision velocimeter optimized for the detection of exoplanets around M dwarfs in radial velocity is *SPIRou* (Un Spectro-Polarimètre Infra-Rouge; [Delfosse et al., 2013b](#); [Artigau et al., 2014](#)). SPIRou is a high-resolution near-infrared velocimeter whose first-light is scheduled on the Canada-France-Hawaii Telescope (CFHT) in 2018. A significant fraction of SPIRou’s allocated time will be spent surveying nearby M dwarfs searching for new exoplanets in a campaign known as the *SPIRou Legacy Survey-Planet Search* (SLS-PS). With the increased sensitivity to cool M dwarfs enabled by nIR detectors, the planet detections resulting from the SLS-PS will be able to constrain the occurrence rate of planets around stars later than  $\sim$  M4.5 and find the closest exoplanetary systems, beyond Proxima Centauri (1.3 pc; [Anglada-Escudé et al., 2016](#)), which may be amenable to direct imaging with the next generation of imagers on-board an Extremely Large Telescope (ELT).

In this study we present a comprehensive simulation of the SLS-PS to estimate its planet yield as well as the bulk properties of the detected SPIRou planet population. These simulations were performed for a variety of survey strategies which enabled the SPIRou science team to establish an experimental setup which optimizes both the detection sensitivity and survey yield given the nominal time allocation of the SLS-PS. The main results of this study are based on the survey version deemed to be optimal for meeting the science goals of SPIRou. However, the results of the various surveys are summarized in the final Sect. 3.13. Comparison of the various survey versions may be useful to inform other up-coming radial velocity planet searches similar to the SLS-PS.

The paper is organized as follows:

- Sect. 3.2 gives an overview of the important aspects of the SPIRou spectro-polarimeter.
- Sect. 3.3 describes the stellar input catalog.
- Sect. 3.4 describes the simulated SLS-PS.
- Sect. 3.5 describes the population of simulated planetary systems.
- Sects. 3.6-3.7 describe how we mitigate the effects of stellar activity and detect planets.
- Sects. 3.8-3.9 describe the results of the survey.
- Sect. 3.10 considers the effect of an increased planet frequency on the survey results.
- Sect. 3.11 describes how well we can measure planet occurrence rates based on the results of the SLS-PS.
- Sect. 3.12 discusses the potential for targeting SPIRou planets in direct imaging campaigns with ELTs
- and Sect. 3.13 compares the merits of various potential versions of SLS-PS with the fiducial version presented throughout this paper.

## 3.2 Un Spectro-Polarimètre Infra-Rouge

SPIRou is an up-coming nIR échelle spectro-polarimeter and high-precision velocimeter whose first light is scheduled for 2018 on the Canada-France-Hawaii Telescope on Maunakea. The instrument is optimized to observe exoplanets via the radial velocity technique around low mass stars and to study the magnetic fields of young embedded protostars (Delfosse et al., 2013b). The design of SPIRou is intended to address its main science goals of detecting and characterizing M dwarf exoplanetary systems and to investigate the role that magnetic fields have on the processes of star and planet formation. SPIRou can be considered a heritage instrument which is built upon the success of the previous generation of optical spectro-polarimeters and high-precision velocimeters such as the ESPaDONs spectro-polarimeter (Donati et al., 2006) as well as the RV spectrographs SOPHIE (Bouchy & Sophie Team, 2006) and HARPS (Mayor et al., 2003).

Here we provide a brief overview of the main instrument specifications as they pertain to the detection of new exoplanetary systems around nearby M dwarfs. Details of the optical and mechanical design of the instrument can be found in Artigau et al. (2014). The instrument itself is a fiber-fed, bench-mounted, double-pass, cross-dispersed, spectro-polarimeter that is cryogenically cooled to an operation temperature of 80 K and provides simultaneous spectroscopic and polarimetric observations. The optical fiber-link connecting the Cassegrain unit—used for polarimetric analysis and guiding—to the calibration module and spectrograph is made from purified fluoride; a special optical material featuring improved transmission at wavelengths  $> 2.0 \mu\text{m}$  thus enabling the inclusion of the  $K$  band. The inclusion of the  $K$  band in the SPIRou spectral coverage is unique among most nIR velocimeters and is highly desirable for the velocimetry of late M dwarfs as a large fraction of the RV information content is contained in the  $K$  band (Artigau et al., 2018). With the broad continuous spectral coverage of SPIRou spanning the nIR  $YJHK$  bands ( $0.98 - 2.35 \mu\text{m}$ ), SPIRou will pioneer infrared planet searches by targeting low mass stars whose flux peaks in the nIR wavelength domain. In order to detect Earth-size planets, SPIRou is required to achieve a long-term RV precision of  $1 \text{ m s}^{-1}$  while simultaneously monitoring the star’s intrinsic activity which is enabled by its high spectral resolution ( $\lambda/\Delta\lambda = 70,000$ ).

Thanks to its spectro-polarimetric capabilities, SPIRou is also optimized for characterizing stellar activity—an obvious asset for studying M dwarfs—and particularly late-type M dwarfs which are known for their significant levels of magnetic activity (West et al., 2015). This will allow users to i) minimize the impact of activity on RV curves and ease planet detections and ii) to characterize the impact of stellar activity on the close-in habitable zone planets that SPIRou will detect.

## 3.3 Stellar Input Catalog

### 3.3.1 Stellar Sample

The SPIRou input catalog used in our simulated SLS-PS (see Sect. 3.4) contains 100 stars visible from CFHT on Maunakea ( $\delta \gtrsim -30^\circ$ ). We note that the stellar sample used in these simulations is intended to be an approximation to the true SPIRou input catalog which has yet to be formalized exactly. The stars chosen were selected based on their high scores in the SPIRou merit function. The merit function is based on being able to detect the RV semi-amplitude from the gravitational pull of an  $3 \text{ M}_\oplus$  planet

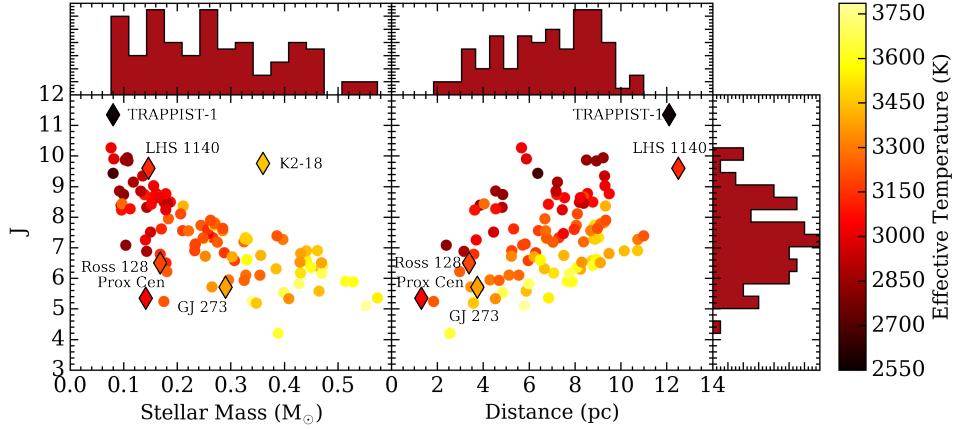


Figure 3.1: Scatter plots and histograms depicting the distribution of SPIRou input catalog  $J$  band magnitudes, stellar masses, distances, and effective temperatures for our fiducial version of the *SPIRou Legacy Survey-Planet Search*. Histograms are in linear units. Nearby M dwarf planetary systems with at least one known HZ planet are depicted with [diamonds](#). K2-18 at 34 pc does not appear in the scatter plot in right panel.

at an equilibrium temperature of 250 K (Malo et al. in prep). Of the optimum selection of 120 stars, approximately 20 were deemed to result in poor detection sensitivities based on the measured fractions of simulated planets detected around those stars in preliminary simulations of the SLS-PS. The primary culprit for the rejection of these stars was their large projected rotation velocities  $v \sin i_s$ , which have a large detrimental effect on the RV measurement precision and hence on our ability to detect planets.

Properties of stars in the SPIRou input catalog are presented in Fig. 3.1. For comparison purposes, in Fig. 3.1 we include six nearby M dwarf planetary systems with at least one known planet in or near the HZ: Proxima Centauri (Anglada-Escudé et al., 2016), Ross 128 (Bonfils et al., 2017), GJ 273 (Astudillo-Defru et al., 2017b), LHS 1140 (Dittmann et al., 2017b), TRAPPIST-1 (Gillon et al., 2017), and K2-18 (Montet et al., 2015; Cloutier et al., 2017b). The global properties of the 100 stars were defined from the Boyajian et al. (2012) relation (effective temperature and radii, when [Fe/H] is fixed to the solar value) and the Delfosse et al. (2000) relation (estimated stellar mass based on absolute  $J$  magnitude). Projected rotational velocities and rotation periods were taken from an exhaustive literature search or derived from the CFHT-CoolSnap program (Moutou et al., 2017). In our sample we consider stars with masses between  $0.08\text{--}0.57\text{ }M_{\odot}$  with  $J$  band magnitudes  $4.2\text{ -- }10.3$  in a range of distances spanning  $\sim 1.8\text{ -- }11\text{ pc}$ .

Rotational information for stars in our sample is an important characteristic for RV modelling as stellar rotation strongly affects the activity arising from rotationally modulated active regions observed in radial velocity (Saar & Donahue, 1997; Meunier et al., 2010; Aigrain et al., 2012; Dumusque et al., 2014a). Rotation timescales also restrict the periodicities at which we can detect planets due to difficulties in detecting RV planets with orbital periods close to the stellar rotation period or its harmonics (Vanderburg et al., 2016). However, the rotational information for our stellar sample is incomplete. For stars with no known available rotation measurements (8/100 stars), or with a  $v \sin i_s$  upper limit only (2/100 stars), we sample  $P_{\text{rot}}$  from a *modified* empirical distribution of M dwarf rotation periods from ground-based photometry as a function of stellar mass (Newton et al., 2016). The nature of this

‘modification’ is discussed in the subsequent paragraph. The corresponding  $v \sin i_s$  is then computed from the sampled value of  $P_{\text{rot}}$ , the known stellar radius, and the inclination of the stellar spin-axis to the line-of-sight  $i_s$  which we draw from a geometrical distribution (i.e. uniform in  $\cos i_s$ ). For stars which only have a measured upper limit on  $v \sin i_s$ , the *modified* empirical distribution from which  $P_{\text{rot}}$  is sampled is truncated at the minimum  $P_{\text{rot}}$  corresponding to the upper limit on  $v \sin i_s$ .

The necessary modification to the empirical  $P_{\text{rot}}$  distribution arises from an observational bias in the [Newton et al. \(2016\)](#) sample which favors rapidly rotating stars. The detection of short photometric rotation periods (i.e. rapid rotators) is attained more easily than long rotation periods because full phase coverage is more readily obtained over many rotation cycles. Furthermore, in the case of early M dwarfs ( $M_s > 0.25 M_\odot$ ), there is evidence for a positive correlation between the star’s rotation rate and the amplitude of its photometric variability ([Newton et al., 2016](#)). Rapid rotators therefore tend to exhibit larger amplitudes of variability thus making the signal more easily detectable. The raw empirical distribution therefore does not represent the true underlying distribution of M dwarf rotation periods in the Solar neighborhood. We attempt to account for this bias in a simplified way by modifying the empirical  $P_{\text{rot}}$  distribution by insisting that only  $\sim 25\%$  of sampled rotation periods can be  $< 10$  days, as estimated from the volume-limited sample of field M dwarfs with measured rotation velocities from [Delfosse et al. \(1998\)](#). This constraint reduces the fraction of fast rotators with  $P_{\text{rot}} < 10$  days by a factor of  $\sim 2$ .

The modification to the empirical  $P_{\text{rot}}$  distribution is visualized in Fig. 3.2. Here we compare the empirical  $P_{\text{rot}}$  distribution, based on the full stellar sample with detected  $P_{\text{rot}}$  from [Newton et al. \(2016\)](#), with the modified distributions of stars in the SPIRou input catalog with either a  $v \sin i_s$  upper limit only or no available rotation data. The latter two distributions are nearly equivalent as the majority of  $v \sin i_s$  upper limits do not provide substantial new information regarding the star’s  $P_{\text{rot}}$  and in both cases we insist that only  $\sim 25\%$  of sampled  $P_{\text{rot}}$  can be  $< 10$  days. We impose this condition by noting that  $\sim 60\%$  of stars in the empirical distribution have  $P_{\text{rot}} < 10$  days and resample a particular fraction of those stars from the subset of the empirical distribution restricted to  $P_{\text{rot}} \geq 10$  days. The fraction of stars with  $P_{\text{rot}} < 10$  days that get resampled is  $1 - 0.25/0.6 \approx 0.58$ .

### 3.3.2 Physical Models of Stellar RV Activity

#### SOAP 2.0: activity simulations

Active regions (ARs) in the stellar photosphere (e.g. star spots and faculae) and in the hot chromosphere (e.g. plages) are expected to be present in M dwarfs. These surface inhomogeneities have characteristic temperatures that differ from the star’s effective temperature and therefore disrupt the symmetry of the visible stellar disk as they rotate in and out of view at the stellar rotation period. One resulting source of RV activity from ARs, known as the *flux effect* ([Dumusque et al., 2014a](#)), results in an anomalous RV signal as the ARs block a fraction of Doppler-shifted photons from the rotating stellar limbs. The strong local magnetic fields associated with ARs at the stellar photospheric boundary also act to inhibit the upward flow of hot convective material in an effect known as the suppression of *convective blueshift* ([Dravins et al., 1981](#)).

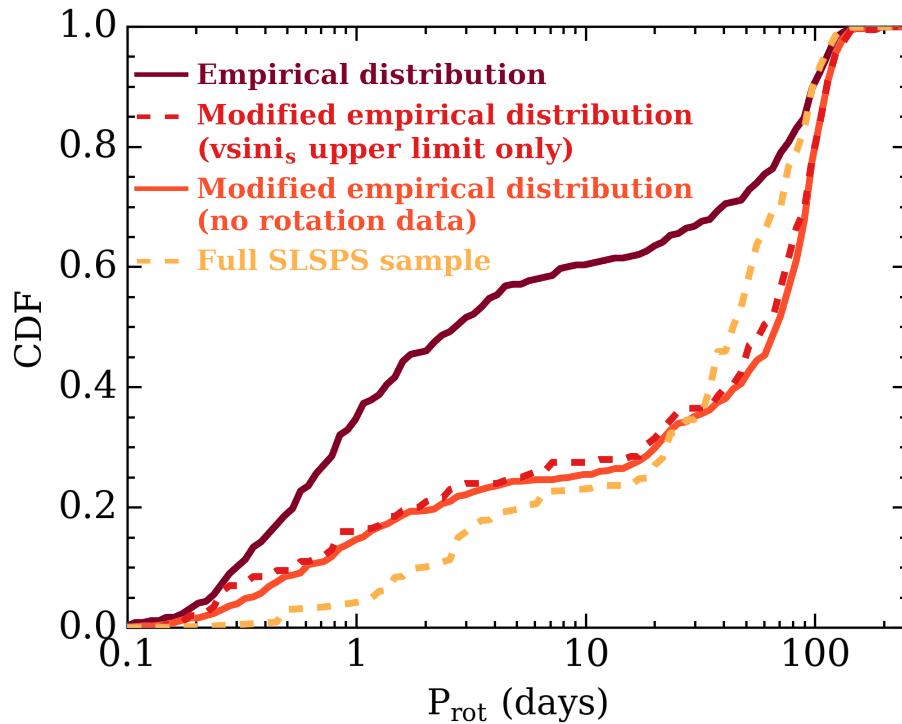


Figure 3.2: Cumulative distribution functions of the [empirical distribution] of M dwarf rotation periods from [Newton et al. \(2016\)](#), the modified empirical distribution of SPIRou stars with a  $[v \sin i_s]$  upper limit measured, the modified empirical distribution of SPIRou stars with [no rotation data] available, and the [full SLS-PS sample].

In the limit of simple distributions of ARs, the observed RV structure from the two aforementioned effects is dependent on the fractional coverage of the visible stellar disk by the ARs and its first time derivative (Aigrain et al., 2012). For each simulated RV time series, we sample the relevant physical parameters of the ARs (i.e. AR sizes and spatial distribution) and simulate the corresponding RV activity, full width at half maximum (FWHM), bi-sector inverse slope (BIS), and photometric time series arising from both the flux effect and from the suppression of convective blueshift using the **SOAP** 2.0 code (Dumusque et al., 2014a). The FWHM and BIS time series are shape parameters of the cross-correlation function between the observed stellar spectra and the template spectrum used to measure the stellar RVs in an observing campaign. These ancillary time series are sensitive to the presence of ARs but not to planets making them useful diagnostics for distinguishing activity-induced RV signals from planetary signals. In particular, the FWHM time series will be used in Sect. 3.6 to train our RV activity model and disentangle RV activity signals from planetary signals.

The **SOAP** 2.0 code outputs time series that are phase-folded to the input stellar rotation period. These time series are initially treated as strictly periodic and interpolated to the epochs of observation. In this way we ignore any contribution from differential rotation whose amplitude has been shown to decrease with decreasing stellar mass (Donati et al., 2008; Morin et al., 2008; Kitchatinov & Olemskoy, 2011) before evoking rigid-body rotation in fully convective M dwarfs ( $M_s \lesssim 0.2 M_\odot$ ). In Sect. 3.3.2, the strictly periodic condition is relaxed to account for the finite lifetimes of ARs and the existence of long-term magnetic activity cycles.

Unfortunately, very little is presently known about the physical nature of ARs on M dwarfs but we do observe quasi-periodic photometric variability which arises from evolving ARs (O’Neal et al., 2005). We use the empirical distribution of photometric amplitudes as a function of stellar mass from Newton et al. (2016) to sample the *average* photometric variability amplitudes<sup>1</sup>  $A$  for each star in the simulated SLS-PS. The sampled value of  $A$  is interpreted as an average value because the phase in the star’s magnetic activity cycle at the time of the Newton et al. (2016) observations is unknown. We then use  $A$  to constrain the size of up to four ARs with each AR being treated as either a cool spot or bright plage. The spatial distribution of ARs is determined from random draws in latitude and longitude as Doppler imaging provides evidence for more uniformly distributed ARs on M dwarfs than on Sun-like (FGK) stars (Barnes & Collier Cameron, 2001; Barnes et al., 2004), whose ARs tend to be more localized along the stellar equator. We note however that such Doppler imaging observations are limited to rapid rotators which are largely avoided in the selection of the SPIRou input catalog. The temperature contrast between the ARs and the stellar effective temperature is fixed to 200 K in all realizations (Berdyugina, 2005). The **SOAP** 2.0 code is designed to model Sun-like stars at an optical wavelength of  $\lambda \sim 529$  nm. The resulting activity is then scaled from the default **SOAP** 2.0 wavelength to the approximate central  $H$  band wavelength of  $\lambda' \sim 1.6 \mu\text{m}$  via the ratio of blackbody emission at  $\lambda$  to  $\lambda'$  with a characteristic temperature of  $T_{\text{eff}}$ . This scaling decreases the amplitude of the flux and convective blueshift effects by a typical factor of a few in the nIR compared to at optical wavelengths (Martín et al., 2006; Huélamo et al., 2008; Prato et al., 2008; Reiners et al., 2010; Mahmud et al., 2011).

---

<sup>1</sup>For photometric variations measured in the near-IR with MEarth over a custom passband spanning  $\sim 0.7 - 1 \mu\text{m}$  known as the  $i + z$  band (Nutzman & Charbonneau, 2008).

### Zeeman broadening

Unlike the flux and convective blueshift effects, the RV activity due to Zeeman broadening tends to *increase* towards the nIR. Zeeman broadening of spectral features in unpolarized light occurs in the presence of strong magnetic fields that cause Zeeman splitting; an effect that grows with wavelength. [Reiners et al. \(2013\)](#) and [Hébrard et al. \(2014\)](#) used polarized radiative transfer at nIR wavelengths to compute the effect of Zeeman splitting from both atomic and molecular sources on stellar line profiles. [Reiners et al. \(2013\)](#) report the following simplified model for the RV signal resulting from Zeeman broadening in M dwarfs ( $T_{\text{eff}} \in [2800, 3700]$  K)

$$\text{RV}_Z(t) = 300 \text{ m s}^{-1} f(t) \left( \frac{B}{1 \text{ kG}} \right)^2 \left( \frac{\lambda}{1 \mu\text{m}} \right)^a, \quad (3.1)$$

where  $f$  is the filling factor or fraction of the visible stellar disk that is spanned by ARs,  $B$  is the local magnetic field strength within the AR, and  $\lambda$  is the wavelength of observation. The powerlaw index  $a \in [0, 2]$ , describes the increase of  $\text{RV}_Z$  with  $\lambda$  and is variable as a result of the apparent distribution of molecular Landé g-values in cool stars. Albeit only the FeH and CO bands are considered in the stellar atmospheric model from which Eq. 3.1 is derived ([Reiners et al., 2013](#)).

Computing  $\text{RV}_Z$  to add to our complete physical RV activity model requires knowledge of the local  $B$  field strength within ARs. The empirical distribution of this quantity in M dwarfs is incomplete despite contributions from various observing campaigns (e.g. [Reiners & Basri, 2007](#); [Shulyak et al., 2014](#); [Hébrard et al., 2016](#); [Moutou et al., 2017](#); [Shulyak et al., 2017](#)). Instead of sampling  $B$  from an empirical distribution, we use an ad hoc method of sampling  $B$  which exploits what is known about small-scale  $B$  fields in M dwarfs as a function of spectral type and rotation. Namely, the fraction of M dwarfs that are magnetically active as a function of rotation period differs between early-type and late-type M dwarfs (M5-M8) as later M dwarfs are able to remain magnetically active late into their lives even after considerable spin-down ([West et al., 2015](#)). For each star we assign an activity flag indicative of being an *active* or *inactive* star where the probability of being flagged as an active star is equal to the measured activity fraction of M dwarfs from [West et al. \(2015\)](#) and is dependent on the star's spectral type and  $P_{\text{rot}}$ . If magnetically inactive, we sample the localized magnetic field strength from  $B \sim \mathcal{U}(0.1, 1)$  kG ([Moutou et al., 2017](#)). If magnetically active, instead we draw from  $B \sim \mathcal{U}(1, 3.1)$  kG ([Moutou et al., 2017](#)). We then calculate the value of  $a$  based on the star's sampled  $B$  and spectral type before evaluating the Zeeman broadening model (Eq. 3.1) at the central  $H$  band wavelength of  $1.6 \mu\text{m}$  as a function of the time-evolving filling fraction which is known from our **SOAP 2.0** simulations.

### Active region lifetimes

The ARs giving rise to stellar activity in our simulations are short-lived compared to the baseline of our observations. RV observations of M dwarfs have suggested that the lifetimes of individual ARs may persist from one to a few stellar rotations and up to  $\gtrsim 10$  (e.g. [Bonfils et al., 2007](#); [Forveille et al., 2009](#); [Hébrard et al., 2016](#)). RV observations have also elucidated that M dwarfs undergo long-term magnetic activity cycles similarly to the Sun (e.g. [Gomes da Silva et al., 2012](#); [Route, 2016](#)). Following the prescription of [Dumusque \(2016\)](#) for Sun-like stars, we proceed in deriving the temporal variation of AR sizes by scaling the total RV activity signal according to each AR's appearance rate  $\lambda(t)$ . [Dumusque \(2016\)](#) also included the time-dependent latitude of ARs which we neglect here due to the more uniform

distribution of ARs observed on M dwarfs (Barnes & Collier Cameron, 2001; Barnes et al., 2004). Furthermore, we assume that the appearance rate for both star spots and bright plages are consistent.

The probability that an AR appears at a time  $t$  is governed by the Poisson distribution

$$P(t) = \frac{e^{-\lambda(t)\tau}(\lambda(t)\tau)^k}{k!} \quad (3.2)$$

where  $\tau$  is the time step in days and  $k = 0, 1, 2, 3$  as we only consider a maximum of four ARs. Next, as a function of time we draw from the probability distribution in Eq. 3.2 which dictates at which epochs an AR is formed. For each newly formed AR we insist that it spends the first third of its lifetime evolving linearly to its maximum size before shrinking towards zero over the remaining two thirds (Dumusque, 2016). Each AR’s lifetime is sampled from a truncated Gaussian distribution with mean  $3P_{\text{rot}}$  and standard deviation  $P_{\text{rot}}$ . The Gaussian distribution is truncated at  $P_{\text{rot}}$  such that all sampled ARs persist for a minimum of one stellar rotation (e.g. Bonfils et al., 2007; Forveille et al., 2009; Hébrard et al., 2016). Also recall that the maximum size of the AR is determined by the star’s sampled amplitude of photometric variability.

The AR appearance rate per unit time is

$$\lambda(t) = (\lambda_{\max,\text{act}} - 0.5) \left[ -0.5 \cos \left( \frac{2\pi t}{P_{\text{cycle}}} + \phi \right) + 0.5 \right] + 0.5 \quad (3.3)$$

where  $\lambda_{\max,\text{act}}$  describes the maximum appearance rate during the maximum of the star’s magnetic activity cycle whose period is  $P_{\text{cycle}}$ . We set  $\lambda_{\max,\text{act}} = 10$  ARs per day and sample  $P_{\text{cycle}}$  from  $\mathcal{U}(6, 10)$  years (Suárez Mascareño et al., 2016; Wargelin et al., 2017) which is a factor of two or more greater than the baseline of the observations. The added term of 0.5 ARs per day to Eq. 3.3 ensures that we maintain a low but non-zero probability of forming an AR close to the minimum of the stellar activity cycle.

Recall that the sampled amplitude of photometric variability sets the size of ARs in our simulations. Furthermore, because the phase within a star’s magnetic activity cycle at the time of photometric observations is unknown, we treat the observed photometric variability amplitude as an average value. To account for this approximately, we rescale our derived AR lifetime scaling to the interval 0.1–2—instead of 0–1—to account for the varying levels of stellar activity up to a factor of two greater than the maximum value and down to a minimum value slightly greater than zero. We then use this scaling to rescale the injected activity in both the RVs and in the ancillary time series.

In general, our rescalings were optimized such that the root-mean-square (rms) of the injected RV activity is roughly consistent with the upper envelope of the RV activity rms observed with HARPS. Fig. 3.3 depicts the distribution of RV activity rms in our simulated stellar sample as a function of  $P_{\text{rot}}$  and directly compares it to the values of M dwarfs observed with HARPS. The HARPS M dwarfs with measured RV activity rms in Fig. 3.3 include Proxima Centauri (Anglada-Escudé et al., 2016), GJ 3293, GJ 3341, GJ 3542 (Astudillo-Defru et al., 2015), GJ 1132 (Berta-Thompson et al., 2015), GJ 876 (Correia et al., 2010), GJ 674 (Bonfils et al., 2007), Gl 205, Gl 358, Gl 388, Gl 479, Gl 526, Gl 846 (Bonfils et al., 2013a), GJ 163 (Bonfils et al., 2013b), Gl 433, Gl 667C (Delfosse et al., 2013a), Gl 176 (Forveille et al.,

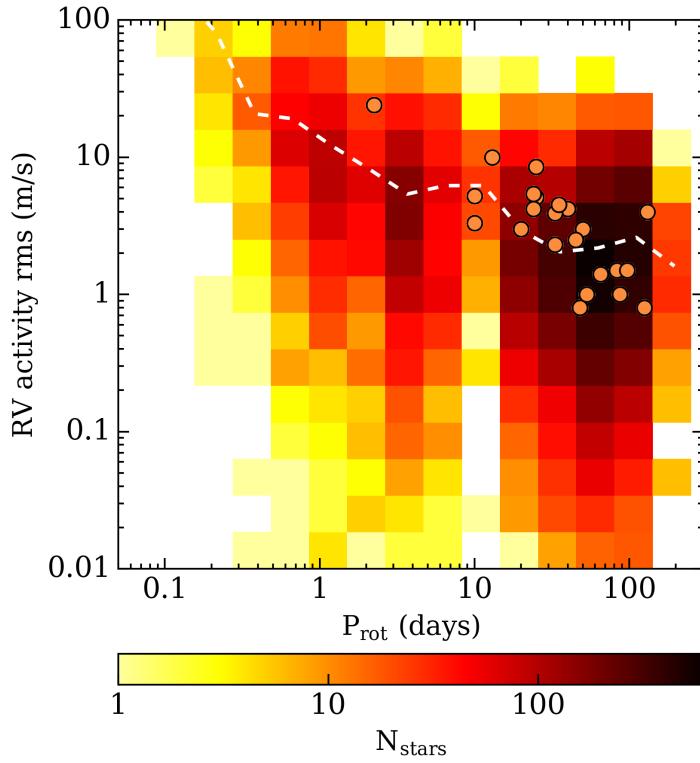


Figure 3.3: The RV activity rms as a function of stellar rotation period for stars in the simulated SLS-PS compared to a population of M dwarfs observed with the HARPS spectrograph. The dashed white curve depicts the mean RV activity rms in each  $P_{\text{rot}}$  bin.

2009), GJ 205, GJ 358, GJ 410, GJ 479, GJ 846 (Hébrard et al., 2016), and GJ 436 (Lanotte et al., 2014).

We note that the comparison depicted in Fig. 3.3 is not one-to-one as RV activity is an intrinsically chromatic effect and our simulated time series are computed at nIR wavelengths whereas the HARPS measurements are taken in the visible. The effect of temperature contrast between ARs and the stellar photosphere on observed RV activity is known to decrease towards longer wavelengths. Conversely, activity from Zeeman broadening increases towards longer wavelengths and is known to be an important source of activity in the nIR (Hébrard et al., 2014; Moutou et al., 2017). Our adopted activity scaling is chosen to best match the HARPS observations at the most frequently sampled rotation periods in our stellar sample of  $\sim 50 - 120$  days. In Fig. 3.3 it is clear that the mean RV activity rms of our sample closely matches the HARPS stars, albeit with a large dispersion. However at smaller  $P_{\text{rot}}$  ( $\lesssim 5$  days), the mean RV activity rms in our simulated sample becomes slightly under-estimated relative to the small sample of HARPS stars at those rotation periods. A larger sample of M dwarfs with measured RV activity rms at small  $P_{\text{rot}}$  will be required to precisely characterize the correlation between the activity rms and  $P_{\text{rot}}$ .

### Ignored sources of RV activity

Various other minor sources of stellar activity in M dwarfs are ignored. These include radial pulsations which do not appear to persist with any significant amplitude within the interiors of M dwarfs (Rodríguez-

(López et al., 2015). Secondly, we disregard the effect of granulation whose amplitude scales with the velocity of convective cells which itself decreases towards later spectral types (Dumusque et al., 2011b; Meunier et al., 2017). Lastly, we ignore flares in our activity model because their distinctive spectral signature allows them to be flagged and removed from the subsequent analysis (Schmidt et al., 2012; Anglada-Escudé et al., 2016). However, the high occurrence rate of flaring events on many M dwarfs may prove costly when attempting to construct large time series of uncontaminated RVs.

## 3.4 Simulated Survey

In this study we conduct a detailed Monte-Carlo (MC) simulation of the SLS-PS for the purpose of predicting the SPIRou planet detection yield. To do so, we must construct a statistically significant number of unique RV time series for each of the 100 stars that we target throughout the simulated SLS-PS. In practice we simulate 100 unique RV time series for each star totalling  $10^4$  realizations in the full simulated SLS-PS. Each RV time series contains unique signals from planets, which are sampled from their known occurrence rates around early M dwarfs (see Sect. 3.5), from physical models of stellar activity (as described in Sect. 3.3.2), and from instrumental noise. Each contribution is sampled in time using a unique window function spanning  $\sim 300$  nights over  $\sim 3$  years, which represents the typical time baseline of observations for a single star in accordance with the expected subset of SPIRou’s time allocation that will be dedicated to the discovery of new exoplanetary systems in the SLS-PS. We note however that at this time the exact number of available nights dedicated to the SLS-PS, nor the duration of the full SLS-PS, have been established absolutely. In addition to the RVs, we also derive various spectroscopic activity indicators arising from stellar activity which are contemporaneous with the RVs. One such ancillary time series is the full width at half maximum (FWHM) of the cross-correlation function which will be used to train non-parametric Gaussian process models of the RV activity based on its common covariance structure with the FWHM (see Sect. 3.6). SPIRou is also unique in that it simultaneously operates its spectroscopic and polarimetry modes thus providing a contemporaneous diagnostic of the star’s magnetic topology. Such time series may also be used to model RV signals from magnetically ARs (Hébrard et al., 2016) although we do not consider such time series in this study.

In each simulated RV time series we attempt to recover the injected planets to form an estimate of the expected planet population that will be discovered with SPIRou. This is facilitated by the joint modelling of stellar activity and planetary signals. This allows for the self-consistent characterization of each RV signal and the detection of a subset of planets which are nominally hidden by their host star’s intrinsic RV activity.

### 3.4.1 Window Functions

In each MC realization, the unique window function  $\mathbf{t}$  is the vector of length  $n_{\text{obs}}$  containing the epochs of observation in barycentric julian dates (BJD). The window function describes the time sampling of our time series. To derive a set of window functions for each star in the SPIRou input catalog we run a separate MC simulation of stellar observing sequences for all targets taking into account when each star is visible from CFHT on Maunakea with an airmass of  $< 2.5$  based on its celestial coordinates. During every available night, all visible stars are observed up to two times, each with an integration time required to achieve a S/N per resolution element of 150—at the central  $J$  band wavelength of 1.25

$\mu\text{m}$ —with a minimum integration time of 15 minutes. The imposed lower limit on the integration time may also be necessary to mitigate the effects of granulation (Lovis et al., 2005) which is expected to be low on M dwarfs. Integration times required to achieve at least the target S/N are computed for each star based using their  $YJHK$  magnitudes. An overhead of 5 minutes is added to each integration for guiding and setup purposes. The output from these simulations is a set of window functions each pertaining to a star in the SPIRou input catalog. Multiple MC simulations are run for various observing sequences and thus provide unique window functions for the simulated SLS-PS. We then sample from these window functions for the purpose of investigating the sensitivity of our planet detection results to the exact form of the window function.

The available nights for observation with SPIRou are limited by two important considerations. The first being the effect of stochastic weather which limits the number of epochs in our derived window functions according to the CFHT observatory’s weather statistics. The second effect is somewhat unique to CFHT as the telescope hosts a suite of instruments that do not operate simultaneously. In particular, the telescope’s wide-field optical imager requires dark-time to conduct its observations whereas the SPIRou spectro-polarimeter does not. As such, we proceed with constructing SPIRou window functions that only include non-dark-time observation and thus correspond to higher levels of lunar contamination. This represents a worst case scenario for SPIRou as aliases from the window function will undoubtedly arise at periodicities close to the cadence of the non-dark-time observing sequences. This cadence evolves with a period close to the period of the lunar cycle at  $\sim 30$  days. Because the SPIRou time sampling occurs in windows separated by  $\sim 30$  days, at that period and its first harmonic at 15 days, significant aliases in the Lomb-Scargle (LS) periodogram (Scargle, 1982) of window function can arise as is shown in Fig. 3.4. The effect of these aliases are detrimental to the detection of periodic planetary signals at these periods because in the LS periodogram of the RVs, one cannot distinguish a-priori between these periodic signals as a planet or as an alias of the time sampling (Dawson & Fabrycky, 2010). This effect has already been shown to mimic planetary signals (e.g. Rajpaul et al., 2016) and is particularly detrimental to finding HZ planets around  $\sim$  M2-M4 dwarfs whose HZ span  $\sim 30$  day orbital periods. We quantify the magnitude of this aliasing effect on the SPIRou planet detection sensitivity in Sect. 3.8.

For the purpose of performing accurate statistical inference of the SPIRou planet population following the full SLS-PS (see Sect. 3.11.1), we attempt to split the total observing time evenly among the stars in the SPIRou input catalog. Doing so, while taking into account stochastic weather and non-dark-time restrictions, limits the size of our window functions over an a-priori 3 year long survey to an average of 198.1 RV measurements per star for 100 stars over  $\sim 300$  nights. The variance in the number of RV measurements per star is relatively small and ranges from 181-212. In the other versions of the SLS-PS containing either more or less stars than in our fiducial survey version (see Sect. 3.13), our MC calculations of the window functions consequently contain less and more RV measurements per star, respectively.

From preliminary simulations of the SLS-PS using 10 unique window functions per sampled planetary system per star, we found that the net planet detection results are largely independent of the exact window function used. Note that all sampled window functions contained the same restrictions discussed previously but do not include the exact same epochs of observation for each star. Explicitly, the SPIRou planet yield was found to vary by only  $\sim 1\%$  across the various window functions used. This dispersion

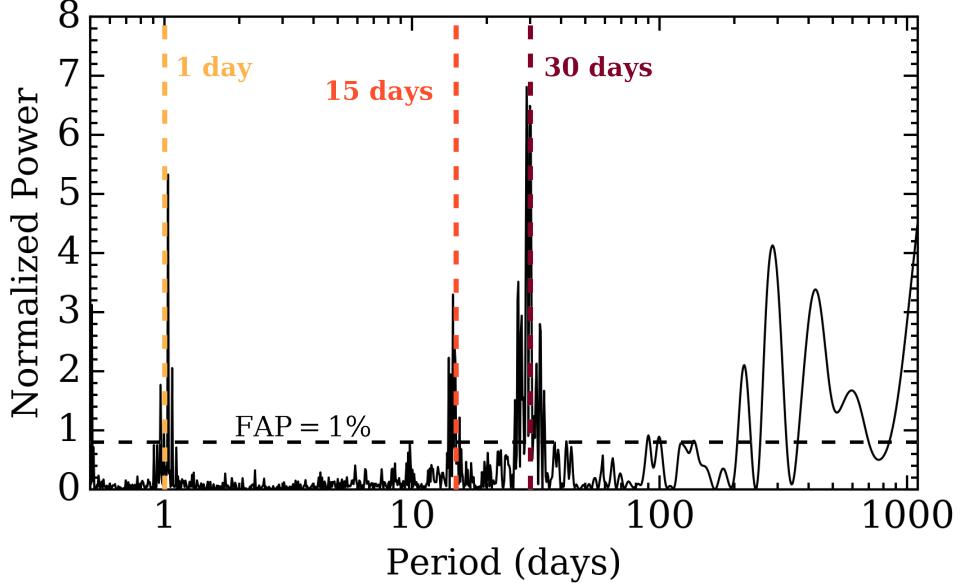


Figure 3.4: Lomb-Scargle periodogram of an example window function from the simulated SLS-PS. The periodogram exhibits strong aliases from our time sampling at the periodicities highlighted by the vertical lines. The power corresponding to a 1% false alarm probability in the periodogram is highlighted by the horizontal dashed line.

is much less than the uncertainties on the resulting estimates of the SPIRou planet yield due solely to uncertainties in the input planet occurrence rates. Therefore we conclude that using multiple window functions for each planetary system in our simulated SLS-PS is an unnecessary computational expense that can be mitigated by considering a single window function per planetary system and not significantly affect the results of our study. However each unique window function considered was still assumed to be restricted to non-dark-time observations thus preserving the aliasing effect on planets with orbital periods of  $\sim 30$  days.

### 3.4.2 Radial Velocity Time Series Construction

In each MC realization, the RV contribution from  $N_p$  injected planets in the simulated planetary system (see Sect. 3.5) is calculated via the superposition of  $N_p$  Keplerian orbital solutions. Each Keplerian  $RV_{kep}$  is parameterized by the planet’s orbital period  $P$ , time of inferior conjunction  $T_0$ , orbital eccentricity  $e$ , argument of periastron  $\omega$ , and RV semi-amplitude  $K$  according to

$$RV_{kep}(t) = K[\cos(\nu(t) + \omega) + e \cos \omega], \quad (3.4)$$

where  $\nu(t)$  is the true anomaly and is computed by solving Kepler’s equation and the eccentric anomaly as a function of time  $t$  contained in the window function  $\mathbf{t}$ . The RV semi-amplitude  $K$  is computed from the planet’s minimum mass  $m_p \sin i$ , the stellar mass  $M_s$ ,  $P$ , and  $e$  using the standard formula

$$K = 1.05 \text{ m s}^{-1} \left( \frac{m_p \sin i}{2 M_\oplus} \right) \left( \frac{P}{20 \text{ days}} \right)^{-1/3} \left( \frac{M_s}{0.3 M_\odot} \right)^{-2/3} \frac{1}{\sqrt{1 - e^2}}. \quad (3.5)$$

The Keplerian RV approximation is valid in all single-planet systems and the majority of multi-

planet systems considered and is best motivated by its ability to negate the need to perform costly numerical integrations of each sampled planetary system. However, we note that this approximation naturally excludes certain dynamical effects in multi-planet systems such as planet-planet interactions and mean-motion resonances which can affect the planet-induced periodicities within the RV time series and therefore also affect our ability to detect those planets. Although the former effect is not accounted for in our simulated time series, its amplitude in real systems is typically small compared to the RV measurement uncertainty as a result of dynamical restrictions on multi-planetary systems making tightly-packed systems less stable over long time-scales and therefore rarely seen in nature. The significance of the latter effect is also expected to be small given the dearth of multi-planet systems at low-order period ratio commensurabilities ([Lissauer et al., 2011](#); [Fabrycky et al., 2014](#)).

All RV components are evaluated at  $\mathbf{t}$  and contain the additive i) RV activity  $\text{RV}_{act}$  derived from physical models (see Sect. 3.3.2) ii)  $N_p$  Keplerian models, and iii) a white noise term with an rms equal to the median RV measurement uncertainty  $\sigma_{\text{RV}}$  expected for the host star. Each star's value of  $\sigma_{\text{RV}}$  is calculated using the nIR RV information content calculations from [Figueira et al. \(2016\)](#) corrected using the empirical spectra from [Artigau et al. \(2018\)](#). We use the results from the condition 3 in [Figueira et al. \(2016\)](#) to estimate  $\sigma_{\text{RV}}$ : the photon-noise contribution to the spectrum being amplified by the limited spectral window due to atmospheric telluric transmission. The information content calculations for each star are computed using SPIRou's spectral resolution, spectral coverage, and the expected S/N per resolution element obtained during an integration of the star.

The RV activity model contains additive contributions from the flux effect, the suppression of convective blueshift, and Zeeman broadening from ARs. The complete RV model is therefore

$$\text{RV}_{model}(t) = \text{RV}_{act}(t) + \sum_{i=1}^{N_p} \text{RV}_{kep,i}(t) + \mathcal{N}(0, \sigma_{\text{RV}}). \quad (3.6)$$

The RV measurement uncertainties pertaining to each measured RV is contained in the vector  $\boldsymbol{\sigma}_{\text{RV}}$  and is modified from the scalar value of  $\sigma_{\text{RV}}$  based on the variable absorption by terrestrial water vapor. The water vapor correction at each epoch in  $\mathbf{t}$  is the product of the airmass and the zenith water column from the well-documented CFHT observing condition statistics throughout the calendar year. The distribution of median  $\sigma_{\text{RV}}$  in our MC simulation for our full stellar sample is shown in Fig. 3.5 as a function of the stellar  $J$  band magnitude and  $v \sin i_s$ . A noise floor is imposed at the expected long-term RV precision limit of SPIRou at  $1 \text{ m s}^{-1}$ . This results in a median  $\sigma_{\text{RV}} = 1.52 \text{ m s}^{-1}$ . The maximum  $\sigma_{\text{RV}}$  is  $6.58 \text{ m s}^{-1}$ . Not depicted in Fig. 3.5 is the dependence of  $\sigma_{\text{RV}}$  on spectral type as the RV measurement uncertainty tends to decrease towards later spectral types due to the increased number of available spectral features and corresponding increase in RV information content.

Similarly to the construction of the RV activity time series at the epochs in  $\mathbf{t}$ , the FWHM and BIS ancillary times-series are constructed from our **SOAP** 2.0 simulations (see Sect. 3.3.2) at the same epochs as the RVs. These contemporaneous time series will be used in Sect. 3.6 to model the RV activity and help to detect underlying planetary signals.

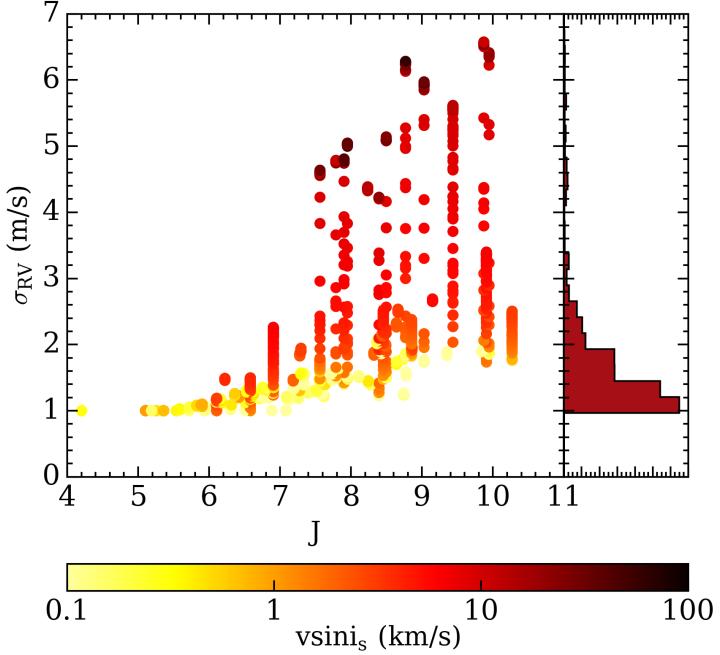


Figure 3.5: The median RV measurement uncertainty in each simulated RV time series  $\sigma_{\text{RV}}$  as a function of the star’s  $J$ -band magnitude and projected stellar rotation velocity  $v \sin i_s$ . A noise floor is imposed at  $1 \text{ m s}^{-1}$ . The histogram shown in the right panel is in linear units.

## 3.5 Planet Sample

### 3.5.1 Planet Sample from known Occurrence rates

In each MC realization we populate the simulated M dwarf planetary system with planets according to their known occurrence rates from the *Kepler* transit survey (Dressing & Charbonneau, 2015). In practice we sample from a grid in orbital period  $P$  and planetary radius  $r_p$  according to the measured planetary occurrence rate at each point in the parameter space  $f(P, r_p)$ . We assume that each point in the parameter space is uncorrelated aside the dynamical stability constraints that we will apply to closely packed systems (see Sect 3.5.2).

The occurrence rates  $f(P, r_p)$  (or equivalently  $f$ ) from which we sample planets have associated asymmetric uncertainties which are derived from data-driven posterior distributions rather than from analytically-defined distributions which are more easily sampled from. Dressing & Charbonneau (2015) only present the modes and  $1\sigma$  dispersions of their  $f$  posteriors which is insufficient information to reconstruct their  $f$  posteriors in each  $P$  and  $r_p$  bin. To account for their reported uncertainties on  $f$  we must assume a functional form of the  $f$  posteriors. For each bin in  $P$  and  $r_p$  we sample planets with a probability drawn from  $\mathcal{N}(\mu, \bar{\sigma})$  where  $\mu$  is the most likely value of  $f$  and  $\bar{\sigma} = \text{mean}(\sigma_{\text{upper}}, \sigma_{\text{lower}})$  where  $\sigma_{\text{upper}}$  and  $\sigma_{\text{lower}}$  represent the asymmetric uncertainties on  $f$ . The values of  $\mu$ ,  $\sigma_{\text{upper}}$ , and  $\sigma_{\text{lower}}$  are provided in Dressing & Charbonneau (2015) (see their Table 4) for all  $P$  and  $r_p$  bins for which  $f > 0$ . For bins in which  $f$  is consistent with zero, we set  $\mu = 0$  and  $\bar{\sigma}$  equal to the  $1\sigma$  upper limit on  $f$ . For bins in which  $f$  is completely unconstrained as a result of a low detection sensitivity, we assume that  $f(P, r_p)$  evolves smoothly such that we can extrapolate the values of  $f$  from surrounding bins with constraints in order

to estimate the values of  $f$  where it is unconstrained by the data (e.g.  $r_p \leq 1 R_{\oplus}$  and  $P \geq 18.2$  days). The extrapolated values are treated as  $1\sigma$  upper limits and  $\mu$  is set to zero such that when integrating the most likely values of  $f(P, r_p)$  over the range of periods and planet radii considered by Dressing & Charbonneau 2015 ( $P \in [0.5, 200]$  days,  $r_p \in [.5, 4] R_{\oplus}$ ) we recover their cumulative planet occurrence rate of  $2.5 \pm 0.2$  planets per M dwarf. However our dynamical stability constraints will reject a subset of sampled planetary systems thus effectively reducing the cumulative planet occurrence rate to  $< 2.5$ .

The dispersion in mutual inclinations among planets in multi-planet systems is related to the dispersion in eccentricities by  $\langle i^2 \rangle \sim \langle e^2 \rangle / 4$  (Stewart & Ida, 2000; Quillen et al., 2007) with the mean orbital inclination to the plane of the sky  $i$  being set to the value of  $i_s$  obtained for the host star (i.e. from a geometrical distribution). Although the distribution of spin-orbit angles for small planets around M dwarfs has yet to be established, the observed low dispersion in orbital eccentricities (Van Eylen & Albrecht, 2015) and mutual inclinations (Figueira et al., 2012; Fabrycky et al., 2014) in these types of planetary systems suggests that they are dynamically cold. If this is indeed the case then the normal vector to the mean planetary orbital plane is expected to be close to parallel to the stellar spin axis;  $i \approx i_s$ . In each planetary system we sample each planet's orbital eccentricity  $e$  from the  $\beta$  probability distribution describing the high detection significance sample of RV planets reported in Cloutier et al. (2015) (see also Kipping, 2013).

### 3.5.2 Modifications to the Planet Sample

Our adopted approach for sampling planetary systems is accompanied by four important caveats.

*Converting planet radii to masses.* Firstly, the Kepler-derived  $f$  directly samples planetary radii whereas RV surveys are only sensitive to the planetary minimum masses. Therefore an assumption must be made regarding the planetary mass-radius relation required to convert the sampled planetary radii into masses. We opt for the following empirical mass-radius relation derived in Weiss & Marcy (2014) from an unbiased sample of known transiting planets:

$$\frac{m_p}{M_{\oplus}} = \begin{cases} 0.440 \left( \frac{r_p}{R_{\oplus}} \right)^3 + 0.614 \left( \frac{r_p}{R_{\oplus}} \right)^4, & r_p < 1.5 R_{\oplus} \\ 2.69 \left( \frac{r_p}{R_{\oplus}} \right)^{0.93}, & r_p \geq 1.5 R_{\oplus}. \end{cases} \quad (3.7)$$

This piece-wise mass-radius relation distinguishes between small rocky planets with  $r_p < 1.5 R_{\oplus}$  and larger gaseous planets (e.g. Rogers, 2015; Dressing et al., 2015; Fulton et al., 2017). The intrinsic dispersion about the *mean* mass-radius relation in Eq. 3.7 has characteristic rms values of

$$\sigma_{\text{rms}} = \begin{cases} 2.7 M_{\oplus}, & r_p < 1.5 R_{\oplus} \\ 4.7 M_{\oplus}, & r_p \geq 1.5 R_{\oplus}. \end{cases} \quad (3.8)$$

To include this intrinsic dispersion in our sampled planet population, for each sampled  $r_p$  we compute the mean  $m_p$  using Eq. 3.7 and add an additional offset drawn from  $\mathcal{N}(0, \sigma_{\text{rms}})$  where the value of  $\sigma_{\text{rms}}$  is given by Eq. 3.8. We reject planets with unphysical negative masses which naturally biases our sample to larger planet masses than are predicted by the *mean* mass-radius relation. We then apply a unique correction factor to each radius bin  $r_p$  with a width of  $0.2 R_{\oplus}$  such that we recover the *mean*

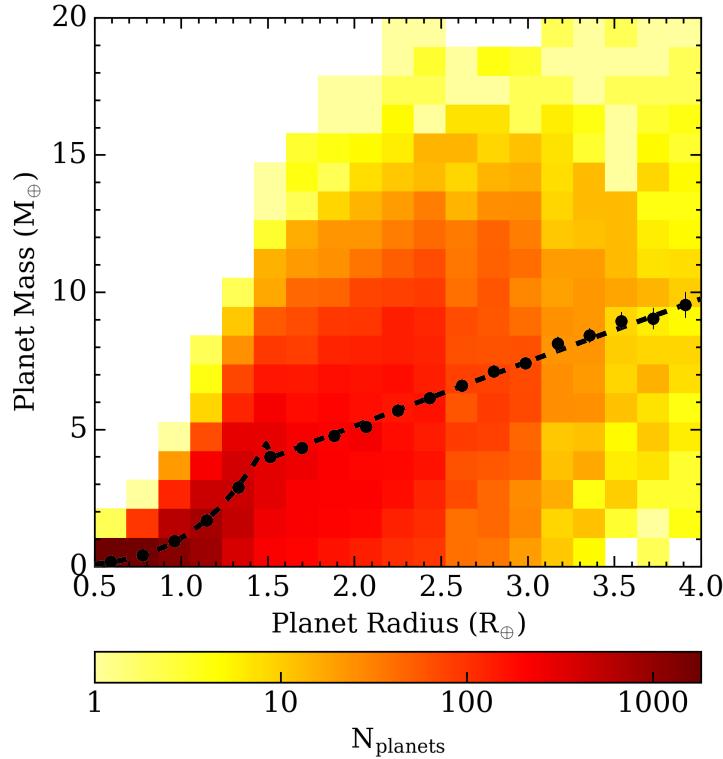


Figure 3.6: 2D histogram depicting the masses and radii of the sampled small planet population with  $r_p \leq 4 R_{\oplus}$ . Planetary radii are drawn from the planet occurrence rates derived from Kepler and are mapped to planetary masses using the mass-radius relation and Gaussian dispersion given in Eqs. 3.7 and 3.8. The mean mass-radius relation is over-plotted as the [dashed curves](#). The mean planet mass and standard deviation of the mean in each planetary radius bin are plotted as [black circles](#) to show their close correspondence with the mean mass-radius relation.

mass-radius relation given in Eq. 3.7 in our final planet sample. Our Kepler-derived planet sample in the mass-radius plane is shown in Fig. 3.6 along with the *mean* mass-radius relation. Over-plotted are the mean planet masses in each  $r_p$  bin demonstrating the accuracy of our unique correction factors used to recover the mean mass-radius relation. Note the large uncertainties in the mean planet masses resulting from the large dispersion observed in the empirical mass-radius relation (Weiss & Marcy, 2014).

*Restricted range of  $P$  and  $r_p$ .* The second caveat with the Kepler-derived  $f$  is that we are restricted in our range of sampled orbital periods and planetary radii as imposed by the size of the parameter space for which  $f$  is robustly measured with Kepler. Specifically,  $f(P, r_p)$  from Dressing & Charbonneau (2015) is restricted to the orbital period domain of  $P \in [0.5, 200]$  days and the planetary radius domain of  $r_p \in [0.5, 4] R_{\oplus}$  which includes small Mars-sized planets up to about the size of Neptune. However, given that the duration of the SLS-PS is longer than 200 days it is important to also consider planets with orbital periods  $> 200$  days. To inform the inclusion of such long period planets we use the RV M dwarf survey detection of the non-zero frequency of giant planets ( $m_p \sin i \geq 100 M_{\oplus}$ ) with orbital periods of  $> 200$  days albeit with a low  $f$  ( $\lesssim 5\%$ ; Bonfils et al., 2013a). We therefore supplement the Kepler occurrence rates with draws from the HARPS occurrence rates of giant planets (Bonfils et al., 2013a). Long period planet sampling is carried out similarly to how the Kepler occurrence rates are

sampled. Here we draw the planetary minimum masses rather than planetary radii;  $f \rightarrow f(P, m_p \sin i)$ . We also note that in Bonfils et al. (2013a)  $f$  for giant planets is computed over a coarse grid in minimum mass spanning an order-of-magnitude ( $10^2 \leq m_p \sin i / M_\oplus \leq 10^3$ ) despite the most massive HARPS detection having only  $m_p \sin i \sim 112 M_\oplus$ . Therefore it is possible that giant planets with minimum masses in excess of  $\sim 112 M_\oplus$  ( $0.3 M_{Jup}$ ) do not exist in nature around M dwarfs. We note that throughout this study we will primarily focus on the population of small planets from the Kepler-derived  $f(P, r_p)$  because of their much larger frequency compared to giant planets around M dwarfs.

*Applicability of occurrence rates to the full input catalog.* Robust statistics regarding  $f$  were derived from the sample of small Kepler stars which almost exclusively contained early-to-mid M dwarfs and late K-dwarfs. The assumption that the resulting  $f$  extend to later spectral types is still largely uncertain but the early discovery of seven transiting Earth-sized planets around the ultracool dwarf TRAPPIST-1 (Gillon et al., 2017), from a small sample of observed stars, hints that small planets around late M dwarfs might be as common, and potentially more common than around the early M dwarfs observed with Kepler. Preliminary estimates of  $f$  around late M dwarfs with *K2* suggests a potential lack of super-Earth-sized planets on close-in orbits but have been insufficient to probe the population of Earth-sized planets (Demory et al., 2016). Theoretically, planet formation scenarios have also predicted the existence of many such small planets on close-in orbits around late M dwarfs (e.g. Alibert et al., 2013; Alibert & Benz, 2017). In Sect. 3.10 we will investigate the effect of increasing  $f$  on our planet detections in the SLS-PS.

*Dynamical considerations of multi-planet systems.* The final caveat arises from  $f(P, r_p)$  being derived in uncorrelated bins whereas dynamical constraints will prevent certain types of planetary systems from existing in nature. For example, close pairs of massive planets. To ensure that sampled multi-planetary systems in our simulations are dynamically stable we impose two priors on each system with multiplicity  $> 1$ . The first constraint is the analytic assessment of Lagrange stability (Barnes & Greenberg, 2006) which only depends on the masses of the central star and planets, the planets' semi-major axes, and eccentricities. For adjacent planet pairs that are Lagrange stable, by definition their ordering remains fixed, both planets remain bound to the central star, and the criterion limits permissible changes in planets semi-major axes. Lagrange stability can be thought of as a more stringent extension of Hill stability (Gladman, 1993). However we note that the analytic treatment of Lagrange stability is only applicable to the three-body system. In planetary systems from our simulations with  $> 2$  planets, we apply the Lagrange stability criterion to every adjacent planet pair. We then supplement the Lagrange stability criterion with the heuristic criterion from Fabrycky et al. (2012) which is applicable to systems with multiplicity  $> 2$ . The resulting criterion for stability<sup>2</sup> is derived from the stability analysis of a set of numerical integrations. In our MC simulations we only include multi-planetary systems which satisfy both aforementioned stability criteria.

As mentioned previously, our dynamical considerations cause the injected planet population to not exactly match the adopted Kepler occurrence rates. The modified Kepler planet occurrence rates, or equivalently our injected planetary population, is shown in Fig. 3.7 as a function of the input variables

---

<sup>2</sup>Long-term stability requires that  $a_{\text{out}} - a_{\text{in}} \gtrsim 3.5 R_{\text{Hill}}$  where  $a_{\text{in}}$  and  $a_{\text{out}}$  are the semi-major axes of the inner and outer planet respectively and  $R_{\text{Hill}}$  is their mutual Hill radius.

$P$  and  $r_p$ . We then report the same population after converting the planetary radii to masses using Eqs. 3.7 and 3.8.

The resulting distribution of planetary system multiplicities is shown in Fig. 3.8. Simulated planetary systems that obey our dynamical stability criteria contain 0-7 planets although only 0.04% of simulated planetary systems can survive with 7 planets. Similarly,  $\sim 4\%$  of simulated planetary systems contain no planets at all. The most common planet multiplicity is 2 with  $\sim 32\%$  of simulated planetary systems containing 2 planets. The resulting average planet multiplicity is  $\sim 2.4$  which is slightly less than the cumulative injected multiplicity of  $2.5 \pm 0.2$ . When recovering the planet yield of our simulated survey we will have to correct for this small discrepancy between the cumulative planet multiplicity of our injected population and the *true* multiplicity of 2.5 for planets with  $P \in [0.5, 200]$  days and  $r_p \in [0.5, 4]$  R<sub>⊕</sub> (see Sect. 3.9).

## 3.6 Activity Mitigation

### 3.6.1 Overview of the Gaussian Process Formalism

RV activity signals present in M dwarfs (e.g. Gl176, [Forveille et al. 2009](#); Gl674, [Bonfils et al. 2007](#), Proxima Centauri; [Robertson et al. 2016](#)) will deter our ability to detect new exoplanets in the SLS-PS. A number of correction techniques have been developed by various groups within the field to mitigate these effects and detect planets in RV (see reviews in [Fischer et al., 2016](#); [Dumusque et al., 2017](#)). One particularly promising method is the use of Gaussian process regression to model correlated noise (i.e. activity) in RV time series.

Gaussian processes (GP) belong to a class of *non-parametric* regression models from the field of machine-learning ([Rasmussen & Williams, 2005](#)). Being non-parametric, the functional form of the activity model is unspecified and instead is determined by the data itself. In this way the GP activity model does not assume any physical model of the underlying processes responsible for the observed activity. Instead the GP is used to model the covariance properties of an input time series according to a user-defined covariance function which itself is described by a small number of hyperparameters<sup>3</sup>. By its non-parametric nature, GP regression is an attractive method for modelling the stochastic processes that give rise to observable stellar activity.

The GP prior distribution, which is described by the aforementioned covariance function, is a multivariate Gaussian distribution of functions specified by a mean function  $\mu(\mathbf{t})$ , evaluated at the epochs contained in the window function  $\mathbf{t}$ , and a covariance matrix

$$K_{ij} = \sigma_i^2 \delta_{ij} + k(t_i, t_j). \quad (3.9)$$

which is computed from the vector of measurement uncertainties  $\sigma(\mathbf{t})$ , the Kronecker delta function  $\delta_{ij}$ , and the user-defined covariance function  $k(t_i, t_j)$  describing the covariance between two measurements taken at times  $t_i$  and  $t_j$  for  $i, j \in [1, n_{\text{obs}}]$ . Obtaining the GP model of an arbitrary input time series  $\mathbf{y}$

---

<sup>3</sup>'Small' compared to the size of the input time series which would be required to fully describe the covariance properties of the input dataset.

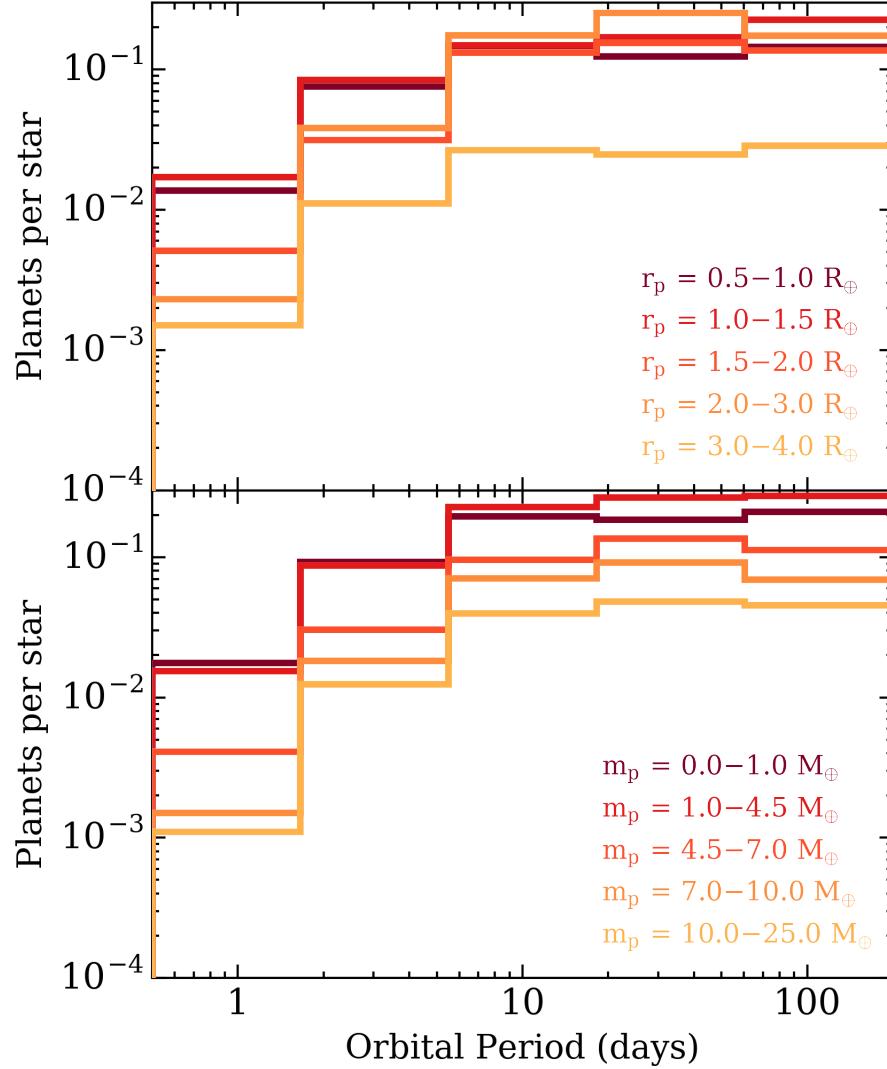


Figure 3.7: *Top:* histograms of the injected small planet population in the simulated SLS-PS as a function of orbital period and planetary radius from Dressing & Charbonneau (2015). *Bottom:* the same planet population as above converted to planetary mass using Eqs. 3.7 and 3.8. For clarity each  $r_p$  bin, and the approximately corresponding  $m_p$  bin, can be viewed individually:  $[r_{p,\min} = 0.5 \text{ } R_\oplus]$ ,  $[r_{p,\min} = 1 \text{ } R_\oplus]$ ,  $[r_{p,\min} = 1.5 \text{ } R_\oplus]$ ,  $[r_{p,\min} = 2 \text{ } R_\oplus]$ ,  $[r_{p,\min} = 3 \text{ } R_\oplus]$ .

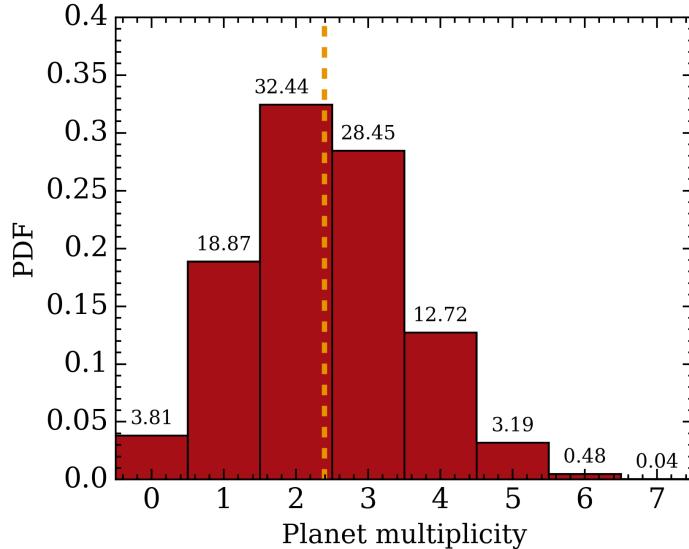


Figure 3.8: The probability density function of planet multiplicities in the simulated SLS-PS. The fraction of simulated systems with a given multiplicity—in percentage—is annotated for each integer multiplicity. We find an average planet multiplicity of  $\sim 2.4$  which is highlighted by the *vertical dashed line* and is slightly less than the cumulative injected multiplicity of 2.5 planets with  $P \leq 200$  days from Dressing & Charbonneau (2015).

and uncertainty vector  $\boldsymbol{\sigma}$  is done by maximizing the Gaussian logarithmic likelihood function

$$\ln \mathcal{L} = -\frac{1}{2} \left( (\mathbf{y} - \boldsymbol{\mu})^T K^{-1} (\mathbf{y} - \boldsymbol{\mu}) + \ln \det K + n_{\text{obs}} \ln 2\pi \right), \quad (3.10)$$

to obtain the ‘best-fit’ values of the GP hyperparameters describing the covariance of the observations  $\mathbf{y}$  through the covariance function  $k(t_i, t_j)$ . In the case of M dwarfs, activity predominantly arises from rotationally modulated ARs which also evolve in time due to their varying lifetimes, spatial distribution, and contrast. Because of this temporal evolution we adopt a quasi-periodic covariance function of the form

$$k(t_i, t_j) = a^2 \exp \left[ -\frac{|t_i - t_j|^2}{2\lambda^2} - \Gamma^2 \sin^2 \left( \frac{\pi |t_i - t_j|}{P_{\text{GP}}} \right) \right], \quad (3.11)$$

described by four hyperparameters:  $a$  the amplitude of the covariance in the units of  $\mathbf{y}$ ,  $\lambda$  the exponential decay of correlations,  $\Gamma$  the coherence scale of correlations, and  $P_{\text{GP}}$  the periodic timescale. From identical reasoning to our own, quasi-periodic covariance functions are commonly adopted in a number of related astrophysical applications such as the recovery of stellar rotation periods (i.e. setting  $P_{\text{GP}} = P_{\text{rot}}$ ) (e.g. Angus et al., 2018) and to the modelling of RV activity from spectroscopic activity diagnostics (e.g. Haywood et al., 2014; Rajpaul et al., 2015; Cloutier et al., 2017b), photometry (e.g. Cloutier et al., 2017a,b), or the raw RVs themselves (e.g. Faria et al., 2016; Donati et al., 2017; Yu et al., 2017) and is applicable when the baseline of the observations spans at least a few rotation periods (Pont et al., 2013).

Following the optimization of the GP hyperparameters one obtains a unique GP prior distribution. Conditioning the GP prior on the dataset  $\mathbf{y}(\mathbf{t})$  results in the GP predictive distribution whose mean

function and posterior variance can be evaluated at previously unseen epochs  $\mathbf{t}^*$  via

$$\boldsymbol{\mu}(\mathbf{t}^*) = K(\mathbf{t}^*, \mathbf{t}) \cdot K(\mathbf{t}, \mathbf{t})^{-1} \cdot \mathbf{y}(\mathbf{t}) \quad (3.12)$$

and

$$C(\mathbf{t}^*) = K(\mathbf{t}^*, \mathbf{t}^*) - K(\mathbf{t}^*, \mathbf{t}) \cdot K(\mathbf{t}, \mathbf{t})^{-1} \cdot K(\mathbf{t}^*, \mathbf{t})^T. \quad (3.13)$$

Here the covariance matrix  $K(\mathbf{t}^*, \mathbf{t}^*)$  is evaluated at unseen epochs such that  $\boldsymbol{\sigma}(\mathbf{t}^*)$  must be set to zero in Eq. 3.9.

### 3.6.2 Modelling RV Activity with GP Regression

As in nature, the RV time series from our MC simulations contain contributions from both planetary companions as well as activity from ARs. These sources of the RV activity signals also have manifestations in other spectroscopic time series, albeit not strictly at the same rotation period (Hébrard et al., 2016). One example is the full width at half maximum (FWHM) of the CCF which we simulate along with the RV signals (see Sect. 3.3.2). As is often the case when searching for small planets in RV, one cannot distinguish a-priori an RV signal from a planet or from activity. However the covariance properties in the FWHM time series will be related to the covariance properties in RV residuals after the removal of planetary sources. We therefore use the FWHM time series to train the hyperparameters of our quasi-periodic GP prior distribution. In principle one could have chosen an alternative activity indicator such as the BIS or contrast of the CCF however recent HARPS observations have demonstrated that the strongest correlations between the RVs and an activity indicator often exist between the RV and FWHM time series (e.g. Astudillo-Defru et al. 2017b, Bonfils et al. 2017b in prep). However this is not universal. Similarly, the contrast or depth of the CCF is not used in place of the FWHM because although the two time series are highly correlated, the CCF contrast is known to be more strongly affected by instrumental noise in practice. Using a spectroscopic training set like the FWHM also has the benefit of being obtained contemporaneously with the RV measurements therefore probing the star’s activity at the same epochs in which we are searching for planets. In this way our GP activity model is also sensitive to variations in the RV activity occurring on long timescales such as from magnetic activity cycles.

To reduce computational wall time of our full MC simulation we do not compute a GP activity model in each MC realization of our simulated SLS-PS. The GP activity model is only computed when the rms of the injected RV activity exceeds the median RV measurement uncertainty. In such cases we run a Markov Chain Monte-Carlo (MCMC) on the FWHM time series to obtain the marginalized posterior probability density functions (PDF) of the four GP hyperparameters:  $a$ ,  $\lambda$ ,  $\Gamma$ , and  $P_{\text{GP}}$ . To sample the posterior PDFs we use the `emcee` affine invariant MCMC ensemble sampler (Foreman-Mackey et al., 2013) coupled with the fast GP package `george` (Ambikasaran et al., 2015) to evaluate the likelihood function in Eq. 3.10. Specifically we initialize 100 walkers with an effective chain length  $\sim 10$  autocorrelation times to ensure convergence of the chains. Samples of the posterior PDFs are only saved following a burn-in phase of  $\sim 10$  autocorrelation times. From a set of supervised preliminary tests of the MCMC procedure, walkers are initialized in the parameter space within Gaussian balls whose variance is chosen such that the mean acceptance fraction among the walkers is 20 – 60%.

Table 3.1: Gaussian Process Hyperparameter Priors Used in Training

Hyperparameter	Prior
Covariance amplitude, $a$ [m s <sup>-1</sup> ]	$\mathcal{J}(10^{-2}, 10^2) \cdot \max  FWHM - \langle FWHM \rangle $
Exponential decay timescale, $\lambda$ [days]	$\mathcal{J}(1, 10 \cdot (\max t - \min t))$
Coherence scale, $\Gamma$	$\mathcal{J}(10^{-2}, 10^2)$
Periodic timescale, $P_{GP}$ [days]	$\mathcal{J}(0.1, 300)^a$ $\mathcal{U}(0.9, 1.1) \cdot P_{rot}^b$

**Notes.**

<sup>a</sup> If  $P_{rot}$  is not detected in the FWHM time series.

<sup>b</sup> If  $P_{rot}$  is detected in the FWHM time series.

We select broad non-informative priors for all GP hyperparameters with the exception of  $P_{GP}$  which is constrained to the narrow uniform range of  $P_{GP}/P_{rot} \in \mathcal{U}(0.9, 1.1)$  either when the photometric rotation period is known for a particular star or if  $P_{rot}$  is detected in a LS periodogram of the FWHM with a false-alarm probability (FAP)  $\leq 1\%$ . Prescription of the GP hyperparameter priors are reported in Table. 3.1. Throughout this paper we calculate FAPs via bootstrapping with replacement using  $10^4$  iterations and each LS periodogram is normalized by its standard deviation. From the marginalized posterior PDFs of the GP hyperparameters, we adopt the *maximum a-posteriori* (MAP) values of each hyperparameter to construct a unique covariance matrix  $K$  and thus our mean GP model and its  $1\sigma$  confidence interval from Eqs. 3.12 and 3.13. Examples of resulting GP models of the FWHM time series for a rapidly and a slowly rotating star are shown in the upper panels in Fig. 3.9.

Following the training phase on the FWHM time series we proceed with modelling the RVs simultaneously with a trained GP activity model plus Keplerian planetary signals. The marginalized posterior PDFs of the GP hyperparameters  $\lambda$ ,  $\Gamma$ , and  $P_{GP}$  from training are used as informative priors in the joint RV analysis which treats the GP amplitude  $a$  as a free parameter. For each assumed mean function  $\mu$ , containing between zero and three Keplerian solutions, we compute the MAP GP activity model from the hyperparameter values sampled using MCMC. Assuming a zero mean function, the resulting mean GP activity models for the two stars shown in Fig. 3.9 are shown in the middle panels of the figure. The residuals following the removal of the mean GP activity model is also shown. In each case the stellar rotation period is detected in the LS periodogram of the FWHM time series and therefore is used to constrain  $P_{GP}$  during training. For the rapid rotator with  $P_{rot} = 2.8$  days, the rms of the injected activity is  $4.43$  m s<sup>-1</sup> and is reduced to  $1.59$  m s<sup>-1</sup> after removing the mean GP activity model ( $\chi^2_{\text{red}} = 1.4$  for four GP hyperparameters). This resulting RV rms is more comparable to the median RV measurement uncertainty  $\sigma_{RV} = 1.35$  m s<sup>-1</sup>. In the slow rotator case ( $P_{rot} = 83.1$  days) the rms of the injected activity is reduced from  $3.60$  m s<sup>-1</sup> to  $1.53$  m s<sup>-1</sup> compared to  $\sigma_{RV} = 1.63$  m s<sup>-1</sup> ( $\chi^2_{\text{red}} = 0.9$ ). In both test cases considered in Fig. 3.9, a planet is detected in the RV residuals and fit simultaneously with the activity assuming a new mean model containing a single Keplerian solution. Details of our planet detection algorithm are discussed in Sect. 3.7.

### 3.6.3 GP Activity Model Performance

As an overview of the performance of our GP activity modelling, we can compare the rms of the known injected RV activity with the rms of the residuals following the removal of our mean GP activity model.

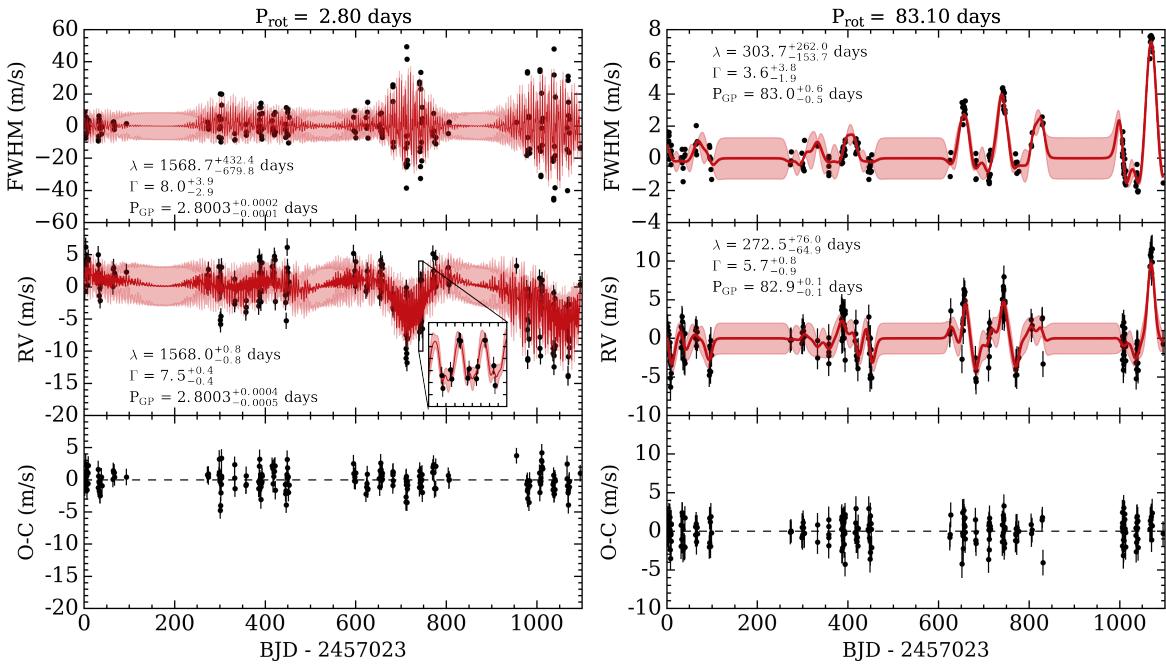


Figure 3.9: Examples of our GP formalism used to mitigate RV activity and detect underlying planetary signals for two simulated systems with either a rapidly ( $P_{\text{rot}} = 2.8$  days; *left column*) or slowly ( $P_{\text{rot}} = 83.1$  days; *right column*) rotating host star. *Top panels*: the FWHM time series used for training along with the mean GP regression model and its  $1\sigma$  confidence interval shown in red. *Middle panels*: the raw RV time series along with the mean GP activity model (in the absence of planets) and its  $1\sigma$  confidence interval. *Bottom panels*: the RV residuals after removal of the mean GP activity model shown in the middle panels. Each system contains 3 planets which contribute to the  $O - C$  residuals albeit with semi-amplitudes which are all  $\lesssim \sigma_{\text{RV}}$ . The two upper rows report the MAP values of GP hyperparameters along with their 16<sup>th</sup> and 84<sup>th</sup> percentiles.

The residual rms should never exceed the rms of the injected activity otherwise our GP formalism would be adding additional noise into the RV time series rather than modelling and reducing it as is its intended purpose. Similarly, optimal GP fits will result in a residual rms that is close to the median RV measurement uncertainty of the time series.

Fig. 3.10 compares the rms of the *injected* RV activity to the rms of the *residual* RV activity after removing the mean GP activity model computing assuming a zero planet model. Recall that we only compute a GP activity model when the injected rms is  $> \sigma_{\text{RV}}$  such that the injected activity rms in units of  $\sigma_{\text{RV}}$  is always greater than or equal to unity in Fig. 3.10. We note that the residual activity rms never exceeds the injected activity rms as expected; i.e. the residual rms always lies beneath the line  $y = x$  in Fig. 3.10. Secondly, there appears to be a positive correlation between  $P_{\text{rot}}$  and the relative reduction of the activity rms, which is analogous to GP performance. That is that within our sample of RV time series which are modelled with a GP activity model, the activity rms is maximally reduced in systems with the longest rotation periods. Conversely, the GP performance when applied to rapid rotators ( $P_{\text{rot}} \gtrsim 2$  days) is often marginal in comparison. The exact cause of these effects above may be related to the poor time-sampling of our observations compared to the short stellar rotation period but is ultimately beyond the scope of this paper and is reserved for a future study.

We note that in many slow rotator cases, the dimensionless residual rms is often less than unity. This suggests that in such cases the non-parametric GP is actually modelling the noise and not just the stellar activity signal. Unfortunately when the GP activity model is over-fitting the RV noise, planetary signals can be absorbed into the mean GP model and thus avoid detection. Furthermore, this apparent over-fitting is a common feature in many of our simulations and does not appear to depend on whether or not the stellar rotation period is detected a-priori in the FWHM time series. However it is true that on average, the GP modelling out-performs cases in which  $P_{\text{rot}}$  is known from the FWHM compared to cases in which  $P_{\text{rot}}$  remains undetected thus providing very weak constraints on  $P_{\text{GP}}$ . This highlights the importance having a-priori knowledge of  $P_{\text{rot}}$  from any of the ancillary spectroscopic times-series, polarimetric time series (Hébrard et al., 2016), or from previously obtained long-baseline photometry (e.g. Newton et al., 2016).

## 3.7 Automated Planet Detection

Due to the large number of planetary systems in our simulated SLS-PS, we must detect planets in an automated way. The steps in our automated planet detection algorithm represent computationally tractable calculations given the large number of planetary systems for which each step must be performed. We note however that other—potentially more robust—planet detection algorithms may be adopted in the real SLS-PS which are likely to include more human intervention than the automated techniques described in the following subsections.

### 3.7.1 Establishing Putative Planetary Detections

We proceed by searching for planetary periodicities in an iterative manner using the LS periodograms of the RVs following the removal of various periodic signals. An example of this iterative process is visualized in Fig. 3.11 for a 3 planet system ( $P_b, P_c, P_d = 1.85, 7.17, 14.66$  days,  $K_b, K_c, K_d = 4.7, 2.3, 0.07$

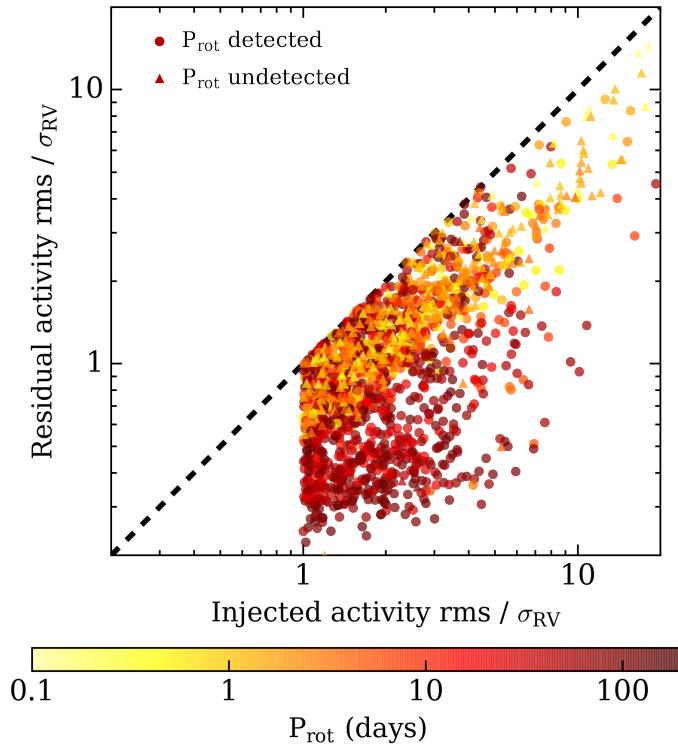


Figure 3.10: Comparison of the rms of the *injected* RV activity to the rms of the *residual* RV activity following the removal of the mean GP activity model which is computed in the absence of planets. Each rms value is normalized by the RV measurement uncertainty  $\sigma_{RV}$  of its time series. The *dashed line*  $y = x$  is indicative of cases wherein the GP activity model does not reduce rms due to activity. Time series in which  $P_{rot}$  is detected in the FWHM time series are depicted as [circles]. Whereas time series in which  $P_{rot}$  is not detected in the FWHM times-series are depicted as [triangles].

$\text{m s}^{-1}$  respectively) around a moderately active star with a measured photometric rotation period  $P_{\text{rot}} \sim 8.8$  days and an RV rms in the absence of planets of  $\sim 4.2 \text{ m s}^{-1}$ .

For MC realizations in which the rms of the injected RV activity exceeds the median RV measurement uncertainty—as is the case for the system shown in Fig. 3.11—we compute two versions of our initial periodogram: one of the raw RVs *only* and a second of the RV residuals after fitting the data with the trained GP activity model and zero mean function (i.e. no planet model). This GP model predicts the RV activity in the absence of planetary signals. For the remaining MC realizations containing quiet stars we only compute the LS periodogram of the raw RVs thus neglecting any modelling of correlated RV residuals. For cases in which we use a trained GP to model RV activity, it is beneficial for the periodic term of the assumed quasi-periodic covariance kernel to be constrained by our training set. For the example shown in Fig. 3.11, the top panel shows the LS periodogram of the FWHM in which  $P_{\text{rot}}$  is detected and is subsequently used in our RV modelling (see Sect. 3.6.2).

The second panel in Fig. 3.11 depicts the LS periodogram of the raw RVs only whereas the third panel depicts the raw RVs corrected by the trained GP activity model with  $P_{\text{GP}} = P_{\text{rot}} = 8.8$  days. Comparing these two periodograms it is apparent that some power at  $P_{\text{rot}}$  is diminished when removing the mean GP activity model along with power at long periods due to the large exponential timescale found during training. More importantly, the strongest periodic signal in each periodogram is at the orbital period of the innermost planet at  $\sim 1.85$  days suggesting the presence of a planet. In these periodograms, and in all subsequent iterations, we claim a putative planet detection for periodicities with i)  $\text{FAP} \leq 1\%$  ii) is within 2% of an injected planet’s orbital period iii) is not associated with significant periodic signals seen in the FWHM time series iv) is not an alias of the time-sampling (see example in Fig. 3.4) and v) is  $> 2\%$  away from the stellar rotation period and any of its first four harmonics. The signals at  $\sim 1.85$  days in the second and third panels of Fig. 3.11 obey these criteria and therefore constitute a putative planet detection. We proceed by referring to this putatively detected planet as ‘b’.

We note that the second condition for a putative planet detection above cannot be utilized in a real survey because any planetary periodicities are not known a-priori. Instead we invoke this condition to accelerate our automated planet detection algorithm relative to the steps that must be taken in a real survey to secure planet detections (e.g. Bayesian model comparison; [Ford & Gregory, 2007](#)). Without a-priori knowledge of planet orbital periods, LS periodicities at high significance that obey all of the remaining aforementioned putative planet criteria, may represent false positive signals if not carefully modelled. In our simulations we find that on average our time series generate  $\sim 0.5$  such false positives in their LS periodograms following the removal of planet models. Although a fraction of these false positive periodicities are likely to be aliases of each other making the above estimate an upper limit. Ideally the determination of these signals as false positives or as true planetary signals would be solved via a formal model comparison in a subsequent analysis. However, such calculations are not guaranteed to converge to the correct solution (see [Dumusque et al., 2017](#)).

The second periodogram iteration requires that we fit the putative planetary signal and search for additional signals in the LS periodogram of the residuals. To model planet ‘b’ we adopt the maximum-likelihood Keplerian model parameters at the detected periodicity and re-compute the GP activity model

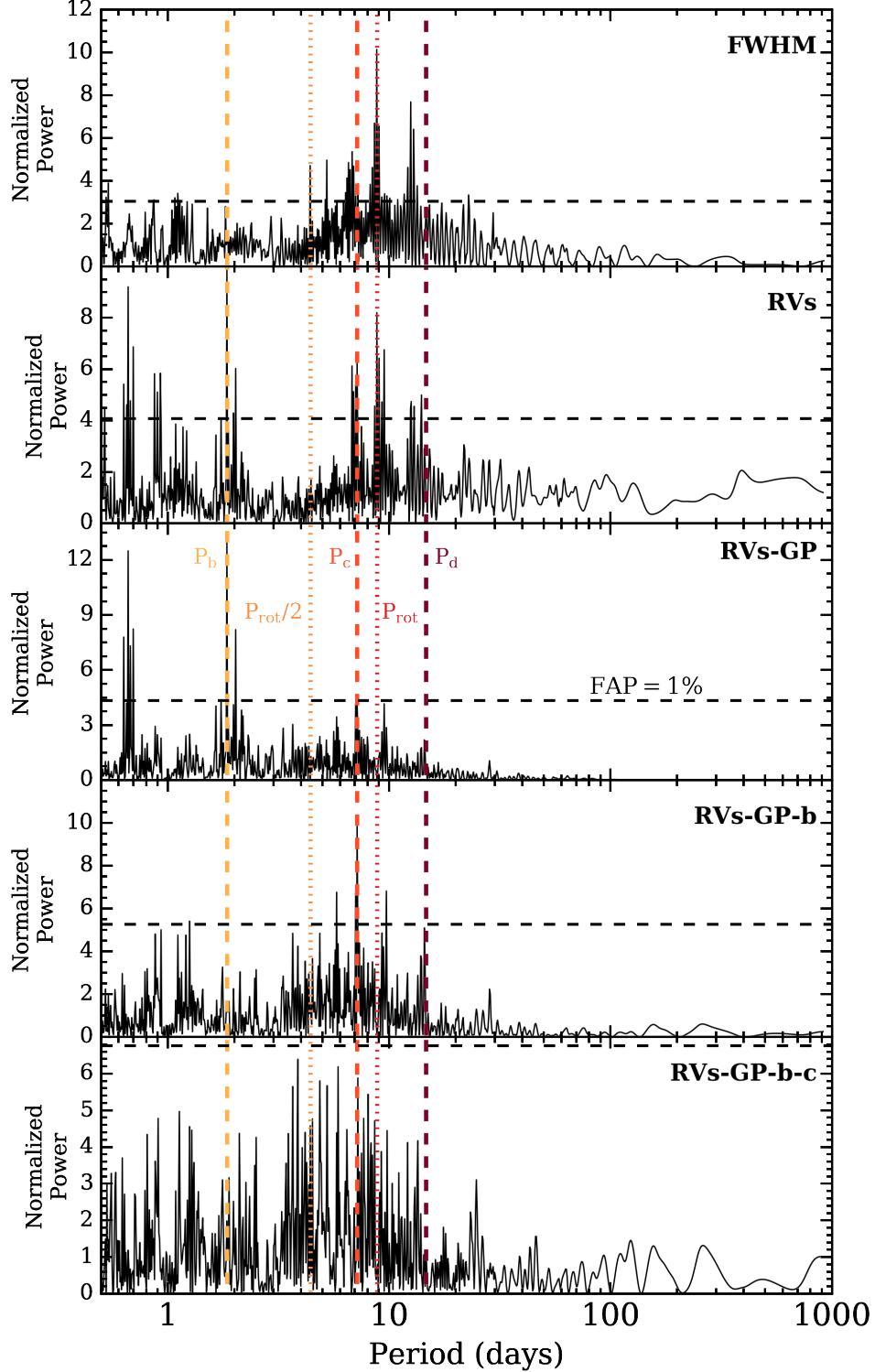


Figure 3.11: *Top to bottom:* Lomb-Scargle periodograms of the full width at half maximum (FWHM) of the cross-correlation function, raw radial velocities (RVs), RVs corrected for activity (RVs-GP), RVs corrected for activity and a planet (RVs-GP-b), and the RVs corrected for activity and two planets (RVs-GP-b-c). Periodicities equal to the stellar rotation period, its first harmonic, and the orbital periods of the three planets in the system are highlighted with vertical lines. The power corresponding to a 1% false alarm probability in each periodogram is highlighted by a horizontal dashed line in each panel.

which is modified due to the new mean function (i.e. a one planet model rather than the previously assumed no planet model). Note that the GP activity model is re-computed after each iteration due to the changing mean function. The fourth panel in Fig. 3.11 shows the resulting periodogram after correcting the raw RVs with the new GP activity model and a Keplerian solution for planet ‘b’. The strongest residual periodic signal is at the orbital period of the middle planet at  $\sim 7.17$  days which we claim as a second putative planet detection because the periodicity obeys the aforementioned criteria. This second putative planet is referred to as ‘c’.

The bottom panel of Fig. 3.11 depicts the LS periodogram of the RVs after being corrected for activity and the superposition of the two maximum-likelihood planet models. Now the strongest residual periodicity has a FAP  $> 1\%$  implying that our automated planet detection algorithm has ceased to detect planetary signals. Therefore in the example shown in Fig. 3.11, the third injected planet at  $\sim 14.66$  days remains undetected. An unsurprising result given the small RV semi-amplitude of the planet ( $K_d = 7 \text{ cm s}^{-1}$ ) compared to the time series’ median RV measurement uncertainty  $\sigma_{\text{RV}} = 1.8 \text{ m s}^{-1}$ . In each MC realization we perform this iterative procedure until no putative planets are detected and up to a maximum of three planets despite many planetary systems having  $> 3$  injected planets (see Fig. 3.8). In this way we are at least sensitive to the expected number of planets per M dwarf ( $2.5 \pm 0.2$ ; [Dressing & Charbonneau, 2015](#)) and limit the computational expense of detecting planets dominated by repeatedly computing GP activity models and LS periodogram FAPs.

### 3.7.2 Model Selection

In Sect. 3.7.1 we established putative planet detections based on low FAP LS periodogram periodicities. However the robust detection of a planet with a particular set of Keplerian model parameters must be favoured over competing models that lack such a planet. The proper diagnostic for model selection is the ratio of Bayesian model evidences which are notoriously difficult and time-consuming to calculate ([Ford & Gregory, 2007](#)). As an alternative model selection technique we turn to time series cross-validation (CV). This technique is a specialized version of general K-fold CV and is suitable to data featuring strong correlations in time as is the case with RV time series ([Arlot & Celisse, 2010](#)).

For MC realizations featuring at least one putative planetary detection we perform time series CV on models that contain an increasing number of planets, including the null hypothesis i.e. no planets. The latter model has zero Keplerian parameters whereas a model containing  $N_p \geq 1$  planets contains  $3N_p$  model parameters where the three parameters per planet are its orbital period, time of inferior conjunction, and RV semi-amplitude. For the purpose of model selection we will assume circular orbits for all planets to limit the size of the parameter space.

The CV algorithm proceeds by first splitting the RV time series  $y_1, \dots, y_{n_{\text{obs}}}$  into training and testing sets. For some  $t > 1$ , each competing model is fit to the training set  $y_1, \dots, y_t$  using a Levenberg-Marquardt optimization routine. The optimized model is then evaluated at the next epoch  $t + 1$  ( $\mu_{t+1}$ ) and the ln likelihood of the testing set  $y_{t+1}$  given the optimized model is computed using Eq. 3.10. When computing the ln likelihood, we adopt a white covariance matrix for systems wherein the GP analysis is not used but otherwise assume the MAP GP hyperparameters from the iterative procedure in Sect. 3.7.1. These steps are repeated for  $t = N_{\text{min}}, \dots, N - 1$  where the minimum size of the training

set  $N_{\min}$  is set to 20. The favoured model is determined by which of the competing models has a largest median ln likelihood per measurement among the  $N - N_{\min}$  CV iterations. In cases wherein two models are consistent within their median absolute deviations, the model containing less planets is accepted as an imposition of Occam’s razor.

### 3.7.3 Vetting of Planet Detections

A consequence of our automated planet detection methodology is various non-deterministic effects which can result in planet detections that are highly unlikely to be favored by model comparison in the real SLS-PS, yet are marginally detected in our simulations. Such planets are commonly those whose RV semi-amplitude is close to the rms of the RV time series. These planets would likely be rejected by any human vetting which we do not conduct in our simulations. We therefore undergo a vetting procedure in an attempt to restrict the detected planet population to be maximally realistic. Our adopted vetting procedure is based on the methods of Cumming et al. (2008) from the Keck Planet Search. For vetting we define the condition that a bona fide planet detection must satisfy  $K/\sigma_K \geq 3$ . *That is that a true planet detection is one in which the planet’s semi-amplitude  $K$  is detected with a minimum expected detection significance of  $3\sigma$ .*

To estimate the expected uncertainty in the RV semi-amplitude  $\sigma_K$  we compute the *Fisher information matrix*  $B$  which quantifies the information content in an observable time series  $\mathbf{y}(\mathbf{t})$  regarding unknown model parameters  $\boldsymbol{\theta}$ . The model parameter covariance matrix is related to the Fisher information via  $C = B^{-1}$ . We can therefore use the Fisher information matrix to analytically predict the measurement uncertainty of the RV model parameters of interest given an input time series with  $n_{\text{obs}}$  measurements  $\mathbf{y} = (y_1, \dots, y_{n_{\text{obs}}})$  obtained at the epochs  $\mathbf{t} = (t_1, \dots, t_N)$  and with measurement uncertainties  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$ .

The Fisher information matrix is a Hessian matrix of the ln likelihood of a single Keplerian model with respect to its model parameters  $\boldsymbol{\theta} = \{P, T_0, K\}$ :

$$B_{i,j} = -\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j}. \quad (3.14)$$

The Fisher information matrix is symmetric and is  $3 \times 3$  in our case as we assume circular orbits.

In order to simplify the calculation of  $B$  we consider planets individually and account for the residual RV signal from additional planets through an “effective” RV uncertainty  $\sigma_{\text{eff}}$  in place of the RV measurement uncertainties  $\sigma_i$  in Eq. 3.9. The effective RV uncertainty is the rms of the RVs after removal of the Keplerian signal from the planet being considered. It therefore contains contributions from any additional planets, stellar activity, and systematic errors. Because we do not fit for each planet’s orbital eccentricity the Keplerian model simplifies to  $\mu(t_k) = -K \sin \phi_k$  where  $\phi_k = \frac{2\pi}{P}(t_k - T_0)$ . Using this mean model in the ln likelihood (Eq. 3.10) along with  $K_{ij}$  approximated by a white covariance matrix  $K_{ij} = \sigma_{\text{eff}} \delta_{ij}$ , we can compute each element of  $B$  analytically (see Appendix 3.14.1). The Fisher information matrix is then inverted to obtain the covariance matrix of the model parameters  $C$ . The measurement uncertainty of the semi-amplitude is then  $\sigma_K = \sqrt{C_{K,K}}$ .

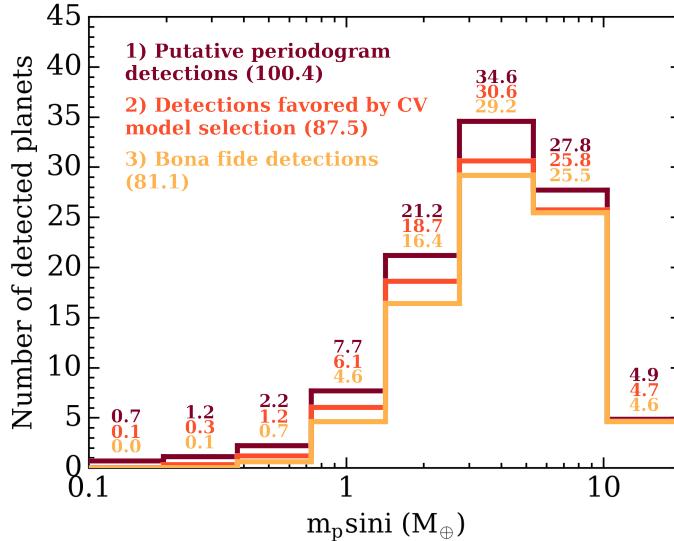


Figure 3.12: The number of planets detected as a function of  $m_p \sin i$  after each of the three steps in our automated planet detection algorithm. The number of detected planets in each  $m_p \sin i$  bin is annotated above the bin to help quantify the decrease in planet detections following each step. The distribution of detected planet minimum masses following each step can be viewed independently for clarity ([1], [2], [3]).

### 3.7.4 Summary of the Automated Planet Detection Algorithm

To recapitulate our process of claiming planet detections in our simulated SLS-PS we recall the three steps discussed throughout this section. Firstly we search for putative planetary signals in the LS periodogram of either the raw RVs or the RVs corrected for activity using a zero-mean GP activity model. We are careful to ensure that significant periodicities have a planetary origin and are not associated with stellar activity signals or our time sampling. Secondly we compute model ln likelihoods using time series cross-validation and compare models with and without the putative planet. We only retain planets which are favored by this model selection technique. The third and final step consists of vetting our planet detections by insisting that they must have an RV semi-amplitude detection significance greater than  $3\sigma$  where the detection significance is estimated from the planet's known semi-amplitude  $K$  and an analytical estimate of the measurement uncertainty on  $K$  from the Fisher information. Planet detections which pass our vetting procedure are treated as bona fide detections.

The distributions of detected planet minimum masses after each step in our automated planet detection algorithm are shown in Fig. 3.12. Here we only include planet detections with  $0.1 \leq m_p \sin i / M_\oplus \leq 20$ . In this SLS-PS there are a small number of giant planets with  $m_p \sin i / M_\oplus > 20 M_\oplus$  detected thus resulting in an underestimated total planet yield annotated in Fig. 3.12 (see Sect. 3.9 for a full description of the detected planet population).

It is clear from Fig. 3.12 that each step in our automated planet detection algorithm reduces the number of detected planets somewhat. In each  $m_p \sin i$  bin other than the most massive bin ( $9.4 \leq m_p \sin i / M_\oplus \leq 20$ ), our CV model selection technique rejects between  $\sim 1 - 3$  planets or  $\sim 13\%$  of all putative planet detections in the range of minimum planet masses considered. Similarly our vetting procedure reduces the number of detected planets in each  $m_p \sin i$  bin by  $\sim 1 - 2$  planets for intermediate

minimum masses;  $0.7 \lesssim m_p \sin i / M_\oplus \lesssim 6$ . Our vetting procedure therefore does not reject a significant number of detected planets at the lowest masses ( $m_p \sin i \lesssim 0.7 M_\oplus$ ) nor at the highest ( $m_p \sin i \gtrsim 6 M_\oplus$ ). The former being the result of the small number of putative low mass planets detected and the latter being due to the large RV semi-amplitude of the most massive planets thus resulting in a typically large detection significance. Vetting rejects  $\sim 7\%$  of planet detections favored by CV. Therefore  $\sim 80\%$  of putative planet detections from the periodogram analysis materialize into bona fide planet detections.

### 3.8 SLS-PS Sensitivity

The detection sensitivity is defined as the recovery fraction of injected planets in our simulated SLS-PS. Because we have a-priori knowledge of the injected planet population we can compute the detection sensitivity for each star in our sample by simply dividing the number of detected planets by the number of injected planets over any desired range of planet properties. We perform this calculation over the discretized parameter space in  $P$ ,  $m_p \sin i$  and in  $S$ ,  $m_p \sin i$ . Here  $S = L_s / 4\pi a^2$  is the insolation received by the planet where  $L_s$  is the stellar luminosity<sup>4</sup> and  $a$  is the planet's semi-major axis. We focus on the following ranges of parameter values which encompass the vast majority of the injected planet population:  $P \in [0.5, 200]$  days,  $S \in [0.01, 100]$   $S_\oplus$ , and  $m_p \sin i \in [0.4, 15] M_\oplus$ .

We note that in this study the recovery fraction is uniquely determined by the performance of our automated planet detection algorithm (see Sect. 3.7). Conversely, the actual SLS-PS will have a much higher degree of human intervention on the data analysis effort. This is afforded by the relatively small size of the RV datasets compared to large surveys (e.g. Kepler and TESS) which benefit greatly from automated detection algorithms. Therefore the detection sensitivity in the actual SLS-PS may not correspond exactly to what is presented here although the automated algorithm used in this study is designed to closely mimic the analysis that will be conducted on the actual SLS-PS data.

The detection sensitivity to planets varies from star-to-star due to their changing stellar properties which can affect our ability to detect planets in radial velocity (e.g. apparent magnitude, stellar mass, stellar rotation, etc). Computing the detection sensitivity for each star individually is necessary for calculating planet occurrence rates (see Sect. 3.11). To improve the detection statistics across the full range of planetary parameters considered we augment the MC realizations for each star with an additional set of planetary systems with logarithmic  $P$ ,  $S$ , and  $m_p \sin i$  sampled uniformly rather than from the planet occurrence rates. The individual detection sensitivity maps for each star can then be combined to obtain the average sensitivity maps for the full SLS-PS as a function of  $P$ ,  $S$  and  $m_p \sin i$  as shown in Fig. 3.13. In this way we marginalize over the stellar properties of our sample stars including the aforementioned parameters which are known to influence the detection sensitivity for each individual star. Hence our sensitivity results might be scaled to various stellar samples provided that its global properties are consistent with our current sample. In Fig. 3.13 the uncertainties in the detection sensitivity within each grid cell come from counting or Poisson statistics and therefore benefit from a large number of simulated planetary systems.

---

<sup>4</sup>The stellar luminosity is computed from the evolutionary models of Baraffe et al. (1998) based on the stellar mass on the main sequence at 2 Gyrs.

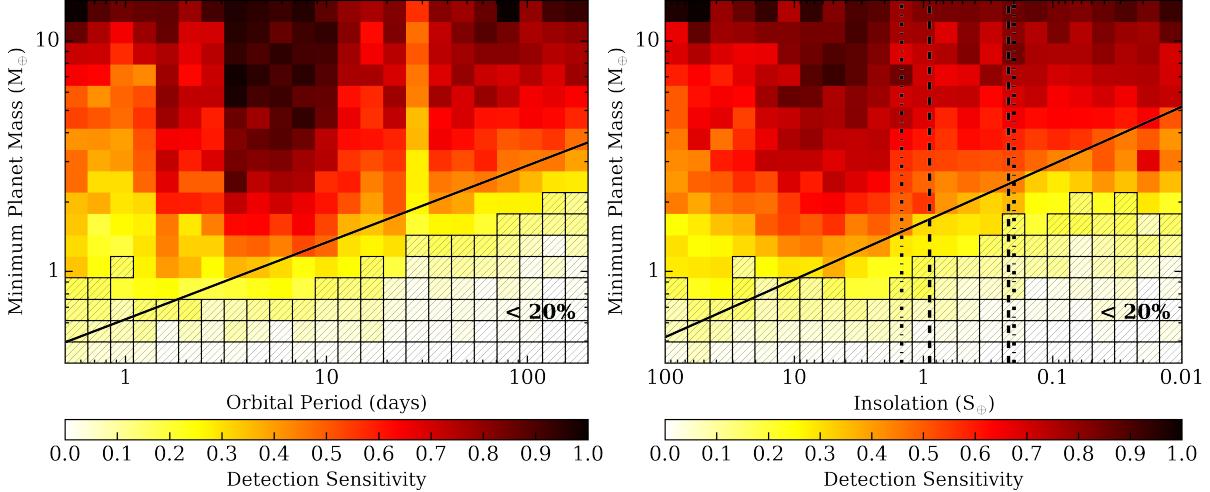


Figure 3.13: Binned maps of the detection sensitivity of the full SPIRou input catalog derived using our automated planet detection algorithm as a function of minimum planet mass and orbital period (left) or insolation (right). The dashed vertical lines in the insolation panel indicate the approximate ‘water-loss’ and ‘maximum-greenhouse’ insolation limits of the HZ from Kopparapu et al. (2013). The dashed-dotted vertical lines indicate the more conservative ‘recent-Venus’ and ‘early-Mars’ HZ limits (Kopparapu et al., 2013). Shaded bins highlight regions of the parameter space wherein our detection sensitivity is < 20%. The solid lines highlight the curve with  $K = 1 \text{ m s}^{-1}$  for a star with our sample’s median stellar mass of  $0.25 M_{\odot}$ .

Given the binning in Fig. 3.13, nowhere do we achieve a 100% detection sensitivity. This is true even for the most massive close-in planets whose RV semi-amplitudes are typically much greater than the characteristic RV measurement uncertainty so long as the system is not orientated close to face-on (i.e.  $\sin i \sim 0^\circ$ ). The geometric effect as well as potential aliasing of periodic signals arising from stellar rotation or the window function, also prevent the detection of certain types of planets. For example, in Fig. 3.2 we see that in our stellar sample there is a dearth of  $P_{\text{rot}} \sim 3 - 10$  days. Thus we run into minimal aliasing from  $P_{\text{rot}}$  at those orbital periods and achieve an increased detection sensitivity relative to planets with equivalent  $m_p \sin i$  but with smaller  $P$  (left panel Fig. 3.13).

Similarly we can see a steep decrease in detection sensitivity at periods of  $\sim 30$  days. Recall that in our window functions we are restricted by the telescope’s observing schedule to only observe outside of dark-time whose cycle follows the lunar cycle with a cadence of  $\sim 30$  days. Consequently the origin of the decrease in detection sensitivity at  $\sim 30$  days can be traced back to aliasing from our time sampling at that period; an unfortunate circumstance as the range of orbital periods corresponding to the HZ around M2-4 dwarfs, spans 30 days.

Due to the aforementioned aliasing effects from stellar rotation and our window functions, the SPIRou detection sensitivity is maximized for the most massive planets with  $3 \lesssim P \lesssim 10$  days. Another pertinent effect at orbital periods close to one day is the potential for signal aliasing due to the Earth’s rotation. This—in part—is responsible for the low detection sensitivity at orbital periods close to a day in Fig. 3.13; another unfortunate circumstance given the general interest in quantifying the occurrence rate of close-in planets (Mulders et al., 2015).

The average detection sensitivity across the full range of  $P$  and  $m_p \sin i$  in Fig. 3.13 is  $44.8 \pm 0.5\%$ . This average detection sensitivity is sufficiently high such that the resulting planet detections from the SLS-PS will be able to place strong constraints on the cumulative occurrence rate of planets around SPIRou stars (see Sect. 3.11).

When considering the detection sensitivity as a function of insolation, we see the same qualitative structure as is seen as a function of orbital period (*right panel* Fig. 3.13). For example, the increased sensitivity between  $\sim 3 - 10$  days has a broad manifestation at  $S \in [1, 10] S_\oplus$ . Similarly, in both cases it is unsurprising to see that the detection sensitivity increases towards more massive planets. In both cases the SPIRou detection sensitivity reaches its lowest values for the least massive planets on wide-orbits. Notably, we achieve a detection sensitivity of  $\lesssim 20\%$  for all planets with  $m_p \sin i \lesssim 1 M_\oplus$  thus making it difficult to detect Earth-mass planets and smaller in the SLS-PS. The average detection sensitivity across the full range of  $S$  and  $m_p \sin i$  considered here is  $47.6 \pm 0.5\%$ .

### 3.8.1 Detection Sensitivity to HZ Planets

It is also important to consider our detection sensitivity to HZ planets as these targets are often flagged for various observational follow-up campaigns. In this study we adopt the ‘water-loss’ and ‘maximum-greenhouse’ definitions as our fiducial HZ limits from Kopparapu et al. (2013) which are derived from a 1D radiative-convective climate model in the absence of clouds. Following Kasting et al. (1993), the inner edge of the HZ is defined by the ‘water-loss’ limit which arises from the photolysis of water in the upper atmosphere and subsequent hydrogen escape. The outer edge of the HZ is determined by the ‘maximum-greenhouse’ limit wherein an increase in atmospheric CO<sub>2</sub> levels will result in a net cooling as the increased albedo from Rayleigh scattering begins to dominate over the increasing greenhouse effect. The insolation levels approximately corresponding to our adopted HZ definition are  $S \in [0.22, 0.90] S_\oplus$  for our stellar sample. For comparison we also consider the more conservative ( $S \in [0.20, 1.48] S_\oplus$ ; Kopparapu et al. 2013) ‘recent-Venus’ and ‘early-Mars’ HZ limits which assume that both Venus and Mars were habitable early-on in the lifetime of the Solar System.

From the fiducial definition of the HZ we find an average detection sensitivity to HZ planets of  $47.3 \pm 1.0\%$  which is consistent with the average detection sensitivity over the full  $S, m_p \sin i$  grid in the right panel of Fig. 3.13 ( $46.2 \pm 0.5\%$ ). Adopting the more generous HZ limits, we find a comparable average detection sensitivity of  $48.7 \pm 0.9\%$ . However, the average detection sensitivity to Earth-like planets ( $m_p \in [1, 5] M_\oplus$ ) in the HZ is significantly reduced to  $36.7 \pm 1.2\%$ . Here we have defined Earth-like planets in terms of their absolute mass where the mass upper limit is approximately equal to the planet mass obtained when evaluating the mean mass-radius relation (Eq. 3.7) at the proposed maximum radius of a rocky planet;  $\sim 1.5 - 1.8 R_\oplus$  (Weiss & Marcy, 2014; Rogers, 2015; Fulton et al., 2017).

## 3.9 SLS-PS Predicted Yield

For each of the 100 stars in the SPIRou input catalog, we can compute the number of planets detected as a function of  $P$ ,  $S$ , and  $m_p \sin i$  given the input occurrence rates and calculations of each star’s detection sensitivity. Here we must correct the reduced injected cumulative planet occurrence rate of 2.4 planets per star over our grid of  $P \in [0.5, 200]$  days and  $m_p \sin i \in [0.4, 15] M_\oplus$  to be equal to the intended 2.5

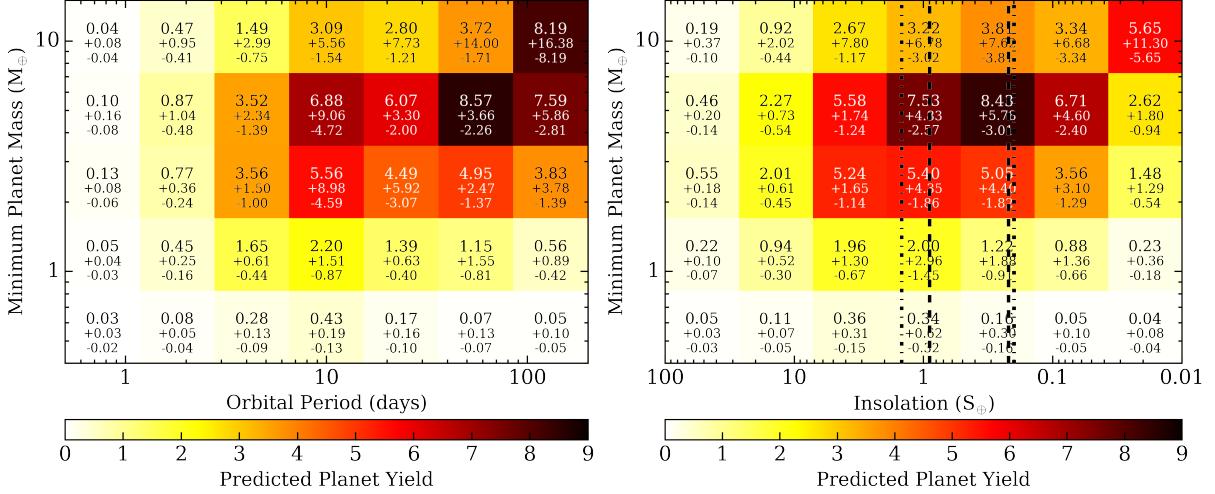


Figure 3.14: Coarsely binned maps of the predicted planet yield from the SLS-PS as a function of minimum planet mass and orbital period (left) or insolation (right). The *dashed vertical lines* in the insolation panel indicate the approximate ‘water-loss’ and ‘maximum-greenhouse’ insolation limits of the HZ from Kopparapu et al. (2013). The *dashed-dotted vertical lines* indicate the more conservative ‘recent-Venus’ and ‘early-Mars’ HZ limits (Kopparapu et al., 2013). The *annotated numbers* in each bin report the predicted number of detected planets and the uncertainties on the prediction which are dominated by uncertainties in the input planet occurrence rates.

planets per star over the Dressing & Charbonneau (2015) grid ( $P \in [0.5, 200]$  days,  $r_p \in [0.5, 4]$   $R_{\oplus}$ ). Then after dividing out the number of simulated planetary systems per star, we compute the total planet detection yield predicted by our simulated survey. The predictions are shown in Fig. 3.14 over a more coarsely binned map than in Fig. 3.13, due to the small number of planets detected in each bin.

The cumulative planet yield over the orbital period domain and minimum mass range considered in Fig. 3.14 is  $85.7^{+29.3}_{-12.5}$  out of  $\sim 180$  injected planets. Of these,  $\sim 53.7$  planets (62.6%) are the only planet detected in the system while  $\sim 26.8$  planets (31.3%) are detected in a 2-planet system. The remaining  $\sim 5.2$  planets (6.1%) are found in systems with 3 detected planets; i.e. we expect to detect 1 – 2 3-planet systems in the SLS-PS.

The number of simulated planetary systems per SPIRou star is large in our simulations. The result is that the uncertainties in the predicted yield are dominated by uncertainties in the input planet occurrence rates from Kepler. Due to Kepler’s low detection sensitivity to small planets on wide-orbits the planet occurrence rate and hence the predicted SPIRou yield is poorly constrained at large orbital periods/low insolation levels. Approximately half of our planet detections are super-Earths with minimum masses  $m_p \sin i \in [3, 7] M_{\oplus}$  owing to their assumed frequency, which is intrinsically high, and the moderately high detection sensitivity achieved across the range of orbital periods considered ( $\sim 30 - 85\%$ ). Considering the ‘water-loss’ and ‘maximum-greenhouse’ definitions of the HZ limits from (Kopparapu et al., 2013) we detect  $22.0^{+18.4}_{-7.9}$  out of  $\sim 47$  injected HZ planets in the SLS-PS over the range of minimum masses considered in Fig. 3.14. These include  $9.0^{+8.5}_{-3.6}$  Earth-like HZ planets out of  $\sim 25$  injected planets with  $m_p \in [1, 5] M_{\oplus}$ . When adopting the more generous ‘recent-Venus’ and ‘early-Mars’ HZ limits, these numbers get inflated to  $31.5^{+26.8}_{-11.4}$  and  $13.4^{+12.7}_{-5.3}$  respectively.

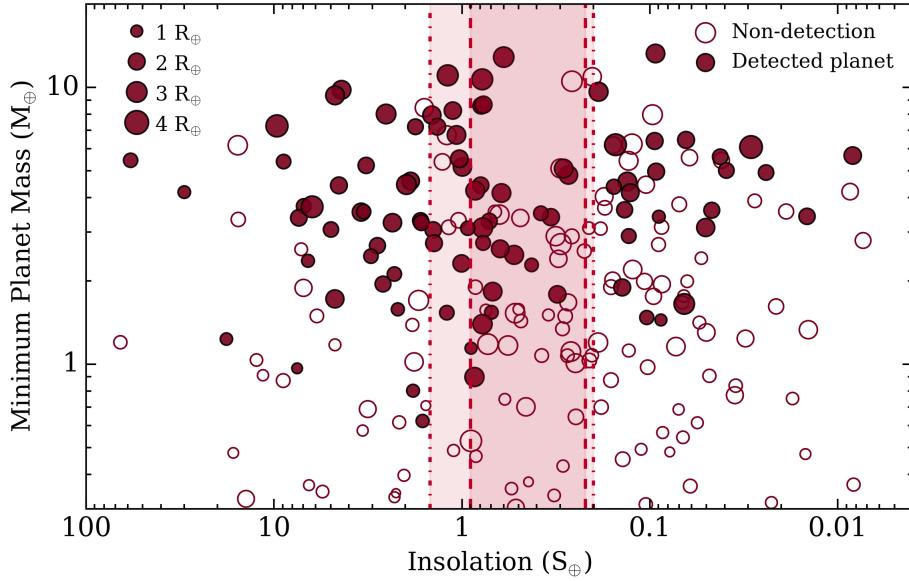


Figure 3.15: A random subset of the simulated SPIRou planets representative of the underlying planet population investigated in the SLS-PS in the insolation/minimum planet mass plane. SPIRou planet detections are marked by *solid circles* whereas *open circles* represent injected planets that remain undetected by SPIRou. We detect 90 planets around 100 stars in the subset of simulated planets shown here. The size of each planet’s marker is proportional to its radius. The *inner shaded region* highlights the approximate ‘water-loss’ and ‘maximum-greenhouse’ limits of the HZ whereas the *outer shaded region* highlights the ‘recent-Venus’ and ‘early-Mars’ HZ limits (Kopparapu et al., 2013).

The population of SPIRou planets can be visualized slightly differently in the insolation/minimum planet mass plane as shown in Fig. 3.15. Here we present a random subset of all simulated planetary systems. The size of this subset is chosen such that the integer number of detected planets in the subset is consistent with the predicted planet yield of 89.9 planets. The resulting total number of planets shown in Fig. 3.15 is 250 with 90 planets detected. Because the subset of planets shown in Fig. 3.15 is random, it does not preserve the number of detected planets in various subsets of the planet parameter space. For example, in the full simulated SLS-PS we detected 9 Earth-like planets whereas the random subset shown in Fig. 3.15 only depicts 4 Earth-like planet detections.

The stellar parameters, planetary parameters, and simulated time series for each MC realization are made available to the community on github<sup>5</sup> in the form of `python pickles`. We also provide the combined results of all realizations for interested users to analyze the full SPIRou input catalog and each star’s suite of simulated planetary systems. These data may be used, for example, to reconstruct most of the figures shown in this paper. Instructions and examples of how to read and access these data are also provided along with all the required `python` scripts.

<sup>5</sup>[https://github.com/r-cloutier/SLSPS\\_Simulations](https://github.com/r-cloutier/SLSPS_Simulations)

### 3.9.1 Giant Planet Detections

Although the vast majority of our predicted SPIRou planet population are derived from the Kepler occurrence rates of small planets ( $r_p \leq 4 R_{\oplus}$  or  $m_p \sin i \lesssim 15 M_{\oplus}$ ), the non-zero frequency of giant planets from RV surveys results in some giant planet detections. From our simulations we find that 4.1 SPIRou detections are giant planets with  $m_p \sin i \geq 20 M_{\oplus}$ . Thus the total SPIRou planet yield including all planets becomes  $89.9^{+30.7}_{-13.0}$ . Furthermore, we find that the average multiplicity of simulated planetary systems containing a giant planet is 2.2 despite our imposed dynamical stability criteria which disfavor giant planets in multi-planet systems. This suggests that M dwarf systems containing both small and giant planets could be discovered with SPIRou. Such systems—should they exist—would be crucial for informing planet formation scenarios around M dwarfs.

## 3.10 The effect of an increased planet frequency around late M dwarfs

Recall that thus far we have assumed that the Kepler occurrence rates, computed using stars with  $T_{\text{eff}} \gtrsim 3200$  K (i.e. spectral types of M4.5 and earlier; [Luhman et al., 2003](#)), are also applicable to the later M dwarfs in our stellar sample. However there are some lines of evidence which suggest that the cumulative planet occurrence rate will increase towards later M dwarfs. For example, planet formation models around very low-mass stars predict many small planets ( $r_p \sim 1 R_{\oplus}$ ) at short orbital periods (e.g. [Alibert et al., 2013](#); [Alibert & Benz, 2017](#)) and the detection of seven Earth-sized planets around the ultra-cool dwarf TRAPPIST-1 from a small sample of ultra-cool dwarfs ([Gillon et al., 2017](#); [Luger et al., 2017](#)) are suggestive of an increased cumulative planet occurrence rate around late M dwarfs compared to early M dwarfs. If this is true then two possible outcomes on the SPIRou SLS-PS planet yield are imaginable. Either the predicted planet yield will increase as a result of the greater number of potential planets to detect, or the predicted planet yield will decrease because adding additional planets will contribute to the observed RV rms thus deterring our ability to detect individual planets. Because the latter effect will modify our detection sensitivity we cannot simply estimate the resulting planet yield from the product of a scaled-up planet occurrence rate with our nominal detection sensitivity from Fig. 3.13. Instead, to address this caveat in a simplified way, we simulate a new version of the SLS-PS in which we artificially *increase* the planet occurrence rates by a scaling factor and recompute the SLS-PS detection sensitivity and planet yield. All inputs in this simulation other than the planet occurrence rates are identical to the fiducial survey presented throughout this paper.

To increase the planet occurrence rates we will use a simple scaling factor. That is because in practice, the planet occurrence rates around late M dwarfs are not well-characterized so we adopt the same Kepler occurrence rates  $f(P, r_p)$  but increase it by a factor of 2 (i.e. a revised cumulative planet occurrence rate of  $5 \pm 0.4$  planets per M dwarf). We then conduct the simulated SLS-PS identically as before. However, the dynamical stability considerations (see Sect. 3.5.2) that restrict the types of simulated multi-planet systems can have a more pertinent effect when the number of sampled planets per planetary system is doubled. We find that in practice these considerations cause the resulting planet occurrence rates to only be increased by a factor of  $\sim 1.5$  instead of 2 such that the resulting cumulative planet occurrence rate is 3.56 rather than 5 planets per star.

Fig. 3.16 compares the detection sensitivity and planet yield recovered by the fiducial version of the SLS-PS—with 2.4 planets per star—with those obtained after increasing the cumulative planet occurrence rate to 3.56 planets per star. The detection sensitivity as a function of  $m_p \sin i$  is only slightly modified by the addition of, on average, 1.2 planets per planetary system. That is that, the fiducial version of the survey presented throughout this paper has only a slightly greater detection sensitivity to planets with  $m_p \sin i \in [0.4, 20] M_{\oplus}$  because the fewer planets on average per planetary system contribute to a lower RV rms than when the planet occurrence is increased. Explicitly, the average detection sensitivity over the range of  $m_p \sin i$  considered in Fig. 3.16 is 47.5% for our fiducial survey version compared to 41.9% when the cumulative planet occurrence rate is increased.

The net effect in the new version of the SLS-PS of having more planets per planetary system and a comparable average detection sensitivity is that more planets are detected overall. In each  $m_p \sin i$  bin shown in Fig. 3.16, more planets are detected when the cumulative planet occurrence rate is increased resulting in 1.3 times more SPIRou planet detections ( $110.7^{+28.5}_{-12.6}$  planets) than when assuming the nominal Kepler planet occurrence rates.

Despite there being evidence for an increased planet occurrence rate  $f$  around late M dwarfs than around the early-to-mid M dwarfs observed by Kepler, consideration of an SLS-PS in which the cumulative planet occurrence rate is increased is not intended to be interpreted as a new estimate of the SLS-PS planet yield. This is because our simplified methodology for scaling-up  $f$  is not intended to replicate the true  $f$  around late M dwarfs. We have adopted a simple scaling of  $f$  as measurements of  $f$  around late M dwarfs are currently poorly constrained. Furthermore, there exist some lines of empirical and theoretical evidence (Gillon et al., 2017; Pan & Schlichting, 2017) that the fraction of multi-planetary systems that form low order mean-motion resonant chains is larger around late M dwarfs than is seen in the Kepler sample (Lissauer et al., 2011; Fabrycky et al., 2014). In this case, the Keplerian RV modelling used in this study would not be applicable which may affect the resulting planet yield. However, this exercise is intended to demonstrate that *if* the cumulative planet occurrence rate increases towards later M dwarfs, what is the nature of that effect on our detection sensitivity and ultimately on the SPIRou planet yield. Fortunately, the effect seems to be a positive one for the detection of small planets around late M dwarfs with SPIRou.

## 3.11 Measuring the RV Planet Frequency

### 3.11.1 Recovering planet frequency

With consistent RV monitoring of the target stars in the SLS-PS the resulting detections from the survey will be able to provide independent constraints on the occurrence rate of RV planets around M dwarfs. Of course this calculation has previously been done using either RV or transit survey data (e.g. Bonfils et al., 2013a; Dressing & Charbonneau, 2015), although those studies were limited to early M dwarfs whereas SPIRou is uniquely designed to study mid-to-late M dwarfs whose planet occurrence rates are less certain (Demory et al., 2016). Furthermore, SPIRou is expected to uncover a larger set of planet detections than the HARPS M dwarf sample presented in Bonfils et al. (2013a) and thus provide

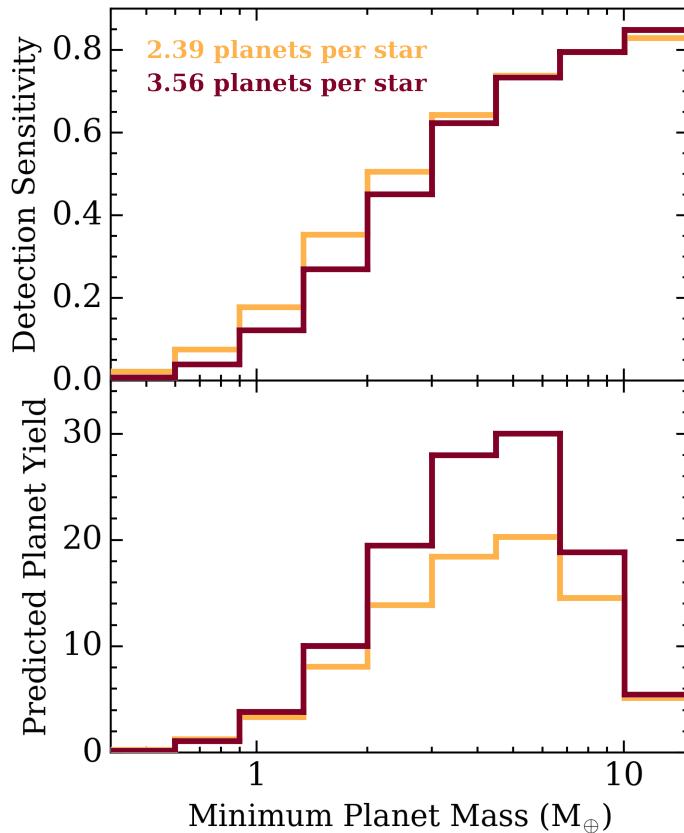


Figure 3.16: *Top:* the detection sensitivity as a function of minimum planet mass for the nominal fiducial version of the SLS-PS containing [2.39 planets per star] and for a modified version of the SLS-PS that is nearly identical except for the increased cumulative planet occurrence rate of [3.56 planets per star]. *Bottom:* the predicted planet yield as a function of minimum planet mass for the two aforementioned survey versions. Over the range of  $m_p \sin i \in [0.4, 15] M_{\oplus}$ , a factor of 1.3 more planets are detected in the survey with the increased cumulative planet occurrence rate.

stronger constraints on the full M dwarf planet occurrence rate as a function of minimum planetary mass.

Here we wish to estimate the precision with which we expect to measure the planet frequency based on the expected results of the SLS-PS. The planet frequency  $f$  differs somewhat from the planetary occurrence rate in that the planetary frequency does not represent the number of a particular type of planet per host star but instead is the fraction of stars which host a particular type of planet and is therefore only defined on the closed interval  $f \in [0, 1]$ . To compute the planet frequency as a function of  $P$  (or similarly  $S$ ) and  $m_p \sin i$ , we will adopt the formalism from [Carson et al. \(2006\)](#); [Lafrenière et al. \(2007b\)](#). In our survey of  $N$  stars denoted by  $j = 1, \dots, N$ , we wish to compute the fraction of M dwarfs that host a planet within a particular range of orbital periods and minimum masses;  $f \rightarrow f(P, m_p \sin i)$ . Firstly, we must estimate the probability that such a planet will be detected around the  $j^{\text{th}}$  star  $p_j$ , based on the star's known detection sensitivity (see Sect. 3.8). The probability that a particular planet will be detected orbiting the  $j^{\text{th}}$  star is then  $fp_j$  whereas the probability of a non-detection is  $1 - fp_j$ . Given the resulting planet yield from our simulated SLS-PS we can identify around which stars a particular planet is detected. For a given range of  $P$  and  $m_p \sin i$ , we denote planet detections within that range around the  $j^{\text{th}}$  star by  $d_j$  which equals 1 if such a planet is detected and 0 otherwise. Now we can write down the likelihood of our planet detections given  $f$  as

$$\mathcal{L}(d_j|f) = \prod_{j=1}^N (1 - fp_j)^{1-d_j} (fp_j)^{d_j}. \quad (3.15)$$

In order to compute the value of  $f(P, m_p \sin i)$  in various  $P, m_p \sin i$  bins, we will invoke Bayes theorem:

$$P(f|d_j) = \frac{\mathcal{L}(d_j|f)p(f)}{\int_0^1 \mathcal{L}(d_j|f)p(f)df} \quad (3.16)$$

where  $p(f)$  is the prior probability of measuring  $f$  and  $P(f|d_j)$  is the posterior PDF of measuring a frequency  $f$  given the observations  $d_j$ . To compute  $f(P, m_p \sin i)$  over the full grid of  $P$  and  $m_p \sin i$ , the above formalism is applied independently to each logarithmically spaced bin in  $P$  and  $m_p \sin i$ . The fraction of M dwarfs with a particular planet is defined as the MAP value of the posterior PDF with its uncertainties characterized by the 16<sup>th</sup> and 84<sup>th</sup> percentiles of the distribution.

Before computing  $f(P, m_p \sin i)$  from the results of our simulated SLS-PS, we must first assign planet detections around each star in the target sample to integer values rather than the statistical averages used to present the results of the full survey. To do so we round the number of detected planets in each  $P, m_p \sin i$  bin to the nearest integer for each star individually and for the full distribution of detected planets from the SLS-PS (see Fig. 3.14). Rounding to integers alters the total number of detected planets so we use a small multiplicative correction factor on each star's detected planet population such that we recover the correct total planet yield of 85 planets. By combining the results of all realizations, the number of planets detected in each  $P, m_p \sin i$  bin over the full simulated survey always exceeds the rounded average value shown in Fig. 3.14. To account for this when assigning the 85 planet detections, we assign planet detections to stars based on their sensitivity to that particular type of planet. Specifically, in each  $P, m_p \sin i$  bin in which we detect at least one planet, we identify the subset of stars which have at least one planet detection in that bin and sort those stars by their detection sensitivities within

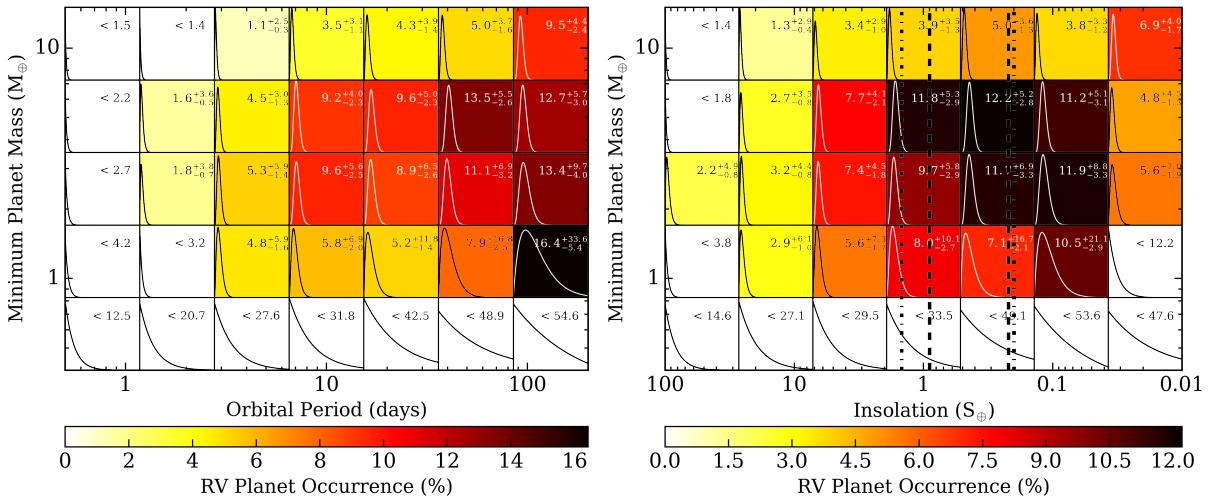


Figure 3.17: Coarsely binned maps of the RV planet frequency derived from the SLS-PS as a function of minimum planet mass and orbital period (left) and insolation (right). The *dashed vertical lines* in the insolation panel indicate the approximate ‘water-loss’ and ‘maximum-greenhouse’ insolation limits of the [HZ] from Kopparapu et al. (2013). The *dashed-dotted vertical lines* indicate the more conservative ‘recent-Venus’ and ‘early-Mars’ HZ limits (Kopparapu et al., 2013). Over-plotted in each bin are the [planet frequency posterior PDFs] derived from the SLS-PS planet detections (Fig. 3.14) and sensitivity (Fig. 3.13) using Eq. 3.16. When the planet frequency PDF is consistent with 0 we report the 68<sup>th</sup> percentile as an upper limit. When a non-zero planet frequency is detected we report the MAP value along with the 16<sup>th</sup> and 84<sup>th</sup> percentiles.

that bin. We then select the  $n$  stars with the highest detection sensitivity where  $n$  is the rounded total number of planets detected within that bin. Those stars have  $d_j$  set to unity within that bin whereas  $d_j = 0$  for all remaining stars. We note that by this routine stars with the highest detection sensitivities are frequently chosen so we limit the number of planets that can be detected around a single star to be  $\leq 3$ ; the maximum number of planets detected by our automated detection algorithm in the simulated SLS-PS. We also note that by selecting stars with the highest detection sensitivity when assigning planet detections, we are maximizing the likelihood (Eq. 3.15) and consequentially computing the maximally constrained planet frequency values.

The planet frequency derived using the predicted SPIRou planet detections spanning  $P \in [0.5, 200]$  days,  $S \in [0.01, 100] S_{\oplus}$ , and  $m_p \sin i \in [0.4, 15] M_{\oplus}$  are shown in Fig. 3.17. For bins in which the MAP value of the  $f$  posterior PDF is non-zero, we report the MAP value and its 16<sup>th</sup> and 84<sup>th</sup> percentiles in Fig. 3.17. For bins in which we detect only a small number of planets or have a low detection sensitivity, we find a MAP  $f = 0$  and report the 68<sup>th</sup> percentile as an upper limit.

By considering the SPIRou planet detections and our detection sensitivity within various ranges of planet parameters, we can estimate the frequency of such planets and the precision of the measurement. We remind the reader that it is the precision of the measurement that is meaningful here as the MAP values of  $f$  are simply the result of combining the known Kepler occurrence rates, an empirical mass-radius relation, and our detection sensitivity and therefore does not provide any new information regarding the planet frequencies themselves. However, our adopted formalism (see Eqs. 3.15 and 3.16) computes the planet frequency over a specified range of planetary parameters. Recall that the planet frequency is not

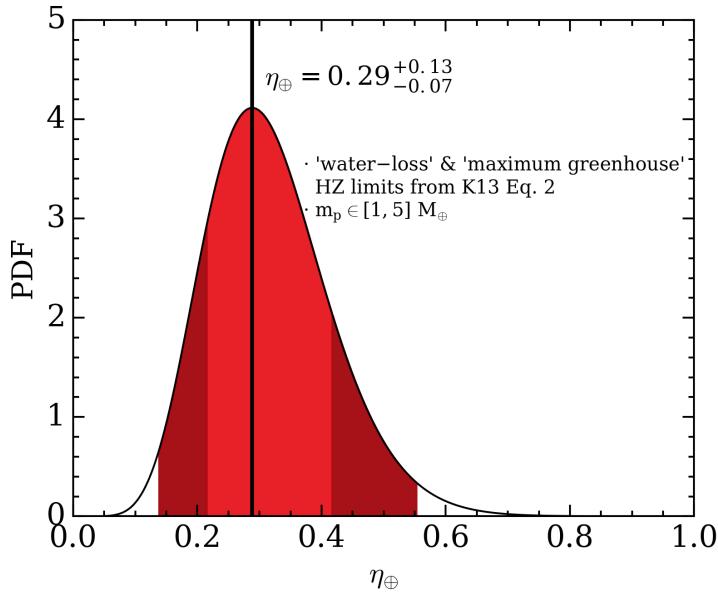


Figure 3.18: The probability density function of the RV value of  $\eta_{\oplus}$  derived from the simulated SLS-PS. Here Earth-like HZ planets are defined according to the ‘water-loss’ and ‘maximum-greenhouse’ HZ limits defined in Kopparapu et al. 2013 (K13) and have absolute planet masses  $\leq 5 M_{\oplus}$ . The *inner shaded* region corresponds to 16<sup>th</sup> and 84<sup>th</sup> percentiles whereas the outer regions mark the 2<sup>nd</sup> and 98<sup>th</sup> percentiles.

equivalent to the planet occurrence rate when the occurrence rate is greater than unity. Therefore, over the full range of  $P$ ,  $S$ , and  $m_p \sin i$  presented in Fig. 3.7, we cannot use this formalism to compute the cumulative SPIRou planet occurrence rate which is  $> 1$  at  $\sim 2.5$  planets per star. Instead we must take the ratio of the SPIRou planet detection map shown in Fig. 3.14—rounded to integer values—with the sensitivity map shown in Fig. 3.13. The resulting map depicts the SPIRou-derived planet frequency, as a function of  $P$ ,  $S$ , and  $m_p \sin i$ , which we can then integrate over to estimate the cumulative planet occurrence rate of  $1.8 \pm 0.2$  planets per M dwarf.

### 3.11.2 Measuring $\eta_{\oplus}$

Using the formalism from Sect. 3.11.1 to compute the planet frequency, we can estimate the frequency of any subset of planets. Of particular interest is the frequency of potentially habitable planets around M dwarfs;  $\eta_{\oplus}$ . We will define potentially habitable—Earth-like—planets as those with  $m_p \in [1, 5] M_{\oplus}$  and within the fiducial HZ period limits defined by the equations in Kopparapu et al. (2013) for the ‘water-loss’ and ‘maximum-greenhouse’. Here the absolute planet mass is inferred from the detected population of minimum planet masses by correcting for the geometrical effect of randomly orientated orbits following a geometrical distribution. The upper limit on  $m_p \sim 5 M_{\oplus}$  approximately corresponds to the expected mass of a  $1.5 R_{\oplus}$  planet which marks the approximate radius boundary between rocky/Earth-like and gaseous planets (e.g. Valencia et al., 2013; Lopez & Fortney, 2014; Fulton et al., 2017). By our definition we expect to detect 9 potentially habitable planets in the full SLS-PS of 100 stars. Based on our derived detection sensitivity to such planets, we measure  $\eta_{\oplus} = 0.28^{+0.12}_{-0.07}$  derived from its posterior PDF shown in Fig. 3.18.

Focusing solely on late M dwarfs (M5–M9) then we expect to detect  $3.4^{+3.2}_{-1.4}$  potentially habitable

planets. As a result we measure a new  $\eta_{\oplus} = 0.29^{+0.24}_{-0.10}$  whose measurement uncertainty is effectively doubled relative to the uncertainty on the value of  $\eta_{\oplus}$  derived from the full SPIRou input catalog. This is because only 37/100 stars from the input catalog are classified as a late M dwarf with spectral type later than M5. The result of a decreased sample size and the, on-average, lower detection sensitivity around dim late M dwarfs, is a greater uncertainty on  $\eta_{\oplus}$  around late M dwarfs. However recall that existing estimates of  $\eta_{\oplus}$  around M dwarfs have been limited to early-to-mid M dwarfs making the SPIRou estimate of  $\eta_{\oplus}$  around late M dwarfs potentially the first of its kind.

### 3.12 Direct Imaging of Nearby Planetary Systems

Assessing the habitability of exoplanets relies heavily on probing the planet's atmospheric conditions. One potential avenue for studying the atmospheres of non-transiting exoplanets is via high-contrast imaging wherein photons from the spatially resolved planet are directly detected following the suppression of quasi-static speckles associated with the bright host star. Various observational techniques such as adaptive optics/coronagraphy and post-processing techniques (e.g. ADI; [Marois et al. 2006](#), LOCI; [Lafrenière et al. 2007a](#), and KLIP; [Soummer et al. 2012](#)) have enabled the direct detection of a number of exoplanets via high-contrast imaging. However, the planet population for which this technique is currently amenable is limited to self-luminous sub-stellar objects at large angular separations from their host star e.g. young gas giants on wide orbits with planet-to-star contrasts of  $\mathcal{O}(10^{-4})$ . However the large apertures on-board the up-coming generation of Extremely Large Telescopes (ELTs) will offer sufficiently high spatial and spectral resolution to reach the small planet-to-star contrasts in the nIR ( $\sim \mathcal{O}(10^{-7})$ ) required to directly image a small number of small HZ planets around the closest stars. Time-resolved rotational color variations of small planets may even permit the detection of the number, reflectance spectra, sizes, and longitudinal positions of major surface features such as continents and/or liquid oceans, and measure cloud properties ([Ford et al., 2001](#); [Fujii et al., 2010, 2011](#); [Cowan et al., 2009](#); [Cowan & Strait, 2013](#)).

To obtain the set of direct images of HZ exoplanets, we must first find the closest habitable worlds. Of particular focus are M dwarfs in the Solar neighbourhood because of their abundance and the favourable contrasts of their HZ planets compared to HZ planets around Sun-like stars ([Crossfield, 2013](#)). The closest HZ exoplanet has already been discovered around Proxima Centauri (1.3 pc; [Anglada-Escudé et al., 2016](#)) but many more M dwarf HZ planets likely remain undetected within  $\lesssim 10$  pc. The SLS-PS will uncover many of these systems and potentially with a lesser total observation time per planetary system than that which is required to detect M dwarf planetary systems using optical velocimeters. To identify the subset of SPIRou detections which are amenable to direct imaging, we adopt the nIR contrast performance expected for a number of dedicated imagers and compare the detected SPIRou planet population to these contrast curves in the planet's projected separation/contrast space.

With the up-coming ELT imagers optimized at nIR wavelengths, targeted planets are observed in reflected light such that the expected planet-to-star contrast is

$$C = 1.81 \times 10^{-7} A \left( \frac{r_p}{1 R_{\oplus}} \right)^2 \left( \frac{a}{0.1 \text{ AU}} \right)^{-2} \quad (3.17)$$

where  $r_p$  is the planet's radius,  $a$  is the semi-major axis, and  $A$  is the geometric albedo which we assume is 0.3 for all detected planets in the simulated SLS-PS. At such low planet-to-star contrasts, no existing high-contrast imager is presently capable of imaging HZ planets. However there are proposed techniques to achieve such small nIR contrasts which involve coupling high-contrast imaging capabilities to spatially resolve the targeted planet followed by the use of high-dispersion spectroscopy to filter out the stellar component as a result of the differentially Doppler-shifted planet and stellar spectra (e.g. Snellen et al., 2015; Lovis et al., 2017). This technique was recently used with CRIRES on the VLT to measure carbon monoxide (Snellen et al., 2010; Brogi et al., 2012; de Kok et al., 2013) and water (Birkby et al., 2013) in the atmospheres of hot Jupiters. This technique has also become considerably more topical since the discovery of a HZ planet in the closest exoplanetary system—Proxima Centauri—and the prospect of detecting the potential biosignature O<sub>2</sub> in this system using the VLT (Lovis et al., 2017).

Similar promising techniques have led to a suite of contrast curve predictions which indicate the types planets which may be detected at a particular detection significance given the capabilities of the imager and the expected S/N achievable based on the target properties and observational strategy. Of these, we will consider the expected geometric mean of the E-ELT EPICS<sup>6</sup> (Kasper et al., 2010) and TMT PFI<sup>7</sup> (Macintosh et al., 2006)  $5\sigma$   $H$  band contrast curves. We also consider the space-based hybrid Lyot coronagraph on-board WFIRST (Trauger et al., 2015).

A random subset of planets from the simulated SLS-PS are shown in Fig. 3.19. Comparing this population to the expected performance of various ELT imaging instruments, we expect  $46.7^{+16.0}_{-6.0}$  SPIRou planets to be imagable<sup>8</sup>. Here we have defined an imagable planet as one whose expected projected angular separation and contrast lie above the geometric mean of the EPICS and PFI  $5\sigma$   $H$  band contrast curves in Fig. 3.19. Note that this definition does not impose a minimum inner working angle although such a cut at say  $3\lambda/D$ , would decrease the total number of imagable SPIRou planets by a factor of  $\sim 2.7$ . The subset of imagable planets represents  $\sim 55\%$  of all SPIRou planets. Among the imagable planets are  $13.7^{+11.5}_{-4.9}$  HZ planets and  $4.9^{+4.7}_{-2.0}$  HZ planets with  $m_p \in [1, 5] M_{\oplus}$ ; the so-called Earth-like planets. These SPIRou planets along with the known Proxima Centauri b, Ross 128b (Bonfils et al., 2017), and GJ 273b (Astudillo-Defru et al., 2017b) will represent the best potential targets for imaging of small HZ exoplanets with ELTs.

### 3.13 Comparison of Different Versions of the SLS-PS

Table 3.2 summarizes the main results of six simulated versions of the SLS-PS including the fiducial version presented throughout this paper which we now refer to as the *optimized* version in Table 3.2. Brief descriptions and motivations for each additional version of the SLS-PS are given below.

1. *Optimized*: the SLS-PS version used throughout this study containing 100 stars in the input catalog. This version approximately represents the optimal compromise between maximizing detection

---

<sup>6</sup>European-Extremely Large Telescope-Exoplanet Imaging Camera and Spectrograph. Since renamed the Planetary Camera and Spectrograph; PCS.

<sup>7</sup>Thirty Meter Telescope-Planet Formation Imager.

<sup>8</sup>Here the term ‘imagable’ need not correspond exactly to an imaging observation as achieving the small planet/star contrasts exhibited by SPIRou planets will likely require high contrast imaging coupled with high-dispersion spectroscopy (e.g. Snellen et al., 2015).

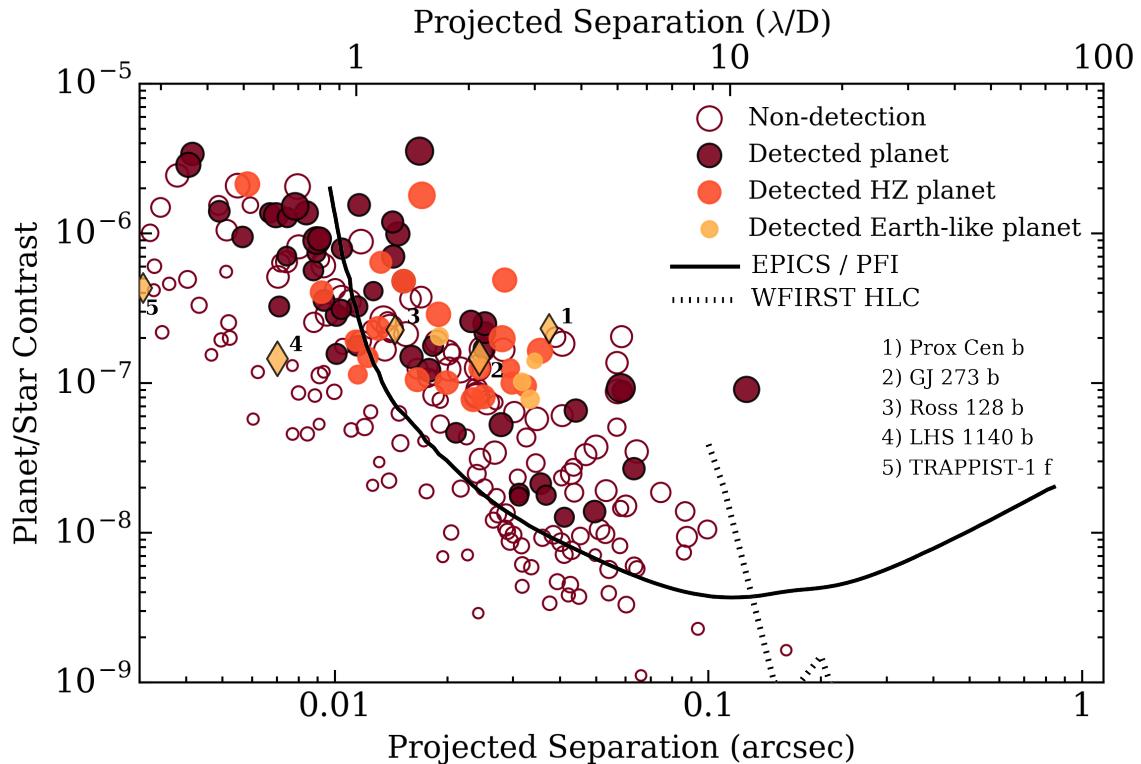


Figure 3.19: The same random subset of the simulated SPIRou planets from Fig. 3.15 in the projected angular separation/reflected light contrast plane. The geometric albedo  $A$  is set to 0.3 for all planets. *Yellow circles* highlight detected Earth-like planets ( $m_p \in [1, 5] M_{\oplus}$ ), *orange circles* highlight the remaining detected HZ planets, and *red circles* highlight all non-HZ detected planets. *Open circles* represent planets that remain undetected by SPIRou. *Yellow Diamonds* labelled 1-5 depict known, likely rocky planets at or near the HZ. The size of each planet's marker is proportional to its radius. The planet population is compared to the geometric mean of the predicted  $5\sigma$   $H$  band contrast curves for EPICS and PFI (solid curve) and the predicted performance for the WFIRST hybrid Lyot Coronagraph (dotted curve). The projected separation is also depicted in units of  $\lambda/D$  in the  $H$  band ( $\lambda = 1.66 \mu\text{m}$ ) for a  $D = 30 \text{ m}$  telescope.

sensitivity and producing a satisfactory number of planet detections, including a sizable fraction of planets that may be amenable to direct imaging with ELTs.

2. *Closest*: contains the 50 closest stars ( $d \lesssim 6.8$  pc) from the *optimized* input catalog. Here we target the closest M dwarfs to the Solar System in an effort to focus observational resources on a small number of target stars thus maximizing detection sensitivity and thus the number of detections of small planets that may be imagable with ELTs.
3. *Large*: contains 360 stars in the input catalog where 360 is the number of targeted stars in the original [SPIRou Science Case](#) proposal from 2013. Targeting a large sample of M dwarfs may result in the greatest planet yield which is desirable for putting tight constraints on the cumulative planet occurrence rates around M dwarfs and  $\eta_{\oplus}$ .
4. *Short*: has an identical input catalog to *optimized* with 100 stars but includes only half of total available time on-sky; we obtain half as many RVs per star as in *optimized*. This experimental setup is undesirable but the simulation is used to demonstrate by how much the SLS-PS detection sensitivity and yield suffer given fewer measurements.
5. *Degraded*: has an identical input catalog and window functions to *optimized* with 100 stars but imposes a degraded noise floor on the RV measurement precision of  $2 \text{ m s}^{-1}$  compared to the  $1 \text{ m s}^{-1}$  noise floor assumed in *optimized*. This survey version is used to access the impact of a degradation in the RV measurement precision on the detection sensitivity and planet yield.
6. *Dark*: is identical to the *optimized* survey with 100 stars but whose window functions are *not* restricted to non-dark-time only. Here we include dark-time observations thus alleviating the strong aliasing of planets with  $P = 15$  or  $30$  days as in *optimized*.

### 3.13.1 Optimized: the optimal survey strategy

The *optimized* survey version represents an ideal compromise between i) achieving sufficient planet detection sensitivity to put meaningful constraints on the planet occurrence rates—including  $\eta_{\oplus}$ —and ii) to produce a large planet yield including a set of imagable planets and, in particular, Earth-like imagable planets from the SLS-PS. For other RV planet search campaigns with similar science goals to those aforementioned, we advocate for a similar survey strategy as *optimized*. With that said, at the time of writing of this manuscript, the SPIRou input catalog has not been absolutely defined and will likely be altered between the time of these simulations and the beginning of the actual SLS-PS. The results presented in this paper are therefore intended as a guideline to inform how many stars should be included in the SLS-PS.

Here we detect  $\sim 8$  Earth-like planets with a sensitivity of  $\sim 33.5\%$ . This will result in a constraint on  $\eta_{\oplus}$  at a level of precision of  $\lesssim 45$ . We also detect  $\sim 5$  Earth-like planets that may be amenable to direct imaging with ELTs. This is the largest number of imagable Earth-like planets detected with SPIRou compared to any other simulated survey version other than *dark*, whose idealized experimental setup is not feasible with the suite of instruments on the CFHT.

### 3.13.2 Closest: the closest M dwarfs

The *closest* survey version reduces the size of the SPIRou input catalog to 50 stars but maintains the same volume of nights. Specifically, the 50 closest stars from the *optimized* SPIRou input catalog are retained. By focusing on fewer stars with the same amount of total survey time as in the *optimized* survey, more observing time can be dedicated to each star. The result is a higher detection sensitivity to all planets around each star compared to *optimized*. In this way, we are able to detect  $\sim 5$  Earth-like planets,  $\sim 4$  of which will be amenable to direct imaging with ELTs due to their close proximity to the Solar System.

Although the sensitivity to any given planet is maximized in the *closest* survey version, the stellar input catalog is too small to result in a larger total planet yield than in *optimized*. The small number of planet detections also has a detrimental effect on measuring the cumulative planet occurrence rate around M dwarfs and particularly late M dwarfs.

### 3.13.3 Large: lots of stars

The *large* survey version was originally considered as the tentative survey strategy for the SLS-PS. Its aim is to discover the greatest number of exoplanets by surveying many more stars (360 stars) than in any other considered survey version. Targeting so many stars comes at the expense of a reduced detection sensitivity per star and a particularly low detection sensitivity to Earth-like planets which seem to require  $\gtrsim 150$  RV measurements to detect (e.g. [Astudillo-Defru et al., 2017b](#)). Although the overall planet yield in the *large* survey version is high, many of the most interesting systems—Earth-like planets that may be imagable with ELTs—will remain largely undetected. For example, we only detect  $\lesssim 2$  imagable Earth-like planets in *large* compared to the  $\sim 5$  in *optimized*. Furthermore, the increase in precision with which the cumulative planet occurrence rate can be measured in *large* is only marginal, and in our opinion, not worth the small yield of imagable Earth-like planets.

### 3.13.4 Short: the effect of fewer observations

The *short* survey version features an identical input catalog to the *optimized* version but with half as many RV observations per star. This survey strategy is useful to characterize the loss in detection sensitivity and planet yield if fewer than the total number of possible measurements are obtained throughout the campaign. The loss in detection sensitivity—and hence in planet yield—compared to *optimized* evolves approximately as  $\sqrt{n_{\text{obs}}}$  for all but the smallest, Earth-like planets. The loss in sensitivity for Earth-like planets is worsened by the small number of measured RVs in *small*. The rough scaling of detection sensitivity with  $\sqrt{n_{\text{obs}}}$  is the result of the use of our GP regression activity modelling to model non-white noise in active time series and that that modelling performs well on the majority of applicable systems (see Fig. 3.10).

### 3.13.5 Degraded: a degraded RV measurement precision

In the *optimized* version of the survey we had assumed that SPIRou will operate with a long-term RV precision of  $\sigma_{\text{RV}} = 1 \text{ m s}^{-1}$ . This was imposed as an RV noise floor on all stars for which we are likely to be able to achieve a photon-noise limited RV uncertainty of  $< 1 \text{ m s}^{-1}$ . However, given that the

long-term RV stability of SPIRou has yet to be tested on-sky, it is conceivable that a degraded level of RV precision  $> 1 \text{ m s}^{-1}$  may instead be realized. Here we repeat the simulation of the *optimized* survey but increase the RV noise floor to  $2 \text{ m s}^{-1}$ .

Similarly to in *short*, the degradation in detection sensitivity approximately scales as  $1/\sqrt{\sigma_{\text{RV}}}$ . The results for *degraded* are therefore closely related to the results for *short*.

### 3.13.6 Dark: no window function restrictions

The *dark* survey version relaxes the assumption used in *optimized* that SPIRou observations may only be obtained during non-dark-time. Although such a scenario is unlikely to be realized for SPIRou on CFHT—given the scheduling of other instruments on the telescope—window functions that *do* include dark-time observations may be obtained with other nIR velocimeters like NIRPS on the ESO 3.6m telescope at La Silla (Bouchy et al., 2017). This experimental setup represents a best-case scenario for SPIRou and illustrates how many more planets can be detected with dark-time observations including a number of HZ planets around M2-4 dwarfs.

We can detect  $\sim 3$  more planets in total compared to *optimized*. Most of these planets lie within the HZ. The number of imagable planets detected is only increased by  $\sim 2$  new planets with less than one additional imagable Earth-like planet in *dark* compared to *optimized*. These modest increases in the planet yield result in only slight improvements to the measured cumulative planet occurrence rate and  $\eta_{\oplus}$ .

## 3.14 Appendix

### 3.14.1 Computing $\sigma_K$ from the Fisher Information Matrix

As discussed in Sect. 3.7.3, we vet putative planet detections by insisting that all bona fide planet detections have an RV semi-amplitude detection significance of at least  $3\sigma$ ;  $K/\sigma_K \geq 3$ . In order to estimate the precision with which a planet’s semi-amplitude  $K$  can be measured in a given RV time series  $y_k = y(t_k)$  we must estimate the semi-amplitude measurement uncertainty  $\sigma_K$  from the Fisher information matrix. The Fisher information matrix  $B$  encodes the amount of information about  $K$  contained within the dataset  $y_k$  and is computed analytically from the ln likelihood given in Eq. 3.10 under some simplifying assumptions. Namely, we treat each detected planet in a given system individually thus restricting the Keplerian model  $\mu_k = \mu(t_k)$  to a fixed number of parameters describing a single Keplerian orbital solution. Secondly, we adopt the simplifying assumption that the planet is on a circular orbit such that only three model parameters (i.e.  $\boldsymbol{\theta} = \{P, T_0, K\}$ ) need be constrained by the time series. When considering planets individually we must absorb all noise sources from unmodelled RV activity and additional planets into an effective RV uncertainty  $\sigma_{\text{eff}}$  equal to the rms of the RVs after the removal of the planet’s Keplerian model. Lastly we assume that the noise properties of the time series are Gaussian distributed such that we can write the ln likelihood as

$$\ln \mathcal{L} = -\frac{1}{2\sigma_{\text{eff}}^2} \sum_{k=1}^{n_{\text{obs}}} (y_k - \mu_k)^2 + c, \quad (3.18)$$

where  $c$  is a constant that is independent of the model parameters in  $\boldsymbol{\theta}$  and  $n_{\text{obs}}$  is the number of observations in the time series. Recall that for a circularized planet the Keplerian orbital solution reduces to a simple sinusoid which we write as

$$\mu_k = -K \sin \phi_k \quad (3.19)$$

where  $\phi_k = 2\pi(t_k - T_0)/P$  and  $T_0$  represents the epoch of inferior conjunction for a circularized planet. The elements of the matrix  $B$  are then computed via

$$B_{ij} = -\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j}. \quad (3.20)$$

Taking  $\boldsymbol{\theta} = \{P, T_0, K\}$ , the symmetric Fisher matrix takes the form

$$B = \begin{bmatrix} B_{P,P} & B_{P,T_0} & B_{P,K} \\ B_{T_0,P} & B_{T_0,T_0} & B_{T_0,K} \\ B_{K,P} & B_{K,T_0} & B_{K,K} \end{bmatrix} \quad (3.21)$$

and contains six independent terms. Here we explicitly compute the analytical forms of each independent element of  $B$  calculated using Eqs. 3.18, 3.19, and 3.20. The first partials of the ln likelihood with respect to each of the model parameters in  $\boldsymbol{\theta}$  are

$$\frac{\partial \ln \mathcal{L}}{\partial P} = -\left(\frac{\pi K}{P^2 \sigma_{\text{eff}}^2}\right) \sum_{k=1}^N (K \sin 2\phi_k - 2y_k \cos \phi_k)(t_k - T_0), \quad (3.22)$$

$$\frac{\partial \ln \mathcal{L}}{\partial T_0} = -\left(\frac{\pi K}{P \sigma_{\text{eff}}^2}\right) \sum_{k=1}^N (K \sin 2\phi_k - 2y_k \cos \phi_k), \quad (3.23)$$

$$\frac{\partial \ln \mathcal{L}}{\partial K} = \left(\frac{1}{\sigma_{\text{eff}}^2}\right) \sum_{k=1}^N (K \sin^2 \phi_k - y_k \sin \phi_k). \quad (3.24)$$

We are now in a position to compute the six independent elements of the Fisher matrix using Eq. 3.20 and the first partials given in Eqs. 3.22, 3.23, and 3.24.

$$\begin{aligned}
B_{P,P} &= -\frac{\partial}{\partial P} \left( \frac{\partial \ln \mathcal{L}}{\partial P} \right) \\
&= \frac{2\pi K}{P^3 \sigma_{\text{eff}}^2} \left[ \sum_{k=1}^N \left( 2y_k \cos \phi_k - \frac{2\pi K(t_k - T_0)}{P} \cos 2\phi_k - \frac{2\pi y_k(t_k - T_0)}{P} \sin \phi_k - K \sin 2\phi_k \right) (t_k - T_0) \right]
\end{aligned} \tag{3.25}$$

$$\begin{aligned}
B_{T_0,P} &= -\frac{\partial}{\partial T_0} \left( \frac{\partial \ln \mathcal{L}}{\partial P} \right) \\
&= \frac{\pi K}{P^2 \sigma_{\text{eff}}^2} \left[ \sum_{k=1}^N \left( 2y_k \cos \phi_k - \frac{4\pi K(t_k - T_0)}{P} \cos 2\phi_k - \frac{4\pi y_k(t_k - T_0)}{P} \sin \phi_k - K \sin 2\phi_k \right) \right]
\end{aligned} \tag{3.26}$$

$$B_{K,P} = -\frac{\partial}{\partial K} \left( \frac{\partial \ln \mathcal{L}}{\partial P} \right) = \frac{2\pi}{P^2 \sigma_{\text{eff}}^2} \left[ \sum_{k=1}^N (K \sin 2\phi_k - y_k \cos \phi_k) (t_k - T_0) \right] \tag{3.27}$$

$$B_{T_0,T_0} = -\frac{\partial}{\partial T_0} \left( \frac{\partial \ln \mathcal{L}}{\partial T_0} \right) = \frac{4\pi^2 K}{P^2 \sigma_{\text{eff}}^2} \left[ \sum_{k=1}^N (-K \cos 2\phi_k - y_k \sin \phi_k) \right] \tag{3.28}$$

$$B_{K,T_0} = -\frac{\partial}{\partial K} \left( \frac{\partial \ln \mathcal{L}}{\partial T_0} \right) = \frac{2\pi}{P \sigma_{\text{eff}}^2} \left[ \sum_{k=1}^N (K \sin 2\phi_k - y_k \cos \phi_k) \right] \tag{3.29}$$

$$B_{K,K} = -\frac{\partial}{\partial K} \left( \frac{\partial \ln \mathcal{L}}{\partial K} \right) = -\frac{1}{\sigma_{\text{eff}}^2} \left[ \sum_{k=1}^N \sin^2 \phi_k \right] \tag{3.30}$$

Using the above expressions to compute the elements of  $B$  we can then compute the covariance matrix  $C$  of the model parameters in  $\boldsymbol{\theta}$  via  $C = |B^{-1}|$ . The diagonal elements of the  $3 \times 3$  matrix  $C$  are the estimated measurement variances of the 3 model parameters  $\boldsymbol{\theta}$ . Therefore the measurement uncertainty of the planet's semi-amplitude is  $\sigma_K = \sqrt{C_{K,K}}$ .

Table 3.2: Overview of SLS-PS Versions

	<i>Optimized</i>	<i>Closest</i>	<i>Large</i>	<i>Short</i>	<i>Degraded</i>	<i>Dark</i>
Fraction of total available time on-sky <sup>a</sup>	1	1	1	0.5	1	1
Number of target stars	100	50	360	100	100	100
Average number of RVs per star	198.1	396.0	62.0	99.5	198.1	198.1
Median $\sigma_{\text{RV}}$ [m s <sup>-1</sup> ]	1.33	1.21	1.88	1.33	2.52	1.33
Median expected $K$ measurement uncertainty [m s <sup>-1</sup> ]	0.19	0.14	0.40	0.25	0.26	0.20
Average detection sensitivity <sup>b</sup> [%]	$44.9 \pm 0.5$	$53.6 \pm 1.1$ (1.2) <sup>c</sup>	$19.5 \pm 0.2$ (0.4)	$34.4 \pm 0.4$ (0.8)	$34.6 \pm 0.6$ (0.8)	$47.2 \pm 0.7$ (1.1)
Average detection sensitivity to HZ planets <sup>d</sup> [%]	$43.1 \pm 1.0$	$51.2 \pm 2.1$ (1.2)	$19.6 \pm 0.3$ (0.5)	$34.8 \pm 0.9$ (0.8)	$34.0 \pm 1.2$ (0.8)	$48.1 \pm 1.3$ (1.1)
Average detection sensitivity to Earth-like planets <sup>e</sup> [%]	$33.5 \pm 1.2$	$45.4 \pm 2.9$ (1.4)	$8.1 \pm 0.3$ (0.2)	$21.4 \pm 1.0$ (0.6)	$20.7 \pm 1.3$ (0.6)	$35.2 \pm 1.5$ (1.1)
Average detection sensitivity to imagable planets [%]	$47.7 \pm 0.7$	$52.8 \pm 1.3$ (1.1)	$24.5 \pm 0.4$ (0.5)	$37.1 \pm 0.6$ (0.8)	$36.0 \pm 0.9$ (0.8)	$49.7 \pm 1.0$ (1.0)
Average detection sensitivity to imagable HZ planets [%]	$46.1 \pm 1.3$	$50.3 \pm 2.3$ (1.1)	$27.9 \pm 1.0$ (0.6)	$37.8 \pm 1.2$ (0.8)	$35.2 \pm 1.6$ (0.8)	$51.5 \pm 1.8$ (1.1)
Average detection sensitivity to imagable Earth-like planets [%]	$33.7 \pm 1.6$	$42.5 \pm 3.1$ (1.3)	$11.4 \pm 1.0$ (0.3)	$20.4 \pm 1.2$ (0.6)	$18.1 \pm 1.7$ (0.5)	$35.7 \pm 2.1$ (1.1)
Total planet yield	$85.3^{+29.3}_{-12.4}$	$50.6^{+15.7}_{-6.5}$ (0.6)	$142.7^{+77.7}_{-29.5}$ (1.7)	$65.7^{+27.8}_{-11.1}$ (0.8)	$65.2^{+27.1}_{-11.0}$ (0.8)	$88.7^{+30.6}_{-13.0}$ (1.0)
Total yield of HZ planets	$20.0^{+16.8}_{-7.2}$	$12.1^{+10.1}_{-4.3}$ (0.6)	$35.3^{+29.5}_{-12.6}$ (1.8)	$16.4^{+13.7}_{-5.9}$ (0.8)	$15.9^{+13.3}_{-5.7}$ (0.8)	$22.9^{+19.1}_{-8.2}$ (1.1)
Total yield of Earth-like planets	$8.1^{+7.6}_{-3.2}$	$5.4^{+5.0}_{-2.1}$ (0.7)	$7.6^{+7.2}_{-3.0}$ (0.9)	$5.2^{+4.9}_{-2.1}$ (0.6)	$5.0^{+4.8}_{-2.0}$ (0.6)	$8.6^{+8.1}_{-3.4}$ (1.1)
Total yield of imagable planets	$46.7^{+16.0}_{-6.0}$	$33.5^{+10.4}_{-3.8}$ (0.7)	$38.3^{+20.9}_{-7.0}$ (0.8)	$36.5^{+15.4}_{-5.4}$ (0.8)	$34.4^{+14.3}_{-5.1}$ (0.7)	$48.3^{+16.7}_{-6.2}$ (1.0)
Total yield of imagable HZ planets	$13.7^{+11.5}_{-4.9}$	$9.7^{+8.2}_{-3.5}$ (0.7)	$8.2^{+6.9}_{-2.9}$ (0.6)	$11.3^{+9.5}_{-4.0}$ (0.8)	$10.1^{+8.5}_{-3.6}$ (0.7)	$15.5^{+12.9}_{-5.5}$ (1.1)
Total yield of imagable Earth-like planets	$4.9^{+4.7}_{-2.0}$	$4.0^{+3.7}_{-1.6}$ (0.8)	$1.5^{+1.4}_{-0.6}$ (0.3)	$3.0^{+2.8}_{-1.2}$ (0.6)	$2.5^{+2.3}_{-1.0}$ (0.5)	$5.3^{+5.0}_{-2.1}$ (1.1)
Cumulative planet occurrence rate <sup>b</sup> [planets per star]	$1.8 \pm 0.2$	$1.8 \pm 0.3$	$1.8 \pm 0.2$	$1.8 \pm 0.2$	$1.8 \pm 0.2$	$1.8 \pm 0.2$
Frequency of Earth-like planets, $\eta_{\oplus}$	$0.29^{+0.13}_{-0.07}$	$0.26^{+0.16}_{-0.08}$	$0.32^{+0.18}_{-0.09}$	$0.30^{+0.18}_{-0.09}$	$0.26^{+0.18}_{-0.09}$	$0.27^{+0.11}_{-0.07}$

**Notes.**<sup>a</sup> For a notional survey duration of  $\sim 300$  nights over  $\sim 3$  years.<sup>b</sup> Over range of  $P \in [0.5, 200]$  days and  $m_p \sin i \in [0.4, 20]$  M<sub>⊕</sub>.<sup>c</sup> Numbers in parentheses indicate the fractional value of the corresponding quantity relative to the *optimized* survey version.<sup>d</sup> Based on the ‘water-loss’ and ‘maximum greenhouse’ limits of the HZ from Kopparapu et al. (2013).<sup>e</sup> Earth-like planets are defined by the ‘water-loss’ and ‘maximum greenhouse’ HZ limits (Kopparapu et al., 2013) and have  $m_p \in [1, 5]$  M<sub>⊕</sub>.

## Chapter 4

# Quantifying the Observational Effort Required for the Radial Velocity Characterization of TESS Planets

### 4.1 Introduction

NASA’s *Transiting Exoplanet Survey Satellite* (TESS; [Ricker et al., 2015](#)) launched in April 2018, is conducting a wide-field survey over at least a 2-year long period and is expected to discover approximately 1700 new transiting exoplanet candidates at a 2-minute cadence around nearby stars over nearly the entire sky ([Sullivan et al., 2015](#), hereafter S15). Due to their proximity, many candidate TESS planetary systems, or TESS objects-of-interest (TOIs), will be amenable to precision radial velocity (RV) observations using ground-based velocimeters to establish their planetary nature and to measure the masses of identified planets. The population of TESS planets to-be discovered are on average systematically closer than the 2342 validated Kepler planets<sup>1</sup> of which only 243 ( $\sim 10\%$ ) have been characterized with RVs.

The growing number of precision velocimeters—and their variety—is vast and includes both optical and near-infrared spectrographs (APF; [Vogt et al. 2014](#), CARMENES; [Quirrenbach et al. 2014](#), CORALIE; ESPRESSO; [Pepe et al. 2010](#), EXPRES; [Jurgenson et al. 2016](#), G-CLEF; [Szentgyorgyi et al. 2016](#), GANS; GIANO; [Oliva et al. 2006](#), GIARPS; [Claudi et al. 2016](#), HARPS; [Mayor et al. 2003](#), HARPS-3, [Thompson et al. 2016](#), HARPS-N; [Cosentino et al. 2012b](#), HDS; [Noguchi et al. 1998](#); HIRES; [Vogt et al. 1994](#), HPF; [Mahadevan et al. 2012](#), IRD; [Kotani et al. 2014](#), iLocater; [Crepp et al. 2016](#), iSHELL, [Rayner et al. 2012](#), KPF; [Gibson et al. 2016](#), MAROON-X; [Seifahrt et al. 2018](#), MINERVA; [Swift et al. 2015](#), MINERVA-Red; [Sliski et al. 2017](#), NEID; [Allen et al. 2018](#), NIRPS; [Bouchy et al. 2017](#), NRES; [Siverd et al. 2016](#), PARAS; [Chakraborty et al. 2008](#), PARVI, PEPSI; [Strassmeier et al. 2015](#), PFS; [Crane et al. 2010](#), SALT HRS; [Crause et al. 2014](#), SOPHIE; [Perruchot et al. 2011](#), SPIRou; [Artigau](#)

---

<sup>1</sup>The contents of this chapter are copied verbatim from the published paper ([Cloutier et al., 2018a](#)).

<sup>1</sup>According to the NASA Exoplanet Archive accessed on March 18, 2018.

et al. 2014, TOU; Ge et al. 2016, Veloce, WISDOM; Férész et al. 2016). Given the large number of velocimeters that can be used for RV characterization of TESS planet masses it is useful to understand the observational effort required to do so. That is, how many radial velocity measurements—and total observing time—are required to detect the masses of the TESS planets at a given significance. Furthermore, it is critical to access which spectrographs are best-suited to the efficient mass characterization of each transiting planet found with TESS. To address these questions, here we present an analytical formalism to compute the number of RV measurements required to detect a transiting planet's mass and apply it to the expected TESS planet yield from S15. Combining these calculations with an exposure time calculator provides estimates of the total observing time required to measure all TESS planet masses and to complete a variety of interesting science cases that will be addressed by TESS.

This paper is structured as follows: Sect. 4.2 describes our model used to compute the total observing time for all TESS objects-of-interest, Sect. 4.3 compares our model to results from existing RV follow-up campaigns of known transiting planetary systems, Sect. 4.4 describes the application of our model to the expected TESS planet population, and Sect. 4.5 reports the results for all TESS planets and for various science cases. We conclude with a discussion and conclusions in Sect. 4.6. Lastly, in the Appendix 4.7.1 we describe our freely available web-tool that utilizes the model from Sect. 4.2 to calculate the total observing time required to detect any transiting planet with a user-defined spectrograph.

## 4.2 Modelling the total observing time required to measure a transiting planet's mass

Here we derive equations to calculate the number of RV measurements  $N_{\text{RV}}$ , of an arbitrary star—in our case a TESS object-of-interest (TOI)—required to measure the mass of its transiting planet at a given detection significance; i.e. with a particular RV semi-amplitude measurement uncertainty  $\sigma_K$ . Together with calculations of the exposure time,  $N_{\text{RV}}$  can be used to compute the total observing time required to detect each TESS planet with RVs.

### 4.2.1 Calculating $\sigma_K$ from the Fisher Information

Given an RV time series of a TOI  $\mathbf{y}(\mathbf{t})$  taken  $t$  times  $\mathbf{t}$ , the amount of information present in the data with regards to the value of the planet's RV semi-amplitude  $K$  is contained in the Fisher information. The amount of information regarding  $K$  is the  $K$  measurement uncertainty  $\sigma_K$ , that is calculated by evaluating the Fisher information given a model of the observed stellar RV variations due to the planet.

The Fisher information matrix  $B$  is defined as the Hessian matrix of the ln likelihood of the data given a model where the model is parameterized by a set of  $n$  parameters  $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_n\}$ . Explicitly,

$$B_{ij} = -\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \quad (4.1)$$

where the indices  $i, j = 1, \dots, n$  and

$$\ln \mathcal{L} = -\frac{1}{2} [(\mathbf{y}(\mathbf{t}) - \boldsymbol{\mu}(\mathbf{t}))^T C^{-1} (\mathbf{y}(\mathbf{t}) - \boldsymbol{\mu}(\mathbf{t})) + \ln \det C + N_{\text{RV}} \ln 2\pi] \quad (4.2)$$

is the generalized ln likelihood of  $\mathbf{y}(\mathbf{t})$  given a model  $\boldsymbol{\mu}(\mathbf{t})$ . The RV time series and model are each 1D vectors containing  $N_{\text{RV}}$  measurements. The matrix  $C$  is the  $N_{\text{RV}} \times N_{\text{RV}}$  covariance matrix of the residual time series  $\mathbf{y}(\mathbf{t}) - \boldsymbol{\mu}(\mathbf{t})$ . Once constructed, the Fisher information matrix can be inverted to return a new covariance matrix; the covariance matrix of the  $n$  model parameters and whose diagonal elements are equal to the measurement variances in the model parameters.

The model of observed stellar RV variations due to a single orbiting planet is a Keplerian solution. Its general form is written as

$$\boldsymbol{\mu}(\mathbf{t}, P, T_0, K, e, \omega) = K[\cos(\boldsymbol{\nu}(\mathbf{t}, P, T_0, e, \omega) + \omega) + e \cos \omega], \quad (4.3)$$

in terms of the star's orbital period  $P$ , time of inferior conjunction  $T_0$ , RV semi-amplitude  $K$ , orbital eccentricity  $e$ , argument of periaapsis  $\omega$ , and true anomaly  $\boldsymbol{\nu}$ . If we assume that the planet's orbit is circular—as was done in S15—then our Keplerian model reduces to

$$\boldsymbol{\mu}(\mathbf{t}, P, T_0, K) = -K \sin \boldsymbol{\nu}(\mathbf{t}, P, T_0) \quad (4.4)$$

where the true anomaly can be expressed as  $\boldsymbol{\nu}(\mathbf{t}, P, T_0) = 2\pi(\mathbf{t} - T_0)/P$ .

### Calculating $\sigma_K$ with white RV noise

Using the Keplerian model given in Eq. 4.4 we can derive a simple analytical expression for  $\sigma_K$  in terms of  $N_{\text{RV}}$  from the Fisher information under a few more simplifying assumptions. From the resulting expression one can fix  $\sigma_K$  to a desired measurement value and calculate  $N_{\text{RV}}$  required to measure  $K$  at that precision. As a initial assumption, where we will assume that the observed RV noise is Gaussian distributed, or white. This assumption is used in the majority of RV analyses in the literature and is especially applicable to planets with semi-amplitudes much greater than the measured point-to-point RV rms. Assuming white noise, the covariance matrix  $C$  in Eq. 4.2 is diagonal with RV measurement variances  $\sigma_{\text{RV}}^2(\mathbf{t})$  along the diagonal. The resulting ln likelihood reduces to

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^{n_{\text{RV}}} \left( \frac{y(t_i) - \mu(t_i)}{\sigma_{\text{RV}}(t_i)} \right)^2, \quad (4.5)$$

modulo a constant offset that is independent of the model parameters.

The second simplifying assumption is rather than considering the full measurement uncertainty time series  $\boldsymbol{\sigma}(\mathbf{t})$ , we will assume that the RV measurement uncertainty is well-characterized over time by a scalar value  $\sigma_{\text{RV}}$ ; a common assumption when deriving model parameter uncertainties from time series observations (e.g. Gaudi & Winn, 2007; Carter et al., 2008). Thirdly, in our Keplerian model we will assume that the values of  $P$  and  $T_0$  are known a-priori with absolute certainty from the planet's TESS transit light curve. Although this is not strictly true,  $P$  and  $T_0$  are often measured at high precision—compared to  $K$ —when multiple transit events are detected. Thus we can treat  $P$  and  $T_0$  as constants rather than as model parameters such that the set of model parameters in our Keplerian RV solution reduces to a single value;  $\boldsymbol{\theta} = \{K\}$ . The Fisher information matrix then reduces to the scalar value

$$B = -\frac{\partial^2 \ln \mathcal{L}}{\partial K^2}, \quad (4.6)$$

$$= \frac{1}{\sigma_{\text{RV}}^2} \sum_{i=1}^{N_{\text{RV}}} \sin^2 \nu(t_i, P, T_0), \quad (4.7)$$

$$= \frac{N_{\text{RV}}}{2\sigma_{\text{RV}}^2}, \quad (4.8)$$

where in the final step we have assumed that the  $N_{\text{RV}}$  measurements are uniformly sampled over the planet's orbital phases such that the summation term in Eq. 4.7 averages to one half of  $N_{\text{RV}}$ .

The inverse of the expression in Eq. 4.8 is the  $K$  measurement variance or

$$\sigma_K = \sigma_{\text{RV}} \sqrt{\frac{2}{N_{\text{RV}}}}. \quad (4.9)$$

This remarkably simple expression for  $\sigma_K$  as a function of the RV measurement uncertainty and number of RV measurements can be rearranged to calculate the value of  $N_{\text{RV}}$  that is required to detect  $K$  of any transiting planet, with a precision of  $\sigma_K$ , when the RV noise can be accurately treated as white.

### Calculating $\sigma_K$ when relaxing the white RV noise approximation

In deriving Eq. 4.9 we must assume that the RV time series noise was Gaussian distributed. However, numerous analyses of transiting systems have shown that there exist cases in which this is a poor assumption. Instead the RV residuals—after the removal of planetary models—can be temporally correlated often owing to the presence of RV signals arising from stellar activity (e.g. Haywood et al., 2014; Grunblatt et al., 2015; López-Morales et al., 2016; Cloutier et al., 2017b; Dittmann et al., 2017b). In such cases, a correlated ‘noise’<sup>2</sup> model must be fit simultaneously with the planetary models to account for all suspected RV signals and any potential correlations between model parameters. One popular choice of correlated noise activity model is a Gaussian process (GP) regression model with a quasi-periodic covariance kernel of the form

$$k(t_i, t_j) = a^2 \exp \left[ -\frac{(t_i - t_j)^2}{2\lambda^2} - \Gamma^2 \sin^2 \left( \frac{\pi|t_i - t_j|}{P_{\text{GP}}} \right) \right], \quad (4.10)$$

and covariance matrix elements

$$C_{ij} = k(t_i, t_j) + \delta_{ij} \sqrt{\sigma_{\text{RV}}^2(t_i) + \sigma_{\text{jitter}}^2}, \quad (4.11)$$

where  $t_i$  is the  $i^{\text{th}}$  observation epoch in  $\mathbf{t}$  and  $\delta_{ij}$  is the Kronecker delta function. The covariance of the GP model function is parameterized by five hyperparameters: the amplitude of the correlations  $a$ , the exponential timescale  $\lambda$ , the coherence parameter  $\Gamma$ , the periodic timescale  $P_{\text{GP}}$ , and an additive scalar jitter  $\sigma_{\text{jitter}}$ . When attempting to measure the semi-amplitude of a known transiting planet, along with a quasi-periodic GP activity model, the full set of model parameters becomes  $\boldsymbol{\theta} = \{K, a, \lambda, \Gamma, P_{\text{GP}}, \sigma_{\text{jitter}}\}$ . Hence our new Fisher information matrix will be  $6 \times 6$  and takes into account the dependence of the  $K$

---

<sup>2</sup>Note that we use the term correlated ‘noise’ whereas—if arising from temporally correlated stellar activity—then this is a signal rather than noise but it is not the planetary signal that we interested in.

measurement precision on the remaining hyperparameters.

The elements of the Fisher information matrix are computed identically as before from Eqs. 4.1 and 4.2 but now using a non-diagonal covariance matrix  $C$  (Eq. 4.11). The derivation of the elements of  $B$  are provided in Appendix 4.7.2. The general effect of computing  $\sigma_K$  from this new Fisher information matrix is to decrease its expected value given a time series of fixed  $N_{\text{RV}}$  compared to the value obtained when using Eq. 4.9. The resulting  $N_{\text{RV}}$  is typically larger than in the white noise limit. However unlike computing  $\sigma_K$  in the white noise limit, the value of  $\sigma_K$  when including a GP correlated noise activity model is dependent on the values of the model parameters themselves and on the time series due to the  $t_i - t_j$  terms in the covariance kernel (Eq. 4.10) and the appearance of  $\mathbf{y}(\mathbf{t})$  and  $\boldsymbol{\sigma}(\mathbf{t})$  in the ln likelihood such that

$$B \rightarrow B(\mathbf{t}, \mathbf{y}, \boldsymbol{\sigma}_{\text{RV}}, \boldsymbol{\theta}). \quad (4.12)$$

Therefore estimating  $\sigma_K$  from the Fisher information including correlated noise requires time series as input and numerical values for all model parameters in  $\boldsymbol{\theta}$ .

We note that when including a GP correlated noise activity model, a simple analytical expression for  $\sigma_K$  in terms of  $\sigma_{\text{RV}}$  and  $N_{\text{RV}}$  cannot be derived. In this case  $\sigma_K$  must be derived as a function of  $N_{\text{RV}}$  from time series of varying  $N_{\text{RV}}$  used to compute the full Fisher information using Eq. 4.1 before calculating the covariance matrix of the model parameters  $C' = B^{-1}$  and ultimately  $\sigma_K$  from its corresponding diagonal matrix element:

$$\sigma_K = \sqrt{C'_{11}}. \quad (4.13)$$

#### 4.2.2 Calculating $\sigma_{\text{RV}}$ for TOIs in the photon-noise limit

Calculating  $\sigma_K$  from RV time series with either white noise or correlated noise is dependent in-part on the RV measurement uncertainty. Here we estimate the photon-noise limited RV measurement precision  $\sigma_{\text{RV}}$  following the formalism from [Bouchy et al. \(2001\)](#). Their formalism is used to calculate  $\sigma_{\text{RV}}$  given the RV information content contained within a stellar spectrum over a particular wavelength range of interest.

From [Bouchy et al. \(2001\)](#) the RV measurement precision  $\sigma_{\text{RV}}$  is shown to be

$$\sigma_{\text{RV}} = \frac{c}{Q \cdot \text{S/N}}, \quad (4.14)$$

where  $c$  is the speed of light,  $Q$  is known as the quality factor of the spectrum, and S/N is the signal-to-noise ratio achieved over the full spectral range considered. The S/N contains contributions from the total number of photoelectrons  $N_{e^-}$  obtained from the source and a contribution from readout noise that begins to dominate the noise budget for the faintest TESS stars. Our S/N prescription is

$$\text{S/N} = \frac{N_{e^-}}{\sqrt{N_{e^-} + N_{\text{ron}}^2}}. \quad (4.15)$$

Throughout this study we assume a fixed readout noise per pixel of 5 e<sup>-</sup> and a 4 pixel PSF sampling in

each orthogonal direction on the detector. The corresponding readout noise is therefore  $N_{\text{ron}} = 20 \text{ e}^1$ . The quality factor

$$Q = \frac{\sqrt{\sum_i W_i}}{\sqrt{\sum_i A_i}} \quad (4.16)$$

is calculated from the noise-free stellar spectrum  $A_i$ —given in photoelectrons and evaluated at the wavelengths  $\lambda_i$ —and from the optimum weighting function given by

$$W_i = \left( \frac{\lambda_i^2}{A_i} \right) \left( \frac{\partial A_i}{\partial \lambda_i} \right)^2. \quad (4.17)$$

The quality factor represents the density of the RV information content in the spectrum  $A_i$ .

When computing  $\sigma_{\text{RV}}$  from Eq. 4.14 we use model stellar spectra from the PHOENIX-ACES library ([Husser et al., 2013](#)). The spectrum for each TOI is retrieved based on the star's effective temperature  $T_{\text{eff}}$  and surface gravity (see [S15](#)) assuming solar metallicity. The native cgs units of flux density for each model spectrum are converted to photoelectrons using the photon energy over the wavelength range provided by the PHOENIX models ( $\lambda \in [0.05, 5.5] \mu\text{m}$ ) and assuming a fixed nominal instrumental throughput of 5%. The spectrum is then segregated into the spectrograph's various spectral bands whose central wavelengths and spectral coverage are summarized in Table 4.1. In each spectral band we mask wavelengths at which the telluric transmission is  $< 98\%$ , where the spectral telluric absorption model is calculated at an airmass of 1 from Maunakea at  $R = 100,000$ . The aforementioned model is obtained from the TAPAS web-tool ([Bertaux et al., 2014](#)). The remaining spectrum in each band that is largely uncontaminated by telluric absorption is resampled assuming a fixed 3 pixel PSF sampling of each resolution element. This spectrum is then convolved with a Gaussian kernel whose full width at half maximum is  $\text{FWHM} = \lambda_0/R$  where  $R$  is the spectral resolution of the spectrograph. After convolving each spectrum with a Gaussian instrumental profile the spectrum is convolved with the rotation kernel presented in [Gray \(2008\)](#) (c.f. Eq. 17.12) that emulates the effect of rotational broadening for stars with a non-zero projected rotation velocity  $v \sin i_s$ . When computing the rotation kernel we adopt a linear limb-darkening coefficient of  $\epsilon = 0.6$ . For each star we compute  $v \sin i_s$  from the known stellar radius (see [S15](#)), the inclination of the stellar spin-axis to the line-of-sight  $i_s$ —drawn from a narrow geometric distribution centered on  $90^\circ$ —and the stellar rotation period  $P_{\text{rot}}$ , that we sample following the methodology described in Sect. 4.2.4.

After the aforementioned convolutions the PHOENIX model's wavelength grid is resampled to a constant  $\delta\lambda = \lambda_0/R$  whose value is specified at the center of a reference band. For the optical spectrograph that will be considered in this study we fix the reference band to be the  $V$  band ( $\lambda_0 = 0.55 \mu\text{m}$ ) whereas the reference band is fixed to the  $J$  band ( $\lambda_0 = 1.25 \mu\text{m}$ ) with the near-IR spectrograph considered.

[Artigau et al. \(2018\)](#) compared the value of  $\sigma_{\text{RV}}$  derived from stellar spectra—as we do here—to the value derived from empirical spectra from HARPS, ESPaDOnS, and CRIRES. They find small discrepancies between these values at optical wavelengths but claim that the RV precision derived from model spectra can be over-estimated in the near-IR YJ bands by  $\sim 2$  and under-estimated in the HK bands by  $\sim 0.5$ . In deriving our own RV measurement precisions from spectral models in the optical or near-IR, we

Table 4.1: Adopted Spectral Bands

Spectral Band	Central Wavelength [μm]	Effective Band Width [μm]	Zero-point Flux Density [erg/s/cm²/μm]
<i>U</i>	0.3531	0.0657	$3.678 \times 10^{-5}$
<i>B</i>	0.4430	0.0973	$6.293 \times 10^{-5}$
<i>V</i>	0.5537	0.0890	$3.575 \times 10^{-5}$
<i>R</i>	0.6940	0.2070	$1.882 \times 10^{-5}$
<i>I</i>	0.8781	0.2316	$9.329 \times 10^{-6}$
<i>Y</i>	1.0259	0.1084	$5.949 \times 10^{-6}$
<i>J</i>	1.2545	0.1548	$2.985 \times 10^{-6}$
<i>H</i>	1.6310	0.2886	$1.199 \times 10^{-6}$
<i>K</i>	2.1498	0.3209	$4.442 \times 10^{-7}$

**Notes.**

Values presented here were obtained from the [SVO Filter Profile Service](#). Explicitly, values pertaining to any of the *UBVRI* bands were obtained from the Generic/Johnson filter set whereas the *YJHK* bands were obtained from the CFHT/Wircam filter set.

apply the multiplicative correct factors derived in [Artigau et al. \(2018\)](#). We note that these corrections were derived based on observations of a single star (i.e. Barnard’s star) whereas the correction factors for other star’s with unique effective temperatures and metallicities may differ from those used here. Despite that, there exists a clear discrepancy between RV measurement precisions derived from model and empirical spectra. This is at least true in a subset of spectral bands. Precise disagreements between model and observationally-derived  $\sigma_{\text{RV}}$  in various bands for *all* spectral types is beyond the scope of this paper but may affect the photon-noise limited RV precision by a factor  $\mathcal{O}(2)$ —depending on the spectrograph—as it does for Barnard’s star from the  $\sim J$  to *K* band. Note that the corresponding effect on total observing times will be  $\mathcal{O}(< 2)$  due to the other contributors to the RV uncertainty in addition to the photon-noise limit (see Sect. 4.2.3).

### 4.2.3 Additional sources of RV noise

Often when searching for transiting planets in radial velocity the residual rms following the removal of the maximum a-posteriori Keplerian planet solution exceeds the characteristic RV measurement uncertainty. This implies the existence of additional sources of RV noise. The effect of these additional noise sources is detrimental to our ability to precisely measure planet masses if not properly modelled. Therefore this excess noise should be taken into account when attempting to estimate  $\sigma_K$  using either Eq. 4.9 or 4.13. These additional sources of noise may be attributed to stellar activity arising from dark spots, plages, and/or faculae, whose corresponding RV signals are modulated by the stellar rotation period and its harmonics, modulo the amplitude of any differential rotation (e.g. [Forveille et al., 2009](#); [Bonfils et al., 2013a](#); [Delfosse et al., 2013a](#)). Another source of dispersion in observed RVs may be from additional planets not seen in-transit (e.g. [Christiansen et al., 2017](#); [Cloutier et al., 2017b](#); [Bonfils et al., 2018](#)); an effect that is especially pertinent when searching for small planets whose RV semi-amplitudes are less than the characteristic RV uncertainty of the time series (e.g. [Astudillo-Defru et al., 2017b](#)).

In order to compute  $N_{\text{RV}}$  for each TOI using our white noise model we will absorb the aforementioned additional noise sources into an effective RV uncertainty  $\sigma_{\text{eff}}$  rather than the previously assumed

measurement uncertainty derived in the photon-noise limit. The effective RV uncertainty is written as the quadrature sum of the systematic RV noise floor of the spectrograph  $\sigma_{\text{floor}}$  (see Table 4.4), the photon-noise limited RV precision, and the RV jitter arising from additional sources of RV noise such as activity and unknown planets:  $\sigma_{\text{eff}} = \sqrt{\sigma_{\text{floor}}^2 + \sigma_{\text{RV}}^2 + \sigma_{\text{act}}^2 + \sigma_{\text{planets}}^2}$ . Eq. 4.9 can then be rearranged for  $N_{\text{RV}}$  in terms of the effective RV uncertainty:

$$N_{\text{RV}} = 2 \left( \frac{\sigma_{\text{eff}}}{\sigma_K} \right)^2. \quad (4.18)$$

In empirical time series with white RV residuals  $\sigma_{\text{eff}}$  can be estimated from the rms of the residual dispersion following the removal of all modelled planets. Eq. 4.18 is not applicable to empirical time series that require a GP correlated noise activity model.

In Sect. 4.3 we will compare the results from real RV campaigns to our analytic estimates as a test of their validity. However unlike in actual RV time series, the RV jitter rms resulting from activity and unknown planets is not known a-priori for any of the TOIs in the S15 synthetic catalog. We therefore need to employ generalized statistical arguments to estimate the expected RV jitter from activity and unknown planets for each TOI. These estimates are described in Sects. 4.2.4 and 4.2.5 and are based on the empirical distributions of RV jitter from each of these two physical effects.

#### 4.2.4 Estimating RV noise due to stellar activity

Here we will consider estimates of the expected RV jitter due to rotationally modulated stellar activity;  $\sigma_{\text{act}}$ . The arguments presented here are intended to be representative of field stars in the solar neighbourhood. Firstly, for each TOI we draw a rotation period  $P_{\text{rot}}$  as a function of the stellar mass from either the Pizzolato et al. (2003) empirical distribution for FGK dwarfs ( $T_{\text{eff}} > 3800$  K) or from the Newton et al. (2016) empirical distribution for M dwarfs ( $T_{\text{eff}} \leq 3800$  K). The corresponding stellar equatorial velocity is calculated using the stellar radius from S15. The projected stellar rotation velocity  $v \sin i_s$  is then calculated after drawing the inclination of the stellar spin-axis from a geometrical distribution. The value of  $v \sin i_s$  acts as a first-order estimate of the star's activity level (e.g. West et al., 2015; Moutou et al., 2017).

For active stars ( $P_{\text{rot}} \lesssim 10$  days), Oshagh et al. (2017) showed through simultaneous K2 photometry and HARPS spectroscopy that monotonic correlations exist between the measured RVs and numerous spectroscopic activity indicators (e.g.  $\log R'_{\text{HK}}$ , FWHM, BIS). Meanwhile quiet stars ( $P_{\text{rot}} \gtrsim 10$  days) appear to lack such strong correlations with spectroscopic activity indicators yet do correlate strongly with  $F_8$ , the photometric flicker or photometric RMS on timescales  $< 8$  hours (Bastien et al., 2013). The  $F_8$  parameter has been shown to correlate with asteroseismic stellar surface gravity measurements (Bastien et al., 2013) that itself correlates with RV jitter (Bastien et al., 2014). Thus for inactive FGK stars ( $P_{\text{rot}} \geq 10$  days) we adopt the following temperature dependent relation from Cegla et al. (2014) for the expected RV dispersion due to stellar activity:

$$\sigma_{\text{act}} = 1 \text{ ms}^{-1} \times \begin{cases} 84.23F_8 - 3.35, & T_{\text{eff}} \geq 6000 \text{ K}, \\ 18.04F_8 - 0.98, & T_{\text{eff}} < 6000 \text{ K}. \end{cases} \quad (4.19)$$

Typical values of  $\sigma_{\text{act}}$  used to derive Eq. 4.19 from [Saar et al. \(2003\)](#) range from  $\sim 0.5 - 10 \text{ m s}^{-1}$  but with a relatively small median value of  $\lesssim 2 \text{ m s}^{-1}$ . We sample  $F_8$  values—measured in parts-per-thousand—from the empirical *Kepler* distribution that has been corrected for their intrinsic *Kepler* magnitude ([Bastien et al., 2013](#)). After sampling  $F_8$  and its uncertainty we use Eq. 4.19 to map to the distribution of  $\sigma_{\text{act}}$  for the inactive FGK stars in the sample of planet-hosting TOIs.

For active FGK stars ( $P_{\text{rot}} < 10$  days) we revert to the  $\log R'_{\text{HK}}$  activity indicator ([Noyes et al., 1984](#)) whose distribution among nearby field FGK stars has been well-characterized ([Henry et al., 1996](#); [Santos et al., 2000](#); [Wright et al., 2004](#); [Hall et al., 2007](#); [Isaacson & Fischer, 2010](#); [Lovis et al., 2011](#)). To estimate  $\sigma_{\text{act}}$  for active FGK stars we compute the corresponding  $\log R'_{\text{HK}}$  from  $P_{\text{rot}}$  and  $B - V$  using the formalism from [Noyes et al. \(1984\)](#). The following formulation from [Santos et al. \(2000\)](#) is then used to map  $\log R'_{\text{HK}} \rightarrow \sigma_{\text{act}}$ :

$$\sigma_{\text{act}} = 1 \text{ ms}^{-1} \times \begin{cases} 9.2R_5^{0.75} & \text{for F dwarfs} \\ 7.9R_5^{0.55} & \text{for G dwarfs} \\ 7.8R_5^{0.13} & \text{for K dwarfs,} \end{cases} \quad (4.20)$$

where  $R_5 = 10^5 R'_{\text{HK}}$ . The rms of the fits in Eq. 4.20 are 0.17, 0.18, and 0.19 dex for FGK stars respectively. In deriving Eq. 4.20 as a function of spectral type, [Santos et al. \(2000\)](#) computed spectral types for each star in their sample based on their CORALIE spectra. However we lack such spectra and instead define the boundaries between FGK stars based on  $T_{\text{eff}}$  given the limited information available for the TOIs. The assumed ranges are  $T_{\text{eff,F}} \in (6000, 7500] \text{ K}$ ,  $T_{\text{eff,G}} \in (5200, 6000] \text{ K}$ , and  $T_{\text{eff,K}} \in (3800, 5200] \text{ K}$ . By computing  $\log R'_{\text{HK}}$  and its uncertainty from sampled values of  $P_{\text{rot}}$  and  $B - V$ , we can map from  $\log R'_{\text{HK}}$  to the distribution of  $\sigma_{\text{act}}$  for the active FGK stars in the sample of planet-hosting TOIs.

Lastly, for M dwarfs there exists a clean relation between  $\log R'_{\text{HK}}$  and  $P_{\text{rot}}$  ([Astudillo-Defru et al., 2017a](#)). The correlation saturates at a maximum mean value of  $\log R'_{\text{HK}} = -4.045$  for rapid rotators with  $P_{\text{rot}} < 10$  days and falls off with rotation period out to the slowest rotating observed M dwarfs with  $P_{\text{rot}} \gtrsim 100$  days. Explicitly, [Astudillo-Defru et al. \(2017a\)](#) find the best-fit step-wise powerlaw to the correlation:

$$\log R'_{\text{HK}} = \begin{cases} -1.509 \log P_{\text{rot}} - 2.550, & P_{\text{rot}} > 10 \text{ days} \\ -4.045, & P_{\text{rot}} \leq 10 \text{ days.} \end{cases} \quad (4.21)$$

The dispersion in the relation for slow rotators is characterized by the uncertainty in the slope and intercept of 0.007 and 0.020 respectively whereas the dispersion in  $\log R'_{\text{HK}}$  for rapid rotators is 0.093. Sampled values of  $P_{\text{rot}}$  for M dwarfs are used to map to  $\log R'_{\text{HK}}$  using Eq. 4.21 from which  $\sigma_{\text{act}}$  values are estimated using a relation similar to Eq. 4.20 but extrapolated to M dwarfs with  $T_{\text{eff}} \leq 3800 \text{ K}$ . The adopted coefficient and powerlaw index for M dwarfs ( $2 \text{ m s}^{-1}$  and 0.1 respectively) were derived from the set of 23 M dwarfs with  $2 \lesssim P_{\text{rot}} \lesssim 150$  days (c.f. Fig. 3 [Cloutier et al. 2018b](#)) whose RV activity rms was characterized with HARPS (X. Delfosse private communication). Additional empirical data to further calibrate these models for M dwarfs are part of an ongoing study with HARPS (Delfosse et al. in prep.) and with a subset of active M dwarfs from CARMENES recently reported ([Tal-Or et al., 2018](#)).

We note that the empirical distributions of RV activity used in this study were derived from observations using optical spectrographs. However, RV activity signals are known to be chromatic as they largely depend on the temperature contrast between an active region and the stellar photosphere where the contrast effect is known decrease from the optical to the near-IR (e.g. [Martín et al., 2006](#); [Huélamo et al., 2008](#); [Prato et al., 2008](#); [Reiners et al., 2010](#); [Mahmud et al., 2011](#)). Meanwhile, Zeeman broadening of spectral features increases with wavelength ([Reiners et al., 2013](#)). The dominant source for RV activity as a function of wavelength and spectral type is not yet fully understood ([Moutou et al., 2017](#)) and so we choose to remain agnostic and set the near-IR RV activity equal to that which is derived in the optical.

#### 4.2.5 Estimating RV noise due to unseen planets

The occurrence rates of planets of various sizes around FGKM stars was well studied with the primary *Kepler* mission (e.g. [Fressin et al., 2013](#); [Dressing & Charbonneau, 2015](#)). For example, the cumulative occurrence rate of planets with radii  $r_p \in [0.8, 22] R_\oplus$  around FGK stars out to 418 days is  $\gtrsim 0.87$  planets ([Fressin et al., 2013](#)). For M dwarfs, small planets with  $r_p \in [0.5, 4] R_\oplus$  with  $P \leq 200$  days appear to be more common with at least 2.5 such planets per M dwarf. In the TESS simulations of [S15](#) up to one transiting planet is detected although the multiplicity of each simulated planetary system is reported. Here we use the number of additional planets around each TOI—along with the known occurrence rates of planets—to estimate the RV contribution due to these planets whose transits are not characterized with TESS.

For TOIs with a reported multiplicity  $N_p > 1$ , we sample the radius  $r_p$  and orbital period  $P$  of the  $N_p - 1$  additional planets from the *Kepler*-derived occurrence rates from [Fressin et al. \(2013\)](#) for FGK dwarfs or from [Dressing & Charbonneau \(2015\)](#) for M dwarfs. Because at most only one planet is detected in transit for each TOI and the transit probability  $\propto P^{-2/3}$ , we draw the orbital periods of additional planets from values greater than the reported orbital period of the known TESS planet. In this way the TESS planet is always the innermost planet in the system and therefore most likely to transit. However, inner non-transiting planets have been detected in known transiting systems as a result of a potentially small mutual inclination ( $\Delta i \sim 1^\circ$ ; [Cloutier et al., 2017b](#)).

When sampling the planet occurrence rates as a function of  $P$  and  $r_p$ ,  $f(P, r_p)$ , a few caveats arise. Firstly,  $f(P, r_p)$  are reported over a coarse grid. Therefore when drawing a planet with a range of potential orbital periods and radii we sample the exact value of  $P$  and  $r_p$  each from a uniform distribution bounded by the edges of that bin. Secondly, due to the poor detection sensitivity to the smallest planets at large orbital periods,  $f(P, r_p)$  is poorly constrained there. To quantify the values of  $f(P, r_p)$  in this regime we assume that  $f(P, r_p)$  evolves smoothly such that in bins where  $f(P, r_p)$  is poorly constrained, we can average the measured values in surrounding bins to populate the previously vacant bin. We restrict all pairs of planets in multi-planet systems to remain Lagrange stable according to the analytic condition from [Barnes & Greenberg \(2006\)](#) while assuming circular orbits for all planets. Lastly, the sampled radii for all additional planets are converted to a planetary mass  $m_p$  according to the empirically

derived mean mass-radius relations from [Weiss et al. \(2013\)](#) or [Weiss & Marcy \(2014\)](#),

$$\frac{m_p}{M_\oplus} = \begin{cases} 0.440 \left( \frac{r_p}{R_\oplus} \right)^3 + 0.614 \left( \frac{r_p}{R_\oplus} \right)^4, & r_p < 1.5 R_\oplus \\ 2.69 \left( \frac{r_p}{R_\oplus} \right)^{0.93}, & 1.5 \leq r_p/R_\oplus < 4 \\ \left( 0.56 \left( \frac{r_p}{R_\oplus} \right) \left( \frac{S}{336.5 S_\oplus} \right)^{0.03} \right)^{1.89}, & 4 \leq r_p/R_\oplus < 13.7 \\ \mathcal{U}(150, 2000), & r_p \geq 13.7 R_\oplus \end{cases} \quad (4.22)$$

where  $S$  is the irradiance received by the planet. By adopting the mean mass-radius relation this formalism neglects to reflect the diversity of exoplanet masses for a given planet radius.

The sinusoidal Keplerian solution with unit amplitude has an rms value of  $\sqrt{2}/2 \sim 0.707$ . Therefore for each additional planet  $i = 1, \dots, N_p - 1$ , we can calculate  $\sigma_{\text{planets},i} = 0.707 K_i$  where  $K_i$  is the planet's semi-amplitude computed from its sampled  $P_i$ ,  $m_{p,i}$ , and host stellar mass. The total value of  $\sigma_{\text{planets}}$  is calculated by the quadrature addition of  $\sigma_{\text{planets},i}$  from each additional planet whose  $K_i < \sigma_{\text{RV}}$ . This latest condition is imposed assuming that additional planets not seen in-transit but whose semi-amplitudes are large compared to the RV measurement precision will be accurately modelled in the RV analysis and therefore not contribute to the residual RV rms.

#### 4.2.6 Exposure time calculator

Together with estimates of  $N_{\text{RV}}$ , the exposure time  $t_{\text{exp}}$  per TOI can be used to calculate the total observing time required to detect a TESS planet in radial velocity. For a given star, the exposure time required to achieve a desired S/N per resolution element will depend on the properties of the spectrograph and telescope used as well as on the star's magnitude in the spectral bands spanned by the spectrograph. For each TOI in this study we calculate the exposure time that is required to reach a S/N of at least 100 at the center of a reference band;  $V$  ( $\lambda = 0.55 \mu\text{m}$ ) or  $J$  ( $\lambda = 1.25 \mu\text{m}$ ) for optical and near-IR spectrographs respectively. The zero point flux densities in each spectral band used to calculate the  $\text{S/N} = \sqrt{N_{e^-}}$  per resolution element are reported in Table 4.1.

The benefit of integrating longer is only to achieve a photon-noise limited RV precision less than a few  $\sigma_{\text{act}}$ , is one of diminishing returns because the effective RV precision becomes dominated by activity and cannot be reduced by increasing  $t_{\text{exp}}$ . Although, it is important to note that increasing  $t_{\text{exp}}$  permits a better sampling for the spectral CCF thus improving one's ability to accurately characterize the activity and mitigate its effects. In cases for which the calculated  $t_{\text{exp}}$  results in  $\sigma_{\text{RV}} > \sigma_{\text{act}}$  or  $>$  the expected  $K$ , we claim that the exposure time is under-estimated. To remedy this we scale-up  $t_{\text{exp}}$  to achieve  $\sigma_{\text{RV}} \lesssim \sigma_{\text{act}}$  and  $\lesssim$  the expected  $K$ . In our exposure time calculator we do impose restrictions on the range of  $t_{\text{exp}}$  that can be considered. Specifically, we restrict  $t_{\text{exp}} \in [10, 60]$  minutes. The shortest permissible exposure time is required to help mitigate the effects of stellar pulsations and surface granulation which evolve on timescales  $\lesssim 10$  minutes ([Lovis et al., 2005](#); [Dumusque et al., 2011b](#)). The upper limit of 60 minutes is applied in order to limit the total observing time dedicated to a single star. In practice, the majority of stars requiring  $> 60$  minutes to achieve the target S/N per resolution element and beat the RV activity rms, will result in a correspondingly low  $\sigma_K$  and will therefore not be amenable to RV characterization within a reasonable timespan. We do note that this upper limit is chosen somewhat

arbitrarily and some observers may wish to increase the maximum exposure time to accommodate certain high-value targets such as temperate Earth-like planets within or near their host star’s habitable zone<sup>3</sup>.

### 4.3 Model Comparison to Observations

In Sect. 4.2 we derived Eq. 4.18 for the number of RV observations required to measure  $K$  of a transiting planet with a precision of  $\sigma_K$  in an RV time series with white noise and an effective RV uncertainty  $\sigma_{\text{eff}}$ . Similar calculations can be made of  $\sigma_K(N_{\text{RV}})$  in the presence of correlated noise using the formalism discussed in Sect. 4.2.1 and Eq. 4.13. Here we compare our analytic estimates of  $N_{\text{RV}}$ —under the applicable noise condition—to observational results from existing RV time series to ensure that our model provides an accurate approximation to  $N_{\text{RV}}$  when applied to the TOIs. We consider two sets of RV time series of transiting planetary systems featuring either a white or correlated noise model. The latter models being restricted to a quasi-periodic GP treatment of correlated RV residuals as was assumed in Sect. 4.2.1. All systems considered must also obey the assumptions imposed when deriving  $\sigma_K$ . To recapitulate, those assumptions are:

1. the planet’s orbital solution is well-approximated as circular.
2. The value of the TESS planet’s  $P$  and  $T_0$  are known to ultra-high precision compared to  $K$  such that correlations between the measured values of  $P$ ,  $T_0$ , and  $K$  are unimportant.
3. The window function of the RV time series is (approximately) sampled uniformly over the planet’s full orbital phase.
4. The white RV time series have a characteristic scalar RV uncertainty equal to the rms of the RV residuals.

To make the analytic estimates of  $N_{\text{RV}}$  for each observed planetary system featuring a white noise model, we use the  $1\sigma$  value of  $\sigma_K$  for one planet in the system. By the nature of the Fisher information the value of  $\sigma_K$  for each modelled planet in a multi-planet system should be equal so long as all ephemerides are well-constrained and planet semi-amplitudes are not correlated. Indeed, the measured planet semi-amplitudes should be uncorrelated when all planets’ are weakly interacting with distinct orbital periods. For each RV time series we set the value of  $\sigma_{\text{eff}}$  to the rms of the RV residuals following the removal of all RV signals modelled by the authors. Some authors report their residual rms values explicitly whilst others treat the stellar activity signal as white by fitting an additive scalar jitter parameter that we add in quadrature to the median RV measurement uncertainty of the time series to estimate  $\sigma_{\text{eff}}$ . The planetary systems with white noise models considered in this analysis are summarized in Table 4.2.

To compare our formalism in the presence of correlated noise to observed systems we consider five cases analyzed with a quasi-periodic GP correlated noise activity model. Namely CoRoT-7 (Haywood et al., 2014), K2-18 (Cloutier et al., 2017b), Kepler-21 (López-Morales et al., 2016), Kepler-78 (Grunblatt et al., 2015), and LHS 1140 (Dittmann et al., 2017b). The calculated value of  $N_{\text{RV}}$  for these systems is obtained by evaluating  $\sigma_K$  from Eq. 4.13 using each systems’ unique RV time series  $\mathbf{t}, \mathbf{y}$ , and  $\sigma_{\text{RV}}$  from their respective papers along with the semi-amplitude and GP hyperparameter values plus uncertainties.

---

<sup>3</sup>There are approximately nine temperate Earth-like planets that will be discovered with TESS with  $T_{\text{eq}} \in [185, 300]$  K and  $r_p \leq 1.5 R_{\oplus}$  (S15).

Table 4.2: Summary of RV Observations for Known Transiting Planets with White RV noise

Planetary System	$\sigma_{\text{eff}}$ [m s $^{-1}$ ]	$\sigma_K$ [m s $^{-1}$ ]	Actual $N_{\text{RV}}$	Calculated $N_{\text{RV}}$	Ref.
GJ 1132	3.38	0.92	25	27.0	1
GJ 1214	4.96	1.60	21	19.2	2
HAT-P-4	10.20	3.00	23	23.1	3
HAT-P-8	10.23	4.20	16	11.9	3
HAT-P-10	6.56	2.70	13	11.8	3
HAT-P-12	5.05	1.60	23	19.9	3
HAT-P-18	20.30	5.20	31	30.5	3
HAT-P-22	9.70	3.20	18	18.4	3
HAT-P-24	13.00	3.60	24	26.1	3
HAT-P-26	3.60	0.98	26	27.0	3
HAT-P-29	11.20	4.60	11	11.9	3
HAT-P-33	66.00	17.50	26	28.4	3
HD 97658	2.78	0.39	96	101.6	4
HD 149026	6.18	1.40	42	39.0	3
HD 189733	15.00	6.00	12	12.5	5
HIP 116454	2.01	0.50	33	32.3	6
Kepler-10	3.55	0.34	220	218.0	7
Kepler-78	2.60	0.40	79	84.5	8
Kepler-93 <sup>a</sup>	1.86	0.27	86	94.9	9
TrES-2	18.10	5.70	19	20.2	3
TrES-4	16.90	10.00	6	5.7	3
WASP-1	7.00	3.20	10	9.6	3
WASP-3	17.60	5.90	15	17.8	3
WASP-4	3.25	2.30	5	4.0	3
WASP-16	2.62	1.60	4	5.36	3
WASP-18	8.77	6.20	6	4.0	3
WASP-19	20.35	5.00	34	33.1	3
WASP-24	4.62	3.20	4	4.2	3
WASP-34	3.60	1.70	8	9.0	3
XO-2N	19.00	8.00	10	11.3	10
XO-5	11.20	3.00	24	27.9	3

**Notes.**<sup>a</sup> HARPS-N measurements only.

- (1) [Berta-Thompson et al. \(2015\)](#), (2) [Charbonneau et al. \(2009\)](#), (3) [Knutson et al. \(2014\)](#), (4) [Howard et al. \(2011\)](#), (5) [Bouchy et al. \(2005\)](#), (6) [Vanderburg et al. \(2015\)](#), (7) [Weiss et al. \(2016\)](#), (8) [Howard et al. \(2013\)](#), (9) [Dressing et al. \(2015\)](#), (10) [Burke et al. \(2007\)](#).

Table 4.3: Summary of RV Observations for Known Transiting Planets with Red RV noise

Planetary System	$K$ [m s $^{-1}$ ]	$a$ [m s $^{-1}$ ]	$\lambda$ [days]	$\Gamma$	$P_{GP}$ [days]	$\sigma_{\text{jitter}}$ [m s $^{-1}$ ]	$\sigma_{\text{eff}}$ [m s $^{-1}$ ]	$\sigma_K$ [m s $^{-1}$ ]	Actual $N_{RV}$	Median Calculated $N_{RV}$	Ref.
CoRoT-7	3.42	7.0	20.6	1.0	23.8	3.44	3.93	0.66	71	$61.4 \pm 2.7$	1
K2-18	3.18	2.8	59.1	1.2	38.6	0.25	4.59	0.75	75	$80.3 \pm 43.7$	2
Kepler-21	2.12	6.7	17.0	2.4	12.6	1.98	4.22	0.66	82	$106.7 \pm 20.8$	3
Kepler-78	1.86	5.6	18.5	2.5	13.3	1.10	1.85	0.25	109	$119.4 \pm 15.0$	4
LHS 1140	5.3	9.0	277.9	2.0	134.0	3.0	9.33	1.1	144	$265.5 \pm 64.4$	5

**Notes.**

- (1) Haywood et al. (2014), (2) Cloutier et al. (2017b), (3) López-Morales et al. (2016), (4) Grunblatt et al. (2015), (5) Dittmann et al. (2017b)

The maximum likelihood parameter values are reported in Table 4.3. Because the model parameters are known to a finite precision we Monte-Carlo sample each model parameter from a Gaussian distribution whose mean is equal to its best-fit value and standard deviation equal to the parameter's measured  $1\sigma$  uncertainty. Evaluating the Fisher information matrix with  $10^3$  model parameter draws results in a distribution of  $\sigma_K$  for each planet from which the distribution of  $N_{RV}$  can be calculated using Eq. 4.18 after  $\sigma_{\text{eff}}$  is derived identically to as in the white noise scenario.

Analytic estimates of  $N_{RV}$  are compared to observed values for known planetary systems in Fig. 4.1. As evidenced in Fig. 4.1, the majority of planetary systems have calculated  $N_{RV}$  values in close agreement with observed values for both the white and correlated noise scenarios. This demonstrates that our analytical models for  $N_{RV}$  are valid for the majority of cases with one notable exception. Quantitatively, the rms of the O-C  $N_{RV}$  values is 2.6 for the white noise cases alone and 5.2 for all planetary systems included in Fig. 4.1 with the exception of the curious outlier LHS 1140. Our calculated value of  $N_{RV}$  for LHS 1140 is overestimated relative to the size of the RV time series presented in Dittmann et al. (2017b) from which  $K$  is measured to be  $5.34 \pm 1.1$  m s $^{-1}$ . After Monte-Carlo sampling  $K$  and the GP hyperparameters from their measurement uncertainties we calculate a median  $N_{RV} = 265 \pm 64$  which is  $\sim 1.8$  times greater than the actual time series size ( $N_{RV} = 144$ ) at  $1.9\sigma$ . The exact cause of this anomalous discrepancy is not known but may be related to how the GP covariance function is implemented although this investigation is beyond the scope of this paper.

## 4.4 Overview of computing $N_{RV}$ for the expected TESS planet population

S15 predicted the population of planets that will be discovered with TESS in its 2-minute cadence observing mode. Their results are provided for one realization of their simulations and contains 1984 TOIs, each with a single transiting planet. The properties of their stellar sample is copied in Table 4.6 for easy reference. This realization contains more detected planets than the average of their simulations;  $\sim 1700$ . We treat each TOI in the S15 sample as a bona-fide exoplanet and not as a false positive. However, some number of TOIs will ultimately be identified as false positives as historically the false positive rate of transit surveys like Kepler have yielded higher false positive rates than initially anticipated (Sliski & Kipping, 2014; Morton et al., 2016). The properties of the adopted planet population were derived from planet occurrence rates measured with Kepler circa 2015. Some planetary properties, particularly the planetary radii, have since been modified slightly following the reanalysis of Kepler-planet host star

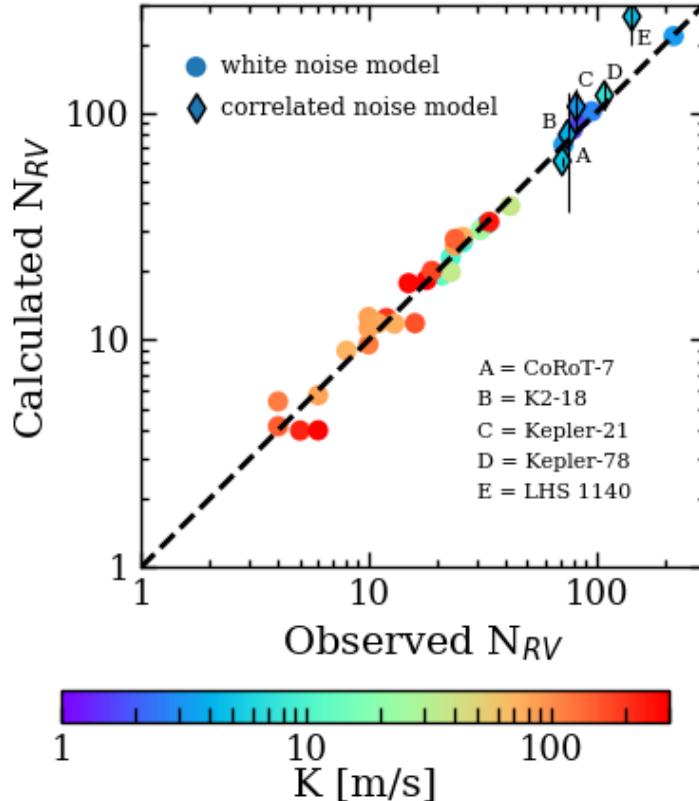


Figure 4.1: Observed values of  $N_{RV}$  compared to the calculated values including either white or correlated RV noise for a suite of known transiting planetary systems. The colorbar indicates each planet’s RV semi-amplitude. The *dashed line* depicts the line  $y = x$  wherein calculated  $N_{RV}$  correspond exactly to the observed  $N_{RV}$ . The region above the line depicts where the observed  $N_{RV}$ —for a given  $K$  measurement precision—are less than the value predicted by the Fisher information and would be considered anomalous. The region below the line depicts where the observed  $N_{RV}$  are larger than the value predicted by the Fisher information implying that the  $K$  measurement uncertainty may be under-estimated.

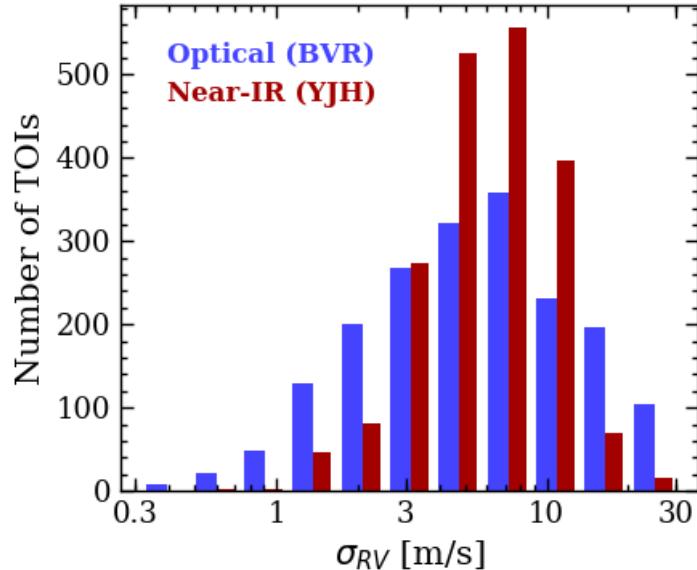


Figure 4.2: The distributions of the TOI photon-noise limited RV measurement precisions derived from PHOENIX stellar models and using the formalism discussed in Sect. 4.2.2. The values of  $\sigma_{RV}$  are computed for each of our fiducial spectrographs in the optical and near-IR with spectral bands *BVR* and *YJH* respectively (see Table 4.4).

properties (e.g. [Fulton et al., 2017](#)).

Here we compute analytical estimates of  $N_{RV}$  required to measure a planet’s RV semi-amplitude—at a given precision—for the entire synthetic catalog of TESS planets from [S15](#). Together with the exposure time calculator described in Sect. 4.2.6 we calculate the total observing time required to detect each planet. These calculations require estimates of RV noise sources from the RV noise floor of the employed spectrograph, photon noise, activity, and additional unseen planets. Photon noise is dependent on the RV information content contained within the stellar spectrum and varies across spectral bands. Therefore we consider RV follow-up observations taken with either a fiducial optical spectrograph or a fiducial near-IR spectrograph. The specifications corresponding to our adopted optical spectrograph are modelled after the HARPS spectrograph on the 3.6m ESO telescope at La Silla observatory ([Mayor et al., 2003](#)). The specifications corresponding to our adopted near-IR spectrograph are modelled after the up-coming NIRPS spectrograph that will join HARPS at the 3.6m ESO telescope at La Silla observatory in 2019 ([Bouchy et al., 2017](#)). The adopted specifications for our two fiducial spectrographs are given in Table 4.4. Using the formalism discussed in Sect. 4.2.2 we calculate  $\sigma_{RV}$  for each TOI with both spectrographs. For reference, the distributions of TOI  $\sigma_{RV}$  for both spectrographs are shown in Fig. 4.2.

In Sect. 4.7.1 we will present a web-based  $N_{RV}(\sigma_{RV})$  calculator that can be used to repeat these calculations but using *any* spectrograph defined by the user.

As discussed in Sects. 4.2.3, 4.2.4, and 4.2.5, estimates of  $N_{RV}$  are also sensitive to additive RV noise sources such as stellar activity and unseen planets. However, these astrophysical noise sources are not known a-priori for the [S15](#) TOIs and therefore must be sampled from known distributions of

Table 4.4: Fiducial Spectrograph Specifications

	Optical	Near-IR
Spectral bands	<i>BVR</i>	<i>YJH</i>
Spectral resolution, $R$	115,000	100,000 <sup>a</sup>
Telescope aperture, $D$ [m]	3.6	3.6
RV noise floor, $\sigma_{\text{floor}}$ [m s <sup>-1</sup> ]	0.5	1.0
Throughput, $\epsilon$	0.05	0.05

**Notes.**

<sup>a</sup> We consider the higher resolution achievable in the NIRPS High Accuracy Mode relative to its High Efficiency Mode at  $R = 75,000$ .

applicable values of  $\sigma_{\text{act}}$  and  $\sigma_{\text{planets}}$ . A Monte-Carlo sampling routine is used for each TOI to sample the aforementioned quantities from the distributions discussed in Sects. 4.2.4 and 4.2.5. The value of  $N_{\text{RV}}$  for each TOI also depends on the desired  $K$  measurement uncertainty that is set by the nature of the follow-up science that one wishes to conduct once the planet’s mass has been characterized with RVs. For example, conventionally the *detection* of a planet’s  $K$  requires a  $3\sigma$  detection significance;  $3 = K / \sigma_K$ . Conversely, targets that will be amenable to atmospheric characterization via transmission spectroscopy will benefit from a more precise measurement of the planet’s bulk density that in-turn requires a mass detection significance  $> 3\sigma$ .

In the following Sect. 4.5 we present results of  $N_{\text{RV}}$  for the full TESS sample in the limit of correlated RV noise (see Sect. 4.2.1). We also consider four subsamples of TESS planets each pertaining to a unique science case that will be addressed by TESS. Each science case merits a unique choice of  $\sigma_K$  for a particular subset of TOIs. Total observing times are then calculated as  $N_{\text{RV}} \cdot (t_{\text{exp}} + t_{\text{overhead}})$ . Because the exact value of  $t_{\text{overhead}}$  varies between observatories, we will set  $t_{\text{overhead}} = 0$  such that its effect can easily be added to the total observing times later-on for non-zero values. To estimate  $N_{\text{RV}}$  in the presence of correlated RV noise we construct time series of increasing  $N_{\text{RV}}$  from  $10 - 10^3$  in steps of 90. For our time series we adopt a simple uniform window function  $\mathbf{t}$  spanning 100 days which is sufficient to sample the full orbit of  $\sim 99\%$  of TOIs. Such simplistic time sampling is admittedly unrealistic given the expected number and frequency of nights lost due to poor observing conditions (clouds, poor seeing, etc). To that end we tested more complex window functions which included longer baselines and gaps due to observing seasons. We found results roughly consistent with the uniform window functions although this need not be true for *any* window function with arbitrary complexity such as those which are often obtained in practice over many observing seasons. For each of the 12 time series with a unique  $N_{\text{RV}}$  we compute  $\sigma_K$  before interpolating  $N_{\text{RV}}(\sigma_K)$  to the desired value of  $\sigma_K$ . The initial guesses of the GP hyperparameters  $\{\lambda, \Gamma, P_{\text{GP}}\}$  are adopted from [Dittmann et al. \(2017b\)](#) (see Methods section *Radial-velocity analysis with Gaussian process regression.*) with the remaining GP hyperparameters set to  $a = \sqrt{2}\sigma_{\text{act}}$  and  $\sigma_{\text{jitter}} = \sigma_{\text{planets}}$ . The RV time series  $\mathbf{y}(\mathbf{t})$  contains Keplerian contribution from the TESS planet and other sampled planets if applicable, plus correlated noise from a sample of the GP prior distribution, and white noise featuring contributions from the RV noise floor of the spectrograph (Table 4.4) and the photon-noise limited measurement precision:  $\sqrt{\sigma_{\text{floor}}^2 + \sigma_{\text{RV}}^2}$ . The RV measurement uncertainty time series  $\boldsymbol{\sigma}_{\text{RV}}(\mathbf{t})$  contains the aforementioned value repeated  $N_{\text{RV}}$  times.

## 4.5 Results for the TESS sample

### 4.5.1 Detecting all TESS planet masses at $3\sigma$

Here we present the results of attempting to detect the masses of *all* TESS planets at  $3\sigma$  using either the optical or near-IR spectrograph. As such, we do not make a cut in declination and restrict targets to half of the sky. Our fiducial spectrographs are intended to be representative of suites of spectrographs—with comparable on-sky performance—thus providing full sky coverage. Realistically not all TESS planets will be characterized with RVs due to either their small RV semi-amplitude, certain intrinsic stellar host properties that deter RV observations (e.g. a low apparent magnitude, rapid projected stellar rotation, or high levels of stellar activity), or simply due to a lack of available observing time. Despite this fact we present the results for *all* TESS planets.

Detecting a planet’s mass at  $3\sigma$  requires a  $K$  detection significance that is slightly larger than three because the calculation of  $m_p$  from  $K$  is also dependent on other observables such as the orbital period and stellar mass whose measurement uncertainties contribute to the  $m_p$  measurement uncertainty. To calculate the value of  $\sigma_K$  required to achieve  $m_p/\sigma_{m_p} = 3$  we first assume that the orbital period of the planet is known to a sufficiently high fractional precision relative to the other parameters of interest (i.e.  $\sigma_P/P \ll 1$ ) such that its contribution to  $\sigma_{m_p}$  can be effectively ignored. Secondly, we assume throughout this study that all stellar masses are measured with a conservative precision of 10% as many field dwarfs in the solar neighbourhood have their masses measured with a precision of  $\lesssim 10\%$  from mass-luminosity relations (Delfosse et al., 2000; Torres et al., 2010). However, this assumed fractional precision will not hold for all TOIs as a subset will have their masses characterized more precisely using other advanced techniques such as asteroseismology or spectroscopy coupled with precision parallaxes (e.g. Van Eylen et al., 2018; Fulton & Petigura, 2018). We note that the calculations presented here represent conservative values if TOI stellar masses can be determined to a precision higher than 10%. Nevertheless, under our current assumptions a  $3.06\sigma$  detection of  $K$  (i.e.  $\sigma_K = 0.327K$ ) is required to detect  $m_p$  at  $3\sigma$ .

The median results—over Monte-Carlo realizations—of our calculations are reported in Table 4.7 for each TOI. Specifically, we report the median photon noise-limited RV precision in both the optical and near-IR spectrographs,  $\sigma_{\text{act}}$ ,  $\sigma_{\text{planets}}$ ,  $N_{\text{RV}}$ , and total observing times in each spectrograph;  $t_{\text{obs, opt}}$  and  $t_{\text{obs, nIR}}$ . The values of  $N_{\text{RV}}$ —and the corresponding  $t_{\text{obs}}$ —are derived from the general case which includes a GP treatment of correlated RV noise.

In Fig. 4.3 we compare  $t_{\text{obs, opt}}$  and  $t_{\text{obs, nIR}}$  as a function of TOI effective temperature to ascertain which flavor of spectrograph is favorable for efficient RV planet mass characterization. A clear trend is discernible with the ratio of the median optical to the near-IR total observing times decreasing towards earlier spectral types. Efficient characterization of planets around late TOIs with  $T_{\text{eff}} \lesssim 3800$  K is significantly favoured by the use of near-IR spectrographs due to the reduced photon noise exhibited by those stars in the near-IR. Conversely for TOIs with  $T_{\text{eff}} \gtrsim 5500$  K, the optical spectrograph is preferred. For intermediate TOIs the two spectrographs offer nearly consistent performance.

Modulo the effects of rotation and stellar activity the observing time required to detect a transiting planet with RVs is dependent on the host star’s brightness and spectral type which directly effect  $\sigma_{\text{RV}}$ .

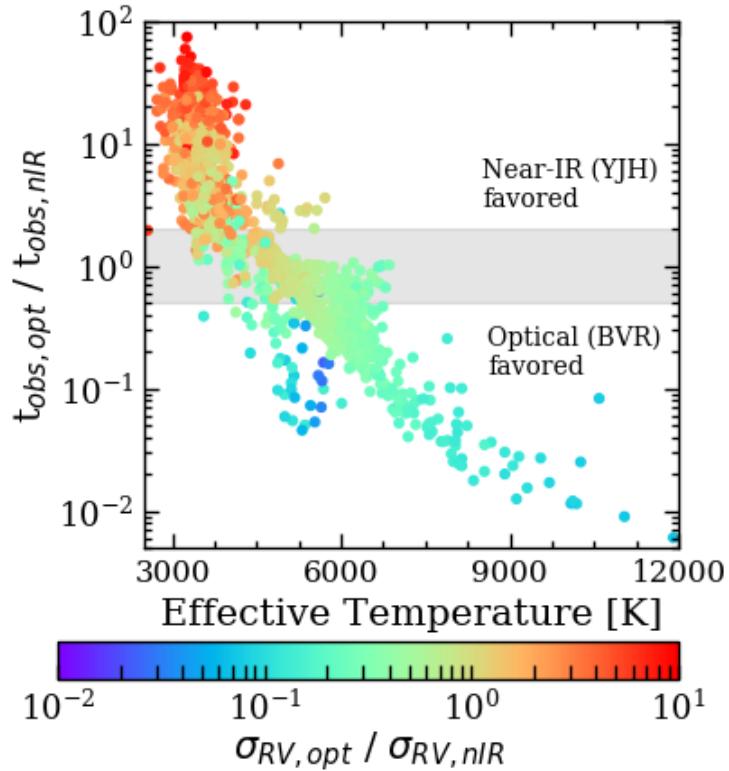


Figure 4.3: Point estimates of the ratio of the median total observing times with an optical spectrograph to with a near-IR spectrograph as a function of TOI effective temperature. Colors are indicative of the ratio of the photon-noise limited RV precision in the optical to the near-IR. Near-IR RV observations are a more efficient means of planet mass characterization when this ratio is a few times unity. Conversely, optical RV observations are a more efficient means of planet mass characterization when this ratio is a few time smaller than unity. The shaded region spanning the ordinate  $\in [1/2, 2]$  approximately depicts where the use of either an optical or near-IR spectrograph offer nearly consistent performance.

Recall that our derived values of  $\sigma_{\text{RV}}$  for each TOI are based on stellar spectral models rather than on empirical spectra such as the M dwarfs observed at high resolution ( $R > 80,000$ ) with the CARMENES visible and near-IR channels (Reiners et al., 2018). Reiners et al. (2018) claim that the RV information content peaks in the RI bands between 700-900 nm for all M dwarfs. This empirical evidence somewhat contradicts theoretical calculations based on model spectra (e.g. Figueira et al. 2016, this study), especially for late M dwarfs whose RV information content is theorized to peak in the HK bands from 1.4-2.4  $\mu\text{m}$ . Convergence towards a more precise scaling of  $\sigma_{\text{RV}}$  with wavelength—for stars of various spectral types—will be achieved in the near future with the onset of multiple new spectrographs from the optical to the near-IR.

### Detecting TESS planet masses versus TOI spectral type

The cumulative median observing time required to detect TESS planets as a function of TOI spectral type is shown in Fig. 4.4 up to  $10^3$  hours. The results are also given in terms of the cumulative number of observing nights assuming a notional value of 7 observing hours per night. We consider spectral type bins with the following adopted definitions: mid-late M dwarfs:  $2500 \leq T_{\text{eff}}/K < 3200$ , early-mid M dwarfs:  $3200 \leq T_{\text{eff}}/K < 3800$ , FGK dwarfs:  $3800 \leq T_{\text{eff}}/K < 7600$ , and BA dwarfs:  $7600 \leq T_{\text{eff}}/K < 12000$ . Spectral type bins are considered separately because of the clear trend exhibited in total observing times with either an optical or near-IR spectrograph with  $T_{\text{eff}}$  as seen in Fig. 4.3. For example, it is clear that all 39 TESS planets around BA stars can be detected with our optical spectrograph in  $\sim 140$  nights whereas only  $\sim 12$  of those planets can be detected with the near-IR spectrograph in a thousand hours (i.e.  $\sim 143$  nights). Planet detections around Sun-like stars (i.e. FGK) are obtained more efficiently with the optical spectrograph with  $\sim 251/964$  optical detections compared to  $\sim 198/964$  near-IR detections in a thousand hours. Efficient M dwarf planet detections favor the near-IR spectrograph wherein a thousand hours of observing time yields  $\sim 165/927$  early-mid M dwarf planets or nearly all  $\sim 54$  mid-late M dwarf planets. These numbers are reduced to  $\sim 60/927$  and  $\sim 21/54$  in a thousand hours with our optical spectrograph.

Further demonstrated in Fig. 4.4 is the first derivative of the total number of planet detections with cumulative observing time;  $dN/dt$ . This quantity describes the efficiency of detecting planets over time as large values of the derivative highlight when planet masses may be detected in a short amount of observing time. We will continue by referring to this quantity as the *detection efficiency*. The detection efficiency can be used to identify after how much total observing time further planet detections become too observationally expensive. That is that when the detection efficiency drops below a set threshold value, any additional planet detections will require too much observing time that may otherwise be spent on potentially more feasible targets. For reference in Fig. 4.4 we highlight the value of the inverse time derivative—which is still a measure of detection efficiency—equal to 20 hours per detection. We suggest this value as a minimum derivative value. With this threshold value planets around BA stars should be observed for up to  $\sim 29$  nights with an optical spectrograph before the detection efficiency drops below this threshold value. Similarly, planets around mid-late M dwarf TOIs should be observed for  $\sim 43$  nights with a near-IR spectrograph. Observing Sun-like and early-mid M dwarf stars can mostly proceed efficiently beyond  $10^3$  hours when using either spectrograph however observing early-mid M dwarfs slowly approaches 20 hours per detection after  $\sim 110$  nights. Observing all TOIs with the optical spectrograph—or with a network of optical spectrographs of comparable performance—until we reach a detection efficiency of 20 hours per detection would require  $\sim 800$  nights of cumulative observing time.

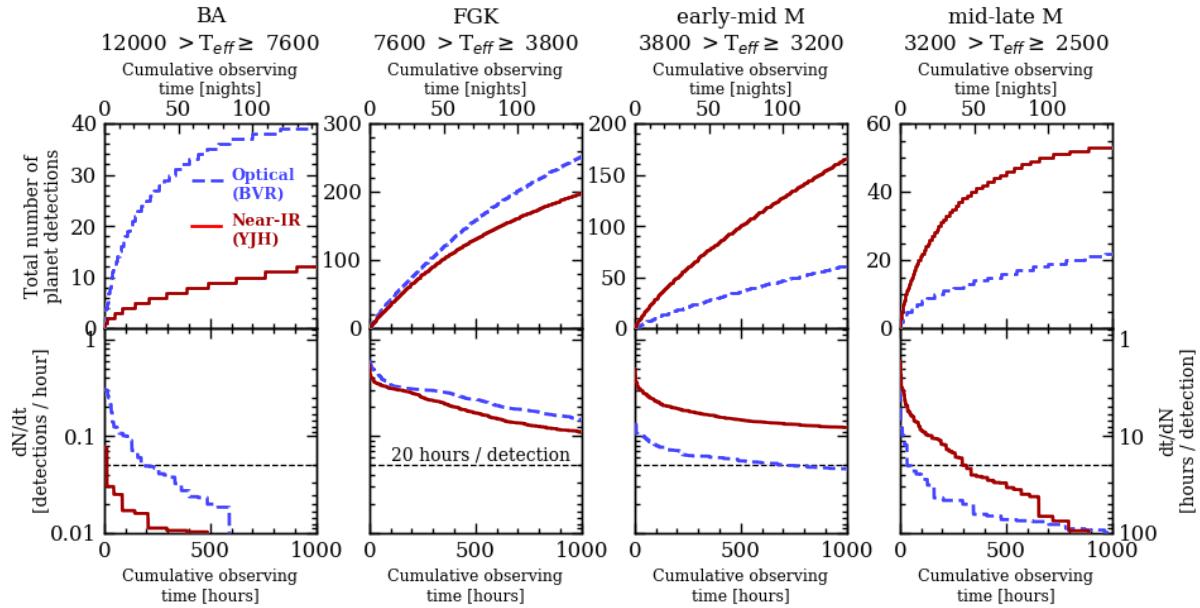


Figure 4.4: *Upper panel:* the cumulative median observing time to measure the  $3\sigma$  RV masses of TESS planets as a function of host star spectral type and up to  $10^3$  hours. The [*dashed blue curves*] represent the results from the optical spectrograph whereas the [*solid red curves*] represent the near-IR spectrograph. *Lower panel:* the time derivative of the cumulative observing time curves used to indicate the RV planet detection efficiency. The *horizontal dashed line* highlights the value of the detection efficiency at 20 hours per detection. Note that unlike the lower panels, the upper panels do not share a common ordinate due to the differing number of planet detections around stars in each spectral type bin.

In that time,  $\sim 620$  planets could be detected around TOIs of any spectral type with  $V \leq 15$ . Repeating this observing campaign with the near-IR spectrograph—or with a network of near-IR spectrographs of comparable performance—would require  $\sim 1600$  nights of cumulative observing time. In that time  $\sim 1030$  planets could be detected around TOIs of any spectral type with  $J \leq 13.6$ .

### Detecting TESS planet masses versus planet type

The cumulative median observing times required to detect TESS planets as a function of planet type are shown in Fig. 4.5 up to  $10^3$  hours. We consider four types of planet defined by their radii to be Earths ( $< 1.25 R_{\oplus}$ ), super-Earths ( $1.25 - 2 R_{\oplus}$ ), Neptunes ( $2 - 4 R_{\oplus}$ ), and giants ( $> 4 R_{\oplus}$ ).

Of the 1984 planets in the S15 TESS sample, 66 are classified as Earths with 26 around stars with  $V \leq \text{median}(V) = 13.5$  and 36 around stars with  $J \leq \text{median}(J) = 10.7$ . Most Earths will be detected in-transit around M dwarfs ( $T_{\text{eff}} \leq 3800$  K) due to their favorable transit depths. With a thousand hours of total observing time we expect  $\sim 26/66$  Earths to be detected with the near-IR spectrograph compared to  $\sim 15$  detections in the optical. We note that detections of the smallest planets can be expensive as the detection efficiency exceeds 20 hours per detection after  $\sim 11$  nights or after just  $\sim 8$  detections in the near-IR. The detection efficiency drops more rapidly in the optical to greater than 20 hours per detection after just  $\sim 6$  nights or  $\sim 4$  detections.

Super-Earths can be detected rather efficiently in the near-IR to beyond  $10^3$  hours and with an

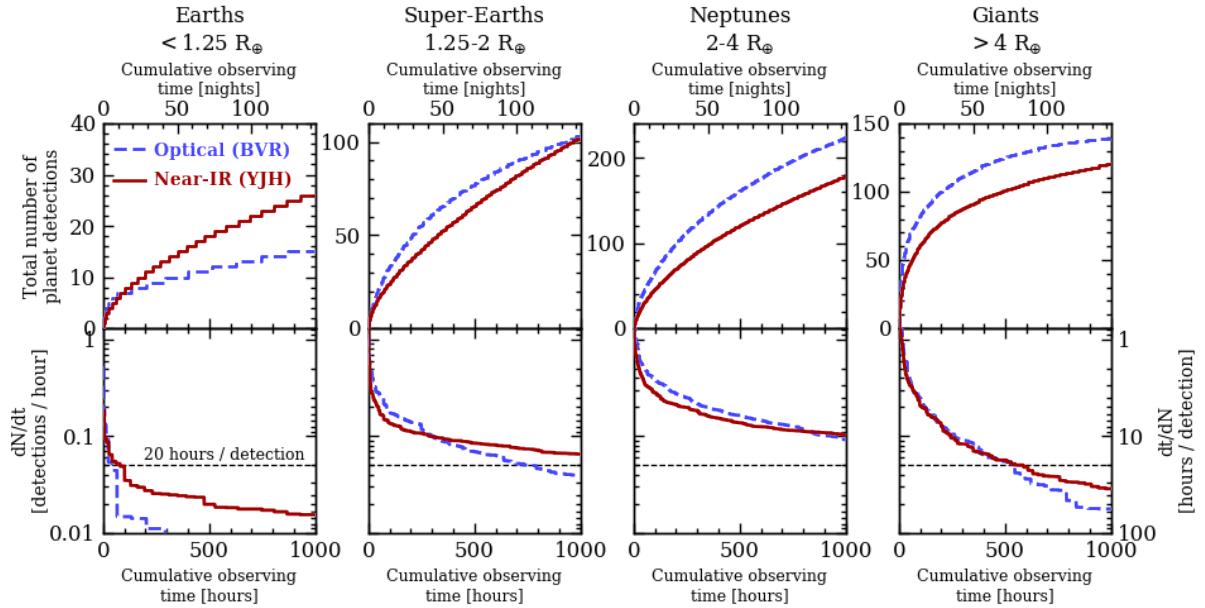


Figure 4.5: *Upper panel:* the cumulative median observing time to measure the  $3\sigma$  RV masses of TESS planets as a function of planet type up to  $10^3$  hours. Planet type definitions are annotated above each column. The [dashed blue curves] represent the results from the optical spectrograph whereas the [solid red curves] represent the near-IR spectrograph. *Lower panel:* the time derivative of the cumulative observing time curves used to indicate the RV planet detection efficiency. The horizontal dashed line highlights the value of the detection efficiency at 20 hours per detection. Note that unlike the lower panels, the upper panels do not share a common ordinate due to the differing number of planet detections around stars in each spectral type bin.

optical detection efficiency better than 20 hours per detection up to  $\sim 100$  nights. We expect to yield  $\sim 100/509$  super-Earths with either the optical or near-IR spectrographs after  $10^3$  observing hours. Neptunes are the most efficiently detected class of planet and with similar detection efficiencies between the two spectrographs.  $\sim 220/1258$  Neptune detections are expected in a thousand hours in the optical compared to  $\sim 180/1258$  in the near-IR. For Neptunes the detection efficiency remains  $< 20$  hours per detection for up to after  $\sim 510$  nights or  $\sim 370$  detections in the optical and after  $\sim 960$  nights or  $\sim 660$  detections in the near-IR. Lastly,  $\sim 130/151$  giant planets are detected in the optical in  $\sim 80$  nights at which point the optical detection efficiency begins to exceed 20 hours per detection. Similarly in the near-IR,  $\sim 110/151$  giants planets are detected in  $\sim 85$  nights.

#### 4.5.2 Science Case 1: mass characterization of 50 TESS planets with $r_p < 4 R_\oplus$

The TESS level one science requirement is to measure the masses of 50 small transiting planets with  $r_p < 4$  Earth radii.<sup>4</sup> To date, the vast majority of masses for planets with  $r_p < 4 R_\oplus$  have been obtained with HARPS, HARPS-N, and HIRES. The onset of many up-coming precision velocimeters will provide many more instruments capable of characterizing such planets. Similarly to Sect. 4.5.1, we define a planet mass measurement requirement for the completion of the TESS level one science requirement of  $5\sigma$ . Given our previous assumptions regarding the measurement precision on  $P$  and  $M_s$  (see Sect. 4.5.1), a  $5\sigma$  mass detection requires a  $5.29\sigma K$  detection (i.e.  $\sigma_K = 0.189K$ ). According to our analytic model

<sup>4</sup><https://tess.mit.edu/followup/>

for  $N_{\text{RV}}$  in the white noise limit, a  $5\sigma$  mass detection requires  $(0.327/0.189)^2 = 2.99$  more observing time than a  $3\sigma$  mass detection.

The cumulative median observing time required to complete the TESS level one science requirement is shown in Fig. 4.6. Here we calculate the cumulative median observing times from various planet samples: i) the 50 small TESS planets sorted in ascending order by total observing times (i.e. the most efficient characterization of 50 small planet masses possible) and ii) for random subsets of the small TESS planets. The latter cases correspond to attempting to conduct RV follow-up observations of *any* subset of small TESS planet up to 50 such planets. In total there are 1833 TESS planets with  $r_p < 4 R_\oplus$  which causes the cumulative observing time to vary drastically depending on whether the input planet set is sorted or random. This is evidenced in Fig. 4.6 wherein it is clear that selecting an optimized set—in terms of shortest median observing times—of 50 small TESS planets is by far the most efficient means of characterizing their masses. Optimized target selection results in the rapid completion of the TESS level one science requirement in only  $\sim 60$  nights with either spectrograph. The performance of the optical and near-IR spectrographs in completing the TESS level one science requirement are seen to be comparable when the ‘best’ TOIs are targeted with 31/50 being most efficiently characterized in the optical and with the remaining 19/50 being done in the near-IR. Selecting the ‘best’ 50 small planets naturally biases the sample towards larger planets with their larger  $K$  values, thus making their mass characterization faster with RVs. Despite this, the set of the ‘best’ small TESS planets with either spectrograph contains 9 super-Earths and 41 Neptunes with the average radius being  $2.8 R_\oplus$  and the smallest planet being likely terrestrial at  $1.37 R_\oplus$  (c.f. inset of Fig. 4.6).

We note that identifying the ‘best’ small TESS planets cannot be done exactly until the conclusion of the full TESS planet search. By not focusing on the ‘best’ 50 small TESS planets and instead opting to obtain RV measurements of any small planet, the total observing time required to complete the TESS level one science requirement will be longer by more than an order of magnitude on average with either spectrograph. Although the exact time allotment will depend on the exact planet sample. However the ‘best’ curves for each spectrograph in Fig. 4.6 are nearly indistinguishable indicating that together optical and near-IR spectrographs will readily complete the TESS level one science requirement and possibly within weeks of relevant TOIs being announced due to the low number of required RV measurements, typically  $N_{\text{RV}} = 32$ . Given the efficiency of measuring the ‘best’ small TESS planets at  $5\sigma$  ( $\sim 5$  hours per detection), the community may opt to focus on a larger subsample of Earth-like planets or to even increase the required mass detection significance to  $> 5\sigma$  thus enhancing the TESS return of planets smaller than  $4 R_\oplus$ .

### A-priori estimate of the ‘best’ targets

Recall that identifying the optimum targets to achieve the TESS level one science requirement in the most efficient manner requires observers either to wait until the conclusion of the 2-year long TESS planet search or to select targets based on a-priori knowledge of the population of the ‘best’ small planets. The latter scenario is favorable as it allows targets to be observed with RVs almost concurrently with reported TESS detections thus leading to the shortest completion time of the TESS level one science requirement; i.e.  $\sim 60$  nights or  $\sim 400$  hours. Based on the predicted TESS planet population (S15) and the results of our study, we can predict the properties of the ‘best’ 50 small TESS planets

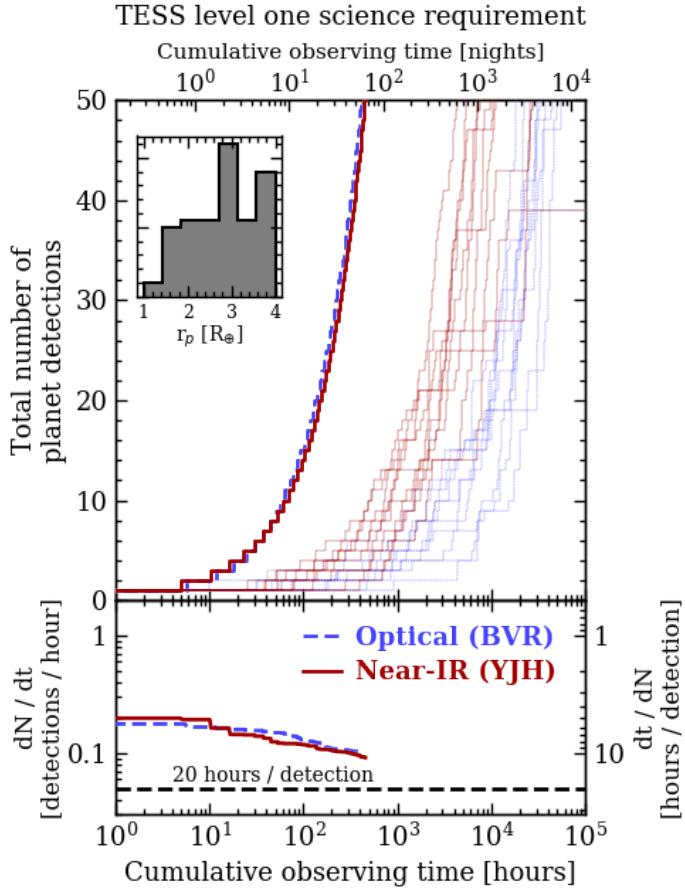


Figure 4.6: *Top panel:* the cumulative median observing time required to achieve the TESS level one science requirement of measuring the masses of 50 planets with  $r_p < 4 R_\oplus$  at  $5\sigma$  with either the optical spectrograph (dashed blue curves) or our near-IR spectrograph (solid red curves). The set of thin curves are calculated from randomly ordered TOI samples whereas the thick curves are calculated from the sorted TOIs thus resulting in the most observationally efficient planet detections. The latter curves for each spectrograph lie almost exactly on top of each other. *Inset:* a histogram showing the joint planet radii distribution of the ‘best’ 50 TESS planets for each spectrograph. *Lower panel:* the time derivative of the thick curves shown in the upper panel. The value of the detection efficiency equal to 20 hours per detection is highlighted by the horizontal dashed line.

and search for trends that will inform their selection throughout the actual TESS planet search. Here we suggest that identification of these planets can be done in an approximate way given some combination of intuitively crucial transit observables: the stellar magnitude,  $r_p$ , and  $P$ . The stellar magnitude constrains the photon-noise limited RV measurement precision while the latter two quantities have a direct effect on  $K$  assuming that the mass-radius relation has a positive, non-zero slope everywhere for planets smaller than  $4 R_\oplus$ . After considering numerous combinations of these parameters we find that the stellar magnitude and the value of the derived-from-transit quantity  $\Omega \equiv r_p^\alpha / P^{1/3}$ —for some value of  $\alpha$ —are good diagnostics for the total observing time required to detect a transiting planet’s mass. The definition of  $\Omega$  was selected to resemble the expected RV semi-amplitude  $K$  assuming a positive scaling between  $r_p$  and  $m_p$  (i.e.  $\alpha > 0$ ) and noting that  $K \propto P^{-1/3}$ . In this way, large values of  $\Omega$  should correspond to large  $K$  values which directly effects the observing time required to achieve a given mass detection significance as larger signals are more easily detected with a given Rv measurement precision. We considered various values of  $\alpha \in (0, 3]$  and found little discrepancy between these values with regards to where in the region of the corresponding magnitude- $\Omega$  parameter space the ‘best’ 50 small planets sit. Given uncertainties and possible discontinuities in the mass-radius relation for small planets we opt for  $\alpha = 1$ .

In Fig. 4.7 we compare the location of the 50 ‘best’ small planets to the remaining 1833 small planets in the apparent magnitude- $\Omega$  parameter space. For considerations with our optical and near-IR spectrographs we use the  $V$  and  $J$  band magnitudes respectively. We note that in what follows we are marginalizing over stellar rotation and the level of stellar activity, both of which have a direct effect on our ability to detect planets in RV.

Unsurprisingly, the 50 ‘best’ small TESS planets are localized around bright TOIs and exhibit an increasing value of  $\Omega$  with stellar magnitude. That is, as the TOIs become dimmer, a larger  $\Omega$  is required for a rapid RV mass detection. To encapsulate the region of the parameter space with the highest likelihood of yielding the most efficient RV planet detections (i.e. the shortest total observing times), we truncate the outer edge of the region at a ‘maximum’ apparent magnitude and derive a lower boundary by fitting a linear function to  $\Omega$  as a function of magnitude for the 50 ‘best’ small TESS planets before translating the lower boundary downwards to encapsulate 96% (i.e. 48 out of 50) of the ‘best’ small TESS planets. The resulting sets of the ‘best’ TOIs are

$$\{\text{TOIs} | V < 10.7, \Omega > 0.09V + 0.28\} \quad (4.23)$$

for follow-up with our fiducial optical spectrograph and

$$\{\text{TOIs} | J < 11.7, \Omega > 0.14J - 0.35\} \quad (4.24)$$

for follow-up with our fiducial near-IR spectrograph. These sets approximately represent the TOIs with the shortest total observing times and should be seriously considered for rapid RV follow-up observations if they are known to not orbit a rapid rotator or an overly active star. Recall that stellar rotation and activity have been marginalized over in the derivation of Eqs. 4.23 and 4.24.

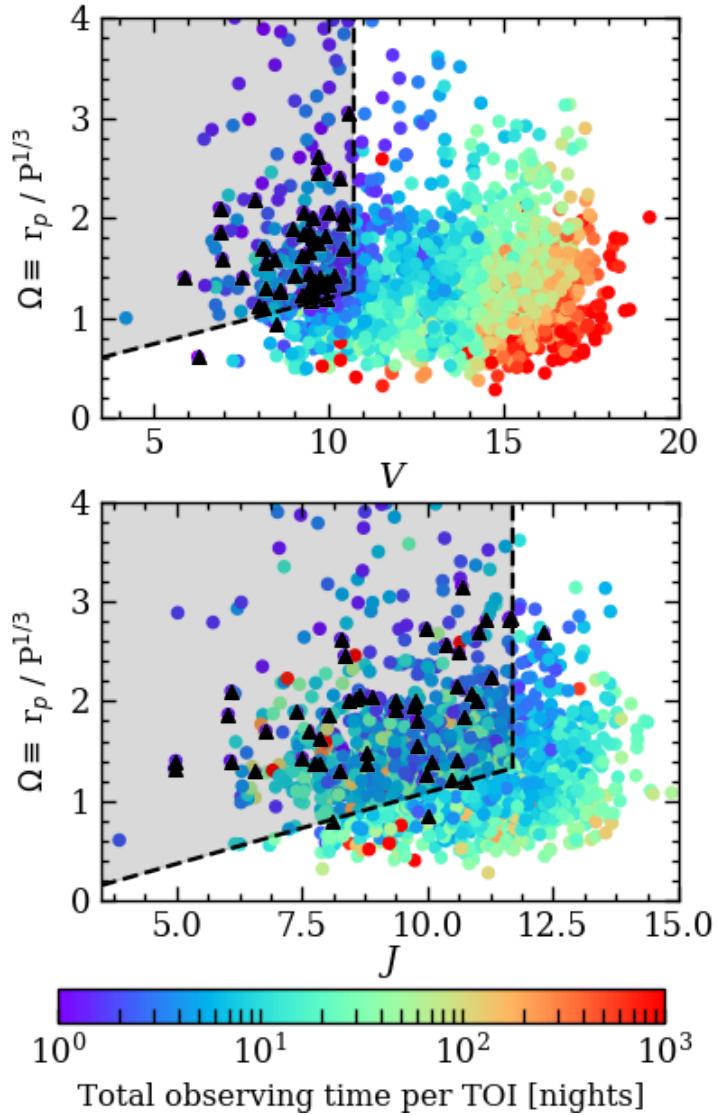


Figure 4.7: The total observing time per TOI required to detect the planet’s mass at  $5\sigma$  with the optical spectrograph (*upper panel*) and the near-IR spectrograph (*lower panel*) as a function of the stellar apparent magnitude and the derived quantity from transit observables:  $\Omega \equiv r_p/P^{1/3}$ . The relevant apparent magnitudes for the optical and near-IR spectrographs are  $V$  and  $J$  respectively. The 50 ‘best’ TOIs with each spectrograph are designated by **black diamonds**. The **shaded regions** bounded by the **black dashed lines** approximate the regions of each parameter space with the highest likelihood of hosting planets amenable to the most efficient RV mass detections. TOIs within the shaded regions and known not to orbit an active host star should be strongly considered for rapid RV follow-up campaigns.

### 4.5.3 Science Case 2: informing the mass-radius relation of planets across the radius valley

Accurate characterization of the empirical mass-radius relation for exoplanets (e.g. Weiss et al., 2013; Rogers, 2015; Wolfgang et al., 2016) is an important step towards understanding the diversity of exoplanet compositions as well as its use as a tool in predictive studies (e.g. S15; Cloutier et al. 2017a, Cloutier et al. 2018b). For example, consideration of small exoplanets ( $r_p \leq 4 R_{\oplus}$ ) with masses measured to better than 20% revealed that a large fraction of planets with  $r_p \lesssim 1.6 R_{\oplus}$  are rocky with bulk compositions consistent with that of the Earth and Venus (Dressing et al., 2015). The transition from bulk rocky compositions to less dense planets with a significant size fraction of volatile-rich envelope gas has also been shown to occur between  $\sim 1.5 - 2.5 R_{\oplus}$  where a paucity of planets exists (Fulton et al., 2017; Van Eylen et al., 2018). The precise characterization of the mass-radius relation in the vicinity of this so-called *radius valley* will elucidate as to whether or not the valley persists in terms of planet bulk densities as the peaks in the bi-modal radius distribution are posited to harbour terrestrial and volatile-rich planets on opposing sides of the radius valley. Characterizing the mass-radius relation in this regime will greatly benefit from the inclusion of relevant TESS planets.

In order to accurately inform the mass-radius relation of planets across the radius valley with TESS planets, we seek a 20% fractional mass uncertainty (i.e.  $5\sigma$  mass detection) following Dressing et al. (2015). We define TESS planets of interest as those spanning the radius valley using the period-dependent locus of planet radii—and its upper and lower bounds—as defined by the powerlaw in Van Eylen et al. (2018) from asteroseismology.

The cumulative median observing time required to detect relevant TESS planets at  $5\sigma$  are shown in Fig. 4.8. To avoid the bias that the most efficiently observed targets have towards larger planets and correspondingly larger  $K$  on average, the ‘best’ planets in this science case are selected equally from two bins on either side of the radius valley with  $r_p \leq 2$  and  $> 2 R_{\oplus}$ . There are 542 TESS planets that span the radius valley. RV mass characterization of all such planets will require  $\gtrsim 5 \times 10^4$  and  $\gtrsim 7000$  observing nights in the optical and near-IR respectively. Evidently, the cumulative observing time for all TESS planet across the radius valley is likely too large to complete even with all available spectrographs. Fortunately, not all 542 planets are required to be measured in order to resolve the radius valley in planet bulk density. If instead we focus on the ‘best’ TESS planets then our detection efficiency with either spectrograph remains less than 20 hours per detection up to  $\sim 80 - 130$  nights. In that time we expect to detect  $\sim 55$  planets that span the radius valley with either spectrograph, if those planets are optimally chosen (c.f. Fig. 4.7). This implies that optical and near-IR spectrographs are equally well-suited to characterizing the ‘best’ TESS planets across the radius valley with near-IR observations only becoming more efficient after  $\sim 60$  planet detections. With this sample of TESS planets, the hypothesized rocky/volatile-rich transition can be resolved and will help in progressing towards potentially resolving the radius/bulk density valley as a function of host spectral type.

### 4.5.4 Science Case 3: characterization of temperate Earths & super-Earths

Temperate planets that orbit close to or within their host star’s habitable zone are of particular interest for the search for life. Here we aim to scrutinize the masses of potentially habitable planets that we

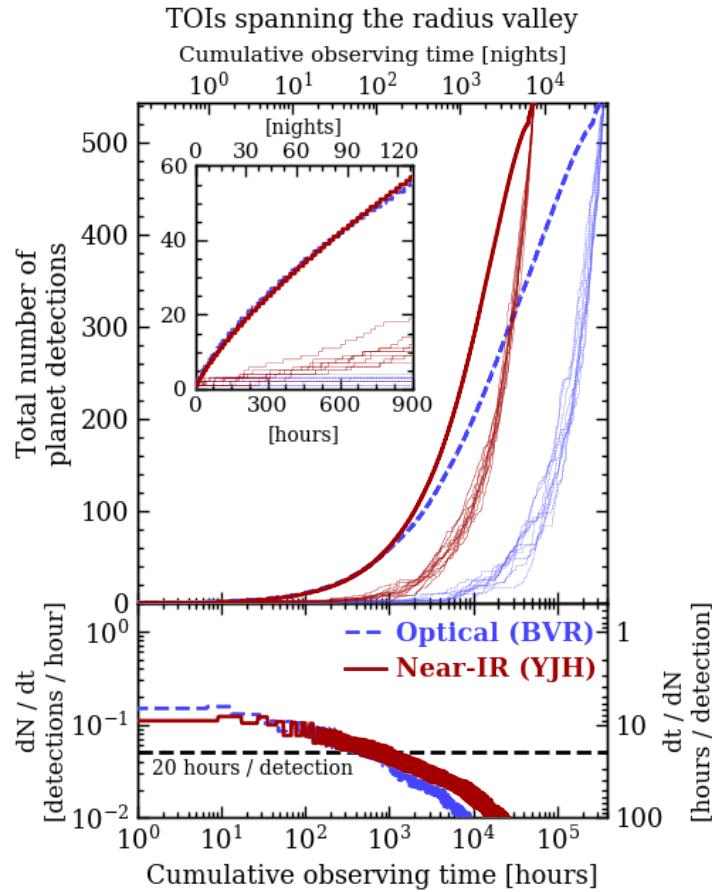


Figure 4.8: *Top panel:* the cumulative median observing time required to measure the  $5\sigma$  RV masses of TESS planets spanning the radius valley ( $1.5 \lesssim r_p/R_\oplus \lesssim 2.6$ ) with either the optical spectrograph (dashed blue curves) or our near-IR spectrograph (solid red curves). The set of thin curves are calculated from randomly ordered TOI samples whereas the thick curves are calculated from the sorted TOIs—with an equal number of planets less than and greater than  $2 R_{\text{oplus}}$ —thus resulting in the most observationally efficient planet detections. *Inset:* focusing on the region up to 900 cumulative observing hours (i.e.  $\sim 130$  nights). *Lower panel:* the time derivative of the thick curves shown in the upper panel. The value of the detection efficiency equal to 20 hours per detection is highlighted by the horizontal dashed line.

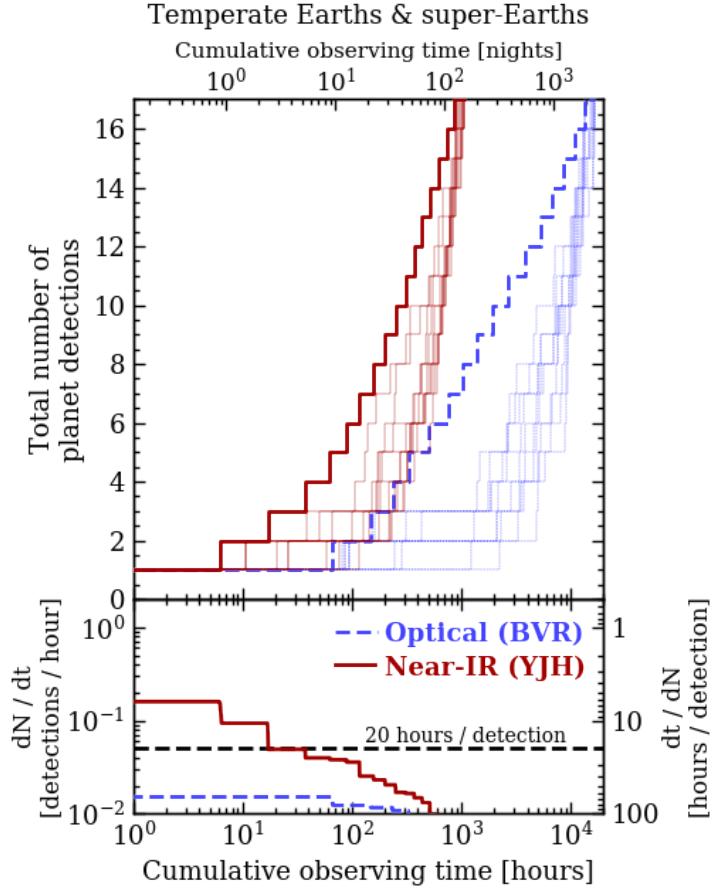


Figure 4.9: *Top panel:* the cumulative median observing time required to measure the RV masses of potentially habitable TESS planets at  $3\sigma$  with either the optical spectrograph (dashed blue curves) or our near-IR spectrograph (solid red curves). Potentially habitable planets are defined as either Earths or super-Earths with  $r_p \leq 2R_\oplus$  and orbit within their host star's habitable zone as defined by (Kopparapu et al., 2013). The set of thin curves are calculated from randomly ordered TOI samples whereas the thick curves are calculated from the sorted TOIs thus resulting in the most observationally efficient planet detections. *Lower panel:* the time derivative of the thick curves shown in the upper panel. The value of the detection efficiency equal to 20 hours per detection is highlighted by the horizontal dashed line.

define herein as Earths and super-Earths ( $r_p \leq 2 R_\oplus$ ) orbiting within the habitable zone (HZ). To define the HZ we adopt the ‘water-loss’ and ‘maximum-greenhouse’ HZ limits from Kopparapu et al. (2013). There are just 17 such planets in the S15 synthetic catalog, three of which are smaller than  $1.5 R_\oplus$  and are likely to be rocky. Fig. 4.9 depicts the cumulative median observing time required to detect the masses of potentially habitable TESS planets at  $3\sigma$ . Due to TESS’s limited observational baselines, the majority of HZ TESS planets—including the full set of 17 potentially habitable TESS planets—orbit M dwarfs and thus favor near-IR RV follow-up. The cumulative observing time required to measure all potentially habitable TESS planet masses at  $3\sigma$  is  $\sim 150$  nights with the near-IR spectrograph. This is about fifteen times shorter than the total time required to complete the same task with the optical spectrograph. Regardless of the spectrograph used, RV follow-up of potentially habitable TESS planets will be an expensive task with just  $\sim 2$  planets detected in  $\sim 2.5$  nights after which the detection efficiency begins to require more than 20 hours per detection.

We caution that with such a small sample of potentially habitable TESS planets that the numbers presented here regarding the cumulative observing time required to detect such planets may be misleading. Due to the small number statistics the results for temperate Earths and super-Earths are highly sensitive to the true properties of those planetary systems. For example, a potentially habitable TESS planet may be detected around an M dwarf with  $J < 10.15$ —the brightest TOI with a potentially habitable planet from S15—thus resulting in a shorter total observing time required to detect one such planet with RVs.

#### 4.5.5 Science Case 4: characterization of favorable JWST follow-up targets

The *James Webb Space Telescope* (JWST) to-be launched in May 2020 will revolutionize our understanding of transiting exoplanet atmospheres (see Beichman et al. 2014 for a summary of science cases). Many TOIs will represent some of the most interesting targets for atmospheric characterization with JWST through transmission spectroscopy observations in particular. To quantify the RV requirement needed to understand the bulk densities of these planets we will consider TOIs that are most amenable to efficient JWST observations. Specifically, TOIs with their expected S/N of transmission features  $\geq 10$ .

For each TOI we calculate the expected S/N in transmission from the expected differential transmission depth  $\Delta D$  of the planet and the photon-noise per spectral bin  $\sigma_{\text{ppm}}$ ;  $S/N = \Delta D / \sigma_{\text{ppm}}$ . The value of  $\sigma_{\text{ppm}}$  is measured in the  $J$ -band with spectral resolution  $R = 50$  (i.e.  $\delta\lambda = 25$  nm), an instrumental throughput of 50%, and an integration time equal to the planet’s full transit duration. Values of  $\Delta D$  for each TESS planet are computed up to five scale heights in a cloud-free atmosphere using the standard equation

$$\Delta D = 15 \text{ ppm} \left( \frac{T_p}{250 \text{ K}} \right) \left( \frac{\rho}{5.55 \text{ g/cm}^3} \right)^{-1} \left( \frac{\mu}{29 \text{ u}} \right)^{-1} \left( \frac{R_s}{0.25 R_\odot} \right)^{-2}, \quad (4.25)$$

where  $T_p$  is the planet’s isothermal atmospheric temperature (calculated assuming uniform heat redistribution over the planetary surface and zero albedo),  $\rho$  is the planet’s bulk density,  $\mu$  is the mean molecular weight of the atmosphere, and  $R_s$  is the stellar radius. The atmospheric mean molecular weights of the TESS planets are not given in S15 so we adopt a very simplistic prescription of  $\mu$  using a step-wise function of H/He-dominated atmospheres ( $\mu = 2$ ) to Earth-like atmospheres ( $\mu = 29$ ) for planets  $\leq 2 R_\oplus$ . However, this simple prescription is known to be inaccurate but an approximation is necessary to facilitate the exercise of estimating  $\Delta D$  for planets yet to be studied in transmission.

According to Eq. 4.25 the interpretation of planetary transmission spectra heavily relies on a-priori knowledge of the planet’s bulk density. The  $\rho$  measurement precision is derived from the measurement precision on both the planet’s radius  $\sigma_{r_p}$ —from its TESS light curve—and on its mass measured from RVs. Due to the cubic dependence of  $\rho$  on  $r_p$  compared to its linear dependence on  $m_p$ , improving a planet’s bulk density measurement precision is most effectively done by reducing  $\sigma_{r_p}$  either through more complete transit data or more precise characterization of the host stellar radius. Because of this, it

is not worthwhile to sit on any TESS planet with RVs to achieve the typical mass detection significance required to precisely measure the planet’s bulk density. For example, given the photometric precision for each TOI from S15 and a notional stellar radius uncertainty of 10% (Carter et al., 2008), achieving a  $3\sigma$  bulk density detection would require a  $\sim 6.9\sigma$  mass detection on average. Such a precise RV mass measurement would require a  $\sim 7.8\sigma$   $K$  measurement or  $(0.327/0.129)^2 = 6.4$  more observing time than a  $3\sigma$  mass detection. We therefore opt for a more reasonable mass detection of  $5\sigma$ , similarly to what was pursued when characterizing 50 TESS planets smaller than  $4 R_{\oplus}$  and planets across the radius gap in Sects. 4.5.2 and 4.5.3.

Fig. 4.10 depicts the cumulative median observing time required to detect the  $5\sigma$  masses of TESS planets that are favorable for JWST follow-up with an expected transmission  $S/N \geq 10$ . There are 1169 such TESS planets. By our simple prescription for  $\Delta D$  for planets smaller than  $2 R_{\oplus}$  and by imposing a minimum expected transmission  $S/N \leq 10$ , the sample of 1169 favourable JWST targets has been restricted to planets larger than  $2 R_{\oplus}$ —with a median value of  $r_p = 2.9 R_{\oplus}$ —due to their systematically larger scale heights compared to Earths and super-Earths. Detecting all 1018 Neptunes and all 151 giant planets would require  $\sim 10^8$  nights with the near-IR spectrograph which is  $\sim 40$  times shorter than the total time required using the optical spectrograph. If follow-up observations are focused on the ‘best’ subset of favorable TESS planets for JWST follow-up, then the detection efficiency of these planets remains less than 20 hours per detection for up to  $\sim 400$  nights wherein  $\sim 220$  planets are measured with the near-IR spectrograph. Conversely, the detection efficiency in the optical drops to 20 hours per detection slightly sooner—after  $\sim 360$  nights but with a similar number of planet detections. Yet again spectrographs in the optical and near-IR demonstrate a comparable performance when characterizing the ‘best’ planets favorable for JWST follow-up with some slight improvement in the optical before its detection efficiency begins to drop off after  $\sim 360$  nights. It is clear that many interesting TESS planets will be readily characterized with RVs thus providing a large sample of TESS planets with precisely characterized masses and bulk densities prior to the launch of JWST.

## 4.6 Discussion and Conclusions

We have presented calculations of the observing time required to measure the masses of the expected TESS planet population using ground-based precision radial velocities. Our calculations are based on analytical estimates (see Sect. 4.2) of the number of RV measurements required to detect a transiting planet’s RV semi-amplitude  $K$  at a given precision. When coupled to an exposure time calculator this yields the total observing time per target. Our main conclusions are summarized below.

1. The number of RV measurements required to detect a transiting planet’s mass is dependent on the desired  $K$  measurement precision, the rms of the RVs observations (this includes contributions from photon-noise, stellar activity, additional unseen planets, and systematic effects), and whether or not the residual RV noise is correlated or not. Eq. 4.18 can be used to calculate the number of required RV measurements if the RV residuals are uncorrelated, otherwise the formalism presented in Sect. 4.2.1 must be used.
2. Efficient characterization of transiting planet masses for a given planet type (e.g. super-Earths) favors targets with small photon-noise limited RV measurement precisions. High precision mea-

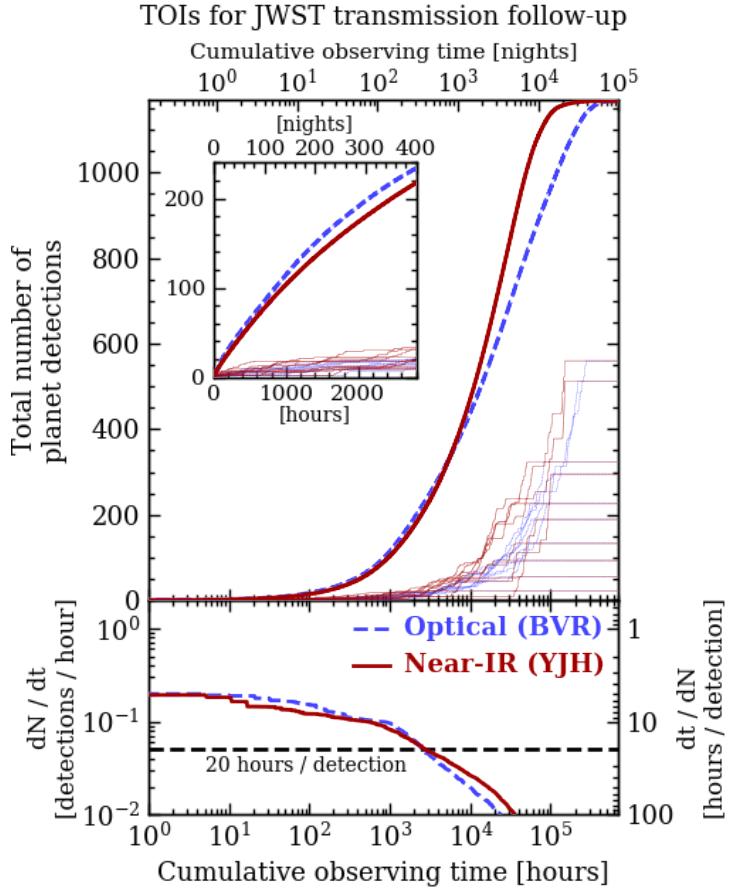


Figure 4.10: *Top panel:* the cumulative median observing time required to measure the RV masses of TESS planets favorable for follow-up transmission spectroscopy observations with JWST (i.e. transmission S/N  $\geq 10$ ) at  $5\sigma$  with either the optical spectrograph (dashed blue curves) or our near-IR spectrograph (solid red curves). The set of thin curves are calculated from randomly ordered TOI samples whereas the thick curves are calculated from the sorted TOIs thus resulting in the most observationally efficient planet detections. *Inset:* focusing on the region up to  $\sim 2800$  cumulative observing hours (i.e.  $\sim 400$  nights). *Lower panel:* the time derivative of the thick curves shown in the upper panel. The value of the detection efficiency equal to 20 hours per detection is highlighted by the horizontal dashed line.

surements are most readily achieved with optical spectral coverage (i.e. BVR bands in this study) for Sun-like stars with  $T_{\text{eff}} \gtrsim 5500$  K whereas M dwarfs with  $T_{\text{eff}} \lesssim 3800$  K are favored by near-IR spectrographs (i.e. YJH bands in this study).

3. Overall, the relative merits of obtaining precise RVs in the optical compared to in the near-IR are nearly equivalent. That is that RV campaigns aiming to characterize TESS planets can largely be done as effectively in either wavelength domain with the exception of the characterization of Earths ( $r_p < 1.25 R_{\oplus}$ ) and temperate TESS planets that are preferentially found around M dwarfs.
4. Not all TESS planets will be amenable to RV follow-up observations and selecting random TOIs for follow-up is an incredibly inefficient method of target selection. Instead, targets should be selected based on their apparent magnitude—to minimize the photon-noise limited RV precision—and the transit-derived quantity  $\Omega = r_p/P^{1/3}$ . The subset of TOIs belonging to either optimal set defined by Eqs. 4.23 and 4.24 should be strongly considered for immediate RV follow-up.
5. The TESS level one science requirement of measuring the masses of 50 planets with  $r_P < 4 R_{\oplus}$  at the 20% level (i.e.  $5\sigma$  mass detection) can be achieved in as little as  $\sim 400$  hours or  $\sim 60$  nights of observation.
6.  $\sim 55$  TESS planets spanning the radius gap for small planets (i.e.  $1.5 \lesssim r_P/R_{\oplus} \lesssim 2.6$ ) can be detected efficiently at  $5\sigma$  in  $\sim 130$  nights before the detection efficiency drops below 0.05 detections per hour (i.e. 20 hours per detection).
7. Only  $\sim 2$  temperate super-Earths can be detected efficiently at  $3\sigma$  in  $\sim 2.5$  nights before the detection efficiency exceeds 20 hours per detection.
8.  $\sim 220$  Neptunes and giant planets amenable to transmission spectroscopy follow-up observations with JWST can be detected efficiently with a  $5\sigma$  mass detection in  $\sim 360 - 400$  nights before the detection efficiency exceeds 20 hours per detection.
9. An online version of the **Radial Velocity Follow-up Calculator** used throughout this paper on the expected TESS planet population is available at <http://maestria.astro.umontreal.ca/rvfc>. This general-usage tool can be used to calculate the number of RV measurements and total observing time required to detect the RV semi-amplitude of any transiting planet to a user-defined detection significance and with a user-defined spectrograph.

The results of this paper have been based on the synthetic planet population presented in S15. In their study they reported the expected TESS planet population (i.e.  $\sim 1700$  planets) recovered from the TESS 2 minute cadence observing mode of the brightest targets. Additionally, deeper full frame images at a 30 minute cadence will be released and result in even more planet candidates, some of which may still be amenable to RV follow-up observations. Another population of TOIs not considered in our calculations is the population of targets featuring only one or two transit-like events. If confirmed, the corresponding planets will be interesting in their own right as their orbits will have systematically longer periods making them cooler and of interest for future habitability studies. Lastly, the TESS mission has the possibility of being extended beyond its nominal 2-year long primary mission. If extended, the extension of the TESS observational baselines will improve the measured ephemerides of confirmed TESS planets, shed light on the nature of systems exhibiting single transit-like events, and expand the

population of planets discovered with TESS (see the overview in [Bouma et al., 2017](#)).

Further caveats to the planet population from [S15](#) used throughout this study were addressed by a variety of studies and corresponding updates to the expected TESS planet population (e.g. [Ballard, 2018](#); [Barclay et al., 2018](#)). For example, [Barclay et al. \(2018\)](#) updated the calculations of [S15](#) by using the TESS Candidate Target List ([Stassun et al., 2017](#)) and consequently updating the number of TESS discoveries including a decreased number of Earths and super-Earths which will have important implications for the corresponding science cases such as the bulk density characterization of TESS planets near the radius valley ( $r_p \sim 2 R_{\oplus}$ ). Furthermore, [Ballard \(2018\)](#) predicted that the planet population around M1-M4 dwarfs was underestimated. The results imply a larger cumulative planet yield around M dwarfs including many systems with multiple transiting planets. This has important implications for RV follow-up campaigns of TOIs as M dwarf planets have the potential to be the most efficiently detected planets in radial velocity. In practice this will depend on the on-sky performance of many of the up-coming generation of near-IR velocimeters which have favorable RV measurement precision when observing M dwarfs. Many up-coming near-IR spectrographs are anticipated to operate with an RV noise floor of  $\sim 1 \text{ m s}^{-1}$  comparable to many high performance optical spectrographs but this may prove challenging as demonstrated by the CARMENES near-IR channel that seems to yield lower precision ( $\sim 2 - 3 \text{ m s}^{-1}$ ) on early to mid-M dwarfs compared to the sub-1  $\text{m s}^{-1}$  performance achieved in the optical ([Reiners et al., 2018](#), c.f. Fig 6). The focus on M dwarf planets with either class of spectrograph is well warranted as these planets represent some of the most interesting planets in terms of potential habitability and for the prospect transmission spectroscopy follow-up with JWST.

In Sect. 4.5.5 we presented the results for the most favorable TESS targets for transmission spectroscopy follow-up with JWST. In addition to identifying the ‘best’ such planets based on the time required to measure their masses (i.e. bulk densities) with RVs, one may also consider metrics describing the ease of detecting atmospheric features in either transmission or thermal emission (e.g. [Kempton et al., 2018](#); [Morgan et al., 2018](#)). Favorable TESS planets based on our calculations of the observing time required to detect their masses in RVs, that also overlap with favorable TESS planets based on the metrics from [Kempton et al. \(2018\)](#), should be strongly considered for rapid RV follow-up. Similar target selection may be done based on the simulated transmission spectra from [Louie et al. \(2018\)](#) and the resulting S/N in transmission.

It is worth reiterating that although TOIs will be frequently reported following the launch and commissioning of TESS, RV follow-up teams should refrain from targeting just any TOI. Many detected planets will orbit stars either too dim or too active for efficient RV characterization. We emphasize that TOIs amenable to RV follow-up can be approximately identified—in real time—if they belong to one of the sets defined in Eqs. 4.23 or 4.24, or equivalently, if they lie in one of the shaded regions of either panel in Fig. 4.7. However the sets of the ‘best’ TOIs are derived by marginalizing over the population of additional unseen planets in the system and the star’s intrinsic RV activity. The former source of RV signals will not be illuminated unless RV follow-up of the system commences but the level of stellar activity can be estimated from the star’s photometric variability in its TESS light curve. In addition to selecting targets based on Eqs. 4.23 and 4.24, stars with high amplitudes of photometric variability or jitter should not be considered for efficient RV follow-up campaigns.

## 4.7 Appendix

### 4.7.1 Radial Velocity Follow-up Calculator

Our models for the number of RV measurements  $N_{\text{RV}}$  required to detect a transiting planet at a given significance—in the presence of white or correlated RV noise—are generalizable to the majority transiting planets observed with any velocimeter. Recall that our models are only applicable to planets on nearly circular orbits, with known ephemerides, and whose orbital phase curves will be (approximately) uniformly sampled in the white noise case. Furthermore, because the results presented throughout this paper have been regarding a hypothetical planet population and with only two fiducial spectrographs, we present to the community an online web-tool version of the generalized *Radial Velocity Follow-up Calculator*<sup>5</sup> (RVFC) used throughout this study.

The RVFC is intended to serve the community by providing rapid calculations of  $N_{\text{RV}}$  and total observing times for an arbitrary user-defined transiting planet with a user-defined spectrograph, given parameters of the planet from its transit light curve, stellar parameters, parameters of the employed spectrograph, RV noise parameters, and a small number of additional simulation parameters. The exact input parameters required by the calculator will depend on which of two possible primary modes-of-operation the user selects. In *option 1*, the calculator is used to calculate the photon-noise limited RV measurement precision  $\sigma_{\text{RV}}$  using the formalism discussed in Sect. 4.2.2 and PHOENIX stellar models. Two suboptions are available for users to either add additional RV noise sources which are sampled from appropriate empirical distributions (*option 1.1*) or for the user to specify verbatim those additional noise sources (*option 1.2*). In *option 2*, the user can input a fixed value of  $\sigma_{\text{RV}}$  thus negating the need for certain input parameters to be specified by the user and speeding up the wall time of the calculation. *Option 2* also features the two suboptions available for *option 1* and additionally has a third option in which the effective RV rms (i.e. the combination of all RV noise sources) is set verbatim if its value is known for the system of interest and only white noise calculations are desired (*option 2.3*). *Option 2* may be viable for users whose employed spectrograph features an independent ETC, the results from which differ from those returned by the built-in RVFC ETC. The input parameters required to run the RVFC are summarized in Table 4.5.

---

<sup>5</sup><http://maestria.astro.umontreal.ca/rvfc>

Table 4.5: Descriptions of RVFC Input Parameters

Parameter	Units	Required for RVFC Option(s)	Usage	Notes
<i>Spectrograph parameters</i>				
Minimum spectrograph wavelength	nm	1	Interpolating $\sigma_{\text{RV}}$ .	
Maximum spectrograph wavelength	nm	1	Interpolating $\sigma_{\text{RV}}$ .	
Spectral resolution, $R$	$\lambda/\Delta\lambda$	1	Interpolating & calculating $\sigma_{\text{RV}}$ .	
Telescope aperture	m	1	Calculating $\sigma_{\text{RV}}$ .	
Throughput	[0-1]	1	Calculating $\sigma_{\text{RV}}$ .	
RV noise floor, $\sigma_{\text{floor}}$	$\text{m s}^{-1}$	1	Contributes to RV rms.	
<i>Planet parameters</i>				
Orbital period, $P$	days	1,2	Calculating the expected $K$ and sampling unseen planets.	
Planetary radius, $r_p$	$R_\oplus$	1,2	Estimating $m_p$ from Eq. 4.22 and sampling unseen planets.	
Planetary mass, $m_p$	$M_\oplus$	1,2	Calculating the expected $K$ .	If unspecified, $m_p$ is estimated from $r_p$ and the mass-radius relation (Eq. 4.22).
<i>Stellar parameters</i>				
Apparent magnitude, $V$ or $J$	-	1	Calculating $\sigma_{\text{RV}}$ .	The spectral coverage of the spectrograph must span either $V$ or $J$ .
Stellar mass, $M_s$	$M_\odot$	1,2	Calculating the expected $K$ , $\log g$ for $\sigma_{\text{RV}}$ , sampling stellar activity, and sampling unseen planets.	
Stellar radius, $R_s$	$R_\odot$	1,2	Calculating $\log g$ for $\sigma_{\text{RV}}$ .	
Effective temperature, $T_{\text{eff}}$	K	1,2	Interpolating $\sigma_{\text{RV}}$ , estimating stellar colors, sampling stellar activity, and sampling unseen planets.	
Metallicity, $Z$	[Fe/H]	1,2	Interpolating $\sigma_{\text{RV}}$ .	
Projected rotation velocity, $v \sin i_s$	$\text{km s}^{-1}$	1	Interpolating $\sigma_{\text{RV}}$ .	
Rotation period, $P_{\text{rot}}$	days	1,2	Calculating $\sigma_{\text{RV}}$ and sampling stellar activity	If unspecified, $P_{\text{rot}}$ is estimated from $R_s$ and $v \sin i_s$ assuming $\sin i_s = 1$ .
<i>RV noise sources</i>				
Photon-noise limited RV precision, $\sigma_{\text{RV}}$	$\text{m s}^{-1}$	2	Contributes to RV rms.	
RV activity rms, $\sigma_{\text{act}}$	$\text{m s}^{-1}$	1,2	Contributes to RV rms.	Can be set to zero or is sampled if left unspecified.
RV rms from unseen planets, $\sigma_{\text{planets}}$	$\text{m s}^{-1}$	1,2	Contributes to RV rms.	Can be set to zero or is sampled if left unspecified.
Effective RV rms, $\sigma_{\text{eff}}$	$\text{m s}^{-1}$	2	Calculating $N_{\text{RV}}$ .	If unspecified, $\sigma_{\text{eff}}$ is computed from above contributing noise sources.
<i>Simulation parameters</i>				
Exposure time, $t_{\text{exp}}$	minutes	1,2	Calculating S/N and the total observing time.	
Overhead time	minutes	1,2	Calculating $\sigma_{\text{RV}}$ .	
Desired $K$ detection significance, $K/\sigma_K$	-	1,2	Calculating $N_{\text{RV}}$ .	
Number of GP trials, $N_{\text{GP}}$	-	1,2	Calculating $N_{\text{RV}}$ in the presence of red noise.	If zero, only do the white noise calculation. If $N_{\text{GP}} > 0$ then we recommend setting $N_{\text{GP}} \gtrsim 10$ for decent sampling. Value returned is the median of all trials if $N_{\text{GP}} \geq 2$ .

One notable bottleneck in the wall-time of running the RVFC is the time required to compute the photon-noise limited RV precision given a unique set of stellar and spectrograph parameters. To facilitate *rapid* calculations with the RVFC we opt to interpolate these values from pre-computed tables rather than perform the calculations explicitly. The tables from which  $\sigma_{\text{RV}}$  values are interpolated from are computed individually for each of the spectral bands shown in Table 4.1 and over five additional parameters: the spectral resolution,  $T_{\text{eff}}$ ,  $\log g$ ,  $Z$ , and  $v \sin i_s$ . Given values for these parameters in the RVFC, the corresponding  $\sigma_{\text{RV}}$  is obtained by interpolating over this grid for each spectral band spanned by the spectrograph’s wavelength domain. The remaining spectrograph parameters and stellar magnitude are then used to scale the interpolated value of  $\sigma_{\text{RV}}$  to the correct S/N per resolution element. Notably, the interpolation of  $\sigma_{\text{RV}}$  necessitates a trade-off between accuracy and computing time. However, the loss in accuracy we deem acceptable given the often inexact values of the other sources of RV noise (i.e. instrument stability, activity, and additional unseen planets).

Lastly, recall that the RVFC can be calculate  $N_{\text{RV}}$  in either white or correlated RV noise limits according to our models discussed in Sects. 4.2.1 and 4.2.1 respectively. As noted in Sect. 4.2.1, the results in the latter scenario are dependent on the time-sampling which has been sampled uniformly in this study over a *fixed* baseline. This is also adopted in the initial version of the RVFC. As such, for users interested in calculating  $N_{\text{RV}}$  in the presence of correlated RV noise we recommend using multiple calculations (e.g.  $N_{\text{GP}} \gtrsim 10$ ) to obtain the most-likely value and spread in  $N_{\text{RV}}$  given a suite of sampled window functions. Users beware that increasing  $N_{\text{GP}}$  will require a correspondingly longer computation time. In the future we would like to implement a way for users to upload custom window functions to avoid this ambiguity.

#### 4.7.2 Fisher information with a quasi-periodic Gaussian process regression model

Here we derive the Fisher information matrix terms for a circular Keplerian RV model plus a quasi-periodic GP correlated noise activity model, including an additional scalar jitter parameter. As discussed in Sect. 4.2.1 the Keplerian model parameter is solely the RV semi-amplitude  $\{K\}$  whilst the GP covariance model has five hyperparameters  $\Theta = \{a, \lambda, \Gamma, P_{\text{GP}}, \sigma_{\text{jitter}}\}$  that describe the quasi-periodic covariance matrix  $C$  commonly used when simultaneously fitting RV planets and stellar activity:

$$k_{ij} = a^2 \exp \left[ -\frac{(t_i - t_j)^2}{2\lambda^2} - \Gamma^2 \sin^2 \left( \frac{\pi|t_i - t_j|}{P_{\text{GP}}} \right) \right], \quad (4.26)$$

$$C_{ij} = k_{ij} + \delta_{ij} \sigma_{\text{RV},i}^2. \quad (4.27)$$

We therefore have six model parameters leading to a  $6 \times 6$  Fisher information matrix  $B$  which is related the model parameter covariance matrix  $C' = B^{-1}$  from which model parameter measurement uncertainties are calculated. Assuming a circular Keplerian orbit for the transiting planet of interest, the residual RV vector is

$$\mathbf{r}(\mathbf{t}) = \mathbf{y}(\mathbf{t}) - (-K \sin(\phi(\mathbf{t}))) \quad (4.28)$$

where  $\mathbf{y}$  are the raw RVs observed at times  $\mathbf{t}$  and  $\phi(\mathbf{t})$  is the planet’s orbital phase centered on mid-

transit. The generalized ln likelihood from which the Fisher information matrix is calculated is then

$$\ln \mathcal{L} = -\frac{1}{2} (\mathbf{r}^T C^{-1} \mathbf{r} + \ln \det C + \text{constant}), \quad (4.29)$$

Before populating the Fisher information matrix, we note two crucial mathematical identities

$$\frac{\partial}{\partial \theta} C^{-1} = -C^{-1} \frac{\partial C}{\partial \theta} C^{-1}, \quad (4.30)$$

$$\frac{\partial}{\partial \theta} \ln \det C = \text{tr} \left( C^{-1} \frac{\partial C}{\partial \theta} \right). \quad (4.31)$$

We can now proceed with calculating the general equation for each of the 21 unique Fisher information matrix entries in terms of partial derivatives of either the RV residual vector or covariance matrix  $C$  with respect to the model parameters instead of the partial derivative of the inverse covariance matrix.

$$B_{ij} = -\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j}, \quad (4.32)$$

$$\begin{aligned} \frac{\partial \ln \mathcal{L}}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i} \left[ -\frac{1}{2} \mathbf{r}^T C^{-1} \mathbf{r} - \frac{1}{2} \ln \det C \right] \\ &= -\frac{1}{2} \left[ \left( \frac{\partial \mathbf{r}}{\partial \theta_i} \right)^T C^{-1} \mathbf{r} - \mathbf{r}^T C^{-1} \frac{\partial C}{\partial \theta_i} C^{-1} \mathbf{r} + \mathbf{r}^T C^{-1} \left( \frac{\partial \mathbf{r}}{\partial \theta_i} \right) + \text{tr} \left( C^{-1} \frac{\partial C}{\partial \theta_i} \right) \right], \end{aligned} \quad (4.33)$$

$$\begin{aligned} \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} &= -\frac{1}{2} \left[ \left( \frac{\partial^2 \mathbf{r}}{\partial \theta_i \partial \theta_j} \right)^T C^{-1} \mathbf{r} - \left( \frac{\partial \mathbf{r}}{\partial \theta_i} \right)^T C^{-1} \frac{\partial C}{\partial \theta_j} C^{-1} \mathbf{r} + \left( \frac{\partial \mathbf{r}}{\partial \theta_i} \right)^T C^{-1} \left( \frac{\partial \mathbf{r}}{\partial \theta_j} \right) \right. \\ &\quad - \left( \left( \frac{\partial \mathbf{r}}{\partial \theta_j} \right)^T C^{-1} \frac{\partial C}{\partial \theta_i} C^{-1} \mathbf{r} - \mathbf{r}^T C^{-1} \frac{\partial C}{\partial \theta_j} C^{-1} \frac{\partial C}{\partial \theta_i} C^{-1} \mathbf{r} + \mathbf{r}^T C^{-1} \frac{\partial^2 C}{\partial \theta_i \partial \theta_j} C^{-1} \mathbf{r} \right. \\ &\quad \left. - \mathbf{r}^T C^{-1} \frac{\partial C}{\partial \theta_i} C^{-1} \frac{\partial C}{\partial \theta_j} C^{-1} \mathbf{r} + \mathbf{r}^T C^{-1} \frac{\partial C}{\partial \theta_i} C^{-1} \left( \frac{\partial \mathbf{r}}{\partial \theta_j} \right) \right) \\ &\quad \left. + \left( \frac{\partial \mathbf{r}}{\partial \theta_j} \right)^T C^{-1} \frac{\partial \mathbf{r}}{\partial \theta_i} - \mathbf{r}^T C^{-1} \frac{\partial C}{\partial \theta_j} C^{-1} \frac{\partial \mathbf{r}}{\partial \theta_i} + \mathbf{r}^T C^{-1} \frac{\partial^2 \mathbf{r}}{\partial \theta_i \partial \theta_j} \right] \end{aligned} \quad (4.34)$$

Each matrix entry is calculated using Eq. 4.34. In Eq. 4.34 there are two first order and two second order partial derivatives that must be computed with respect to each of the six model parameters:  $\Theta = \{K, a, \lambda, \Gamma, P_{GP}, \sigma_{\text{jitter}}\}$ . These being

$$\frac{\partial \mathbf{r}}{\partial \theta_i}, \frac{\partial K}{\partial \theta_i}, \frac{\partial^2 \mathbf{r}}{\partial \theta_i \partial \theta_j}, \text{ and } \frac{\partial^2 K}{\partial \theta_i \partial \theta_j} \quad (4.35)$$

and can be computed analytically or symbolically using the open-source `sympy` package in `python` given the analytical expressions for the residual vector (Eq. 4.28) and the covariance matrix (Eq. 4.26 & 4.27).

Table 4.6: Stellar parameters from the S15 synthetic catalog

TOI	$\alpha$ [deg]	$\delta$ [deg]	$P$ [days]	$m_p$ [M $_{\oplus}$ ]	$K$ [m s $^{-1}$ ]	$S$ [S $_{\oplus}$ ]	$M_s$ [M $_{\odot}$ ]	$T_{\text{eff}}$ [K]	Distance [pc]	$B$	$V$	$R$	$Y$	$J$	$H$	med( $v \sin i_s$ ) [km s $^{-1}$ ]
0000	0.44	45.22	9.135	8.19	2.03	361.7	1.40	6531	100.0	8.87	8.47	8.24	7.80	7.63	7.46	2.17
0001	0.48	-66.20	14.200	5.57	3.11	2.1	0.33	3426	60.3	16.52	15.08	14.17	12.05	11.56	10.96	0.23
0002	0.65	42.94	4.957	4.51	1.66	235.0	1.05	5546	97.7	10.86	10.12	9.72	9.09	8.81	8.46	3.24
0003	0.92	-26.07	2.160	4.32	1.95	1240.0	1.17	5984	53.7	8.61	8.06	7.76	7.20	6.98	6.73	5.81
0004	1.31	-24.95	9.753	5.81	2.95	5.9	0.46	3622	66.1	15.59	14.19	13.32	11.47	10.99	10.34	0.57
0005	1.38	10.61	13.990	5.89	3.32	2.1	0.33	3425	58.9	16.48	15.04	14.13	12.01	11.52	10.92	0.18
0006	1.78	-71.93	8.420	8.15	5.23	4.5	0.35	3444	70.8	16.73	15.29	14.38	12.28	11.79	11.19	0.27
0007	1.79	-9.14	5.618	7.03	9.15	2.6	0.15	3228	23.4	16.63	15.05	13.98	11.66	11.13	10.63	0.09
0008	1.95	-17.00	1.344	37.81	16.03	10160.0	1.63	6668	100.0	7.78	7.51	7.37	6.83	6.72	6.63	8.19
0009	2.17	-15.53	17.140	11.57	2.10	341.2	1.63	6668	100.0	7.78	7.51	7.37	6.83	6.72	6.63	7.54

**Notes.**

Only a portion of this table is shown here to demonstrate its form and content. A machine-readable version of the full table is available in the published version of the manuscript.

Table 4.7: Median radial velocity noise sources and follow-up calculations for  $3\sigma$  planet mass detections of the S15 synthetic catalog

TOI	$\sigma_{\text{RV, opt}}$ [m s $^{-1}$ ]	$\sigma_{\text{RV,nIR}}$ [m s $^{-1}$ ]	$\sigma_{\text{act}}$ [m s $^{-1}$ ]	$\sigma_{\text{planets}}$ [m s $^{-1}$ ]	$\sigma_{\text{eff,opt}}$ [m s $^{-1}$ ]	$\sigma_{\text{eff,nIR}}$ [m s $^{-1}$ ]	$N_{\text{RV, opt}}$	$N_{\text{RV,nIR}}$	$t_{\text{obs,opt}}$ [nights]	$t_{\text{obs,nIR}}$ [nights]
0000	1.02	3.14	1.25	0.00	1.61	3.59	14.8	67.6	0.4	1.6
0001	10.18	8.21	7.84	2.00	13.06	11.62	294.1	159.1	42.0	3.8
0002	2.30	3.74	0.97	0.00	2.41	3.96	57.8	155.7	1.4	3.7
0003	1.09	2.77	15.11	0.00	15.16	15.41	87.2	125.7	2.1	3.0
0004	4.15	6.06	8.15	0.77	9.22	10.33	66.9	115.2	8.7	2.7
0005	9.94	8.04	7.56	1.18	12.62	11.17	241.4	111.2	34.5	2.6
0006	11.55	9.33	7.64	1.21	14.40	12.60	128.4	69.8	18.3	1.7
0007	8.99	6.69	7.93	3.37	12.86	11.30	34.6	32.0	4.9	0.8
0008	0.97	4.43	0.10	0.00	2.11	5.47	19.8	20.0	0.5	0.5
0009	0.91	4.11	0.76	0.00	1.78	4.85	14.8	108.5	0.4	2.6

**Notes.**

The resulting number of RV measurements and total observing times reported here are computed in the general case of RVs in the presence of correlated noise (see Sect. 4.2.1). Only a portion of this table is shown here to demonstrate its form and content. A machine-readable version of the full table is available.

# Chapter 5

## Characterization of the K2-18 multi-planetary system with HARPS: a habitable zone super-Earth and discovery of a second, warm super-Earth on a non-coplanar orbit

### 5.1 Introduction

Exoplanets orbiting within their host star's habitable zone may have surface temperatures that allow for the presence of liquid water on their surfaces, depending on the properties of the planetary atmosphere (Kasting et al., 1993). The presence of liquid water is a condition likely required to sustain extraterrestrial life. This implies that habitable zone exoplanets receive comparable levels of stellar insolation to what the Earth receives from the Sun. Habitable zone exoplanets therefore represent superlative opportunities to search for life outside of the solar system via the characterization of their atmospheric structure and composition via transmission spectroscopy for transiting exoplanets.

M dwarf host stars are ideal targets to probe potentially habitable exoplanetary atmospheres (e.g. Kaltenegger et al., 2011; Rodler & López-Morales, 2014). Transmission spectroscopy observations of transiting habitable zone (HZ) exoplanets around M dwarfs are favourable compared to those around Sun-like stars given the increased depth of the transit for a planet of a given size (e.g. Stevenson et al., 2010; Kreidberg et al., 2014). In addition, the orbital periods corresponding to the HZ are less than those

---

The contents of this chapter are copied verbatim from the published paper (Cloutier et al., 2017b).

around Sun-like stars (weeks for those in the HZ of M dwarfs, compared to 12 months) thus increasing the number of accessible transit events within a given observational baseline. M dwarfs are also known to frequently host multiple small planets (typically 2.5 planets per star with  $0.5 \leq r_p/R_\oplus \leq 4$  and within 200 days; Dressing & Charbonneau, 2015; Gaidos et al., 2016) thus enabling direct comparative planetology to be conducted on known multi-planet systems.

Montet et al. (2015) reported the detection of the HZ super-Earth K2-18b originally proposed in the K2 light curve analysis of Foreman-Mackey (2015). In these studies, two transit events were observed in Campaign 1 data from the re-purposed Kepler spacecraft mission K2 whose field coverage only lasted for  $\sim 80$  days. The existence of the planet was confirmed and uncertainties regarding its ephemeris were significantly reduced in Benneke et al. (2017) (hereafter B17) who used follow-up transit observations with the Spitzer Space Telescope to detect an additional transit event. The now confirmed super-Earth K2-18b orbits an M2.5 dwarf with a period of  $\sim 32.9$  days placing it directly within the star's habitable zone (Kopparapu et al., 2013). The measured radius of  $2.38 R_\oplus$  is suggestive of an extended H/He envelope (Valencia et al., 2013; Rogers, 2015; Fulton et al., 2017) that may contain additional molecular species such as water and/or methane that could be detectable with the James Webb Space Telescope (JWST; Beichman et al., 2014). Owing to the proximity of the system ( $\sim 34$  pc,  $V = 13.5$ ,  $I = 11.7$ ,  $K = 8.9$ ; Cutri et al., 2003; Zacharias et al., 2013), K2-18 is truly an attractive target for characterizing the atmosphere of a HZ super-Earth with unprecedented precision in the JWST-era.

In this study we report the first measurement of the planetary mass of K2-18b using precision radial velocity measurements taken with the HARPS spectrograph. In this data we also find strong evidence for an additional planet of similar minimum mass whose orbit is interior to K2-18b but is not found to transit. In Sect. 5.2 we summarize the HARPS spectroscopic and K2 photometric observations used in our analysis, in Sect. 5.3 we analyze the periodic signals in the spectroscopic data and in Sect. 5.4 we discuss our various radial velocity modelling procedures. In Sect. 5.5 we present the results of our radial velocity analysis including the detection of a second super-Earth K2-18c in the system which we show is non-transiting and therefore is not perfectly coplanar with K2-18b in Sect. 5.6. Lastly we perform a dynamical analysis of the two-planet system in Sect. 5.7 to dynamically constrain the orbital eccentricities of the planets before concluding with a discussion in Sect. 5.8.

## 5.2 Observations

### 5.2.1 HARPS spectra

From April 2015 (BJD=2457117.5) to May 2017 (BJD=2457875.5), we collected 75 spectra of K2-18 (EPIC 201912552) with the high-resolution ( $R=115000$ ) HARPS spectrograph (Mayor et al., 2003; Pepe et al., 2004). The majority of exposure times were fixed to 1800 seconds with the exception of the following six epochs whose exposure times were modified to the following reported values: 2400 seconds (BJD-2,450,000 = 7199.503915, 7200.503114), 1200 seconds (BJD-2,450,000 = 7204.491167), and 900 seconds (BJD-2,450,000 = 7810.806284, 7814.760772, 7815.759421). The online HARPS pipeline returned the extracted, wavelength-calibrated spectra (Lovis & Pepe, 2007). Initial radial velocity estimates were computed from the cross-correlation of each spectrum with a numerical mask (Baranne

et al., 1996; Pepe et al., 2002). Using each spectrum’s initial estimate, all spectra were shifted to a common reference frame by their corresponding barycentric correction such that spectral features originating from the target star become aligned while telluric features are shifted by minus the epoch’s barycentric correction. The median combination of these shifted spectra was then used to construct a custom reference spectrum at high signal-to-noise (S/N). A telluric template was then constructed from the median combination of all residual spectra after removal of the high S/N reference stellar spectrum. The process of computing the median reference stellar spectrum was then repeated using the individual spectra with tellurics masked by the median telluric spectrum. We then computed precision radial velocities by performing a  $\chi^2$ -minimization of each spectrum with the reference spectrum (Astudillo-Defru et al., 2015). Radial velocity uncertainties were then estimated directly on the reference spectrum (Bouchy et al., 2001).

From the extracted spectra we also derived a number of activity indicators including the time series of the H $\alpha$  index which is sensitive to chromospheric activity and is computed following the definition in Bonfils et al. (2007). For the M dwarf K2-18 ( $V=13.5$ ; Henden & Munari, 2014) the H $\alpha$  index is favoured over the Ca II H+K Mt. Wilson S index (Wilson, 1968; Baliunas et al., 1995) due to the low S/N obtained in the blue. From the S index we derived  $\log R'_{HK} = -5.247 \pm 0.318$  (Astudillo-Defru et al., 2017b). Additionally we derived the full width at half maximum (FWHM) and bi-sector inverse slope (BIS) shape parameters of the cross-correlation function which are modified by dark and/or bright active regions traversing the visible stellar surface. In Sect. 5.3 we use these ancillary time series to learn about the star’s activity simultaneously with our radial velocity measurements. All spectroscopic time series are reported in Table 5.3.

### 5.2.2 K2 photometry

K2-18 was observed in long-cadence mode during Campaign 1 of the K2 mission as part of the ‘Targeting M dwarfs with K2’ proposal (GO1053<sup>1</sup>, PI: B. Montet). The baseline of the K2 light curve is just 80 days but provides nearly continuous coverage between June 1st, 2014 (BJD=2456810.5) and August 20th, 2014 (BJD=2456890.5).

We obtained the full de-trended light curve from the MAST<sup>2</sup> data retrieval service. As a result of the loss of two reaction wheels on-board the Kepler spacecraft, photometric observations from the K2 mission exhibit a reduced pointing precision, and hence photometric precision, compared to the original Kepler mission. Raw K2 light curves must be de-trended with the variable pointing of the spacecraft throughout the observing sequence. We selected the EVEREST reduction of the K2 light curve which performs this de-trending correction (Luger et al., 2016).

The majority of the residual photometric variability following de-trending of the light curve can be attributed to the intrinsic photometric variability of the star and two observed transits of K2-18b from Montet et al. (2015). Removal of the transit events provides a dataset that can be used to investigate the correlated photometric activity resulting from active regions traversing the visible stellar surface thus

<sup>1</sup>K2-18 was also targeted in the following programmes: GO1006, GO1036, GO1050, GO1051, GO1052, GO1059, GO1063, GO1075

<sup>2</sup><http://archive.stsci.edu/k2/hlsp/everest/search.php>

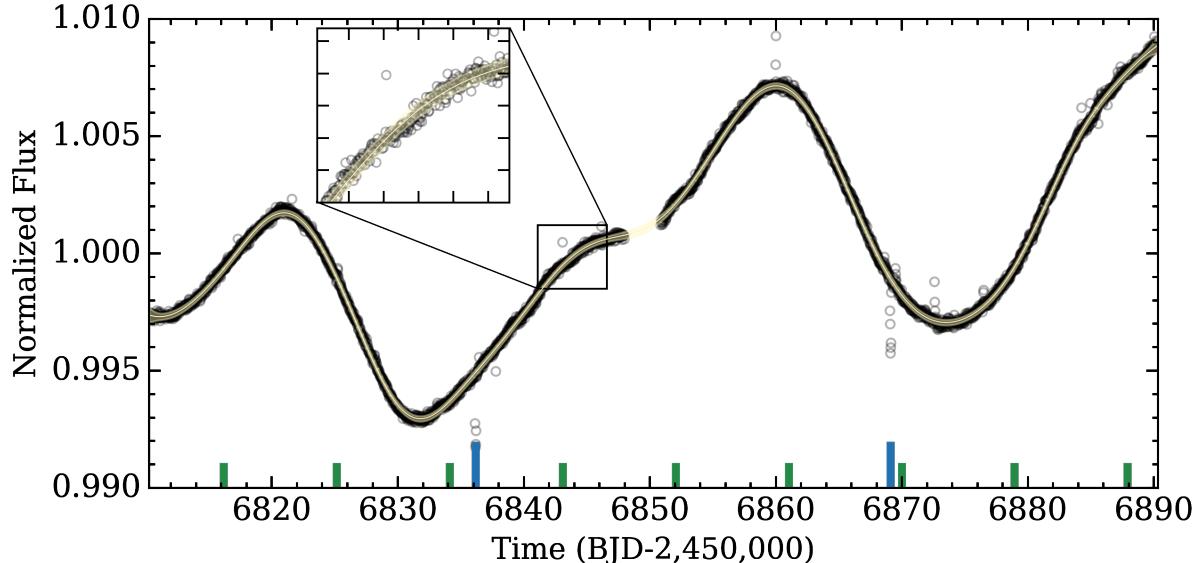


Figure 5.1: K2 photometric light curve after the removal of known unphysical spurious signals. The two [transits] of K2-18b are highlighted by the long blue ticks with the expected times of mid-transit for K2-18c highlighted with short green ticks (see Sect. 5.5). The [solid yellow curve] is the mean of the predictive GP distribution and the surrounding shaded regions mark its 99% confidence intervals. The upper left sub-panel is a magnified view of the highlighted region to aid in the visualization of the data and the GP fit.

giving rise to the star’s observed photometric variability. For reference, the de-trended light curve is shown in Fig. 5.1 along with our Gaussian process fit to the light curve (see Sect. 5.4.1 for an explanation of the fit).

### 5.3 Periodogram analysis

Accurate modelling of the stellar radial velocity (RV) variations requires knowledge of the strong periodicities present in the data. These signals include contributions from both orbiting planets and from the rotation of active regions present on the stellar surface which give rise to a correlated stellar RV signal which is modulated by the stellar rotation period and/or its harmonics. In the top panel of Fig. 5.2 we plot the Lomb-Scargle periodogram (Scargle, 1982) of the raw RVs to determine which periodicities are present at high significance, that is, those with a low false alarm probability (FAP). In all LS-periodograms we calculated FAPs via bootstrapping with replacement using  $10^4$  iterations and individually normalize each periodogram’s power by its standard deviation.

Two important features are detected in the LS-periodogram of the RVs. The first is a forest of peaks ranging from  $\sim 25 - 45$  days with distinct peaks centred on both the orbital period of K2-18b ( $P_b \sim 33$  days B17; FAP = 2.9%) and the approximate stellar rotation period from the K2 photometry (see Sect. 5.4.1 for fitting of the stellar rotation period  $P_{\text{rot}} \sim 38.6$  days; FAP = 0.1%). The second important feature is a pair of closely spaced peaks at  $\sim 9$  days (FAP  $< 0.01\%$ ) which because of their similar period and power likely result from a single source. This feature at approximately nine days constitutes the strongest periodic signal in the periodogram of the raw RVs and is not observed in the periodograms of any of the ancillary time series, nor in the periodogram of the window function, all of

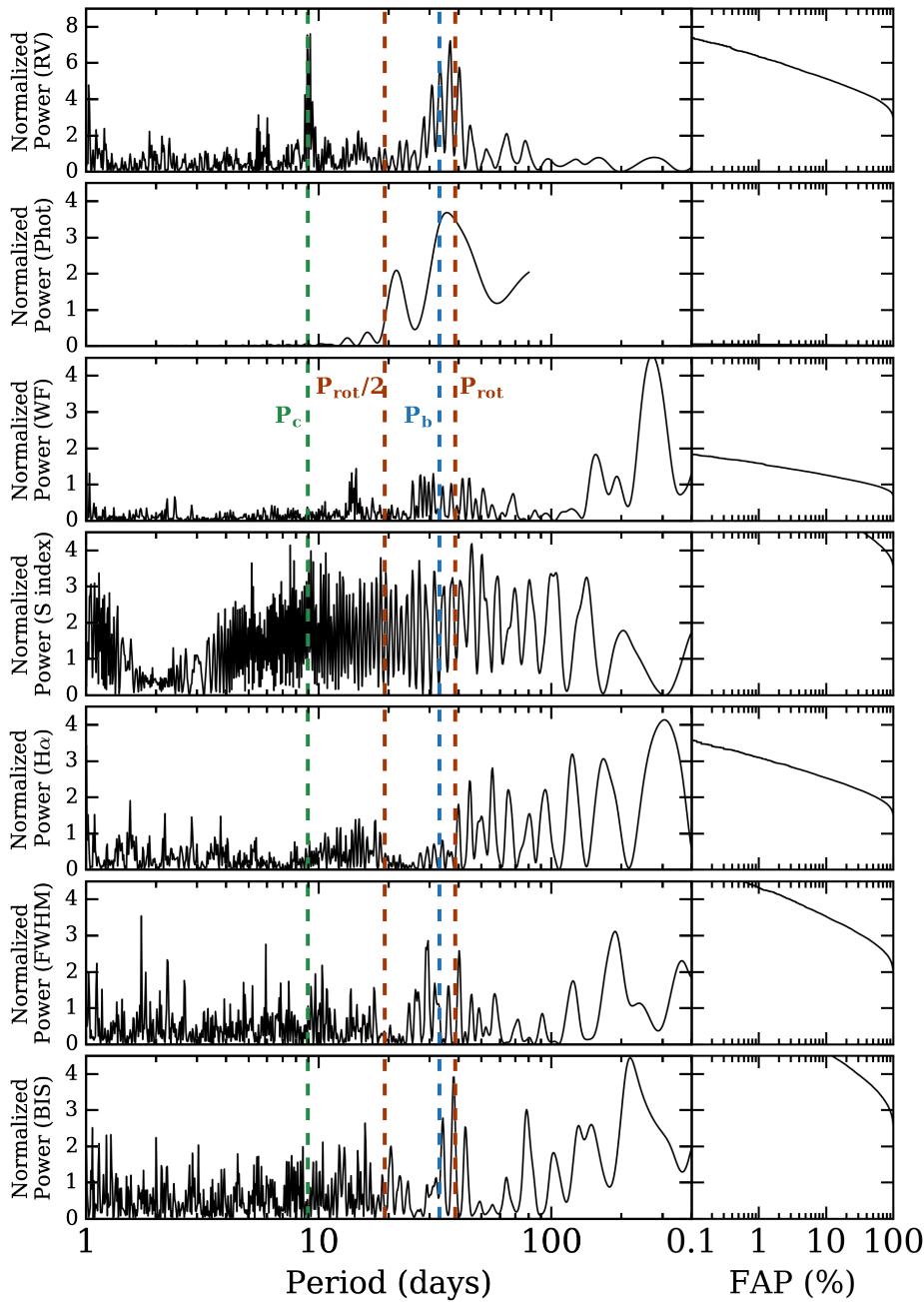


Figure 5.2: Left column (top to bottom): Lomb-Scargle periodograms of the raw radial velocities (RV), the K2 photometry (Phot), the HARPS window function (WF), S index, H $\alpha$ , full width half maximum (FWHM), and bi-sector inverse slope (BIS) time series. The orbital periods of K2-18b, K2-18c, the stellar rotation period, and its first harmonic are highlighted with vertical dashed lines. Right column: the false alarm probability as a function of normalized periodogram power for each time series. The FAP curve for the photometry spans very low power and is barely discernible in its subpanel. Suffice it to say that any signal with normalized power  $> 10^{-2}$  has a FAP  $\ll 0.1\%$ . The FAP curve for the S index exhibits FAP  $\gtrsim 30\%$  for all power visible on its ordinate and therefore is only visible in the upper right of its subpanel.

which are shown in the remaining panels of Fig. 5.2. The aforementioned time series include the K2 photometry (see Fig. 5.1), the window function or time sampling of the HARPS observations, and four spectroscopic activity indicators: the S index, H $\alpha$  index, FWHM, and the BIS of the cross-correlation function. Together the presence of the strong  $\sim 9$  d. signal in radial velocity and its absence elsewhere provides strong initial evidence for a second planet in the K2-18 system at  $\sim 9$  d.

## 5.4 Joint modelling of planets and correlated RV activity

### 5.4.1 Training the GP activity model on ancillary time series

The K2 photometry of K2-18 exhibits quasi-periodic photometric variability with a semi-amplitude of  $\sim 0.008$  mag and a rotation period of  $P_{\text{rot}} \sim 38.6$  days as seen in Fig. 5.1. This makes K2-18 a moderately active early M dwarf in terms of its photometric variability (Newton et al., 2016) the origin of which likely results from the rotation of active regions across the projected stellar disk at or close to  $P_{\text{rot}}$  owing to the characteristically low amplitudes of differential rotation in M dwarfs (Kitchatinov & Olemskoy, 2011). The observed photometric variability—or photometric activity—has a correlated manifestation in the variation of the star’s apparent radial velocity and certain spectroscopic indicators because it is a single physical process that is responsible for the activity in each time series (Aigrain et al., 2012).

In order to obtain accurate and self-consistent detections of the planetary signals in radial velocity we must model the RV activity signal of K2-18 simultaneously with our planet model. The photometric stellar rotation period of  $\sim 38$  days (see second panel in Fig. 5.2) is marginally detected in the LS-periodogram of the RVs. However  $P_{\text{rot}}$  is clearly discernible by eye in the K2 light curve (Fig. 5.1) and has significant power in the LS-periodogram of the K2 light curve (second panel in Fig. 5.2) although the power is spread over a wide range of periodicities. Because of this we consider in our first model—called Model 1—the K2 photometric light curve, less the observed transits of K2-18b, to train our RV activity model whose covariance properties are common with the observed photometric variability. However two important caveats had to be considered when adopting the K2 photometry as our training set. The first being that the baseline of the photometry spans just 80 days implying that any temporal variation whose characteristic timescale is greater than this baseline will remain unconstrained or at best weakly constrained. Secondly the K2 photometry were obtained nearly eight months prior to our HARPS observations such that any evolution in the covariance structure of the stellar activity between observing sequences from say magnetic activity cycles, would not be captured in the training set. For these reasons we also considered the BIS time series as an alternative training set in a second round of modelling called Model 2. Being contemporaneous with the RV measurements, training on the BIS time series mitigates the two aforementioned issues. In place of the BIS we also tested training on the S index, H $\alpha$ , and FWHM time series but find results consistent with training on the BIS. Following Faria et al. (2016) we also considered a joint activity + planet model but neglected any training of the activity model’s covariance structure in a third model; Model 3. Finally for comparison purposes we also considered a fourth model—called Model 4—that neglected any contribution from stellar activity.

To implement this joint modelling procedure we followed Cloutier et al. (2017a) by using a Gaussian process (GP) regression model to model the covariance between adjacent observations in our training

sets where applicable (i.e. in Models 1 and 2). GP regression is an attractive method for modelling the stochastic processes that gives rise to observable RV activity signals as it is non-parametric and therefore independent of an assumed functional form of the signal. The GP prior was represented by a multivariate Gaussian distribution of functions described by a covariance matrix  $K_{ij} = k_{ij}(\boldsymbol{\theta}) + \sigma_i^2 \delta_{ij}$  with a function  $k_{ij}(\boldsymbol{\theta}) = k(t_i, t_j, \boldsymbol{\theta})$  that parameterizes the covariance between values of the observable  $\mathbf{y}(\mathbf{t})$  at the epochs  $t_i$  and  $t_j$  in  $\mathbf{t}$ . The observable  $\mathbf{y}(\mathbf{t})$  has associated uncertainties  $\boldsymbol{\sigma}(\mathbf{t})$  which were added along the diagonal of the covariance matrix  $K$  in quadrature. The set of GP hyperparameters  $\boldsymbol{\theta}$  are unique to the chosen covariance function  $k_{ij}(\boldsymbol{\theta})$  and are solved for in the training step. After solving for the GP hyperparameters and thus obtaining a unique GP prior distribution, the GP prior conditioned on the data  $\mathbf{y}(\mathbf{t})$  becomes the predictive distribution. The mean function of the GP predictive distribution can be evaluated at previously unseen epochs  $\mathbf{t}^*$  using

$$\mu(\mathbf{t}^*) = K(\mathbf{t}^*, \mathbf{t}) \cdot K(\mathbf{t}, \mathbf{t})^{-1} \cdot \mathbf{y}(\mathbf{t}), \quad (5.1)$$

which we took to be our GP activity model of the RVs by evaluating Eq. 5.1 at  $\mathbf{t}$ .

Because the stellar activity, and in particular the long-term photometric variation, is modulated by the stellar rotation period, we included a periodic term in our assumed covariance function  $k_{ij}(\boldsymbol{\theta})$  with period equal to  $P_{\text{rot}}$ . We also included a radial component due to the stochastic temporal evolution of starspot lifetimes, spatial distributions, and contrasts thus forcing the covariance to not be strictly periodic. Explicitly the adopted covariance structure is parameterized by a quasi-periodic covariance kernel of the form

$$k_{i,j}(\boldsymbol{\theta}) = a^2 \exp \left[ -\frac{|t_i - t_j|^2}{2\lambda^2} - \Gamma^2 \sin^2 \left( \frac{\pi|t_i - t_j|}{P_{\text{GP}}} \right) \right], \quad (5.2)$$

which is parameterized by four hyperparameters  $\boldsymbol{\theta} = (a, \lambda, \Gamma, P_{\text{GP}})$ :  $a$  the amplitude of the correlations,  $\lambda$  the exponential timescale,  $\Gamma$  the coherence scale of the correlations, and  $P_{\text{GP}}$  the periodic timescale of the correlations which we interpret as  $P_{\text{rot}}$ . We also included an additional scalar jitter parameter  $s$  which is added in quadrature to the diagonal elements of the covariance matrix  $K$  such that  $\boldsymbol{\theta}$  becomes  $(a, \lambda, \Gamma, P_{\text{GP}}, s)$ .

Using the Markov chain Monte-Carlo (MCMC) ensemble sampler `emcee` (Foreman-Mackey et al., 2013) we sample the marginalized posterior probability density functions (PDFs) of the five hyperparameters assuming uniform priors on the logarithm of each hyperparameter and maximizing the Gaussian logarithmic likelihood function

$$\ln \mathcal{L} = -\frac{1}{2} (\mathbf{y}^T K^{-1} \mathbf{y} + \ln \det K + N \ln 2\pi), \quad (5.3)$$

where  $\mathbf{y}$  is the vector of  $N$  observations in the training set. In Model 1  $\mathbf{y}$  = the binned photometric data points<sup>3</sup> whereas  $\mathbf{y}$  = BIS time series in Model 2.

The MCMC was initialized with 200 walkers and hyperparameter values  $(a, \lambda, \Gamma, P_{\text{GP}}) = (\max(\mathbf{y} - \langle \mathbf{y} \rangle), 10^2 \text{ days}, 1, 36 \text{ days})$ . We sampled the logarithmic hyperparameters up to  $\approx 10$  autocorrelation

---

<sup>3</sup>Binning the K2 photometry in one day bins results in  $N = 78$  compared to the 3439 unbinned photometric observations thus drastically increasing the computational efficiency of the evaluating Eq. 5.3 in each step of the Markov chains.

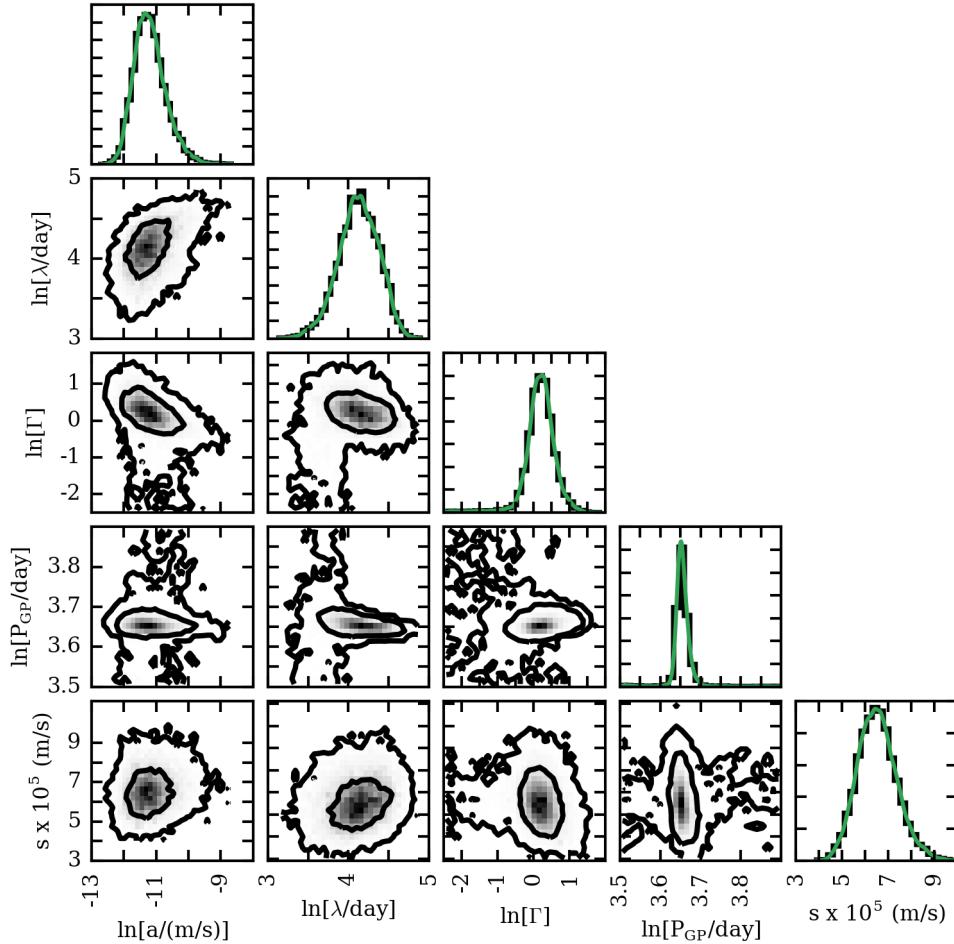


Figure 5.3: Marginalized and joint posterior PDFs of the logarithmic GP hyperparameters used to model the K2 photometry shown in Fig. 5.1. Kernel density estimations of each model parameter’s posterior are overlaid on the histograms with solid green lines.

times to ensure adequate convergence of the chains. We also monitored the acceptance fraction for each walker and insist that it lies within 20 – 50%. The sampling of each hyperparameter’s marginalized posterior PDF commences following a burn-in phase of the same duration. The resulting marginalized and joint posterior PDFs are shown in Fig. 5.3 along with kernel density estimations of each marginalized 1D distribution. From the posterior PDF of  $P_{\text{GP}}$  we measured a stellar rotation period of  $P_{\text{rot}} = 38.6^{+0.6}_{-0.4}$  days.

#### 5.4.2 Joint modelling of RVs

We proceeded with modelling the RVs jointly with Keplerian solutions for both K2-18b and c plus a trained quasi-periodic GP to model the correlated RV residuals attributed to stellar activity. The marginalized posterior PDFs of the GP hyperparameters  $\lambda$ ,  $\Gamma$ , and  $P_{\text{GP}}$  from training were then used as informative priors in the joint RV analysis which treat the remaining GP hyperparameters  $a$  and  $s$  as free parameters. We sampled the informative priors using the kernel density estimations of each hyperparameter’s PDF obtained during training. This methodology allows the model to learn the covariance

structure of the stellar activity through observations which are independent of planetary sources and then apply that knowledge to the joint modelling of the RVs thus distinguishing between stellar activity and planet-induced Doppler shifts.

The RV modelling is again performed using `emcee`. In Models 1, 2, and 3 our RV model consisted of 16 parameters including the five GP hyperparameters discussed in Sect. 5.4.1, the systemic velocity of K2-18  $\gamma_0$ , the orbital periods of the two planets  $P$ , their times of inferior conjunction  $T_0$ , their RV semi-amplitudes  $K$ , and the MCMC jump parameters  $h = \sqrt{e} \cos \omega$  and  $k = \sqrt{e} \sin \omega$  describing each planet's orbital eccentricity  $e$  and argument of periastron  $\omega$ . This parameterization was chosen to minimize the correlation between  $e$  and  $\omega$  as well as reduce the tendency for the MCMC sampler to favour high-eccentricity solutions (Ford, 2006). In Model 4 we only considered 11 model parameters as no GP activity model was included. Table 5.1 summarizes the adopted priors on each RV model parameter in each of the models considered in this study. We adopted non-informative priors for all Keplerian parameters other than the orbital period and time of mid-transit of K2-18b which were well-constrained by the transit light curve modelling in B17.

## 5.5 Results

### 5.5.1 Results from RV Data Analysis

Here we compare results from the four considered RV models. Each model contains Keplerian solutions for each of the two planets. Additionally Model 1 models the RV residuals with a quasi-periodic GP regression model that is trained on the K2 photometry in which the stellar rotation period is clearly detected (see Sect 5.4.1). In this model the stellar rotation period, and hence the GP periodic term, is sufficiently distinct from the orbital period of K2-18b such that the two signals are not confused in our joint modelling and the measured semi-amplitude of K2-18b is not mis-estimated. Model 2 models the RV residuals with a quasi-periodic GP regression model that is trained on the BIS time series which is contemporaneous with the RVs. Model 3 models the RV residuals with an effectively unconstrained quasi-periodic GP and Model 4 neglects any modelling of the RV residuals therefore assuming that they are uncorrelated.

The maximum a-posteriori (MAP) values of each model parameter along with the 16<sup>th</sup> and 84<sup>th</sup> percentiles of the marginalized posterior PDFs are reported in Table 5.2. The marginalized and joint posterior PDFs of the model parameters used in Model 1 are shown in Fig. 5.4. In Sect. 5.5.2 we shall see that Model 1 is the best predictor of the observed RVs and therefore we only show the results from Model 1 in Fig. 5.4.

We emphasize that all Keplerian model parameters for the two planets are consistent at  $1\sigma$  across all four models. Recall that the stellar rotation period is only well-constrained in Model 1 via the K2 photometry and yet the measured semi-amplitudes of K2-18b are consistent in each of the four models. This further demonstrates that there is minimal confusion between the RV signals at the stellar rotation period (38.6 days) and at the orbital period of K2-18b ( $\sim 32.93$  days) both of which are not distinctly detected in the periodogram of the raw RVs (see top panel of Fig. 5.2) but appear to be hidden within a

Table 5.1: Summary of RV models and adopted priors

Parameter	Prior
Model 1 (2 planets + GP trained on $\mathbf{y} = \text{K2 photometry}$ )	
<i>GP hyperparameters</i>	
Covariance amplitude, $a$ [m s <sup>-1</sup> ]	$\mathcal{J}(0.1, 30)^{\bullet}$
Exponential timescale, $\lambda$ [days]	$p(\lambda \mathbf{y})$
Coherence, $\Gamma$	$p(\Gamma \mathbf{y})$
Periodic timescale, $P_{\text{GP}}$ [days]	$p(P_{\text{GP}} \mathbf{y})$
Additive jitter, $s$ [m s <sup>-1</sup> ]	$\mathcal{J}(10^{-2}, 10)$
<i>Keplerian parameters</i>	
$\gamma_0$ [m s <sup>-1</sup> ]	$\mathcal{U}(620, 670)$
$P_b$ [days]	$\mathcal{N}(32.939614, 10^{-4})^{\circ}$
$T_{0,b}$ [BJD-2,450,000]	$\mathcal{N}(7264.39144, 6.3 \times 10^{-4})^{\circ}$
$K_b$ [m s <sup>-1</sup> ]	$\text{mod}\mathcal{J}(1, 20)^{*}$
$h_b = \sqrt{e_b} \cos \omega_b$	$\mathcal{U}(-1, 1)^{\dagger}$
$k_b = \sqrt{e_b} \sin \omega_b$	$\mathcal{U}(-1, 1)^{\dagger}$
$P_c$ [days]	$\mathcal{U}(8, 10)$
$T_{0,c}$ [BJD-2,450,000]	$\mathcal{U}(7259, 7269)$
$K_c$ [m s <sup>-1</sup> ]	$\text{mod}\mathcal{J}(1, 20)$
$h_c = \sqrt{e_c} \cos \omega_c$	$\mathcal{U}(-1, 1)^{\dagger}$
$k_c = \sqrt{e_c} \sin \omega_c$	$\mathcal{U}(-1, 1)^{\dagger}$
Model 2 (2 planets + GP trained on $\mathbf{y} = \text{BIS}$ )	
see Model 1	
Model 3 (2 planets + untrained GP)	
see Model 1 with the following modifications:	
ln Exponential timescale, ln $\lambda$ [days]	$\mathcal{U}(-10, 10)$
ln Coherence, ln $\Gamma$	$\mathcal{U}(-3, 3)$
ln Periodic timescale, ln $P_{\text{GP}}$ [days]	$\mathcal{U}(3.2, 4)$
Model 4 (2 planets)	
see Keplerian parameters in Model 1	

**Notes.** <sup>(•)</sup> $\mathcal{J}$  refers to a non-informative Jeffreys prior which is scale invariant; equal probability per decade which is necessary to sample multiple orders of magnitude (Gregory, 2005).

<sup>(◦)</sup> based on the transit light curve measurements from B17.

<sup>(\*)</sup>  $\text{mod}\mathcal{J}(k, l)$  m s<sup>-1</sup> refers to a modified Jeffreys prior on a parameter  $A$  which behaves like a uniform prior for  $A \ll k$  m s<sup>-1</sup> and behaves like a Jeffreys prior at  $A \gg k$  up to  $l$ . We use a modified Jeffreys prior on the RV semi-amplitudes  $K$  to sample multiple decades as a Jeffreys prior but also include  $K = 0$  m s<sup>-1</sup> which a Jeffreys prior does not (Gregory, 2005).

<sup>(†)</sup> We further insist that  $e = h^2 + k^2 < 1$ .

forest of peaks spanning periodicities between  $\sim 25 - 45$  days. The consistency of all measured Keplerian parameters in each model also suggests that the RV residuals, following the removal of the two MAP Keplerian solutions, are weakly correlated because nearly identical RV solutions are obtained with and without a GP treatment of the RV residuals following the removal of our planet models. That is that K2-18 appears to be a spectroscopically quiet star with the majority of its observed RV variation being attributable to planetary companions. Being spectroscopically quiet is promising for the prospect of transmission spectroscopy of K2-18b; an observation that is significantly complicated by the presence of stellar activity. The quiet nature of K2-18 is highlighted by its low measured value of  $\log R'_{\text{HK}} = -5.247$ .

### 5.5.2 RV model comparison

A formal model comparison between the four considered models was performed using time series cross-validation to compute the likelihood of each model given various training and testing subsets of the observed RVs (Arlot & Celisse, 2010). We split the RVs into chronological training sets with sizes ranging from 20 measurements to the size of the full dataset less one (i.e. 74 measurements). The model parameters for each of the four considered models were optimized on the training set and the likelihood of the corresponding model is evaluated on the testing set. The testing set was simply the next observation chronologically following the final observation in the training set. The resulting median likelihood and median absolute deviation for each model is reported at the bottom of Table 5.2 and was used to distinguish which of our four RV models performs optimally on the prediction of unseen RV measurements and thus best fits the data without over-fitting. Through time series cross-validation we find that Model 1 is the best predictor of the observed RVs. In the remainder of this study we consider the results from Model 1 to be the measured values of the planets K2-18b and c.

To confirm that we have detected a second planet K2-18c in our RV data, we performed a second round of time series cross-validation calculations. In these calculations we compared three RV models each containing 0, 1, or 2 planets. We considered K2-18b to be the lone planet in the one planet model. In each model we also considered a GP activity model that was trained on the K2 photometry as was Model 1 above. Following the same methodology as previously discussed we find median logarithmic likelihoods of  $\ln \mathcal{L}_0 = -2.693 \pm 0.056$ ,  $\ln \mathcal{L}_1 = -2.642 \pm 0.047$ , and  $\ln \mathcal{L}_2 = -2.566 \pm 0.026$ . From this we find that  $\ln \mathcal{L}_2 - \ln \mathcal{L}_1 = 0.076 \pm 0.054 > 0$  therefore arguing that the two planet model is the best predictor of unseen RV measurements and confirming that our two planet model containing both K2-18b and c is the RV model most favoured by the data.

The contributions to the observed RVs from stellar activity and each planet were depicted in Fig. 5.5. Together these physical models account for all significant periodicities in the observed RVs. In Fig. 5.5 we show the raw RVs as well as the RVs corrected for the individual RV component and compare each time series to its MAP model. In panel b we can see the relative importance of the GP activity model at modelling the RV residuals following the removal of our two planet model. Owing to the spectroscopically quiet nature of K2-18, the contribution to the observed RVs from activity is relatively small yet still holds a periodic manifestation at the stellar rotation period, albeit with a small amplitude. The residual rms following the removal of all modelled contributions is  $2.89 \text{ m s}^{-1}$ . This value is less than the median photon noise limit of the measured RVs of  $3.56 \text{ m s}^{-1}$  suggesting that we have modelled all significant RV contributions. For comparison, the residual rms achieved in Model 4, which neglects

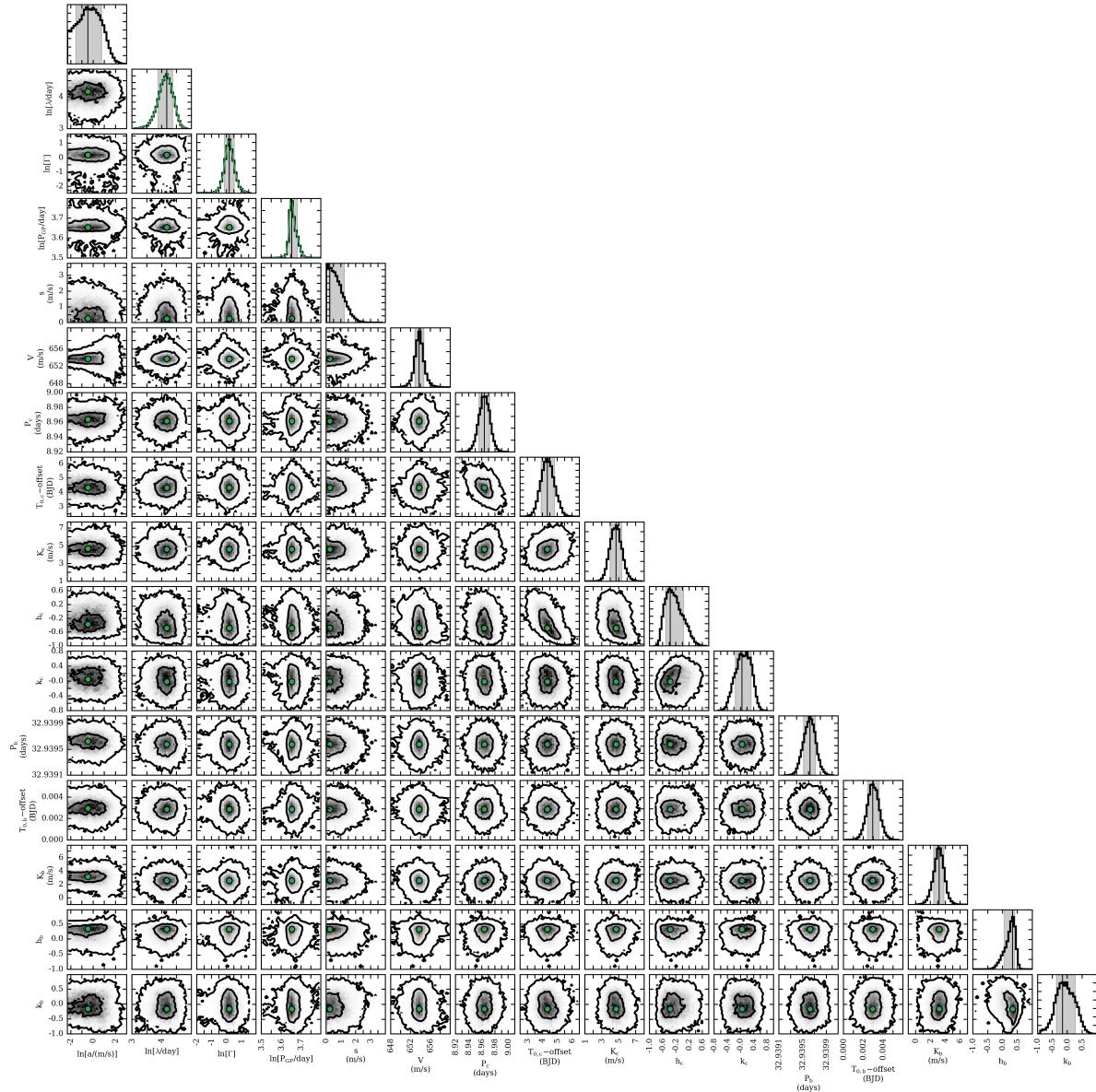


Figure 5.4: The marginalized and joint posterior PDFs of the model parameters from Model 1 of the observed RVs. Model 1 of the observed RVs models the two planets with Keplerian orbital solutions and the residual RV activity signal with a GP regression model trained on the K2 photometry in Fig. 5.1. Kernel density estimations of the trained posteriors are shown in the histograms of the logarithmic GP hyperparameters  $\lambda$ ,  $\Gamma$ , and  $P_{GP}$  (columns 2,3,4).

any red noise modelling following the removal of the two Keplerian solutions, is  $3.16 \text{ m s}^{-1}$ . This value is also less than the median photon noise limit suggesting that the GP regression modelling alone in Models 1, 2, and 3 does not result in over-fitting of the data.

The Model 1 MAP Keplerian orbital solutions for K2-18b and c, but with eccentricities fixed to zero, are shown in Fig. 5.6. Here we report circular orbital solutions given that with these data we can only place upper limits on each planet’s eccentricity rather than detect it directly. The RV data are phase-folded to each planet’s MAP orbital period and time of inferior conjunction and are corrected for stellar activity using the mean GP activity model trained on the K2 photometry and assuming a mean model equal to the superposition of the two MAP Keplerian solutions.

## 5.6 Searching for transits of K2-18c

From our RV analysis in Sect. 5.5 we derived the approximate linear ephemeris of K2-18c. We can therefore predict the passage of K2-18c at inferior conjunctions within the mostly continuous K2 photometric monitoring shown in Fig. 5.1. In Fig. 5.1 we indicate the nine such passages of K2-18c. Given the comparable minimum masses of K2-18b and c, it is reasonable to expect that the two planets also have comparable radii (recall  $r_{p,b} \sim 2.38 R_\oplus$ ). Furthermore (Ciardi et al., 2013) argued that Kepler multi-planet systems with planet radii  $\lesssim 3 R_\oplus$  do not exhibit a size—semi-major axis correlation such that the inner K2-18c is not expected to have undergone significant atmospheric escape compared to K2-18b. However the two  $10\sigma$  transits of K2-18b are clearly discernible by-eye in the K2 photometry whereas the predicted transits of K2-18c are not. This suggests that either K2-18c is much smaller than K2-18b such that its resulting transit depth is below the threshold for detection, or that the orbit K2-18c is mutually inclined with that of K2-18b such that it misses a transit configuration.

Here we attempt to confirm that K2-18c is indeed not transiting in the K2 light curves. To do so we perform an MCMC sampling of the K2-18c transit model parameters ( $P_c$ ,  $T_{0,c}$ ,  $r_{p,c}/R_s$ ,  $a_c/R_s$ , and impact parameter  $b_c$ ) using the K2 photometry and following the removal of the known transits of K2-18b. A quadratic limb-darkening law is assumed with fixed parameters in the Kepler bandpass:  $a = 0.3695$  and  $b = 0.3570$ . These values are interpolated from the tables of Claret & Bloemen (2011) based on the known K2-18 surface gravity and effective temperature. In each MCMC step we compute the corresponding transit model using the `batman` implementation (Kreidberg, 2015) of the Mandel & Agol (2002) transit model. We assumed a circular orbit of K2-18c and adopt the same MCMC methodology utilized on the RV data in Sect. 5.4.2. The orbital period and time of inferior conjunction (i.e. time of mid-transit) were sampled from their joint RV posterior which maintains their apparent correlation (see Fig. 5.4). Priors on the scaled planet radius, and impact parameter are assumed uniform. In this way the scaled planetary radius is uncorrelated with its measured minimum mass and the impact parameter is constrained to be  $|b_c| < 1$  as is required for a transit to occur.

Based on our MCMC analysis we find that the values of  $r_{p,c}/R_s$  are consistent with zero, that is, no transit is detected in the K2 data. Assuming the most likely value of  $R_s = 0.411 R_\odot$ , we calculate a planet radius upper limit of  $r_{p,c} < 0.52 R_\oplus$  at 99% confidence assuming that K2-18c is indeed transiting. However, if K2-18c were this size and transiting, albeit unknowingly due to its small size, the planet

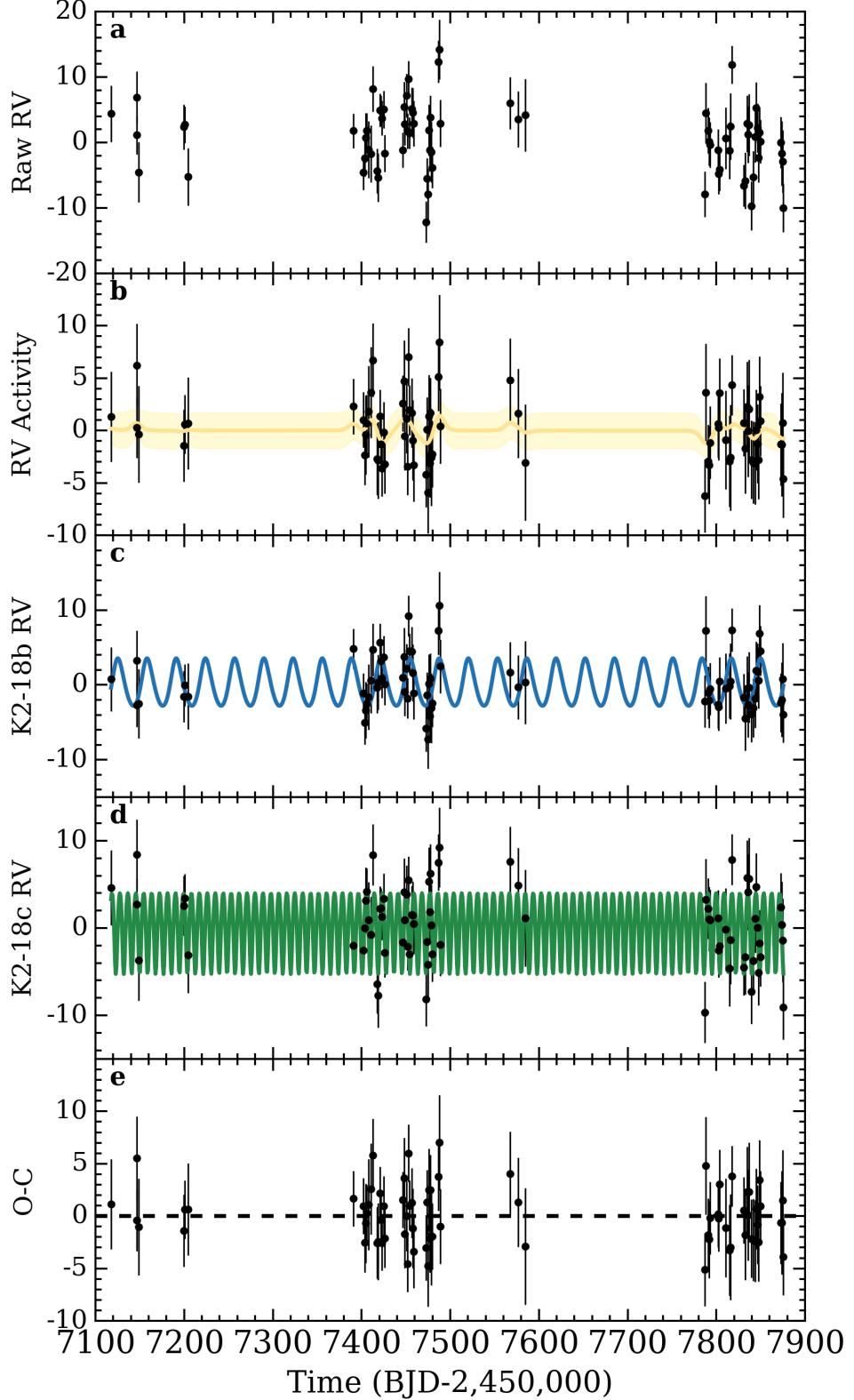


Figure 5.5: Panel a: Raw RVs less the systemic velocity of K2-18. Panel b: RV contribution from stellar activity; raw RVs corrected by the MAP Keplerian orbital solutions for each detected planet. Panel c: RV contribution from K2-18b; raw RVs corrected for activity and K2-18c. Panel d: RV contribution from K2-18c; raw RVs corrected for activity and K2-18b. Panel e: the RV residuals. The solid curves in panels b, c, and d depict the mean GP activity model, and the MAP Keplerian models for K2-18b and c respectively. The surrounding shaded region in panel b is the 68% confidence interval on the mean GP model. All RV units are in  $\text{m s}^{-1}$ .

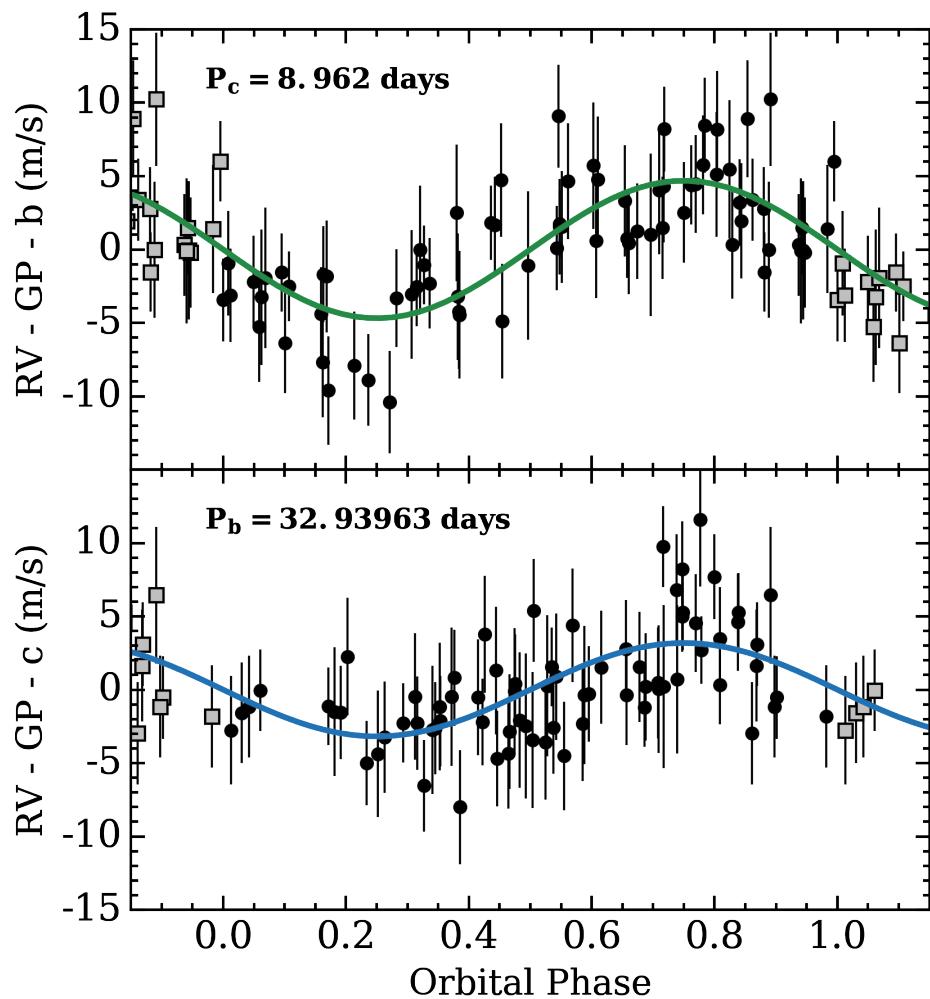


Figure 5.6: Phase-folded RVs for each planet in the K2-18 planetary system (top: K2-18c and bottom: K2-18b). The RVs have been corrected for stellar activity with a quasi-periodic GP model trained on the K2 photometry. The [solid curves](#) depict the maximum a-posteriori Keplerian orbital solutions for each planet with fixed circular orbits.

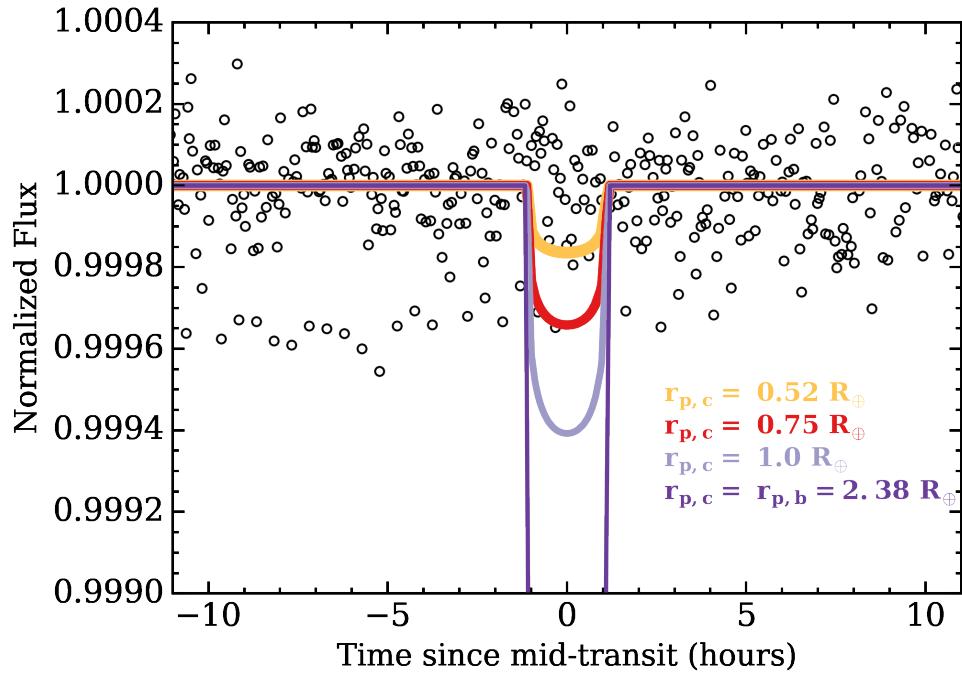


Figure 5.7: K2 light curve of K2-18 phase-folded to the maximum a-posteriori orbital period and time of mid-transit of K2-18c from Model 1. Four transit light curve models are over-plotted for various illustrative values of the assumed size of K2-18c: the 99% upper limit on  $r_{p,c}$  ( $0.52 R_{\oplus}$ ),  $0.75 R_{\oplus}$ ,  $1 R_{\oplus}$ , and the size of K2-18b ( $2.38 R_{\oplus}$ ). No transit of K2-18c is detected in the data.

would have a bulk density of  $\gtrsim 295 \text{ g cm}^{-3}$  or  $\gtrsim 54 \rho_{\oplus}$ ; an unphysically large value given the compressibility of pure iron. Thus we conclude that K2-18c is not transiting in the K2 data and is therefore non-coplanar with K2-18b despite having a smaller orbital separation.

To visualize the data a selection of light curve models are compared to the phase-folded K2 photometry in Fig. 5.7. Models shown include a suite of  $r_{p,c}$  values including its upper limit derived from MCMC ( $0.52 R_{\oplus}$ ),  $0.75 R_{\oplus}$ ,  $1 R_{\oplus}$ , and the radius of K2-18b ( $2.38 R_{\oplus}$ ) which is detected at the  $10\sigma$  level in the K2 photometry (Montet et al., 2015).

## 5.7 Dynamical stability and eccentricity restrictions

The non-detection of K2-18c in transit (see Sect. 5.6) suggests that its orbital plane is not perfectly coplanar with the outer transiting K2-18b whose semi-major axis is  $\sim 2.4$  times greater than K2-18c's. The orbital inclination of K2-18b is  $89.5785^{+0.0079}_{-0.0088}$  degrees with a corresponding impact parameter of  $0.601^{+0.013}_{-0.011}$  (B17). In order for the orbit of K2-18c to not pass in front of its host star its orbital inclination must be tilted either  $\gtrsim 1.41^\circ$  or  $\lesssim -2.25^\circ$  from the orbit of K2-18b depending on which hemisphere of the stellar disk its transit chord will traverse. Such a mutual inclination is consistent with the peak in the distribution of Kepler multi-planet mutual inclinations (Figueira et al., 2012; Fabrycky et al., 2014). If indeed the planetary angular momentum vectors are within only a few degrees and therefore nearly aligned then we can analytically evaluate their Hill stability given estimates of their orbital eccentricities

and assuming an inclination correction factor that is close to unity (Gladman, 1993). If we assume the simplest case of initially circular orbits then the system is strongly Hill stable given that the two planets are currently separated by  $\sim 23$  mutual Hill radii.

Accurate orbital eccentricities of small planets with precision radial velocities are notoriously difficult to measure. For example, the change in RV semi-amplitude of a circular K2-18b compared to an eccentricity of 0.1 is  $\lesssim 2 \text{ cm s}^{-1}$  ( $\sim 0.5\%$  of  $K_b$ ) or  $15 \text{ cm s}^{-1}$  ( $\sim 5\%$  of  $K_b$ ) for an eccentricity of 0.3. The aforementioned values are both at least an order of magnitude smaller than the characteristic RV uncertainty of the HARPS measurements presented in this work. Given that the system is Hill stable at small eccentricities we can use dynamical simulations to constrain the orbital eccentricities of the planets insisting that the system remain stable throughout its simulated evolution.

To constrain the planet eccentricities we performed a suite of  $10^4$  dynamical integrations wherein we sample linearly each planet's  $e \in [0, 1]$ . In each simulation the orbital inclination of K2-18b is drawn from  $\mathcal{N}(89.5785^\circ, 0.0084^\circ)$  while the system's mutual inclination is drawn from  $\mathcal{N}(\Delta i_{\min,c}, 1.5^\circ)$  such that the planet inclinations remain uncorrelated with the orbital eccentricities thus permitting an unbiased assessment of the system's stability across the Keplerian parameter space. We insist that K2-18c be non-transiting at the start of each simulation by setting  $\Delta i_{\min,c}$  to be the minimum mutual inclination required for  $|b_c| > 1$  and rejecting draws for which this is not true. The dispersion in sampled mutual inclinations is tuned such that the mode of the resulting distribution lies within  $\sim 1 - 2^\circ$  (Fabrycky et al., 2014). Each planet's initial semi-major axis, true anomaly, and absolute mass is drawn from a Gaussian distribution with a mean value equal to the parameter's MAP value from Model 1 in Table 5.2 and with standard deviations equal to its average measurement uncertainty. The stellar mass is drawn from  $\mathcal{N}(0.359, 0.047) \text{ M}_\odot$ . The ascending node longitudes and arguments of periapsis are both drawn from  $\mathcal{U}(-\pi, \pi)$ . The system is then integrated forward in time from the epoch of the first K2 photometric observation (BJD=2456810.26222) for  $10^6$  years using the Wisdom-Holman symplectic integrator WHFast (Rein & Tamayo, 2015) implemented in the open-source REBOUND N-body package (Rein & Liu, 2012). These integrations are not intended to provide a comprehensive overview of the system's dynamical stability but rather are useful to show that the system can remain stable up to at least 1 Myr and provide constraints on the planet eccentricities.

We classify stable integrations as those in which the minimum distance between the planets never becomes less than their mutual Hill radius. The fraction of stable systems as a function of each planet's eccentricity and marginalized over all other dynamical parameters is shown in Fig. 5.8. Strong correlations between the fraction of stable systems and dynamical parameters other than planet eccentricities was not apparent so we focus here on the effect of eccentricities only. At small eccentricities there is a large stable region wherein the fraction of systems that remain stable is  $\gtrsim 80\%$  and the system is known to be Hill stable based on the analytic criterion. As we increase either planet's eccentricity the fraction of stable systems decreases. This is also illustrated by further marginalizing over planet eccentricities and considering the one-dimensional representations of each system's stability fraction in the histograms shown in Fig. 5.8.

The RV analysis discussed Sect. 5.5 and our dynamical simulations provide two independent methods

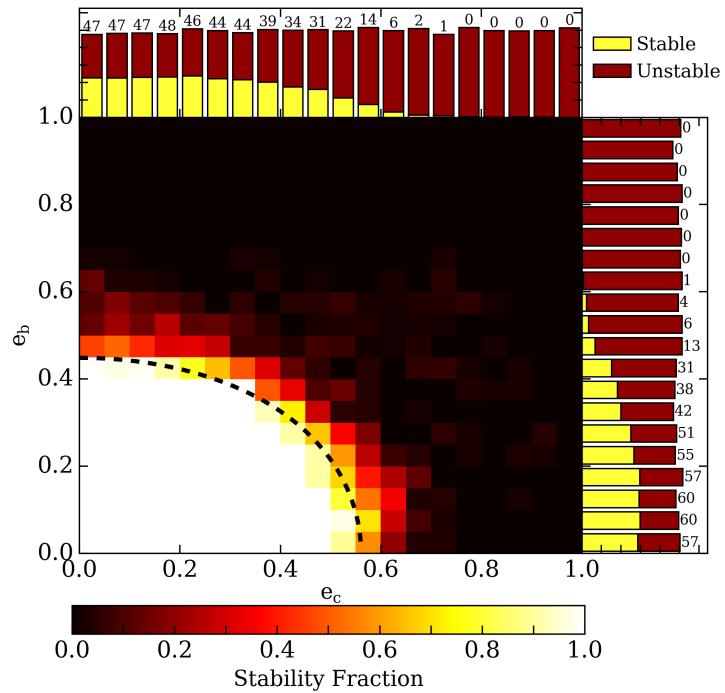


Figure 5.8: Smoothed 2D map depicting the fraction of stable systems as a function of each planet's eccentricity based on a suite of dynamical integrations. The dashed black curve depicts the analytic condition for Hill stability from Gladman (1993) assuming the MAP masses and semi-major axes from Model 1 in Table 5.2. 1D histograms depict the number of stable and unstable (yellow and red respectively) systems in eccentricity bins of width 0.05 and marginalized over all other dynamical parameters. The annotated numbers report each bin's stability fraction in percentages.

for constraining the orbital eccentricities of the K2-18 planets. We can therefore combine these independent results by using the dynamical stability fractions shown in Fig. 5.8 as an additional prior on the  $i^{\text{th}}$  planet's derived eccentricity posterior:  $e_i = h_i^2 + k_i^2$ . To do this we resample each planet's RV eccentricity posterior and accept draws with a probability equal to the stability fraction at that drawn eccentricity value  $\pm 0.025$ . This choice of bin width was varied between 0.01 and 0.1 and was found not to have a significant effect on the results. In this way numerous random samples from each planet's RV eccentricity posterior are rejected due to the low corresponding stability fraction. This is especially true for large eccentricities wherein the system no longer satisfies the Hill stability criterion. From the modified eccentricity posteriors we can calculate the 99<sup>th</sup> percentiles and find that  $e_b < 0.43$  and  $e_c < 0.47$  at that confidence level. These are the eccentricity values reported in Table 5.2 and represent a more stringent evaluation of each planet's eccentricity than considering the RV data alone.

## 5.8 Discussion

With a set of 75 precision radial velocity measurements taken with the HARPS spectrograph we have obtained a robust mass measurement of the transiting HZ super-Earth K2-18b and detected a second super-Earth K2-18c. The orbit of the newly discovered K2-18c lies interior to that of K2-18b and yet the planet is non-transiting. This implies that the orbital planes of the planets are mutually inclined. In order for K2-18c to not be seen in-transit the planetary system requires a mutual inclination of just  $\gtrsim 1.4^\circ$  which is consistent with the observed distribution of mutually inclined multi-planet systems (Figueira et al., 2012; Fabrycky et al., 2014). Dynamical simulations of the system revealed that the oscillation timescale of the planets' orbital inclinations is  $\mathcal{O}(10^6 \text{ years})$  suggesting that it may take many years before K2-18c reaches a transiting configuration. Although exactly how long depends sensitivity on its current inclination which remains unknown. The discovery of RV planets in transiting M dwarf planetary systems further emphasizes the prevalence of multiple Earth to super-Earth-sized planets around nearby M dwarfs and that these additional planets can be uncovered with moderate RV follow-up (Cloutier et al., 2017a). Multi-planet systems such as K2-18 provide unique opportunities to study planet formation processes around M dwarfs via direct comparative planetology.

The presence of a second planet in the K2-18 transiting system will result in mutual planetary interactions thus making the orbit of K2-18b non-Keplerian and possibly resulting in an observable transit timing variation (TTV). Assuming a mutual inclination of K2-18c that just misses a transiting configuration, we estimate the expected TTVs of K2-18b using the `TTVFaster` package (Deck et al., 2014; Agol & Deck, 2016). We adopted the maximum a-posteriori masses and orbital periods of the two planets from Model 1 and uniformly sample their eccentricities up the 99% upper limits reported in Table 5.2. The remaining orbital parameters of the planets that are unconstrained by the RV data are sampled uniformly between 0 and  $2\pi$ . We find a maximum TTV for K2-18b of  $\sim 40$  seconds which is slightly less than, but of the same order as the uncertainty in its measured time of mid-transit ( $\sim 50$  seconds; B17). Thus with photometric monitoring of at least comparable quality to the K2 photometry shown in Fig. 5.1, detecting TTVs in the K2-18 multi-planet system is unlikely to provide any significant new insight into the nature of the system. Indeed no significant TTVs were observed in the Spitzer light curves from B17.

With our measured mass of K2-18b the planet joins a select group of HZ planets with constraints on

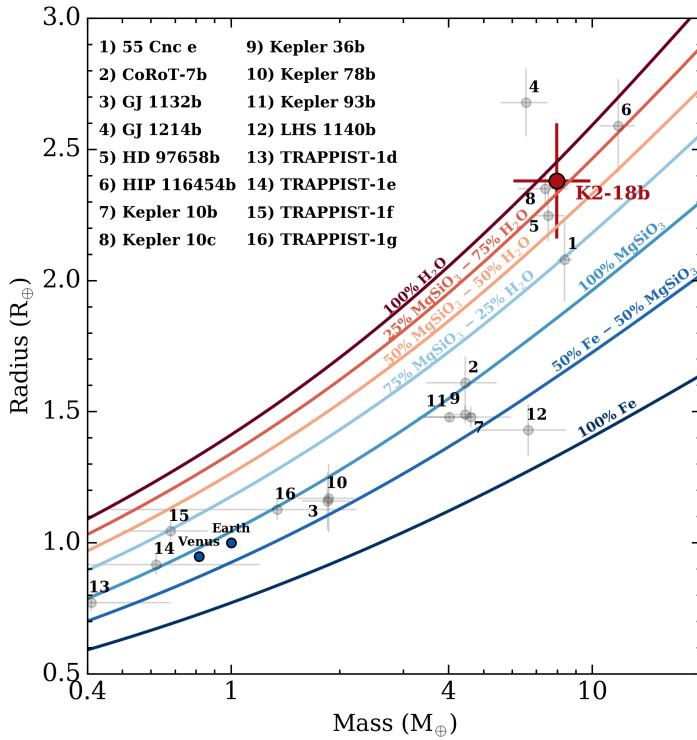


Figure 5.9: K2-18b along with a sample of other small exoplanets in the planetary mass and radius space. Overlaid curves are two-component interior structure models of fully-differentiated solid planets with mass fractions annotated for each curve.

both its mass and radius. This represents a significant step towards searching for potentially habitable planets around stars earlier than  $\sim M4$  (Dittmann et al., 2017b). With its maximum a-posteriori mass of  $7.96 \pm 1.91 M_{\oplus}$  the bulk density of K2-18b ( $\rho_{p,b} = 3.3 \pm 1.2 \text{ g cm}^{-3}$ ) lies between that of an Earth-like rocky planet and a low density Neptune-like planet. The planet is therefore likely too large to be a terrestrial Earth-like planet (Valencia et al., 2007; Fulton et al., 2017). Including K2-18b on the exoplanet mass-radius diagram in Fig. 5.9 we find that the internal structure of K2-18b is consistent with a range of two-component solid-planet models (Zeng & Sasselov, 2013) owing to the uncertainty in its measured radius and mass which is at the level of  $\sim 24\%$ . In particular, the  $1\sigma$  lower mass limit of K2-18b permits it to be a complete ‘water world’ even though its upper mass limit is consistent with a largely rocky interior surrounded by a significant mass fraction of water ice. In this parameter space the physical parameters of K2-18b are most similar to the super-Earths HD 97658b (Van Grootel et al., 2014) and Kepler 10c (Dumusque et al., 2014b; Rajpaul et al., 2017) despite receiving  $\sim 65$  and  $\sim 24$  times less insolation than those two planets respectively. Furthermore K2-18b is of a similar mass to the habitable zone planet LHS 1140b (Dittmann et al., 2017b) and receives a comparable level of insolation despite being  $\sim 1.6$  times larger than LHS 1140b. Analyzing the mass-radius relationship of these small planets over a range of equilibrium temperatures is a critical step towards understanding which of these systems have retained significant atmospheric content thus making them more suitable to extraterrestrial life.

Distinguishing between K2-18b as a pure water-world or a scaled-up Earth with a significant gaseous envelope will likely require transmission spectroscopy follow-up observations either with high-resolution

spectrographs from the ground or from space with JWST. With  $J=9.8$ ,  $H=9.1$ , and  $K=8.9$  ([Cutri et al., 2003](#)), we stress that K2-18 is currently the second brightest M dwarf with a transiting habitable zone planet behind the recently discovered LHS 1140b. In the coming years the sample of habitable zone M dwarf planets is expected to increase dramatically following the launch of TESS ([Ricker et al., 2014](#)), although the majority of TESS planets will be more distant than LHS 1140 ([Sullivan et al., 2015](#)).

Considering the prospect of observational follow-up of K2-18b in transmission spectroscopy, if we consider an atmosphere that is cloud-free and dominated by hydrogen, then spectral features from well-mixed near-IR absorbing species such as water would have amplitudes of  $\Delta F/F \sim 10Hr_p/R_s^2 \sim 230$  ppm where  $H = k_B T_{\text{eq}}/\mu g$  is the atmospheric scale height,  $T_{\text{eq}}$  is the planet's equilibrium temperature set by the stellar insolation,  $\mu$  is the mean molecular weight of the atmosphere, and  $g$  is the surface gravity ([Miller-Ricci et al., 2009](#); [Kaltenegger & Traub, 2009](#)). If instead the atmosphere is dominated by heavier elements, similar Earth's (e.g.  $\text{N}_2 + \text{O}_2$ ,  $\mu = 29$ ), then the transmission signal will be significantly smaller ( $\sim 10$  ppm) though potentially still detectable with JWST with several visits. Because of the brightness of its host star and the low bulk density of K2-18b, the system offers a unique opportunity to study super-Earth atmospheres receiving Earth-like insolation in the JWST-era.

Table 5.2: Model parameters of the K2-18 planetary system

Parameter	Maximum a-posteriori values with 16 <sup>th</sup> and 84 <sup>th</sup> percentiles			
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
<i>Stellar Parameters</i>				
2MASS Photometry	$J=9.763 \pm 0.028$ , $H=9.135 \pm 0.026$ , $K_s=8.899 \pm 0.019$			
Stellar mass, $M_s$ [ $M_\odot$ ]	$0.359 \pm 0.047$			
Stellar radius, $R_s$ [ $R_\odot$ ]	$0.411 \pm 0.038$			
Effective temperature, $T_{\text{eff}}$ [K]	$3457 \pm 39$			
Distance, $d$ [pc]	$34 \pm 4$			
Rotation period, $P_{\text{rot}}$ [days]	$38.6^{+0.6}_{-0.4}$			
Systemic velocity, $\gamma_0$ [m s <sup>-1</sup> ]	$653.7 \pm 0.9^\bullet$			
<i>GP hyperparameters</i>				
Covariance amplitude, $a$ [m s <sup>-1</sup> ]	$2.8^{+2.1}_{-1.6}$	N.S. <sup>◦</sup>	N.S.	—
Exponential timescale, $\lambda$ [days]	$59.1^{+19.1}_{-11.2}$	N.S.	N.S.	—
Coherence, $\Gamma$	$1.2^{+0.6}_{-0.3}$	N.S.	N.S.	—
Periodic timescale, $P_{\text{GP}}$ [days]	$38.6^{+1.1}_{-0.4}$	N.S.	N.S.	—
Additive jitter, $s$ [m s <sup>-1</sup> ]	$0.25^{+0.97}_{-0.06}$	$0.45^{+0.85}_{-0.23}$	$0.22^{+1.09}_{-0.04}$	—
<i>K2-18c</i>				
Period, $P_c$ [days]	$8.962 \pm 0.008$	$8.964 \pm 0.010$	$8.965 \pm 0.010$	$8.966^{+0.005}_{-0.010}$
Time of inferior conjunction, $T_{0,c}$ [BJD-2,450,000]	$7264.55 \pm 0.46$	$7264.55 \pm 0.51$	$7264.49^{+0.58}_{-0.45}$	$7264.48^{+0.61}_{-0.31}$
Radial velocity semi-amplitude, $K_c$ [m s <sup>-1</sup> ]	$4.63 \pm 0.72$	$4.74^{+0.71}_{-0.98}$	$4.52 \pm 0.82$	$4.63^{+0.82}_{-0.58}$
$h_c = \sqrt{e_c} \cos \omega_c$	$-0.37^{+0.41}_{-0.13}$	$-0.46^{+0.45}_{-0.09}$	$-0.39^{+0.51}_{-0.10}$	$-0.51^{+0.57}_{-0.05}$
$k_c = \sqrt{e_c} \sin \omega_c$	$0.05^{+0.28}_{-0.21}$	$-0.04^{+0.35}_{-0.13}$	$0.00^{+0.35}_{-0.17}$	$0.01^{+0.32}_{-0.17}$
Semi-major axis, $a_c$ [AU]	$0.060 \pm 0.003$	$0.060 \pm 0.003$	$0.060 \pm 0.003$	$0.060 \pm 0.003$
Eccentricity, $e_c$	$< 0.47^*$	—	—	—
Minimum planet mass, $m_{p,c} \sin i_c$ [ $M_\oplus$ ]	$7.51 \pm 1.33$	$7.68^{+1.33}_{-1.72}$	$7.33 \pm 1.48$	$7.51^{+1.48}_{-1.15}$
Equilibrium temperature, $T_{\text{eq},c}$ [K] Bond albedo of 0.3	$363 \pm 14$	$363 \pm 14$	$363 \pm 14$	$363 \pm 14$
<i>K2-18b</i>				
Period, $P_b$ [days]	$32.93963 \pm 1.0 \times 10^{-4}$	$32.93962 \pm 1.1 \times 10^{-4}$	$32.93961 \pm 1.0 \times 10^{-4}$	$32.93960 \pm 9.3 \times 10^{-5}$
Time of inferior conjunction, $T_{0,b}$ [BJD-2,450,000]	$7264.39157 \pm 5.9 \times 10^{-4}$	$7264.39133 \pm 6.4 \times 10^{-4}$	$7264.39155^{+0.0006}_{-0.0008}$	$7264.39135^{+0.0007}_{-0.0005}$
Radial velocity semi-amplitude, $K_b$ [m s <sup>-1</sup> ]	$3.18 \pm 0.71$	$3.29^{+0.71}_{-0.64}$	$3.26^{+0.63}_{-0.85}$	$3.25^{+0.60}_{-0.41}$
$h_b = \sqrt{e_b} \cos \omega_b$	$0.33^{+0.08}_{-0.28}$	$0.33^{+0.06}_{-0.30}$	$0.31^{+0.07}_{-0.37}$	$0.32^{+0.06}_{-0.27}$
$k_b = \sqrt{e_b} \sin \omega_b$	$-0.16^{+0.42}_{-0.21}$	$-0.16^{+0.42}_{-0.19}$	$-0.10^{+0.41}_{-0.25}$	$-0.11^{+0.31}_{-0.22}$
Semi-major axis, $a_b$ [AU]	$0.143 \pm 0.006$	$0.143 \pm 0.006$	$0.143 \pm 0.006$	$0.143 \pm 0.006$
Eccentricity, $e_b$	$< 0.43^*$	—	—	—
Planet mass, $m_{p,b}$ [ $M_\oplus$ ] <sup>†</sup>	$7.96 \pm 1.91$	$8.23^{+1.92}_{-1.76}$	$8.16^{+1.73}_{-2.24}$	$8.13^{+1.66}_{-1.25}$
Planet density, $\rho_{p,b}$ [g cm <sup>-3</sup> ] <sup>‡</sup>	$3.3 \pm 1.2$	$3.4^{+1.2}_{-1.2}$	$3.4^{+1.2}_{-1.3}$	$3.3^{+1.2}_{-1.0}$
Surface gravity, $g$ [m s <sup>-2</sup> ] <sup>‡</sup>	$14.0 \pm 4.3$	$14.5^{+4.3}_{-4.1}$	$14.4^{+4.1}_{-4.8}$	$14.3^{+4.0}_{-3.5}$
Escape velocity, $v_{\text{esc},b}$ [km s <sup>-1</sup> ] <sup>‡</sup>	$20.6 \pm 2.7$	$20.9^{+2.6}_{-2.4}$	$20.9^{+2.4}_{-3.0}$	$20.8^{+2.3}_{-1.9}$
Equilibrium temperature, $T_{\text{eq},b}$ [K] Bond albedo of 0.3	$235 \pm 9$	$235 \pm 9$	$235 \pm 9$	$235 \pm 9$
<i>Model diagnostics</i>				
Median ln $\mathcal{L}$ from Cross-Validation $\ln \mathcal{L}_1 - \ln \mathcal{L}_i$	$-1.462 \pm 0.013$ 0	$-1.574 \pm 0.017$ $0.112 \pm 0.021$	$-1.600 \pm 0.018$ $0.138 \pm 0.022$	$-1.548 \pm 0.017$ $0.086 \pm 0.021$

**Notes.** (•) from the Model 1 RV modelling.

(◦) N.S. stands for ‘no solution’ and occurs when the parameter’s marginalized posterior PDF is unconstrained by the data.

(\*) upper limit based on the 99% confidence interval from the RV data analysis and conditioned on the dynamical stability constraints from Sect. 5.7.

(†) assuming the measured orbital inclination of K2-18b from B17;  $i_b = 89.5785^{+0.0079}_{-0.0088}$  degrees.

(‡) assuming the measured radius of K2-18b from B17;  $r_{p,b} = 2.38 \pm 0.22 R_\oplus$ .

$M_\odot = 1.988499 \times 10^{30}$  kg,  $R_\odot = 6.955 \times 10^8$  m,  $M_\oplus = 6.045898 \times 10^{24}$  kg,  $R_\oplus = 6.378137 \times 10^6$  m.

Table 5.3: HARPS time series

BJD-2,450,000	RV [m s <sup>-1</sup> ]	$\sigma$ RV [m s <sup>-1</sup> ]	S index	H $\alpha$	FWHM	BIS
7117.565870	658.15	4.30	0.551	0.06337	3.050	9.397
7146.526948	654.90	2.94	0.851	0.06646	3.068	-3.008
7146.646070	660.66	3.98	0.537	0.06850	3.079	-6.150
7148.518851	649.21	4.62	0.973	0.06667	3.071	8.583
7199.503915	656.11	3.43	0.558	0.06597	3.090	16.922
7200.503114	656.48	2.79	0.290	0.06625	3.080	19.343
7204.491167	648.52	4.38	0.368	0.06409	3.076	-10.068
7390.845075	655.54	2.65	1.040	0.06721	3.106	-0.196
7401.779223	649.19	2.66	0.960	0.06640	3.105	1.064
7403.826871	651.38	2.87	1.287	0.06598	3.106	5.332
7404.814521	654.49	3.78	1.410	0.06679	3.095	6.633
7405.789149	655.54	2.71	1.066	0.06644	3.106	7.934
7407.773473	652.66	4.36	-	0.06680	3.048	0.311
7410.791609	651.99	4.35	1.296	0.07166	3.114	-2.344
7412.810195	661.94	3.50	0.677	0.06618	3.094	-8.771
7417.787334	649.38	3.41	0.724	0.06615	3.092	7.447
7418.799229	648.38	3.67	1.402	0.06842	3.100	2.274
7420.791577	658.68	2.54	1.170	0.06738	3.099	11.285
7421.794046	658.62	2.30	1.509	0.07463	3.099	11.310
7422.781258	657.47	2.66	0.986	0.06505	3.104	7.032
7424.777426	658.79	2.86	0.924	0.06642	3.104	5.314
7425.850669	652.07	2.81	0.871	0.06636	3.088	-0.471
7446.704487	652.60	2.67	0.933	0.06481	3.108	1.511
7447.830725	659.17	3.87	0.862	0.06539	3.085	0.282
7448.686909	656.59	3.26	0.713	0.06470	3.098	0.282
7450.675147	660.91	3.33	0.917	0.06521	3.102	-8.218
7451.677499	655.48	2.70	1.063	0.06575	3.111	-5.083
7452.695705	663.48	2.75	1.112	0.06532	3.093	1.667
7453.701988	655.18	2.38	1.038	0.06689	3.104	10.356
7456.704230	658.87	3.31	0.887	0.06586	3.111	-17.835
7457.683261	658.28	3.81	1.010	0.06615	3.101	12.792
7458.660021	656.62	3.47	1.404	0.06634	3.099	9.160
7472.784787	641.61	3.12	1.058	0.06653	3.094	-1.001
7473.684129	648.26	3.08	0.908	0.06529	3.096	-1.293
7474.737446	645.84	3.90	0.783	0.06621	3.085	-6.991
7475.698658	655.63	3.94	0.331	0.06647	3.094	-0.455
7476.707703	652.59	3.26	0.983	0.06517	3.091	-0.985
7477.674398	657.56	3.37	0.816	0.06554	3.100	-4.990
7478.631994	652.31	4.64	0.682	0.06713	3.093	-9.857
7479.737617	649.94	3.17	1.116	0.06541	3.086	-9.996
7486.661319	666.10	3.24	1.168	0.06527	3.096	-10.681
7487.617699	667.97	4.55	0.573	0.06504	3.091	-0.986
7488.670507	656.68	3.56	0.922	0.06431	3.086	-8.201
7567.516862	659.76	4.00	0.993	0.06530	3.096	10.014
7576.473152	657.32	4.27	0.989	0.06587	3.115	7.131
7584.477527	657.94	5.55	0.725	0.06563	3.080	16.445
7786.842858	645.83	3.49	1.146	0.06430	3.093	-1.668
7787.825672	658.25	4.66	0.944	0.06581	3.092	5.481
7790.828228	655.56	3.49	1.092	0.06500	3.088	1.878
7791.843445	653.87	3.70	0.700	0.06414	3.093	-3.719
7792.815105	653.38	3.47	0.893	0.06861	3.083	6.408
7801.827514	652.57	3.18	0.812	0.06518	3.094	12.389
7802.790293	648.98	3.14	0.860	0.06504	3.094	-3.525
7803.809311	649.61	3.27	0.926	0.06616	3.091	18.332
7810.806284	654.41	4.71	0.905	0.06593	3.098	6.496
7814.760772	652.53	4.35	-	0.06632	3.098	-0.563
7815.759421	656.22	5.06	-	0.06536	3.086	15.990
7817.748614	665.63	2.89	0.998	0.06423	3.103	7.162
7830.668729	647.15	3.18	0.866	0.06476	3.084	-1.992
7832.659387	647.89	4.31	1.029	0.06292	3.079	14.048
7834.636450	656.64	4.31	1.025	0.06407	3.091	9.870
7835.596293	654.97	4.36	1.004	0.06308	3.093	11.521
7836.626075	656.46	4.71	0.792	0.06491	3.093	-2.004
7839.650934	644.07	3.74	0.774	0.06437	3.076	5.253
7841.638147	648.49	3.92	0.810	0.06487	3.089	0.543
7843.648062	654.65	3.92	0.408	0.06404	3.090	12.825
7844.626814	659.07	3.85	0.722	0.06417	3.087	1.584
7846.692642	655.96	3.75	0.923	0.06435	3.086	5.210
7847.693240	651.46	3.77	0.903	0.06363	3.094	8.538
7848.677241	655.24	3.82	0.848	0.06445	3.102	1.280
7849.696944	653.92	3.34	0.853	0.06440	3.082	9.616
7872.656844	653.72	3.92	0.913	0.06371	3.082	-5.325
7873.525484	652.07	4.95	0.813	0.06478	3.094	13.489
7874.671695	650.86	4.80	0.886	0.06543	3.088	1.857
7875.596914	643.77	3.69	0.913	0.06467	3.077	5.739

# Chapter 6

## Confirmation of the radial velocity super-Earth K2-18c with HARPS and CARMENES

### 6.1 Introduction

The nearby M2.5 dwarf K2-18 (EPIC 201912552,  $d \sim 38$  pc,  $J = 9.8$ ) is known to host a transiting sub-Neptune-sized planet at  $\sim 33$  days; K2-18b ([Foreman-Mackey, 2015](#); [Montet et al., 2015](#); [Benneke et al., 2017](#)). Given the planet's orbital separation and corresponding equilibrium temperature, K2-18b is a temperate planet and represents one of the most attractive targets for the atmospheric characterization of a habitable zone exoplanet that was discovered in the pre-TESS era. Indeed K2-18b is already slated for transmission spectroscopy observations as part of the NIRISS GTO program 1201<sup>1</sup>.

Given the requirement for a priori knowledge of a planet's bulk density in order to interpret observations of its atmosphere, multiple groups have endeavoured to measure the mass of K2-18b via ground-based radial velocity (RV) measurements in the visible wavelength domain. Specifically, [Cloutier et al. \(2017b\)](#) (hereafter [C17a](#)) first reported the mass of K2-18b to be  $8.0 \pm 1.9$   $M_{\oplus}$  based on 75 measurements taken with the HARPS spectrograph on the ESO 3.6m telescope at La Silla ([Mayor et al., 2003](#)). Their RV timeseries also exhibited a strong additional signal at  $\sim 9$  days which was not seen in any other contemporaneous activity indicators<sup>2</sup> or in the window function. [C17a](#) presented evidence for the planetary nature of the 9-day signal by simultaneously modelling both planetary signals with Keplerians and the correlated RV residuals using a trained quasi-periodic Gaussian process (GP). Correlated RV residuals, after the removal of planetary signals, are expected to arise from stellar activity whose components can be seen in various activity indicators such as photometry and the aforementioned

---

The contents of this chapter are copied verbatim from the published paper ([Cloutier et al., 2019b](#)).

<sup>1</sup><http://www.stsci.edu/cgi-bin/get-proposal-info?id=1201&observatory=JWST>

<sup>2</sup>For example, the S-index, H $\alpha$  index, full width at half maximum, and the bi-sector inverse slope of the spectral cross-correlation function

spectroscopic indicators. All of these ancillary timeseries were used for training in the multiple analyses presented in C17a. Stellar activity on M dwarfs is largely modulated by stellar rotation (Boisse et al., 2011) and thus produces a quasi-periodic structure in the RVs that is physically motivated. Such a correlated structure is often not strictly sinusoidal as the active regions that give rise to the observed stellar activity have finite lifetimes, spatial distributions, and temperature contrasts that evolve temporally over a few rotation cycles and thus lead to non-sinusoidal structures over the observational baseline.

Recently, Sarkis et al. (2018) (hereafter S18) presented an independent set of 58 RV measurements of K2-18 taken with the visible channel on CARMENES (561–905 nm; Quirrenbach et al., 2014). With these data S18 independently measured the mass of K2-18b to be  $8.9_{-1.6}^{+1.7} M_{\oplus}$ , a result that is consistent with the measured value from C17a. However in their data, with comparable RV precision, the 9-day signal with its proposed planetary origin from C17a was only marginally detected. Furthermore, S18 claimed that the signal was seen to vary in time and that the strength (as measured by the false alarm probability in the generalized Lomb-Scargle periodogram) appeared to vary with wavelength. Given the proximity of the 9-day signal to the fourth harmonic of the photometric stellar rotation period<sup>3</sup> ( $P_{\text{rot}} = 38.6$  days; C17a), S18 interpreted the weak 9-day signal as one whose origin is more likely due to stellar activity than to a second, non-transiting planet in the system.

Based on the strong evidence for the detection of K2-18c with HARPS<sup>4</sup> and the low significance of its periodic signal being seen with CARMENES, here we conduct a systematic re-analysis of all available RV data to confirm or disprove the existence of a stable periodic signal at  $\sim 9$  days in the K2-18 system and ultimately to determine the nature of that signal as planetary or otherwise. In this study we independently analyse the HARPS and CARMENES RV timeseries and their joint timeseries. We include 31 previously unpublished HARPS RVs that aid in the interpretation of the 9-day signal and improve the measurement precision of the planetary parameters. In Sect. 6.2 we present a detailed analysis investigating the effects of time-sampling on the probability of the 9-day signal. In Sects. 6.3 and 6.4 we investigate the proposed chromatic and temporal dependencies of the 9-day signal with HARPS. In Sect. 6.5 we self-consistently analyse all RVs in the presence of a probabilistic correlated noise (i.e. activity) model. Overall we find evidence for the planetary nature of the 9-day signal and conclude with a discussion in Sect. 6.6.

## 6.2 Sub-optimal window functions

One potential reason why the strong 9-day signal was seen in the published HARPS RVs and not with CARMENES may be due to sub-optimal time-sampling (i.e. the window function; WF). For example, the 9-day signal seen with HARPS may arise from a sub-optimal WF and is therefore not associated with an astrophysical source such as a planet or stellar activity. Similarly, if the 9-day signal exists, and if its origin is physical, then it is possible that the CARMENES WF may suppress its signal in a Lomb-Scargle periodogram. Sub-optimal WFs have indeed been shown to lead to inaccurate RV planet

---

<sup>3</sup>Although periodicities at the second and third harmonics are not seen in the CARMENES RVs with comparable significance to that of the 9-day signal.

<sup>4</sup>A strong periodic signal in the periodogram of the HARPS RVs at  $\sim 9$  days, a  $6.3\sigma$  semi-amplitude measurement, the favourability of a two-planet model by cross-validation model comparison (C17a).

masses and false planet detections (e.g. GL 581d; [Hatzes 2016](#),  $\alpha$  Cen Bb; [Rajpaul et al. 2016](#), Kepler-10c; [Rajpaul et al. 2017](#)). Before proceeding we note that neither of these scenarios is expected to significantly enhance or suppress the 9-day signal as investigated by preliminary analyses in [C17a](#) and [S18](#). However, a more subtle effect may be at play here. Specifically, the periodogram of the HARPS WF showed no excess power at 9 days (see Fig. 2 [C17a](#)) such that the signal is unlikely to originate from sub-optimal HARPS sampling. Similarly, [S18](#) created a synthetic RV timeseries with the maximum a posteriori (MAP) solution for K2-18c from [C17a](#), plus white noise, and sampled the Keplerian curve with synthetic RVs using the CARMENES WF. They reported that the  $\sim 9$ -day signal was seen in the periodogram and thus was not suppressed by the CARMENES WF. Here we extend these analyses to establish definitively whether either published WF is responsible for the ambiguity of the  $\sim 9$ -day signal.

### 6.2.1 Detecting the 9-day signal in synthetic RV timeseries

Here we aim to establish the ease with which the K2-18c signal at  $\sim 9$  days can be detected in any of the published HARPS, CARMENES, or joint WFs. Firstly, for each of the three WFs we construct a set of synthetic RV timeseries containing a variety of injected physical signals, plus a white noise term with standard deviation equal to the mean RV measurement precision of that timeseries.<sup>5</sup> We consider four flavours of injected physical signals of increasing complexity: i) K2-18c only, ii) K2-18b and c, iii) both planets plus correlated noise due to stellar activity, and iv) K2-18b and stellar activity. The last timeseries, which does not contain an injected K2-18c signal, is included to test the hypothesis that the  $P_c$  signal could arise from sampling or stellar activity, without K2-18c existing at  $P_c$ , as posited by [S18](#). The test with K2-18b and c only using the CARMENES WF corresponds to the test performed by [S18](#), which showed that  $P_c$  is detected when the MAP value of the K2-18c semi-amplitude  $K_c = 4.63 \text{ m s}^{-1}$  from [C17a](#) was injected. In our analysis, the Keplerian model parameters for each planet are fixed to their average value between the [C17a](#) and [S18](#) results—where applicable—with the exception of  $K_c$ , which is sampled on a logarithmically equidistant grid from 1 to  $10 \text{ m s}^{-1}$ . When including correlated noise models, those models are sampled from a quasi-periodic GP prior distribution which has been shown to be an effective means of describing quasi-periodic stellar activity signals in both Sun-like and M dwarf stars (e.g. [Haywood et al., 2014](#); [Cloutier et al., 2017a](#)). The adopted hyperparameters are given by those measured in Model 1 from [C17a](#) and includes a covariance amplitude of  $2.8 \text{ m s}^{-1}$ . These hyperparameters describe the covariance structure of the stellar activity signal as seen in the star’s K2 photometry and the HARPS RVs.

For each synthetic RV timeseries we compute the Bayesian generalized Lomb-Scargle periodogram (GLSP; [Mortier et al., 2015](#)) from which we isolate the probability of a sinusoidal function with the period of K2-18c ( $P_c = 8.962$  days) being present in our synthetic timeseries;  $p(P_c|\text{RV})$ . The left column of Fig. 6.1 depicts  $p(P_c|\text{RV})$  as a function of the injected K2-18c semi-amplitude for three out of the four RV models. The synthetic timeseries containing K2-18b and stellar activity are not included in Fig. 6.1 as they were consistently seen to result in  $p(P_c|\text{RV}) \ll 1\%$ , thus indicating that the  $P_c$  did not arise with any significance when not explicitly added to the timeseries. The ordinate values in Fig. 6.1 are the median probabilities derived from a set of 50 synthetic timeseries realizations per value of the injected  $K_c$ . In this way, we marginalize over the exact form of the injected white and correlated noise sources which are sampled randomly in each of the 50 iterations. As expected, because the  $P_c$  periodic

---

<sup>5</sup>3.60, 3.08, and  $3.37 \text{ m s}^{-1}$  for HARPS, CARMENES, and their joint timeseries respectively.

signal is injected into each synthetic timeseries, the probability of that signal existing within the data increases with the  $K_c$  from zero probability when  $K_c \sim 1 \text{ m s}^{-1}$  towards  $p(P_c|\text{RV}) = 100\%$  as  $K_c \rightarrow 10 \text{ m s}^{-1}$  for any of the three types of synthetic timeseries. It is true that as the complexity of the synthetic timeseries increases (i.e. as more signals are added) the semi-amplitude  $K_c$  needs to be larger in order to be detected with high probability. It is also clear that detecting the injected  $P_c$  signal is easier with either the HARPS or joint WFs as their probability curves tend to increase more rapidly with  $K_c$  and they approach 100% probability at a lower  $K_c$  than with the CARMENES WF alone. This is particularly true at the MAP value of  $K_c = 4.63 \text{ m s}^{-1}$  (C17a) wherein  $p(P_c|\text{RV})$  is  $\sim 40\%$  larger with the HARPS WF than with CARMENES for any of the synthetic timeseries. This shows that with the CARMENES time-sampling the strength of the  $P_c$  periodic signal is less prominent in the GLSP than with the HARPS or joint time-sampling. With any of the three types of synthetic timeseries, the strength of  $P_c$  is typically lower with CARMENES until  $K_c \sim 10 \text{ m s}^{-1}$  wherein the probability of  $P_c$  with CARMENES becomes consistent with 100%. However, an injected value of  $K_c = 10 \text{ m s}^{-1}$  is inconsistent with the C17a measured value at  $\gtrsim 7\sigma$ .

The systematically lower  $P_c$  probability with CARMENES may be due to sampling, instrumental effects, or to the fact that the CARMENES WF contains fewer RVs: 58 compared to 75 with HARPS. The smaller WF affects the sampling of periodic signals and  $P_c$  may not be strongly detectable with only 58 RVs. To investigate this possibility, we again compute  $p(P_c|\text{RV})$  in our synthetic RV timeseries, but for random subsets of each timeseries and with an increasing number of RV measurements  $N_{\text{RV}} \in [10, N_f]$  where  $N_f$  is the full size of each RV timeseries.<sup>6</sup> When creating these synthetic times-series,  $K_c$  is fixed to its MAP value of  $4.63 \text{ m s}^{-1}$ . The smoothed probability curves for each synthetic timeseries and each WF are shown in the right column of Fig. 6.1. The curves are smoothed to remove the high-frequency noise and make the trends in the curves easier to parse visually. As can be seen in the probability of  $P_c$  as a function of  $K_c$ , when  $K_c$  equals its MAP value, the  $P_c$  signal is detected at a higher probability with the HARPS or HARPS+CARMENES WFs than with CARMENES alone. Here we focus on the probability of  $P_c$  when the HARPS and CARMENES timeseries contain the same number of measurements. When both timeseries are equal to the size of the full CARMENES WF (i.e.  $N_{\text{RV}} = 58$ ), the probability of detecting  $P_c$  is always lowest with the CARMENES WF than with any subset of 58 measurements with either the HARPS or joint WFs. For the most realistic set of synthetic RVs featuring two planets + a stellar activity model, the discrepancy in  $p(P_c-\text{RV})$  is modest with HARPS being  $\sim 29\%$  greater than with CARMENES and their joint WF being  $\sim 63\%$  greater.

Overall we see that the probability of the  $P_c$  periodic signal existing in timeseries with the sampling of HARPS, CARMENES, or their joint timeseries, is systematically lowest with CARMENES. By the nature of this experiment we conclude that the sole reason for the lower CARMENES probability is due to its WF. Although this discrepancy hints at why  $P_c$  may not have been detected in the GLSP of the CARMENES RVs, the relative values of  $p(P_c-\text{RV})$  to surrounding periodicities in these synthetic RVs is considered high and is certainly sufficient to detect  $P_c$ . Next we show that a small subset of anomalous CARMENES observations are likely responsible for the suppression of the  $P_c$  periodic signal in the GLSP.

---

<sup>6</sup> $N_f = 75, 58, \text{ and } 133$  for HARPS, CARMENES, and their joint timeseries, respectively.

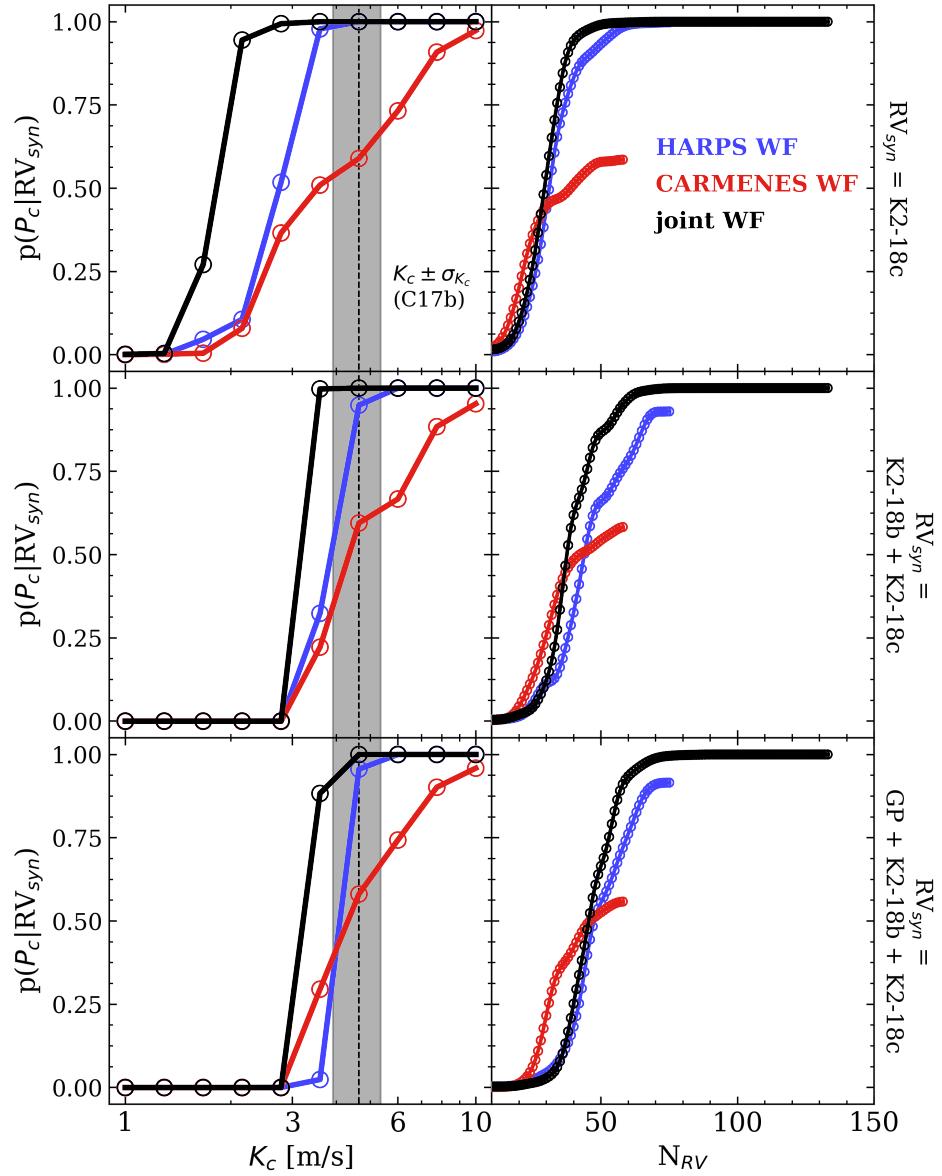


Figure 6.1: *Left column:* Probability of the injected periodic signal at  $P_c = 8.962$  days existing in synthetic RV timeseries, as a function of the injected semi-amplitude  $K_c$  with time-sampling identical to the published HARPS WF (C17a), the published CARMENES WF (S18), or their joint WF. Three sets of synthetic RV timeseries are considered and contain K2-18c only (top row), K2-18b and c (middle row), or both planets plus a GP correlated noise model of stellar activity (bottom row). The shaded vertical region highlights the MAP and  $1\sigma$  measured value of  $K_c = 4.63 \pm 0.72$  m s<sup>-1</sup> from C17a. *Right column:* Probability of the injected periodic signal at  $P_c$  existing in synthetic RV timeseries, with fixed  $K_c = 4.63$  m s<sup>-1</sup>, as a function of the number of RV measurements.

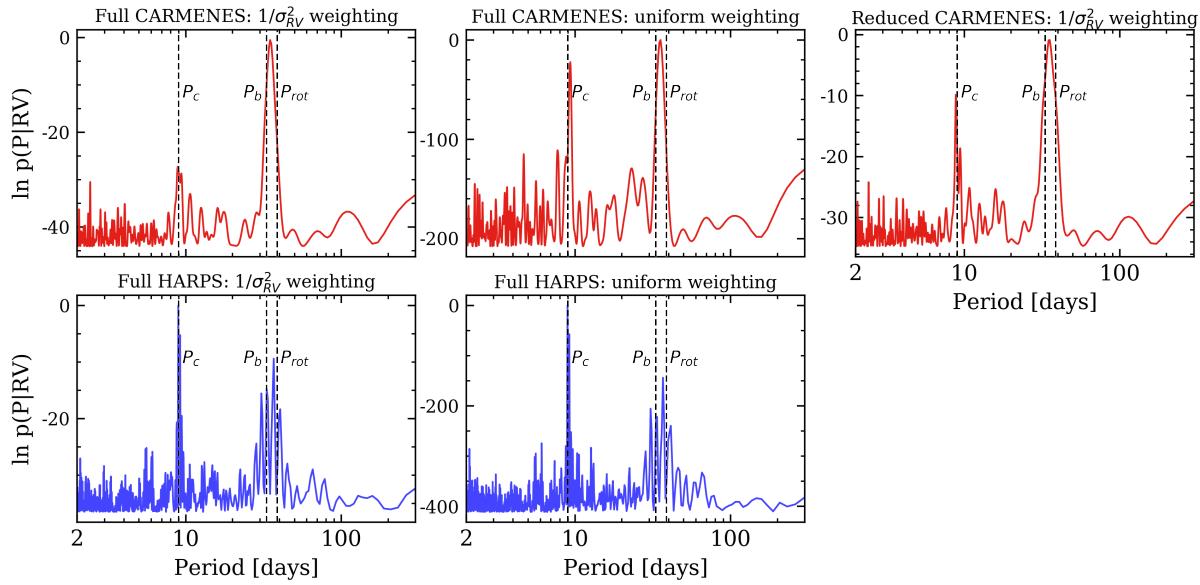


Figure 6.2: Bayesian generalized Lomb-Scargle periodograms for various subsets of the published CARMENES and HARPS RVs with one of a pair of possible weighting schemes. The details of the timeseries shown in each panel are annotated above the panel. The three [dashed vertical lines](#) depict the orbital period of the proposed non-transiting planet K2-18c ( $P_c = 8.962$  days), the orbital period of the known transiting planet K2-18b ( $P_b = 32.93963$  days), and the photometric stellar rotation period ( $P_{\text{rot}} = 38.6$  days). The  $P_c$  signal posited to be due to a second, non-transiting planet is seen at high relative probability in all but the full CARMENES RV timeseries from S18 with a  $1/\sigma_{\text{RV}}^2$  weighting.

### 6.2.2 Identifying anomalous CARMENES observations

We recall that the periodic signal from the proposed planet K2-18c at  $P_c = 8.962$  days was not seen with a low false alarm probability in the GLSP of the full CARMENES timeseries (S18). This is confirmed in the first panel of Fig. 6.2 although a small (albeit non-significant) hint of the  $\sim 9$ -day signal is visible. In computing the GLSP the CARMENES RVs are weighted by the inverse square of their respective measurement uncertainties. As a brief experiment, we considered what the effect of adopting a uniform weighting on each RV (i.e. unweighted) would have on the probability of the 9-day signal. As can be seen in the second panel of Fig. 6.2, the 9-day signal becomes much more significant when using a uniform weighting. For comparison, the probability of the 9-day signal in the HARPS GLSP varies only weakly between the weighted and unweighted conventions (see bottom row of Fig. 6.2). This suggests that the 9-day signal does exist within the CARMENES RV dataset despite only appearing with significance when using an unconventional—and incorrect—method of computing the GLSP.

The sudden appearance of the 9-day periodic signal in the CARMENES RV suggests that some anomalous measurements may be partially responsible for the signal’s suppression to the extent that it becomes buried in the noise of the full CARMENES GLSP. If the number of such anomalous measurements is small compared to the full size of the dataset, then we can justify the removal of those measurements to measure the 9-day signal with CARMENES given our strong prior evidence for the signal from HARPS (C17a). We proceed by calculating the probability of  $P_c$  existing within various subsets of the full CARMENES timeseries via leave-one-out cross-validation. In each of the 58 considered sub-

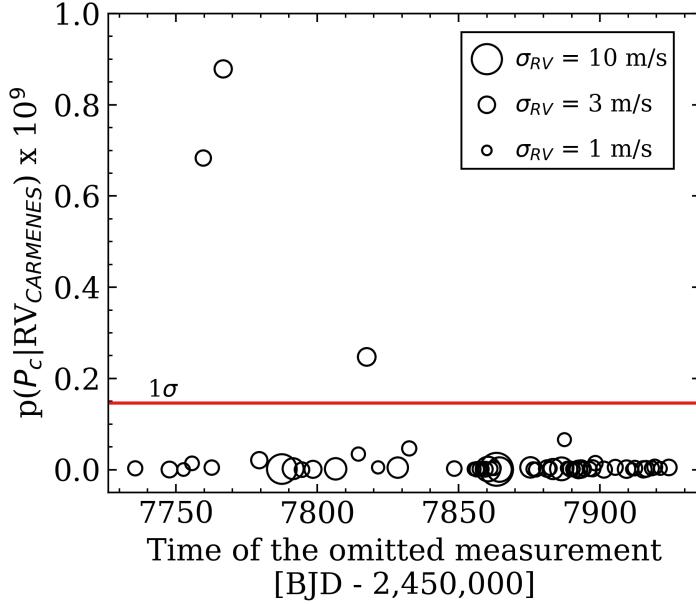


Figure 6.3: Probability of the proposed K2-18c periodic signal  $P_c = 8.962$  days existing within the CARMENES RV dataset from S18 but with a single measurement omitted via leave-one-out cross-validation. The abscissa depicts the observation epoch of each omitted RV measurement. The *solid horizontal line* depicts the  $1\sigma$  dispersion of the probabilities. The three measurements which lie above the  $1\sigma$  line significantly suppress the probability of  $P_c$  and are henceforth treated as anomalous.

sets, we omit a single unique measurement, compute the GLSP of the remaining 57 RVs, and isolate the probability of  $P_c$  existing within the data using an identical method to what was used in Sect. 6.2. The resulting probabilities of  $P_c$  as a function of the epoch of the omitted measurement are shown in Fig. 6.3.

In Fig. 6.3 we identify three anomalous RVs via a visual  $\sigma$ -clip<sup>7</sup>. We note that we refer to these measurements as anomalous as their inclusion versus their omission clearly results in a significant reduction in  $p(P_c | \text{RV})$  which is not seen for the majority of the CARMENES RVs. These measurements have associated RV uncertainties that are comparable to the mean CARMENES RV measurement uncertainty and thus have a significant effect on the probabilities of the periodicities sampled in the GLSP. The removal of these three anomalous measurements and the recalculation of the GLSP using the proper RV weighting is shown in the third panel of Fig. 6.2. The 9-day periodic signal is now clearly seen at high probability. Clearly the strategic removal of just 3 out of 58 CARMENES RVs enhances the  $P_c$  periodic signal. Thus, we have significant preliminary evidence for the existence of the proposed planet K2-18c at  $\sim 9$  days from the GLSP of the remaining 55 CARMENES RVs.

S18 provided their contemporaneous spectroscopic timeseries of the CARMENES ‘full’, blue, and red RVs, as well as timeseries of the chromospheric H $\alpha$  index and the three CaII infrared triplet line indices. Inspection of these timeseries does not reveal any obvious reason for why the three measurements identified in Fig. 6.3 significantly suppress the 9-day signal. We shared this result among the CARMENES team members who were also unable to identify any potential causes of the anomalous nature of these

<sup>7</sup>For RV indices starting at 0, the three anomalous CARMENES RVs have indices 4, 6, and 14 (i.e. BJD-2,450,000 = 7759.69656, 7766.73773, 7817.51320).

measurements after inspecting the measured RVs in individual orders. Therefore, at this time we are unable to explain the cause of the anomalous nature of these three measurements.

An exercise similar to that shown in Fig. 6.3 was also conducted using the HARPS RVs, the results of which are not presented here because the removal of individual HARPS RVs did not result in any significant changes to the probability of  $P_c$  existing within the reduced dataset; i.e. all values of  $p(P_c - \text{RV}_{\text{HARPS}})$  were close to 100% with a small rms of  $\sim 6\%$ . The discrepancy between HARPS and CARMENES in this regard may be because the 9-day signal is less suppressed by the HARPS WF compared to the CARMENES WF (see Fig. 6.1) or because the HARPS WF contains more measurements and is thus less sensitive to the removal of individual measurements. The latter scenario highlights the need to obtain large  $N_{\text{RV}}$  when searching for small planets whose RV semi-amplitudes are comparable to the RV measurement precision. This result has also been noted in simulations of ‘blind’ RV searches (e.g. Cloutier et al., 2018b) that strongly advocate for ‘more RVs per star’ rather than ‘more stars with fewer RVs per star’ in order to maximize future discoveries of small RV planets.

### 6.3 Chromatic dependence of the 9-day signal with HARPS

In addition to the RV variations derived from the 42 CARMENES visible orders, S18 also derived RVs from the first and second halves of these orders spanning 561–689 and 697–905 nm, respectively. Signal variations between these *blue* and *red* RVs may suggest that the nature of these signals as arising from stellar activity or from achromatic dynamical influences from planetary companions. Fluctuations in the strength of the 9-day signal in the CARMENES RVs helped lead S18 to conclude that the signal is due to stellar activity because of its apparent wavelength dependence. However, this evidence does not rule out the possibility that instead the 9-day signal is planetary in nature and appears to vary between the blue and red RVs because its suppression by activity is chromatically variable.

Similarly to the method used by S18, here we compute the chromatic HARPS RVs to investigate the dependence of the 9-day signal strength with wavelength. The method used to derive these RVs at each observation epoch is detailed in Sect. 2.1 of C17a and is based on the methodology from Astudillo-Defru et al. (2015). The HARPS RVs are re-derived in each of the 72 HARPS orders although we restrict our analysis to orders redder than 498 nm where the signal-to-noise ratio (S/N) per spectral order is sufficient to reach a  $\sigma_{\text{RV}}$  per order  $\lesssim 30 \text{ m s}^{-1}$ . The RVs derived from the remaining 34 orders are then grouped into blue and red orders whose weighted mean is used to compute the blue and red HARPS RVs. Our chromatic HARPS RVs span uneven wavelength ranges of 498–594 nm and 618–688 nm such that the resulting median RV measurement precision of  $\sim 7 \text{ m s}^{-1}$  is comparable between the two sets of RVs. We note that the wavelength domain spanned by the red HARPS RVs is approximately equal to the redder half of the blue CARMENES wavelength domain.

The GLSPs of the blue and red HARPS RVs are shown in Fig. 6.4. In both GLSPs the  $\sim 9$ -day signal is discernible along with the forest of peaks around  $P_b$  and the stellar rotation period due to aliasing from the HARPS WF (cf. Fig. 2 in C17a). Most notably, the probability of the 9-day peak is significantly greater in the HARPS red RVs than in the blue. This is expected if the 9-day signal is indeed due to a planet whose signal strength is achromatic, whereas stellar activity arising from the temperature contrast

of active regions is expected to increase bluewards (Reiners et al., 2010), thus degrading the S/N of the planetary signal in the blue RVs relative to the red. As such, if the 9-day signal originated from stellar activity rather than from a planet, one would expect the 9-day periodic signal to be stronger in the blue RVs, which it is not. Instead the rms of the blue RVs is slightly greater than in the red ( $7.8 \text{ m s}^{-1}$  compared to  $6.9 \text{ m s}^{-1}$ ) even though each set of chromatic RVs has a comparable S/N. We note that this excess dispersion in the blue HARPS RVs is only marginal given the star's moderate activity level ( $\sim 2.7 \text{ m s}^{-1}$ ; C17a; S18) which is less than RV measurement precision in either the blue or red HARPS RVs ( $\sim 7 \text{ m s}^{-1}$ ). The stronger activity level seen in the blue is likely responsible for the decreased significance of the 9-day signal and the enhanced probability at the stellar rotation period compared to the red.

Furthermore, we include the GLSP of the blue minus red RVs (see Fig. 6.4). The 9-day signal is significantly suppressed, whereas some residual probability close to the stellar rotation period persists along with some residual probability near  $P_b$  due to the aliasing of  $P_{\text{rot}}$  by the HARPS WF. The suppression of the 9-day signal in the differential RVs is indicative of its achromatic nature (i.e. a dynamical signal), whereas the differing signal strength of RV activity in the blue and red RVs results in some residual power close to  $P_{\text{rot}}$ . This further supports the planetary interpretation of the 9-day signal.

## 6.4 Temporal dependence of the 9-day signal with HARPS

In addition to the proposed chromatic dependence of the 9-day signal, S18 addressed the possibility that the 9-day signal strength also varies with time. This was posited based on the increased strength of the 9-day peak in the GLSP of the second half of the CARMENES RVs compared to the first. However, as was shown in Sect. 6.2.2, three anomalous CARMENES RVs exist in the first half of the CARMENES WF that significantly suppress the 9-day signal in the GLSP. This naturally explains why a stark increase in the 9-day signal strength was seen in the latter half of the CARMENES WF, and shows why it is not due to temporal variability in the stellar activity.

To further investigate the dependence of the 9-day signal on activity with HARPS, we can consider HARPS activity indices and the probability of the 9-day signal in each HARPS observing season separately. To extend the investigation of the temporal dependence of the 9-day signal we obtained 31 additional HARPS spectra of K2-18 (i.e. in addition to the 75 presented in C17a). These new spectra extend the full HARPS baseline from April 2015 (BJD=2457117.5) to July 2018 (BJD=2458307.5). The method used to derive the stellar RVs at each observation epoch is detailed in Sect. 2.1 of C17a. The full set of 106 HARPS RVs are provided in Table 6.4.

The full HARPS timeseries is spanned by three separate observing seasons containing  $N_{\text{RV}} \geq 22$ . The GLSPs of the HARPS RVs in each observing season are shown in Fig. 6.5. Although the 9-day signal is visible in each GLSP, its probability relative to the surrounding continuum is seen to increase with time from early 2016 to mid-2018. If the 9-day signal is planetary in nature rather than due to stellar activity, then we would expect the K2-18 activity level to decrease with time thus enhancing the 9-day signal in the GLSP as the activity level subsides. Next we show that this is indeed the case.

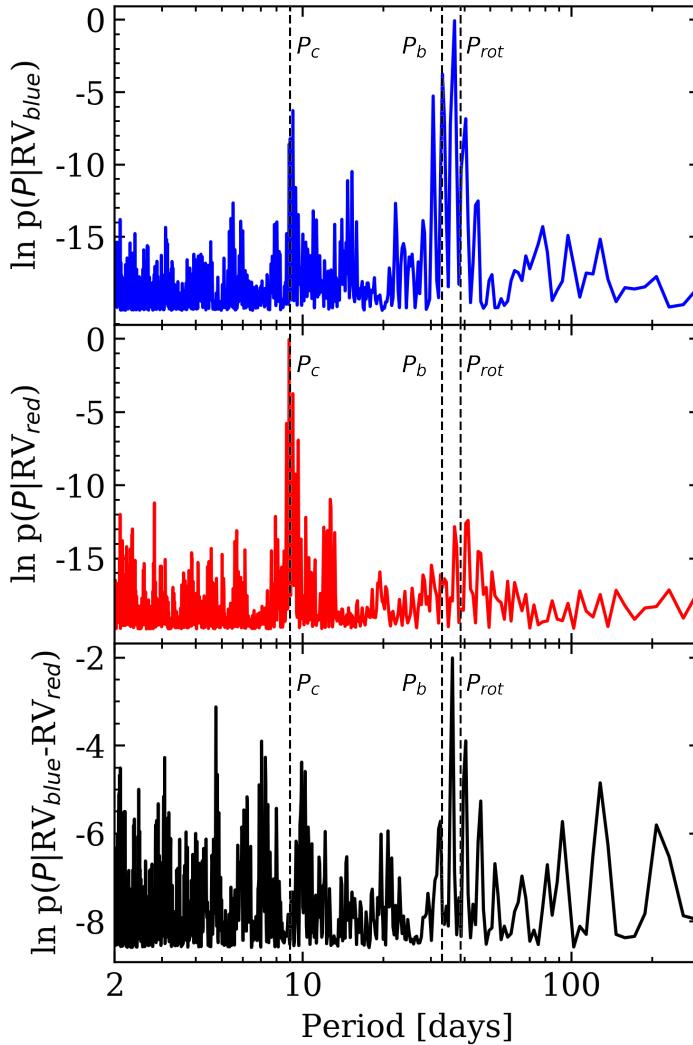


Figure 6.4: Bayesian generalized Lomb-Scargle periodograms of the blue (*top*), red (*middle*), and HARPS RVs and their difference (*bottom*). The three *dashed vertical lines* depict the orbital period of the planets K2-18b and c ( $P_c \sim 9$  days), and the photometric stellar rotation period. The 9-day signal is seen in the first two timeseries, but at a lower probability in the blue likely due to the higher levels of stellar activity in that wavelength regime. The 9-day signal is suppressed in the GLSP of the RV difference, while some residual probability close to  $P_{rot}$  continues to persist due to the incomplete removal of stellar activity.

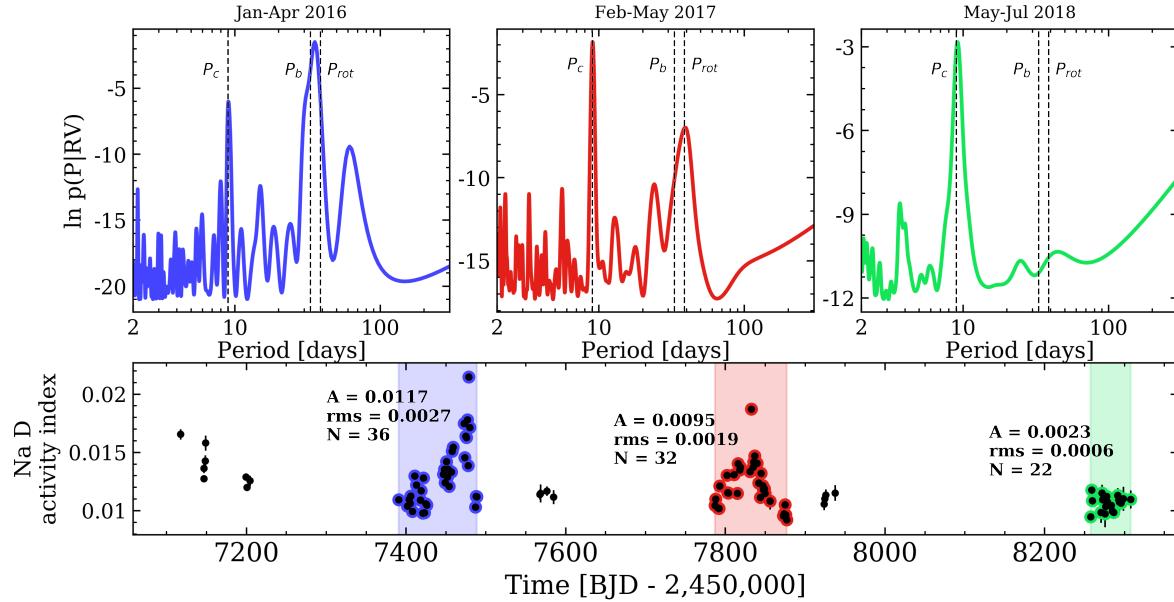


Figure 6.5: *Top row:* Bayesian generalized Lomb-Scargle periodograms of the HARPS RVs in the three observing seasons annotated above each panel. The vertical dashed lines depict the orbital period of the proposed non-transiting planet K2-18c ( $P_c = 8.962$  days), the orbital period of the known transiting planet K2-18b ( $P_b = 32.93963$  days), and the photometric stellar rotation period ( $P_{rot} = 38.6$  days). *Bottom row:* Sodium doublet timeseries as measured by HARPS. The coloured regions/markers are indicative of the epochs used to compute each RV GLSP in the upper row. The annotation group adjacent to each observing season depicts the Na D peak-to-peak amplitude  $A$ , the Na D rms, and the number of measurements within that observing season. These  $A$  and rms diagnostics indicate that the level of stellar activity is decreasing with time, while the  $P_c$  signal is simultaneously becoming more prominent.

To characterize the temporal variability of the K2-18 activity level we compute the strength of the sodium doublet activity index (Na D) in all HARPS spectra following [Astudillo-Defru et al. \(2017a\)](#). The Na D timeseries is shown in the lower panel of Fig. 6.5. In particular we focus on the peak-to-peak amplitude  $A$  and rms of the Na D measurements in each observing season. In doing so we see that the amplitude of the variation in the Na D activity index and its rms both decrease across subsequent observing seasons. Specifically, we find that  $A = 0.0117$  in the first observing season and drops to 0.0023 after  $\sim 26$  months. Similarly, the Na D rms drops from 0.0027 to 0.0006 over the same time interval. These diagnostics indicate that the level of stellar activity is indeed decreasing with time and thus supports the planetary interpretation of the 9-day signal. A similar trend of increasing activity is also observed when considering other activity indicators such as the H $\alpha$  index, although its timeseries is not depicted in Fig. 6.5.

## 6.5 Simultaneous RV modelling of planets and correlated noise

In the era of ultra-precise RV spectrographs whose inherent stability often operates below the photon-noise limit, RV detections of small planets such as K2-18c are limited by nuisance signals from stellar activity. Numerous techniques have been tested to mitigate the effects of stellar activity whose amplitude and quasi-periodic temporal variability can mask and/or mimic planetary signals. Such techniques include linear correlations with contemporaneous activity indicators (e.g. [Boisse et al., 2009](#)), pre-whitening

(e.g. [Queloz et al., 2009](#)), parametric modelling of stellar surface features (e.g. [Dumusque et al., 2014a](#)), and sine wave fitting such as that used in [S18](#). The main issue with this last technique is that the rotationally modulated activity in photometry and in RVs is not strictly periodic as the finite lifetimes of active regions, along with their variable sizes, contrasts, and spatial distributions will introduce a quasi-periodic component. This is especially true when RV timeseries span many stellar rotation cycles. Incomplete models can result in the miscalculation of planetary parameters and the marginalization of coherent signals (e.g. additional planets) that are required to properly interpret the observed RV variations. When modelling RVs it is therefore crucial to include a flexible model that can account for stochastic variations in stellar activity. This is effectively done in a non-parametric way using GP regression simultaneously with planetary models (i.e. Keplerians), thus ensuring self-consistent solutions between planets and stellar activity. Furthermore, GP modelling fits within a Bayesian formalism as a single GP, which describe the temporal covariance between RV measurements with a single set of hyperparameters, is itself a prior distribution of functions whose mean represents the ‘best-fit’ activity model ([Haywood et al., 2014](#); [Faria et al., 2016](#); [Cloutier et al., 2017a](#)). Here we analyse a variety of RV timeseries from either the HARPS ([C17a](#)) or CARMENES ([S18](#)) spectrographs using a model that includes one or two planets plus a correlated noise component from stellar activity in the form of a GP regression model.

Our full two-planet model with observations taken by a single spectrograph contains 16 model parameters: the systemic velocity  $\gamma$ , an additive scalar jitter  $s$ , four quasi-periodic GP hyperparameters  $\{a, \lambda, \Gamma, P_{\text{GP}}\}$ , and five Keplerian parameters per planet  $\{P, T_0, K, h = \sqrt{e} \cos \omega, k = \sqrt{e} \sin \omega\}$ . For cases in which we combine observations from HARPS and CARMENES, we treat their activity models as separate GPs (e.g. [Grunblatt et al., 2015](#)) owing to their unique systematics, the chromatic dependence of stellar activity, and each spectrograph’s distinct wavelength coverage. In this case, all GP hyperparameters are common between the two GP models with the exception of the additive jitter and the covariance amplitude. When modelling the joint HARPS+CARMENES timeseries we therefore have 19 model parameters.

The GP regression models of stellar activity are trained on the star’s precision K2 photometry. The apparent photometric variability, from which the photometric stellar rotation period was measured ( $P_{\text{RV}} = 38.6$  days; [C17a](#)), is sensitive to photospheric active regions which also have an observable manifestation in the RVs with common covariance properties. However, we note that photometry is only weakly sensitive to chromospheric plages which also contribute to RV activity signals, at least in Sun-like stars ([Haywood et al., 2016](#)). We use the K2 photometry to train our GP stellar activity models to ensure that the mean GP model from the simultaneous planet + activity modelling is representative of stellar activity and does not settle into a solution that describes other temporally correlated signals (e.g. non-transiting planets) by restricting the  $P_{\text{GP}}$  to  $P_{\text{rot}}$  or one of its low-order harmonics. By training our GP on ancillary timeseries we empirically constrain the covariance structure of the activity signal and use the posterior probability density functions (PDFs) of the GP hyperparameters from training as priors during the RV modelling stage (see Table 6.1).

In these analyses we sample the posterior PDFs of the RV model parameters given an input dataset via Markov chain Monte Carlo (MCMC) simulations. All simulations are run using the affine-invariant

Table 6.1: Summary of the RV model parameter priors used for all models throughout this study

Parameter	Prior
Systemic velocity, $\gamma$ [m/s]	$\mathcal{U}(\bar{\mathbf{RV}} - 10, \bar{\mathbf{RV}} + 10)$
<i>GP hyperparameters</i>	
Covariance amplitude, $\ln(a/(m/s))$	$\mathcal{U}(-3, 3)$
Exponential timescale, $\ln(\lambda/\text{days})$	$p(\ln \lambda   \text{K2 photometry})$
Coherence, $\ln(\Gamma)$	$p(\ln \Gamma   \text{K2 photometry})$
Periodic timescale, $\ln(P_{\text{GP}}/\text{days})$	$p(\ln P_{\text{GP}}   \text{K2 photometry})$
Additive jitter, $s$ [m s $^{-1}$ ]	$\mathcal{U}(0, 10)$
<i>Keplerian parameters</i>	
$P_b$ [days]	$\mathcal{N}(32.93961, 10^{-4})^{(\bullet)}$
$T_{0,b}$ [BJD-2,450,000]	$\mathcal{N}(7264.3914, 6.3 \times 10^{-4})^{(\bullet)}$
$K_b$ [m s $^{-1}$ ]	$\text{modJ}(1, 20)^{(*)}$
$h_b = \sqrt{e_b} \cos \omega_b$	$\mathcal{U}(-1, 1)^{(\dagger)}$
$k_b = \sqrt{e_b} \sin \omega_b$	$\mathcal{U}(-1, 1)^{(\dagger)}$
$P_c$ [days]	$\mathcal{U}(8, 10)$
$T_{0,c}$ [BJD-2,450,000]	$\mathcal{U}(7259, 7269)$
$K_c$ [m s $^{-1}$ ]	$\text{modJ}(1, 10)^{(*)}$
$h_c = \sqrt{e_c} \cos \omega_c$	$\mathcal{U}(-1, 1)^{(\dagger)}$
$k_c = \sqrt{e_c} \sin \omega_c$	$\mathcal{U}(-1, 1)^{(\dagger)}$

**Notes.**  $(\bullet)$  based on the transit light curve measurements from Benneke et al. (2017).

$(*)$   $\text{modJ}(k, l)$  m s $^{-1}$  refers to a modified Jeffreys prior on a parameter  $A$ , which behaves like a uniform prior for  $A \ll k$  m s $^{-1}$  and like a Jeffreys prior at  $A \gg k$  up to  $l$ . We use a modified Jeffreys prior on the RV semi-amplitudes  $K$  to sample multiple decades as a Jeffreys prior but also include  $K = 0$  m s $^{-1}$  which a Jeffreys prior does not (Gregory, 2005).

$(\dagger)$  We further insist that  $e = h^2 + k^2 < 1$ .

MCMC ensemble sampler `emcee` (Foreman-Mackey et al., 2013). All model parameters are initialized around their MAP values with  $1\sigma$  dispersions from C17a. The adopted model parameters are consistent between the various timeseries considered and are summarized in Table 6.1. In each MCMC simulation we manually monitor the acceptance fraction and ensure that it always lies between 20% and 50% for both the burn-in phase and throughout the actual posterior PDF sampling.

### 6.5.1 CARMENES RVs

Here we model the subset of the CARMENES-visible RVs presented in S18 which are known to not result in the anomalous suppression of the 9-day signal. We consider two RV models, each containing a quasi-periodic GP regression model of stellar activity. The first model contains only one planetary signal from the confirmed transiting planet K2-18b, while the second model includes the second planet K2-18c at  $\sim 9$  days. The RVs and GLSPs are plotted in Fig. 6.7 for both the one- and two-planet models after iteratively removing the MAP models of activity and planetary signals.

In the one-planet model of the 55 CARMENES RVs, the GP activity model has a covariance amplitude of  $7.5 \text{ m s}^{-1}$  that is greater than the sinusoidal amplitude of  $2.7 \text{ m s}^{-1}$  measured by S18 on nearly the same dataset. Based on the GLSP of K2-18b (i.e. with activity removed), it is clear that although the activity model has a large amplitude, it fails to model the 9-day signal. The GLSP of the residuals following the removal of activity and K2-18b ( $K_b = 3.61 \pm 0.82 \text{ m s}^{-1}$ ) clearly exhibits a strong periodic signal at  $\sim 9$  days hinting at the existence of an additional signal that is unmodelled when assuming a one-planet model.

The stellar activity in the two-planet model has a similarly large covariance amplitude of  $8.2 \text{ m s}^{-1}$ . However, the only significant signal in the GLSP of the RV activity is at the stellar rotation period. Similarly, the GLSP of K2-18b ( $K_b = 2.91 \pm 0.88 \text{ m s}^{-1}$ ) only exhibits a significant signal at  $P_b$ , and the GLSP of K2-18c ( $K_c = 2.31 \pm 0.76 \text{ m s}^{-1}$ ) exhibit a strong signal at  $\sim 9$  days with a somewhat weaker signal at  $\sim 5.5$  days. The GLSP of the residuals following the removal of both planets and activity only shows a significant residual probability at  $\sim 5.5$  days, which only arises after the removal of activity and K2-18b (cf. panels of O-C and K2-18c in Fig. 6.7). The nature of this signal is less obvious; unlike the 9-day signal, it does not appear with enough significance in either GLSP of the HARPS or CARMENES RVs prior to the removal of any modelled signals (see Fig. 6.2). One possible explanation is that the  $\sim 5.5$ -day signal arises from an alias of  $P_c$  with the CARMENES WF which exhibits excess power close to the baseline duration of  $\sim 189$  days. Using the standard formula to compute the alias frequency from the signal and WF frequencies (i.e.  $f_{\text{alias}} = f_{\text{signal}} + n f_{\text{WF}}$ ), and setting  $f_{\text{signal}} = 1/8.997 \text{ days}^{-1}$  and  $f_{\text{WF}} = 1/189 \text{ days}^{-1}$ , we find an aliased periodicity at  $\sim 5.56$  days when  $n = 13$ . Given the high-order  $n$  required to identify an aliased periodicity that is seemingly consistent with the excess probability at  $\sim 5.5$  days, we do not claim that this WF alias explains the signal's origin and similarly we cannot discard the possibility that the 5.5-day signal comes from an additional planet that has not yet been detected. More RV data are required to investigate the source of this signal. In Sect. 6.5.5 we perform a model comparison considering the possibility that the 5.5-day signal is due to a third planet in the system.

### 6.5.2 All HARPS RVs

In Sect. 6.4 we presented 31 new HARPS RVs to investigate the temporal variability of the 9-day signal. Hence the full HARPS WF has been extended to over a year past the previously most recent published HARPS measurement for this system (C17a) and now contains 106 RV measurements. Here we model the full HARPS timeseries in the same way as was done for the CARMENES RVs in Sect. 6.5.1. The RVs and GLSPs are plotted in Fig. 6.8.

In the one-planet model the GP activity model has a covariance amplitude of  $2.3 \text{ m s}^{-1}$ , comparable to the MAP  $K_b = 2.75 \pm 0.66 \text{ m s}^{-1}$ . Similarly to the one-planet model of the CARMENES RVs, the activity model fails to account for the high probability of the 9-day signal. The 9-day peak continues to persist following the removal of the K2-18b Keplerian.

In the two-planet model the GP activity model has a somewhat larger covariance amplitude compared to the one-planet model:  $4.18 \text{ m s}^{-1}$ . This amplitude is comparable to the MAP semi-amplitudes of the two planets ( $K_b = 3.32 \pm 0.60 \text{ m s}^{-1}$ ,  $K_c = 3.71 \pm 0.57 \text{ m s}^{-1}$ ) and, given the proximity of the stellar rotation period to  $P_b$  and aliases of the two aforementioned periods with the WF (S18), the activity model only partially suppresses the GLSP probabilities between  $\sim 30$  and 50 days. It is also clear that when the mean activity model and only a single planet are removed, the only remaining signal at high probability is that of the remaining planet at 9 days. Furthermore, it is clear that there are no residual signals at high probability when all modelled signals are removed. Most notably, a probability peak at  $\sim 5.5$  days, as was seen in the CARMENES residuals with a two-planet model (Fig. 6.7), is visible but only at the level of the noise.

### 6.5.3 Joint HARPS+CARMENES RVs

Here we model the joint RV timeseries of the 106 HARPS plus the 55 CARMENES RVs. The RVs and GLSPs are plotted in Fig. 6.9. In the one-planet model the covariance amplitude of the HARPS and CARMENES stellar activity models are  $1.5$  and  $5.5 \text{ m s}^{-1}$ , respectively. These values are each slightly smaller than the covariance amplitudes measured when considering each spectrograph's time-series individually but their ratio is nearly preserved. Similarly to either spectrograph's individual RV analysis in the presence of a one-planet model, the GLSP of the residuals following the removal of K2-18b ( $K_b = 3.00 \pm 0.50 \text{ m s}^{-1}$ ) and activity exhibits a strong periodic signal at  $\sim 9$  days which again hints at the existence of an additional planetary signal.

The stellar activity covariance amplitudes in the two-planet model are comparable to that in the one-planet model, i.e.  $3.0$  and  $5.5 \text{ m s}^{-1}$  for HARPS and CARMENES, respectively. The corresponding GLSP of the RV activity is reminiscent of the one-planet RV activity GLSP with the exception that the inclusion of two modelled planets ( $K_b = 2.75 \pm 0.43 \text{ m s}^{-1}$ ,  $K_c = 2.76 \pm 0.41 \text{ m s}^{-1}$ ) drastically reduces the probability of the 9-day signal. Indeed in the GLSP of K2-18c, the strongest signal is at  $\sim 9$  days with only a hint of the  $\sim 5.5$ -day signal that was seen in CARMENES. In both the GLSP of the HARPS and joint RV residuals following the removal of both planets and activity (see Figs. 6.8 and 6.9) the  $\sim 5.5$ -day signal is not seen at high probability, which suggests that the signal is not physical and instead arises stochastically as a by-product of the CARMENES WF.

### 6.5.4 Overlapping HARPS and CARMENES window functions

For a maximal one-to-one comparison we can compare the RV model analyses and GLSP structures in the subsets of the HARPS and CARMENES RVs that are restricted to the 138 days from February 2 to June 20, 2017. Between these dates the HARPS and CARMENES WFs overlap such that we have approximately contemporaneous RVs taken with each spectrograph. By only considering the observations taken throughout the overlapping time span we minimize our sensitivity to temporal variations in stellar activity whose properties may vary between successive observing cycles. The overlapping WF contains 35 HARPS and 50 CARMENES RVs. One of the CARMENES RVs in the overlapping window was found to anomalously suppress the  $\sim 9$ -day signal in Sect. 6.2.2, so we discarded it and were left with 49 CARMENES RVs. The RVs and GLSPs are plotted in Fig. 6.10.

In the one-planet model the covariance amplitudes are equivalent with each spectrograph (i.e.  $2.0 \text{ m s}^{-1}$ ) and are notably small compared to the previously analysed timeseries. This may be due to the lack of a long-term near-linear trend in the stellar activity over the short time span considered here. The corresponding activity model appears close to flat indicating that the RV activity has only weak structure over this relatively short time span. The low activity amplitude also results in a low probability at  $P_{\text{rot}}$  and the activity GLSP being dominated by the 9-day signal, which is effectively unmodelled when only one planet is considered. We measure  $K_b = 3.96 \pm 0.73 \text{ m s}^{-1}$ , which along with the activity model reveals the residual 9-day signal and the  $\sim 5.5$ -day signal that was seen in the CARMENES residuals.

In the two-planet model the covariance amplitudes are nearly identical to the one-planet model ( $2.0 \text{ m s}^{-1}$ ) and therefore exhibit a similarly featureless structure. The small covariance amplitudes of the activity models result in the activity GLSP containing primarily noise. Comparatively, the GLSPs of the modelled planets ( $K_b = 3.59 \pm 0.62 \text{ m s}^{-1}$ ,  $K_c = 2.65 \pm 0.58 \text{ m s}^{-1}$ ) are dominated by their respective periodicities with the  $\sim 5.5$  days signal appearing in the GLSP of K2-18c, albeit at a much lower probability than the 9-day signal. However in the residual GLSP, the  $\sim 5.5$ -day signal is largely suppressed after removing K2-18c.

### 6.5.5 Model comparison

The detection of exoplanets in RV data is fundamentally based on whether or not the input dataset favours the existence of the planet of interest. This is typically done within a Bayesian framework wherein the fully marginalized likelihoods (i.e. the evidence) of competing models (i.e. one versus two planets) are computed and used for model comparison. In this formalism, a planet is said to be detected if the evidence for the  $(n + 1)$ -planet model is significantly larger than the evidence for a model containing  $n$  planets. Here we calculate the model evidence for the purpose of model comparison and use the resulting values to determine whether the putative RV planet K2-18c is favoured by the various timeseries considered.

Each model's Bayesian evidence is approximated using the estimator from [Perrakis et al. \(2013\)](#) and the marginalized posterior PDFs from our MCMC analyses as importance samplers. The [Perrakis et al. \(2013\)](#) estimator is known to result in quantitatively similar results to other more robust but computationally expensive methods (e.g. nested samplers; [Nelson et al., 2018](#)). Model comparison requires that

all common model parameters between competing models be drawn from identical prior distributions, which are listed in Table 6.1. Our Bayesian evidence estimates are reported in Table 6.2 for both the one- and two-planet models and for all input timeseries considered.

Also included in Table 6.1 are the 2-1 Bayes factors (i.e. evidence ratios) of the two-planet model relative to the one-planet to determine whether the second planet K2-18c is favoured or disfavoured by the corresponding timeseries. Overall, we find that the explicit values of the Bayesian evidence favour the two-planet model for all the timeseries considered. However, the dispersion in calculated evidence values when using various methods of calculation are known to vary by factors of  $\gtrsim 10^2$  depending on the complexity of the model (i.e. the number of planets; [Nelson et al. 2018](#)). We recall that the simplest model considered in this study is not the zero-planet model as we know from the transit light curves that K2-18b exists at  $\sim 33$  days. Effectively, we are therefore only tasked with detecting one new RV planet rather than two. But given the caveat that uncertainties in the calculated evidence can be of the order of  $10^2$ , we require that the evidence ratio of the two-planet model to the one-planet model must be  $\geq 10^2$  for the second planet K2-18c to be detected. Under this condition there are two instances in which K2-18c is not detected. The first occurs with the full set of the 58 CARMENES RVs from [S18](#) in which K2-18c is not detected due to the three anomalous measurements identified in Sect. 6.2.2. This result is consistent with the null detection of K2-18c with these data in [S18](#). However, the 2-1 Bayes factor for CARMENES alone exceeds  $10^2$  following the removal of the three measurements mentioned above. The second occurs because the blue CARMENES RVs only weakly favour a second planet which can be attributed to the increased RV rms at these shorter wavelengths<sup>8</sup>. This trend is seen again in the blue and red HARPS RVs for which a second planet is more strongly favoured by the red RVs where the RV rms is smaller. The increased measurement uncertainty for CARMENES in the blue hides planetary signals and makes the inference of their presence less certain given the correspondingly low data likelihoods.

We recall the  $\sim 5.5$ -day signal seen in the K2-18c and residual GLSPs of the RV timeseries containing CARMENES data in Figs. 6.7, 6.9, and 6.10. As a test of the potential planetary origin of this signal we first ran an MCMC on the CARMENES RVs as it is there that the residual 5.5-day signal exhibited the highest probability in the GLSP following the removal of K2-18b, c, and stellar activity. As we did for the one- and two-planet models, we then estimate the evidence of this three-planet model using the estimator from [Perrakis et al. \(2013\)](#) and compare it to the two-planet model for the same input timeseries. For the third planet we adopt identical priors to that of K2-18c (see Table 6.1) with the exception of the planet's orbital period and time of mid-conjunction which are modified to  $\mathcal{U}(4.5, 6.5)$  days and  $\mathcal{U}(7259, 7265.5)$  BJD-2,450,000, respectively. The resulting ln evidence for the two- and three-planet models are -165.0 and -162.4, respectively. The corresponding 3-2 Bayes factor is  $\sim 7$ , implying that the three-planet model including a planet at  $\sim 5.5$  days is not significantly favoured over the two-planet model. By a similar exercise, using the full joint HARPS+CARMENES timeseries yields a 3-2 Bayes factor of  $\sim 0.8$ . Therefore, by the effective accounting of the 5.5-day periodic signal by our K2-18c models in Figs. 6.8, 6.9, and 6.10, and the disfavourability of the three-planet model compared to the model containing just two planets, we conclude that a third planet at  $\sim 5.5$  days is not detected in the available RV data, but its signal origin may be alluded to with additional RV monitoring.

---

<sup>8</sup> $7.5 \text{ m s}^{-1}$  compared to  $5.14$  and  $5.73 \text{ m s}^{-1}$  in the full and red CARMENES RVs, respectively.

Table 6.2: Marginal likelihood estimations and Bayes factors for various RV datasets and models

Dataset $D$	$N_{\text{RV}}$	Model $\mathcal{M}_i$	ln Model evidence <sup>(•)</sup> ln p( $D \mathcal{M}_i$ )	Bayes factor: 2 to 1 planets <sup>(*)</sup> $p(\mathcal{M}_2 D)/p(\mathcal{M}_1 D)$
HARPS	106	1 planet + GP	-338.5	-
		2 planets + GP	-325.5	$3 \times 10^5$
CARMENES	58	1 planet + GP	-180.6	-
		2 planets + GP	-178.1	7
reduced CARMENES	55	1 planet + GP	-169.5	-
		2 planets + GP	-164.1	143
HARPS + reduced CARMENES	161	1 planet + GP	-489.6	-
		2 planets + GP	-475.8	$6 \times 10^5$
blue HARPS	106	1 planet + GP	-375.5	-
		2 planets + GP	-369.2	336
red HARPS	106	1 planet + GP	-375.5	-
		2 planets + GP	-359.4	$6 \times 10^6$
blue CARMENES	55	1 planet + GP	-186.8	-
		2 planets + GP	-183.0	28
red CARMENES	55	1 planet + GP	-174.0	-
		2 planets + GP	-159.1	$2 \times 10^6$
HARPS (Feb-Jun 2017)	35	1 planet + GP	-121.0	-
		2 planets + GP	-113.6	1018
CARMENES (Feb-Jun 2017)	49	1 planet + GP	-151.7	-
		2 planets + GP	-146.0	181
HARPS + CARMENES (Feb-Jun 2017)	84	1 planet + GP	-256.5	-
		2 planets + GP	-249.6	613

**Notes.** <sup>(•)</sup> Estimates of the model evidence are calculated using the [Perrakis et al. \(2013\)](#) estimator and the marginalized posterior probability density functions from our MCMC runs.

<sup>(\*)</sup> Bayes factors, or evidence ratios, are written as  $\frac{p(\mathcal{M}_2|D)}{p(\mathcal{M}_1|D)} = \frac{p(D|\mathcal{M}_2)}{p(D|\mathcal{M}_1)} \frac{p(\mathcal{M}_2)}{p(\mathcal{M}_1)}$ , where each model prior is  $p(\mathcal{M}_i) = \alpha^i$  for  $\alpha = (\sqrt{5} - 1)/2 \approx 0.618$  such that  $\sum_{i=1}^2 p(\mathcal{M}_i) = 1$ .

## 6.6 Discussion and conclusions

We have conducted a systematic re-analysis of the published HARPS ([C17a](#)) and CARMENES ([S18](#)) RVs of the transiting planet host K2-18 to identify the source of the apparent 9-day signal which prior to this study had only been seen in the HARPS dataset. We have also included an additional set of 31 new HARPS RVs to investigate the temporal dependence of the 9-day signal and to improve the measurement precision of planet parameters. Our main conclusions are the following:

1. The CARMENES window function is somewhat detrimental to the detection of an injected 9-day Keplerian signal compared to the HARPS window function, in that the injected signal is seen at a lower probability in the generalized Lomb-Scargle periodogram (GLSP) when using the CARMENES window function.
2. The cause of the non-detection of the 9-day signal in [S18](#) was shown to result from three anomalous CARMENES measurements; when they are removed the existence of the 9-day signal is revealed in the GLSP of the remaining 55 RVs.
3. We computed two sets chromatic HARPS RVs. The 9-day signal is seen in both timeseries and at a significantly higher probability in the red HARPS RVs where stellar activity is weaker. This supports the planetary interpretation of the 9-day signal.
4. The 9-day signal is retrieved with HARPS in each of its three observing seasons separated by  $\sim 1$  year. The probability of the 9-day signal increases with time simultaneously with a decrease in the

level of stellar activity as probed by the Na D activity index. This further supports the planetary interpretation of the 9-day signal.

5. We adopt a non-parametric stellar activity model to account for stellar variability over the multiple stellar rotation cycles spanned by the observations, and simultaneously model activity and planetary signals. This results in self-consistent planet solutions and the ability to compare one- and two-planet models on equal grounds.
6. In all the considered times-series, the Bayesian model evidence favours a two-planet model over the one-planet model which includes K2-18c at  $\sim 9$  days.

By the points listed above, we have obtained compelling evidence for the planetary nature of the 9-day signal seen in HARPS and in the reduced CARMENES RV timeseries. It is important to highlight the importance of basing RV planet detections on robust Bayesian model comparison tests rather than basing those detections solely on periodogram false alarm probabilities (FAPs) which can vary stochastically and are highly sensitive to variations in the input timeseries (e.g. weighting schemes). Although significant peaks in a GLSP are useful for the initial identification of periodic signals in unevenly sampled timeseries, conclusions regarding their actual existence and origin should not be made solely based on their FAP. Accurate and simultaneous modelling of all signals present in a timeseries is required to determine accurate model parameters of planets and activity. Furthermore, Bayes factors—or the ratio of the competing models’ fully marginalized likelihoods—are robust model comparison tools which marginalize over all prior information about models with competing numbers of planets and penalize overly complicated models. In this way they are optimally suited to the detection confirmation of periodic planetary signals.

In our re-analysis of the joint HARPS+CARMENES RVs we have measured the most likely Keplerian solution to each planet’s orbit. By including all available RV observations of the K2-18 system (excluding those known to be anomalous), we have obtained the most precise planetary solutions for K2-18 to date. The point estimates of the two-planet model parameters resulting from this analysis are presented in Table 6.3. As a sanity check we can compare the resulting marginalized posterior PDFs for parameters of interest between the individual HARPS, CARMENES, and their joint RV timeseries. In this way we can ensure that the planetary solutions from the timeseries of the two spectrographs are consistent with each other and with their joint timeseries. For instance, we compare the resulting marginalized poster PDFs of  $K_b$  and  $K_c$  obtained with each timeseries in Fig. 6.6. It is evident that the MAP  $K_b$  solutions are nearly equivalent when measured with any of the three timeseries. Similarly, MAP  $K_c$  values are consistent at the  $1\sigma$  level, albeit with more dispersion than the  $K_b$  PDFs given the comparatively large uncertainties in the K2-18c ephemeris.

### 6.6.1 Improved stellar parameters based on Gaia DR2

To map the observable transit and RV parameters to physical planetary parameters we must first characterize the host star. Specifically, we can exploit the exquisite precision of the Gaia DR2 to improve the stellar mass and radius of K2-18.

Firstly, the K2-18 stellar mass is computed from the M dwarf mass-luminosity relation (MLR) from [Benedict et al. \(2016\)](#). The analytical MLR based on absolute  $K$ -band magnitudes is favoured over

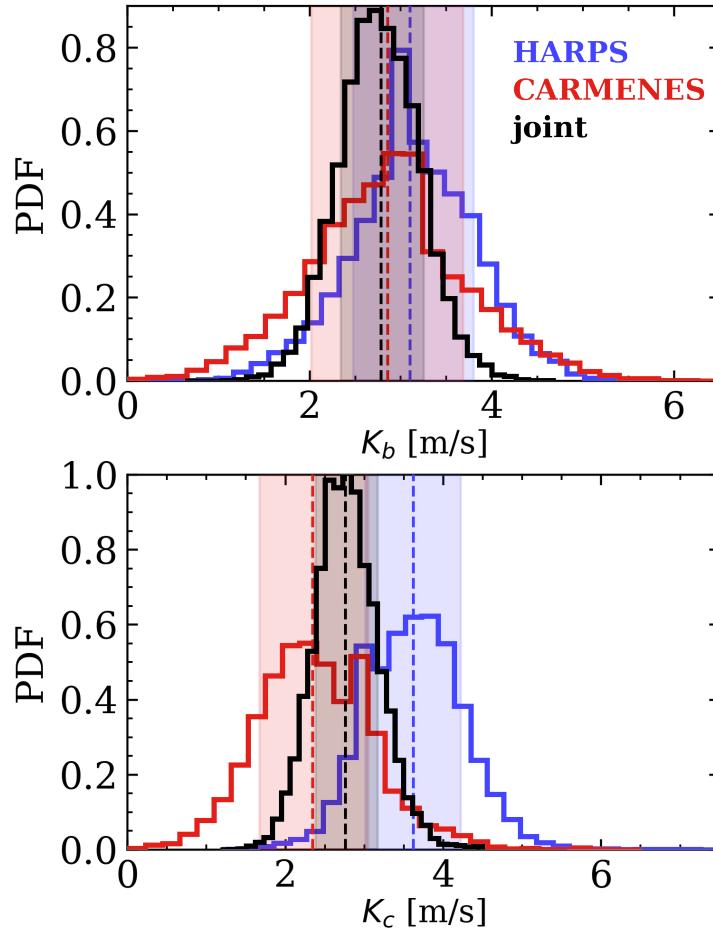


Figure 6.6: The 1D marginalized posterior PDFs of the K2-18b and c semi-amplitudes from analyses of the full HARPS (blue), the reduced CARMENES (red), and their joint (black) RV timeseries. The dashed vertical lines and shaded regions depict the maximum a posteriori values and  $1\sigma$  confidence intervals respectively. All  $K_b$  and  $K_c$  values are consistent at the  $1\sigma$  level which is approximated by each PDF's 16th and 84th percentiles.

the  $V$ -band whose dispersion about the relation is twice that in the  $K$ -band. The distance modulus is calculated from the precision Gaia DR2 stellar parallax ( $p = 26.299 \pm 0.055$  mas; [Gaia Collaboration et al., 2018](#)) to be  $\mu = 2.900 \pm 0.005$ , where we have added the  $30\ \mu\text{as}$  systematic offset in the measured parallax as noted in [Lindegren et al. \(2018\)](#). By propagating errors in the K2-18  $K$ -band magnitude ( $K = 8.899 \pm 0.019$ ; [Cutri et al., 2003](#)), the distance modulus, and the MLR coefficients, we find an absolute  $K$ -band magnitude of  $M_K = 5.999 \pm 0.020$  and a corresponding stellar mass of  $M_s = 0.495 \pm 0.004 M_\odot$ .

From the stellar mass we are able to derive the stellar radius using the empirical mass-radius relationship (MRR) for M dwarfs from [Boyajian et al. \(2012\)](#). By propagating the uncertainties in the M dwarf MRR coefficients we compute the K2-18 stellar radius to be  $R_s = 0.469 \pm 0.010 R_\odot$ . We note that both the updated stellar mass and radius, which are based on the stellar parallax, are considerably larger than the spectroscopically derived values of  $0.359 \pm 0.047 M_\odot$  and  $0.411 \pm 0.038 R_\odot$  ([Benneke et al., 2017](#)). The new mass and radius values are inconsistent with their previous values at the levels of  $2.9\sigma$  and  $1.5\sigma$ , respectively. This is the direct result of the increased K2-18 distance from Gaia ( $38.025 \pm 0.079$  pc) compared to its previously measured distance ( $34 \pm 4$  pc) and will have important implications for the derived physical parameters of both K2-18b and c. We also note the improved fractional uncertainties on the updated stellar mass and radius of 0.8% and 2.1%, respectively, compared to the previous fractional uncertainties of 13.1% and 9.2%.

### 6.6.2 Precise planetary parameters

The improved stellar parameters, along with our joint HARPS+CARMENES RV analysis, provide the most precise set of planetary parameters for the planets K2-18b and c to date. Point estimates of the planetary parameters from our joint HARPS+CARMENES RV analysis are presented in Table 6.3. In particular, we measure the precise mass and minimum mass of K2-18b and c, respectively, to be  $m_{p,b} = 8.64 \pm 1.35 M_\oplus$  and  $m_{p,c} \sin i_c = 5.63 \pm 0.84 M_\oplus$ .

The improved stellar radius also provides a more precise planetary radius given the measured  $r_{p,b}/R_s$  value from [Benneke et al. \(2017\)](#). We find that  $r_{p,b} = 2.711 \pm 0.065 R_\oplus$ . From this we derive a planetary bulk density for K2-18b of  $\rho_{p,b} = 2.4 \pm 0.4 \text{ g cm}^{-3}$  thus making K2-18b inconsistent with either an Earth-like composition or a pure water-world ([Zeng & Sasselov, 2013](#)). Prior to updating the mass and radius of K2-18b, neither of these scenarios could have been ruled out. It is now clear that at minimum,  $\sim 8\%$  of the size of K2-18b (i.e.  $\sim 1382$  km) must be attributed to an optically think gaseous atmosphere as evidenced by its low bulk density. The expected signal amplitude in transmission for a cloud-free hydrogen-dominated atmosphere ( $\mu = 2$ ) is  $\sim 10Hr_{p,b}/R_s^2 \sim 155$  ppm, where  $H = k_B T_{\text{eq}}/\mu m_p g$  is the atmospheric pressure scale height,  $k_B$  is the Boltzmann constant,  $T_{\text{eq}}$  is the planet's equilibrium temperature assuming an Earth-like Bond albedo,  $\mu m_p$  is the assumed mean molecular weight, and  $g$  is the surface gravity ([Kaltenegger & Traub, 2009](#)). For comparison, a well-mixed water-dominated atmosphere ( $\mu = 18$ ) has a transmission signal amplitude of  $\sim 18$  ppm. Given the scale height of its extended gaseous envelop and its proximity to the solar system, K2-18b continues to represent an exciting opportunity to characterize a sub-Neptune-sized exoplanet receiving Earth-like insolation with upcoming space missions such as the James Webb Space Telescope and ARIEL.

## 6.7 Appendix

### 6.7.1 Iterative radial velocity timeseries and GLSP figures from Sect. 6.5

In Sect. 6.5 we considered a variety of RV datasets and models which included either one or two planets along with a GP regression model of stellar activity that had been trained on the star’s K2 photometry. The following figures depict the iterative RVs and GLSPs for each dataset and model. In each iteration we removed one or more coherent signals (i.e. planets or activity) to see if any residual periodicities persist for which additional model components may be required.

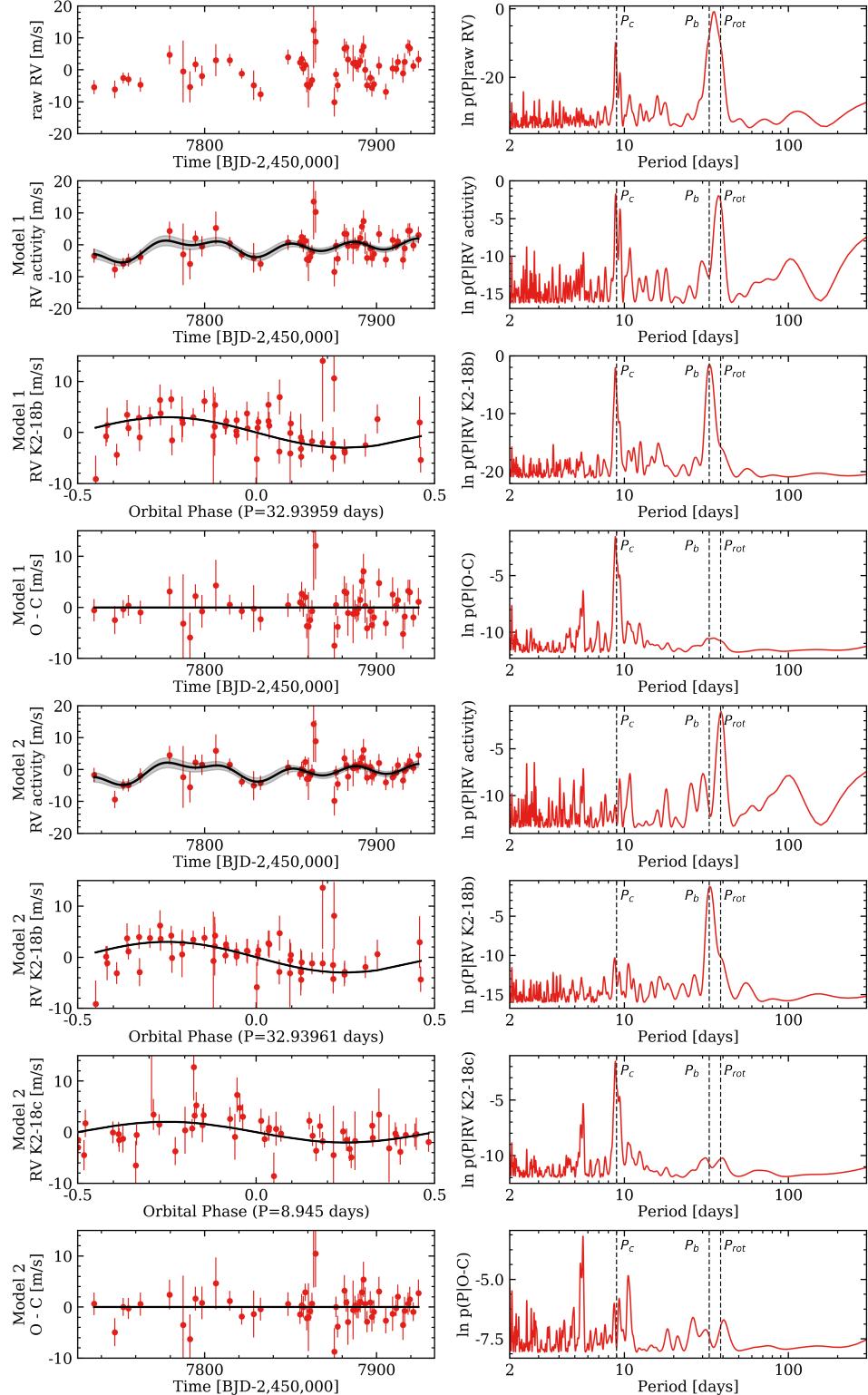


Figure 6.7: Results of our RV analysis of the 55 CARMENES RVs that are known not to significantly suppress the apparent 9-day signal seen with HARPS. The RV timeseries and their corresponding GLSP are plotted in common rows for each coherent RV signal modelled (i.e. planets and stellar activity) in either a one- or two-planet model. The overplotted [RV] models are computed using the MAP model parameters from our MCMC analysis. The [vertical dashed lines] in the GLSPs are indicative of the MAP orbital periods for K2-18b and c and the photometric stellar rotation period. The first row depicts the raw RVs; the next three following rows present the results assuming a one-planet model (i.e. K2-18b); and the final four rows present the results assuming a two-planet model (i.e. K2-18b and c). The residual rms values assuming a one- and two-planet model are  $3.84$  and  $3.57 \text{ m s}^{-1}$ , respectively. We find that the source of the residual  $\sim 5.5$ -day signal in the bottom GLSP is due to an alias rather than an unmodelled physical source (see text).

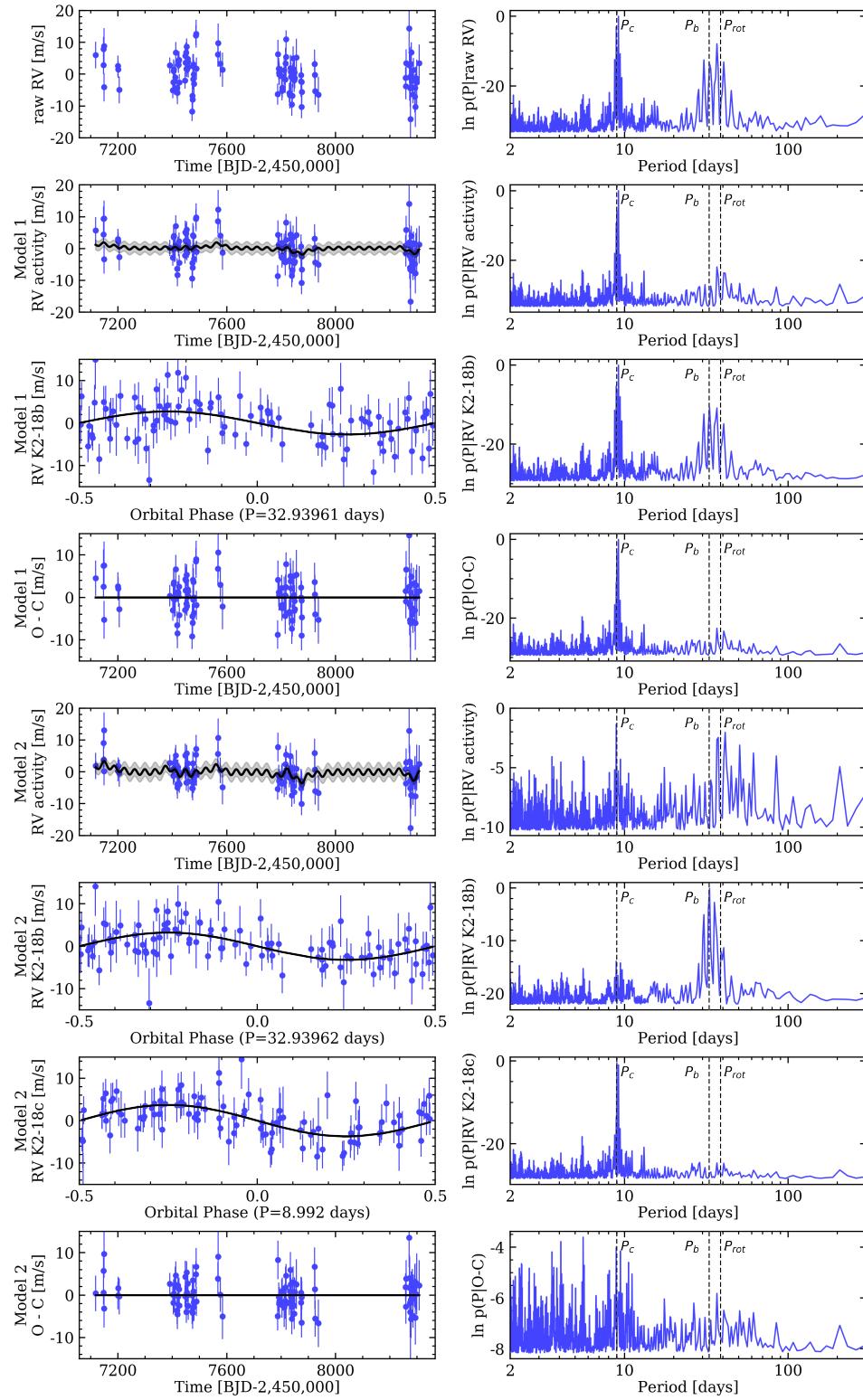


Figure 6.8: Similar to Fig. 6.7, but for the full HARPS WF containing 106 RVs. The rms of the residual timeseries assuming a one- and two-planet model are 4.68 and  $3.93 \text{ m s}^{-1}$ , respectively. MAP RV models. GLSP periodicities.

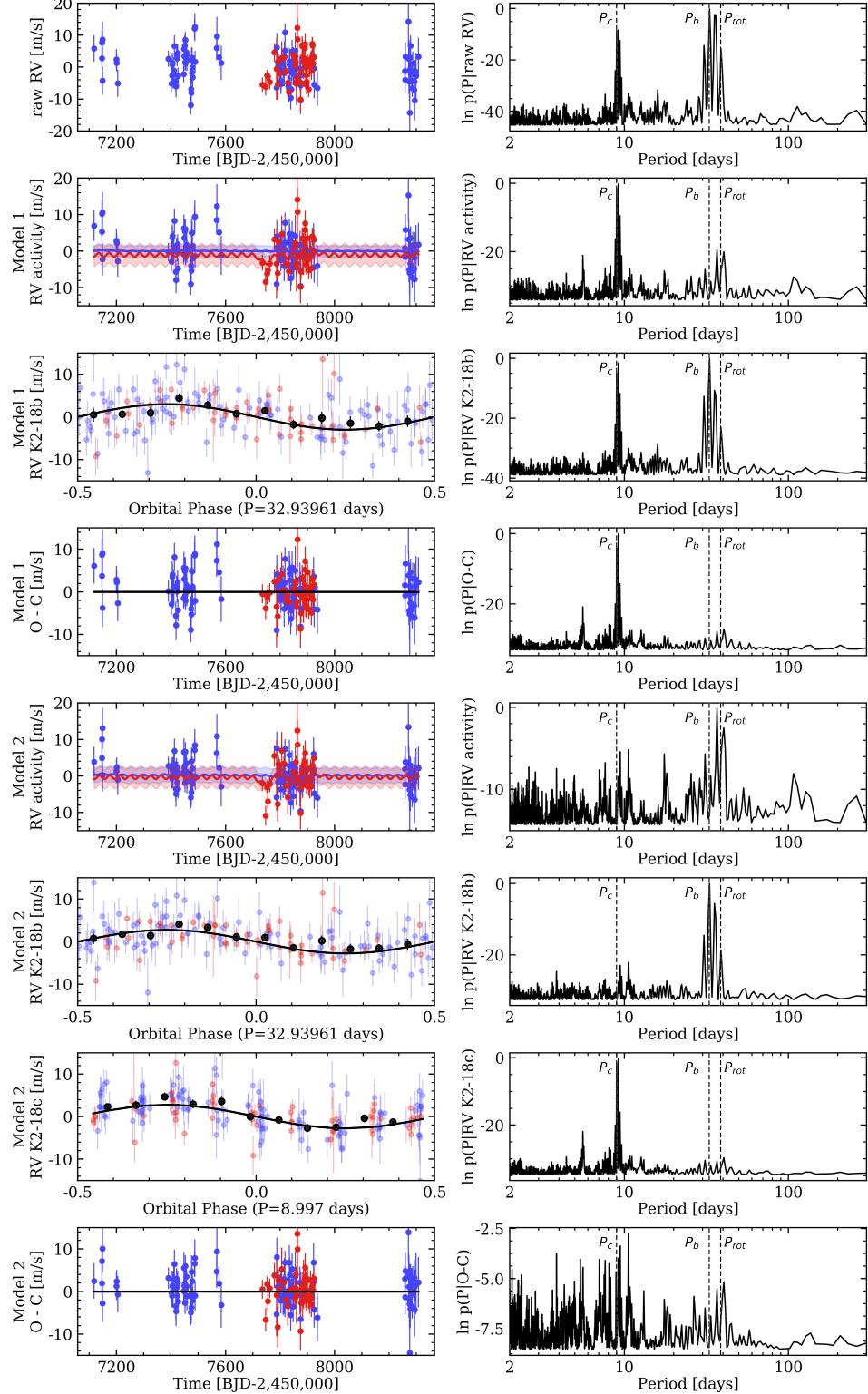


Figure 6.9: Similar to Fig. 6.7, but for the 161 joint HARPS+CARMENES RVs. The HARPS and CARMENES RVs are plotted as blue and red markers, respectively. The phase-folded RVs depicting planetary signals are binned for clarity. The rms of the residual timeseries assuming a one- and two-planet model are 4.51 and 3.82 m s<sup>-1</sup>, respectively. [MAP RV models], [GLSP periodicities].

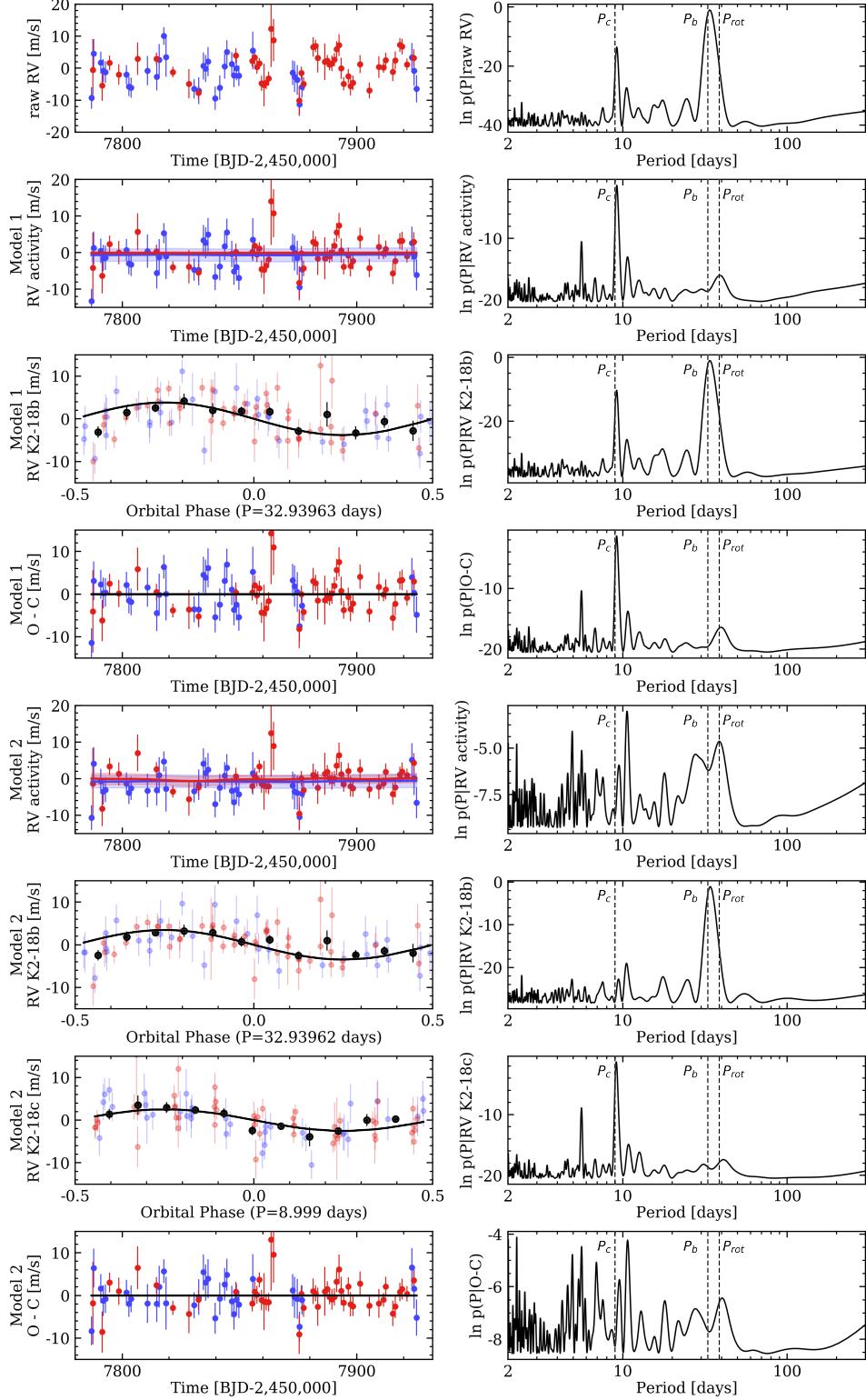


Figure 6.10: Similar to Fig. 6.7, but for the 84 joint HARPS+CARMENES RVs obtained during the time interval in which the two spectrograph WFs overlap (February-June 2017). The HARPS and CARMENES RVs are plotted as blue and red markers, respectively. The phase-folded RVs depicting planetary signals are binned for clarity. The rms of the residual timeseries assuming a one- and two-planet model are 4.26 and 3.62 m s<sup>-1</sup>, respectively. [MAP RV models](#) [GLSP periodicities](#)

Table 6.3: K2-18 model parameters from the HARPS+CARMENES joint RV analysis

Parameter	Point estimate
<i>Stellar Parameters</i>	
2MASS Photometry	$J=9.763 \pm 0.028, H=9.135 \pm 0.026, K_s=8.899 \pm 0.019$
Stellar mass, $M_s$ [ $M_\odot$ ]	$0.495 \pm 0.004$
Stellar radius, $R_s$ [ $R_\odot$ ]	$0.469 \pm 0.010$
Effective temperature, $T_{\text{eff}}$ [K]	$3503 \pm 60$
Stellar parallax, $p$ [mas]	$26.299 \pm 0.055$
Distance, $d$ [pc]	$38.025 \pm 0.079$
HARPS systemic velocity, $\gamma_0, \text{HARPS}$ [ $\text{m s}^{-1}$ ]	$652.51 \pm 1.0$
CARMENES systemic velocity, $\gamma_0, \text{CARMENES}$ [ $\text{m s}^{-1}$ ]	$-2.87 \pm 0.9$
<i>GP hyperparameters</i>	
HARPS covariance amplitude, $a_{\text{HARPS}}$ [ $\text{m s}^{-1}$ ]	$3.0^{+3.4}_{-1.7}$
CARMENES covariance amplitude, $a_{\text{CARMENES}}$ [ $\text{m s}^{-1}$ ]	$5.0^{+5.3}_{-2.9}$
Exponential timescale, $\lambda$ [days]	$448.8 \pm 67.3$
Coherence, $\Gamma$	$0.17^{+0.07}_{-0.04}$
Periodic timescale, $P_{\text{GP}}$ [days]	$37.4^{+0.5}_{-0.3}$
HARPS additive jitter, $s_{\text{HARPS}}$ [ $\text{m s}^{-1}$ ]	$0.48 \pm 0.42$
CARMENES additive jitter, $s_{\text{CARMENES}}$ [ $\text{m s}^{-1}$ ]	$0.58 \pm 0.53$
<i>K2-18c</i>	
Period, $P_c$ [days]	$8.997 \pm 0.007$
Time of inferior conjunction, $T_{0,c}$ [BJD-2,450,000]	$7263.69 \pm 0.44$
Radial velocity semi-amplitude, $K_c$ [ $\text{m s}^{-1}$ ]	$2.76 \pm 0.41$
$h_c = \sqrt{e_c} \cos \omega_c$	$0.00^{+0.24}_{-0.30}$
$k_c = \sqrt{e_c} \sin \omega_c$	$0.15^{+0.23}_{-0.28}$
Semi-major axis, $a_c$ [AU]	$0.0670 \pm 0.0002$
Minimum planet mass, $m_{p,c} \sin i_c$ [ $M_\oplus$ ]	$5.62 \pm 0.84$
Equilibrium temperature, $T_{\text{eq},c}$ [K]	
Bond albedo of 0.3	$409 \pm 8$
<i>K2-18b</i>	
Period, $P_b$ [days]	$32.93962 \pm 1.0 \times 10^{-4}$
Time of inferior conjunction, $T_{0,b}$ [BJD-2,450,000]	$7264.39142 \pm 6.4 \times 10^{-4}$
Radial velocity semi-amplitude, $K_b$ [ $\text{m s}^{-1}$ ]	$2.75 \pm 0.43$
$h_b = \sqrt{e_b} \cos \omega_b$	$0.30^{+0.11}_{-0.24}$
$k_b = \sqrt{e_b} \sin \omega_b$	$-0.05^{+0.26}_{-0.25}$
Semi-major axis, $a_b$ [AU]	$0.1591 \pm 0.0004$
Planet radius, $r_{p,b}$ [ $R_\oplus$ ] <sup>(•)</sup>	$2.711 \pm 0.065$
Planet mass, $m_{p,b}$ [ $M_\oplus$ ] <sup>(*)</sup>	$8.63 \pm 1.35$
Planet density, $\rho_{p,b}$ [ $\text{g cm}^{-3}$ ]	$2.4 \pm 0.4$
Surface gravity, $g$ [ $\text{m s}^{-2}$ ]	$11.5 \pm 1.9$
Escape velocity, $v_{\text{esc},b}$ [ $\text{km s}^{-1}$ ]	$19.9 \pm 1.6$
Equilibrium temperature, $T_{\text{eq},b}$ [K]	
Bond albedo of 0.3	$265 \pm 5$

**Notes.** <sup>(•)</sup> based on the measured  $r_{p,b}/R_s$  of K2-18b from Benneke et al. (2017);  $r_{p,b}/R_s = 0.05295 \pm 0.00060$ .

<sup>(\*)</sup> assuming the measured orbital inclination of K2-18b from Benneke et al. (2017);  $i_b = 89.5785^{+0.0079}_{-0.0088}$  degrees.

Table 6.4: Full HARPS time series from C17a and this work

BJD-2,450,000	RV [m s <sup>-1</sup> ]	$\sigma$ RV [m s <sup>-1</sup> ]	blue RV [m s <sup>-1</sup> ]	blue $\sigma$ RV [m s <sup>-1</sup> ]	red RV [m s <sup>-1</sup> ]	red $\sigma$ RV [m s <sup>-1</sup> ]	NaD	$\sigma$ NaD	H $\alpha$	$\sigma$ H $\alpha$	FWHM	BIS
7117.565870	659.11	4.16	670.79	7.12	652.95	5.75	0.01656	0.00044	0.06339	0.00039	3.067	4.750
7146.526948	656.01	2.82	658.01	4.72	653.12	3.95	0.01275	0.00025	0.06649	0.00027	3.068	-3.030
7146.646070	660.97	3.85	654.96	6.68	661.33	5.28	0.01364	0.00038	0.06854	0.00037	3.079	-6.310
7148.518851	649.04	4.47	645.45	7.99	647.27	6.00	0.01429	0.00047	0.06665	0.00042	3.071	8.580
7148.639664	661.91	5.62	684.89	10.16	652.86	7.44	0.01582	0.00064	0.06578	0.00052	3.069	-12.900
7199.503915	655.69	3.27	657.67	5.47	653.19	4.66	0.01290	0.00031	0.06653	0.00032	3.090	16.920
7200.503114	654.59	2.65	657.05	4.44	657.69	3.77	0.01200	0.00023	0.06628	0.00025	3.080	19.340
7204.491167	648.17	4.25	652.91	7.30	641.60	5.83	0.01257	0.00044	0.06411	0.00040	3.076	-10.070
7390.845075	655.87	2.53	651.28	3.92	658.23	3.85	0.01094	0.00022	0.06724	0.00026	3.106	-0.200
7401.779223	648.08	2.52	641.77	4.03	652.49	3.82	0.01041	0.00021	0.06675	0.00026	3.105	1.060

**Note.** Only the first ten rows of this table are shown to demonstrate its format. The full time series is available in the online published version of this manuscript.

# Chapter 7

## The independent discovery of planet candidates around low mass stars and astrophysical false positives from the first two TESS sectors

### 7.1 Introduction

With our current observational capabilities, nearby transiting planets offer the best targets to characterize exoplanetary systems in detail. By their proximity many of these planets are amenable to follow-up observations to, for example, refine their radii and orbital ephemerides from the ground (Stefansson et al., 2017; Cooke et al., 2018) and from space (Broeg et al., 2013; Gaidos et al., 2017), measure planetary masses via precision radial velocities (Cloutier et al., 2018a), and study their atmospheric compositions, dynamics, and thermal structures (Louie et al., 2018; Kempton et al., 2018). NASA’s *Transiting Exoplanet Survey Satellite* (TESS; Ricker et al., 2015), which launched on April 18 2018, is a purpose-built survey observatory and currently offers the best opportunity to discover nearby transiting planets smaller than Neptune around stars within  $\sim 1000$  pc (Stassun et al., 2017). Indeed TESS has already produced a number of new confirmed planet detections (Esposito et al., 2018; Gandolfi et al., 2018; Huang et al., 2018b; Trifonov et al., 2018; Vanderspek et al., 2018) in addition to its set of TESS Objects of Interest or TESS ‘alerts’.

TESS features four refractive lenses that provide a combined wide field-of-view of  $24^\circ \times 96^\circ$  (i.e.  $\sim 2300$  square degrees) for a single sector. The primary TESS mission splits the sky into 26 equal sectors (13 per hemisphere) anchored on the ecliptic poles and extending towards the ecliptic plane where

---

The contents of this chapter are copied from the current version of the paper Cloutier (2018) that is currently under review by AAS journals. The current version includes some alterations as suggested in the referee’s first appraisal report. Thus, this chapter does not exactly reflect the arXiv version of the Cloutier (2018) manuscript.

fields at the lowest ecliptic latitudes ( $\sim 63\%$  of the sky) will be continuously monitored for  $\sim 27$  days. Conversely, fields centered at the galactic poles ( $\sim 2\%$  of the sky) will be continuously monitored for  $\sim 350$  days and overlap with continuous viewing zone of the *James Webb Space Telescope (JWST)*. In total, TESS will survey  $\sim 85\%$  of the entire sky over its two year-long survey targeting  $2\text{-}4 \times 10^5$  predominantly bright dwarf stars listed in the TESS Input Catalog (TIC; [Stassun et al., 2017](#)) with 2 minute cadence. Full Frame Images for all visible objects within each field will also be released with a 30 minute cadence. From these data products TESS is expected to discover thousands of new transiting exoplanets ([Sullivan et al., 2015](#); [Ballard, 2018](#); [Barclay et al., 2018](#)) plus potentially thousands more from a variety of proposed extended missions ([Bouma et al., 2017](#); [Huang et al., 2018a](#)). The launch of TESS and its recent large data release marks the beginning of a new era of exoplanetary survey science that will carry on the legacy of the famous *Kepler* space telescope which was decommissioned on November 16 2018, after nearly a decade of transformative exoplanet observations and thousands of planet discoveries.

TESS is also unique relative to previous space-based transiting exoplanet survey observatories (e.g. *Kepler*, *CoRoT*) in that its bandpass extends further redward into the near-IR: 600-1000 nm. This enables TESS to access more cool M dwarf stars at high signal-to-noise than previous missions. Systems of sub-Neptune-sized planets are common around M dwarfs ([Dressing & Charbonneau, 2013](#); [Morton & Swift, 2014](#); [Dressing & Charbonneau, 2015](#)) and are required in order to provide a global view of the outcomes of the planet formation process across the Initial Mass Function. Given their lower luminosities relative to Sun-like stars, detecting close-in planets around low mass stars probes a subset of exoplanets with systematically lower equilibrium temperatures including temperate planets orbiting within the habitable zone ([Kasting et al., 1993](#); [Kopparapu et al., 2013](#)). Given their relative abundance within the solar neighbourhood ([Winters et al., 2015](#)), planet masses around nearby low mass stars may be readily characterized with radial velocities to build up a statistically significant view of the mass-radius relationship for small planets ([Weiss & Marcy, 2014](#); [Rogers, 2015](#); [Wolfgang et al., 2016](#); [Chen & Kipping, 2017](#)). The small sizes of low mass stars also works to increase observational signatures of transiting planets thus making their planetary systems of particular interest for the atmospheric characterization of terrestrial to super-Earth-sized planets whose scale heights are expected to be inherently small ( $\lesssim$  [Miller-Ricci et al., 2009](#)) and thus difficult to detect even with state-of-art instrumentation on-board the up-coming *JWST* ([Morley et al., 2017](#)).

The recent public data release from the first two TESS sectors, processed and validated by the TESS Science Processing Operations Center (SPOC; [Jenkins et al., 2016](#); [Twicken et al., 2018](#); [Li et al., 2018](#)), provides an opportunity for members of the extended exoplanet community to pursue a variety of unique science cases. This includes the search for new transiting planets using transit detection algorithms that are independent of those used by the TESS Science Team and on distinct subsets of stars targeted by TESS. In this study, we focus on low mass dwarf stars from the TIC and use Gaia parallaxes to infer precise stellar parameters and refine the sample of probable M dwarfs within the TIC. We then search for transiting exoplanets around these low mass dwarfs in the high cadence TESS light curves using our custom-built transit detection pipeline described herein.

In Sect. 7.2 we present the derivation of our input target list of low mass TIC members. In Sect. 7.3 we present the details of our transit detection pipeline **ORION**. In Sect. 7.4 we present our pipeline results

and our supplementary efforts to classify flagged transit-like events via human vetting and statistical validation before culminating our final list of planet candidates and astrophysical false positives. We conclude with a discussion in Sect. 7.5.

## 7.2 Stellar Sample

### 7.2.1 Initial stellar sample

Our initial stellar sample is retrieved from version 7 of the TESS Input Catalog (TIC-7) which is accessed via the Barbara A. Mikulski Archive for Space Telescope (MAST) Portal<sup>1</sup>. Among other parameters, the TIC-7 table contains estimates of each star’s physical parameters (i.e. effective temperatures  $T_{\text{eff}}$ , surface gravities  $\log g$ , radii  $R_s$ , masses  $M_s$ , etc.), astrometry (either from the *Tycho*-Gaia astrometric solution; [Gaia Collaboration et al. 2016b,a](#) or from *Hipparcos*),  $G$ -band magnitude from the Gaia data release 1, and 2MASS photometry ([Cutri et al., 2003](#)). To identify putative low mass dwarf stars within the TIC-7 we first restrict our sample to sources flagged as dwarf stars based on their 2MASS colors and the reduced proper motion criterion from [Stassun et al. \(2017\)](#), modified from [Collier Cameron et al. \(2007\)](#). We further restrict our sample to stars whose ‘priority’ is  $\geq 10^{-3}$  where the TIC priority metric is based on the relative probability of detecting small planetary transits. As such, the priority is dependent on  $R_s$ , the expected photometric precision, the number of TESS sectors in which the TIC member will be visible, and its contamination ratio: the ratio of contamination to source flux where contamination is computed over ten TESS pixels from the source ( $\sim 3.5$  arcmin).

Next we establish our initial sample of low mass stars based on the physical stellar parameters from the TIC-7 and using the following criteria:

- $T_{\text{eff}} \in [2700, 4200]$  K,
- $\log g > 4$ ,
- $R_s < 0.75 R_{\odot}$ ,
- $M_s < 0.75 M_{\odot}$ .

We note that these criteria are not intended to reflect the exact M dwarf parameter ranges of interest but instead are chosen to be intentionally conservative as to avoid missing any potential M dwarfs prior to their final classification (for use within this study) based on  $T_{\text{eff}}$  and near-IR luminosities ( $M_{K_S} \in [4.5, 10]$ ; [Delfosse et al. 2000](#); [Benedict et al. 2016](#)) that will be refined in Sect. 7.2.2 using Gaia parallaxes. At this stage we find a total of 93090 TIC members that obey our criteria. Of these, 2849 TIC members are observed in one or both of TESS sectors 1 and 2 and are depicted in Fig. 7.1.

### 7.2.2 Refined stellar sample based on Gaia DR2

The stellar parameters used to derive our initial stellar sample were obtained from a variety of sources as outlined in [Stassun et al. \(2017\)](#). Effective temperatures within our sample were predominantly obtained from the cool dwarf catalog ([Muirhead et al., 2018](#)) or alternatively from spectroscopic catalogs or  $V-K_S$

---

<sup>1</sup><https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>

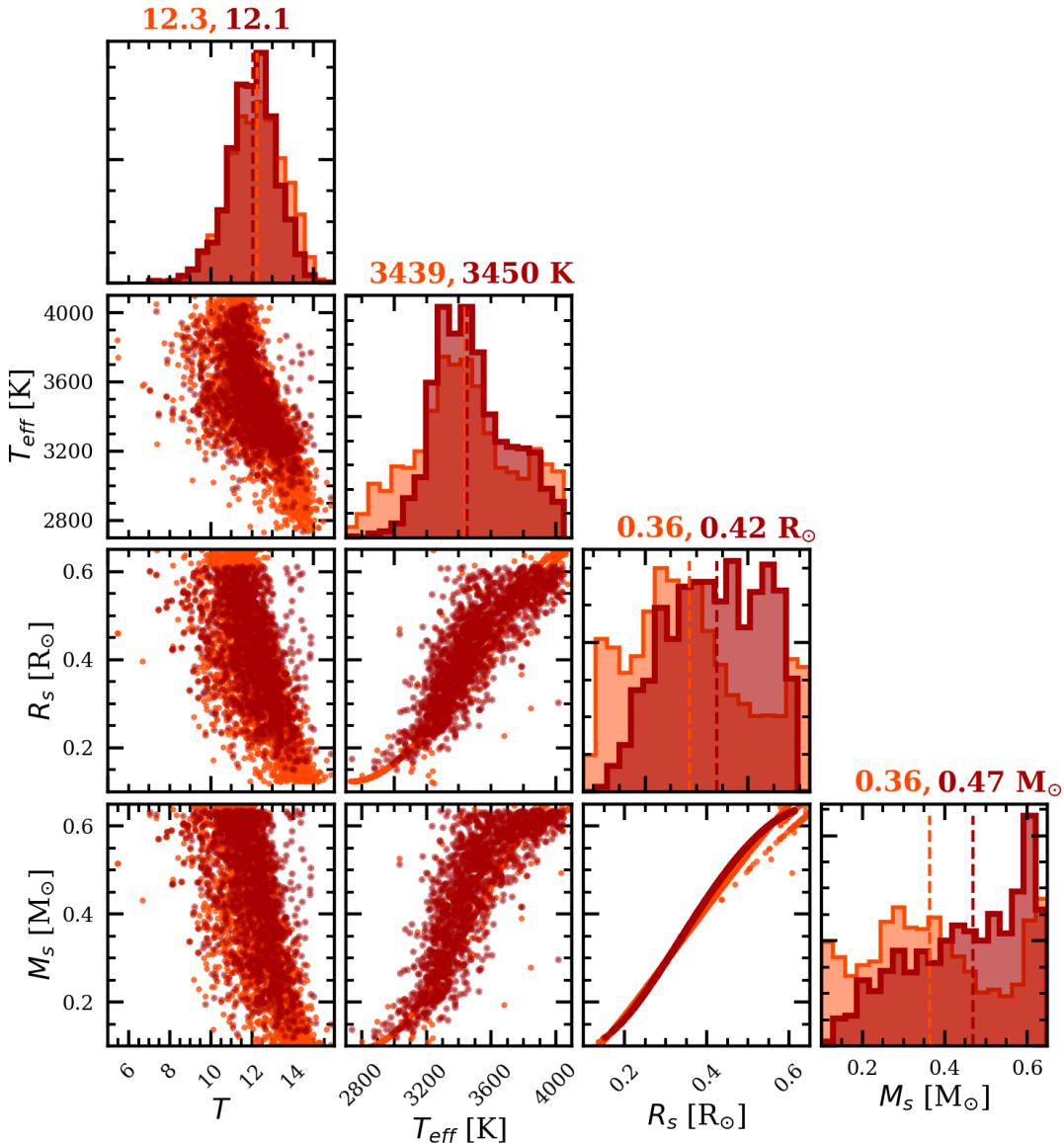


Figure 7.1: Distributions of TESS apparent magnitudes  $T$ , effective temperatures, stellar radii, and stellar masses for our initial (translucent orange markers) and final (dark red markers) stellar samples. Our initial sample contains 2849 low mass stars observed in sectors 1 and/or 2 and are identified by their stellar parameters from the TIC-7. Our final sample contains 1599 low mass stars with refined stellar parameters based on Gaia DR2 parallaxes. The median parameter values for each sample are annotated above each histogram.

colors. Most stellar masses and radii also come from the cool dwarf catalog or from the [Torres et al. \(2010\)](#) spectroscopic relations. Stellar surface gravities follow from measurements of  $R_s$  and  $M_s$ . The TIC-7 stellar radii are typically known at the level of  $\sim 16\%$  which often dominates the error budget of the measured planetary radii from transit observations. Here we aim to produce a homogeneously-derived set of precise stellar parameters by exploiting the exquisite precision of their 2MASS photometry and stellar parallaxes from the Gaia DR2 [Gaia Collaboration et al. \(2018\)](#) available for the majority of stars in our initial sample.

For the accurate and precise characterization of transiting planets we are principally interested in the measurement of host stellar radii  $R_s$ . Additionally, the derivation of other fundamental parameters such as stellar masses and effective temperatures are of importance for a more complete understanding of the effect that stars can have on their host planetary systems. Here, we re-derive the stellar parameters of our initial sample by deriving their near-IR absolute magnitudes coupled with empirically-derived M dwarf radius-luminosity relations ([Mann et al., 2015](#)). We begin by querying the Gaia DR2 archive using the star’s right ascension and declination ( $\alpha, \delta$ ) with a search radius of 10-60 arcseconds. Cross-matching the TIC-7 with the Gaia DR2 data is necessary to obtain each star’s updated parallax from the DR2, additional Gaia photometry (i.e.  $G_{\text{BP}}$  and  $G_{\text{RP}}$ ) which was not included in the TIC-7, and point-estimates of their respective measurement uncertainties that we will approximate as Gaussian distributed. Our querying procedure utilizes the `astropy.astroquery` python package ([Ginsburg et al., 2017](#)). Next we identify source matches according to their predicted photometric colors based on the 2MASS-Gaia color-color relations reported in [Evans et al. \(2018\)](#). Explicitly, we use the quadratic polynomial fits from [Evans et al. \(2018\)](#) to predict each of the colors  $G - K_S$ ,  $G_{\text{BP}} - K_S$ ,  $G_{\text{RP}} - K_S$  and  $G_{\text{BP}} - G_{\text{RP}}$  and then compare those predictions for potential source matches to the measured TIC member colors. Each of the four color-color relations is accompanied by a characteristic scatter of 0.3692, 0.4839, 0.2744, and 0.2144 magnitudes respectively. We claim a source match when all of the calculated colors are within  $3\sigma$  of their predicted values. Based on numerous checks of individual known TIC members we determined that such a tolerance is required to ensure accurate source matches. This dispersion is also expected given that higher order effects not taken into account by the polynomial fits can have stark effects on the accuracy of the photometric predictions.

We proceed with identifying bona-fide low mass TIC members within our initial stellar sample by using the Gaia data of matched sources to refine the stellar parameters that were initially used to flag low mass stars. We will classify low mass stars within this study based on their absolute  $K_S$ -band magnitudes ([Delfosse et al., 2000; Mann et al., 2015; Benedict et al., 2016](#)) and effective temperatures which are derived in the coming sub-sections. We focus our analysis on the  $K_S$ -band in this study due to the reduced effects of dust extinction at near-IR wavelengths compared to in the visible. The absolute  $K_S$ -band magnitude is

$$M_{K_S} = K_S - \mu - A_{K_S}, \quad (7.1)$$

where  $K_S$  is the source’s  $K_S$ -band apparent magnitude,  $\mu = 5 \log_{10} (d/1 \text{ pc}) - 5$  is the distance modulus given the distance to the source  $d$ , and  $A_{K_S}$  is the source extinction in the  $K_S$ -band. Therefore, in order to compute  $M_{K_S}$  for our stellar sample we must first obtain the parameters  $d$  and  $A_{K_S}$ .

### Stellar distances from Gaia

The Gaia DR2 reports precise stellar parallaxes  $\varpi$  for the majority of stars in our initial sample. The typical parallax uncertainty for the stars in our sample is  $\sim 0.2\%$ . As noted by numerous authors (e.g. [Bailer-Jones, 2015](#); [Astraatmadja & Bailer-Jones, 2016](#); [Luri et al., 2018](#)), reliable distances to the majority of stars in the Gaia DR2 cannot be obtained by simply inverting the stellar parallax. Given  $\varpi$  values with posterior probability density functions (PDFs) that are presumed to be Gaussian distributed, making them fully described by their mean values and  $1\sigma$  dispersions, the non-linearity of the transformation from  $\varpi \rightarrow d$  will result in an asymmetric  $d$  posterior PDF whose skewness is dependent on the absolute  $\varpi$  measurement value and its signal-to-noise ([Luri et al., 2018](#)). By the proximity of the majority of sources in the TIC-7, their parallaxes are measured with high precision such that the resulting  $d$  PDF obtained using the standard formula  $(d/\text{pc}) = (\varpi/\text{arcsec})^{-1}$  can be well-approximated as a Gaussian distribution ([Bailer-Jones et al., 2018](#)). The median distance and distance uncertainty for our sample is 53.55 and 0.13 pc ( $\sim 0.2\%$  fractional uncertainty) which are sufficiently small distances such that the transformation to  $d$  from  $\varpi$  results in approximately Gaussian distributions of  $d$ . The maximum a-posteriori (MAP)  $d$  value and its  $1\sigma$  uncertainty are then propagated to the calculation of  $\mu$  which we will ultimately use in Eq. 7.1 to calculate  $M_{K_S}$  after the extinction coefficients are obtained (see Sect. 7.2.2).

There are known systematic effects in the Gaia astrometric solution in the form of a non-zero parallax zero-point that is dependent on the source position,  $G$ -band magnitude, and possibly color ([Lindegren et al., 2018](#)). In computing  $d$  from  $\varpi$  we first apply a simple correction by adding the globally-averaged parallax zero-point of 29  $\mu\text{as}$  ([Lindegren et al., 2018](#)) to the verbatim stellar parallaxes from the Gaia DR2. Although this correction may not be strictly accurate for our M dwarf target sample ([Leung & Bovy, 2019](#)), the effect of differences in the zero-point are marginal compared to the derived stellar parameter uncertainties in the coming sections.

### Source extinction estimates

The source extinction is dependent on the source's location on the sky and particularly on its proximity to the galactic plane where the dust column density is highest. To estimate the  $K_S$ -band extinction for each source we utilize the `mwdust` package ([Bovy et al., 2016](#)) which queries one of three E(B-V) reddening maps (i.e. [Drimmel et al., 2003](#); [Marshall et al., 2006](#); [Green et al., 2015](#)) based on the applicability of each map to the input source position. Given the source's galactic coordinates  $(l, b)$ , Gaia distance, and uncertainties as input, `mwdust` queries the reddening maps and returns the extinction coefficient  $A_\lambda = R_\lambda E(\text{B-V})$  in the desired band using the extinction vector scaling  $R_{K_S} = 0.31$  from [Schlafly & Finkbeiner \(2011\)](#). Uncertainties in  $A_{K_S}$  are derived from the  $d$  measurement uncertainty and from inherent uncertainties in the value of  $R_{K_S}$  ([Green et al., 2018](#), e.g.) which we attempt to account for via the quadrature addition of a 30% fractional uncertainty on  $R_{K_S}$  following the methodology of ([Fulton & Petigura, 2018](#)).

### Deriving the set of refined stellar radii

Combining the retrieved values of  $K_S$ ,  $\mu$ , and  $A_{K_S}$  into Eq. 7.1 returns the distribution of  $M_{K_S}$  for all of the 2489 stars in our initial sample for which 2MASS photometry and Gaia parallaxes are available.

Calculations of M dwarf stellar radii from their bolometric magnitudes would require  $K_S$ -band bolometric corrections which for cool stars are known to frequently suffer from comparatively large inaccuracies ( $T_{\text{eff}} \lesssim 4100$  K; Berger et al., 2018). We therefore adopt the alternative approach from Berger et al. (2018) which uses the empirically-derived M dwarf radius-luminosity relation (RLR) from Mann et al. (2015) to update M dwarf stellar radii in the *Kepler* field using Gaia distances. The fitted RLR employs a quadratic function to map  $M_{K_S}$  to directly measured  $R_s$  from the combination of interferometry and parallaxes. Because we are interested in deriving a self-consistent sample of low mass stellar radii we restrict our analysis to TIC members with  $M_{K_S}$  values that are applicable to the Mann et al. (2015) RLR which is valid for M dwarfs with  $M_{K_S} \in (4.6, 9.8)$ . This condition will be used to establish our final sample of low mass dwarf stars following the derivation of  $T_{\text{eff}}$  within our initial sample. The radii inferred from the RLR have a fractional residual dispersion of 0.0289  $R_s$  which we add in quadrature to the radius uncertainty propagated from  $M_{K_S}$ .

Fig. 7.2 compares the TIC-7 stellar radii (compiled from various input sources) with those derived from Gaia distances and the M dwarf RLR from Mann et al. (2015). The relation is largely one-to-one but with a slight translation of the updated  $R_s$  distribution to larger radii ( $\sim 3.7\%$  median increase). The effect is already known (Berger et al., 2018) and is the result of many sources having their distance measures increased following the release of the Gaia DR2 parallaxes. More notably for the measurement of transit planet radii is the significant reduction in the fractional radius uncertainty as evidence in the histogram included in Fig. 7.2. The typical fractional radius uncertainty  $\sigma_{R_s}/R_s$  within our updated sample is  $\sim 4 - 5$  smaller than in the TIC-7. The median fractional uncertainty on our Gaia-derived stellar radii is  $\sim 3\%$ .

### Deriving the set of refined stellar effective temperatures

Similarly to the RLR, Mann et al. (2015) also parameterized an empirically-derived M dwarf temperature-color-metallicity relation (TCMR). Our sample contains both 2MASS and Gaia photometry so we use these photometric systems to infer  $T_{\text{eff}}$  for the stars in our newly refined sample. Specifically, we adopt the fitted TCMR from Mann et al. (2015) which is cubic in  $G_{\text{BP}} - G_{\text{RP}}$  and quadratic in  $J - H$ . The latter color is used as a proxy for metallicity (Leggett, 1992; Johnson et al., 2012; Mann et al., 2013; Newton et al., 2014). The TCMR used here has a residual temperature dispersion of 49 K which we add in quadrature to the uncertainties in  $T_{\text{eff}}$  propagated from the input photometric uncertainties.

### Deriving the set of refined stellar masses

We revise the stellar masses using the empirically-derived M dwarf mass-luminosity relation (MLR) from Benedict et al. (2016). Their fitted relation uses a quartic function to map  $M_{K_S}$  to directly measured  $M_s$  from dynamical analyses of binary star systems and is valid for M dwarfs with  $M_{K_S} \leq 10$ . This valid range of  $M_{K_S}$  is consistent with the range required for the Mann et al. (2015) empirical relations. The five fitted coefficients that parameterize the MLR all have an associated uncertainty which we sample from, along with  $M_{K_S}$  sampled from a Gaussian distribution, to infer the  $M_s$  PDF. Point estimates of each star's  $M_s$  and uncertainty come from the MAP of the PDF and the average of its 16<sup>th</sup> and 84<sup>th</sup> percentiles.

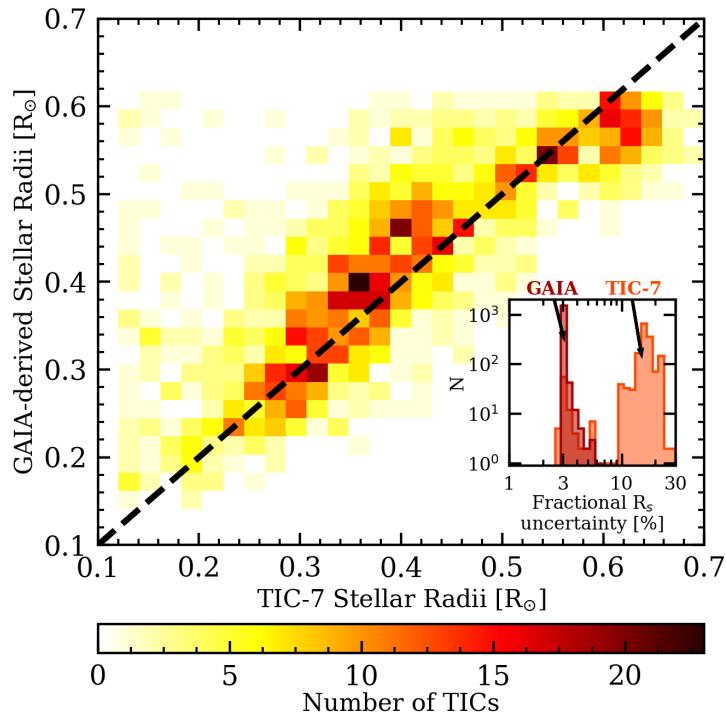


Figure 7.2: A 2-dimensional histogram comparing the stellar radii in our sample of 1599 low mass stars, derived from Gaia parallaxes and the M dwarf radius-luminosity relation from [Mann et al. \(2015\)](#), to those from the TESS Input Catalog (TIC-7) which are compiled from a variety of sources. The subpanel compares the distributions of fractional stellar radius uncertainties in each catalog. The refined stellar radii based on Gaia parallaxes have a typical precision improvement of  $\sim 4 - 5$  compared to the TIC-7 parameters.

### Final stellar sample

Using the refined stellar parameters obtained from cross-matching putative low mass stars from the TIC-7 with the Gaia DR2, we now construct our final stellar sample as stars that obey the following criteria:

- $M_{K_S} \in (4.6, 9.8)$
- $T_{\text{eff}} - \sigma_{T_{\text{eff}}} < 4000 \text{ K}$ ,
- $\log g + \sigma_{\log g} > 4$ ,
- $R_s - \sigma_{R_s} < 0.75 \text{ R}_\odot$ ,
- $M_s - \sigma_{M_s} < 0.75 \text{ M}_\odot$ .

That is that we retain all TIC members whose luminosities, effective temperatures, surface gravities, radii, and masses are consistent with those of late-K to M dwarf parameter ranges ([Pecaut & Mamajek, 2013](#)) at the  $1\sigma$  level.

Our final stellar sample contains 1599 low mass stars with 537, 694, and 368 observed within TESS sectors 1, 2, and both respectively. The final stellar sample is over-plotted in Fig. 7.1. The dispersion in the joint  $T_{\text{eff}}, R_s, M_s$  space for our final stellar sample and the TIC-7 sample are comparable indicating that we have likely derived physically consistent stellar parameters within our final sample. The distribution of  $T$  in our stellar samples spans 7-15 with a median  $T = 12.1$ . The distribution of effective temperatures extends from 2740-4040 K with a median  $T_{\text{eff}} = 3450 \text{ K}$  whose approximate spectral type is M3V ([Pecaut & Mamajek, 2013](#)). The  $T_{\text{eff}}$  distributions from the TIC-7 and our Gaia-derived values are roughly consistent. The stellar radii span  $0.15\text{-}0.61 \text{ R}_\odot$  with a median  $R_s = 0.42 \text{ R}_\odot$ . The stellar masses span  $0.12\text{-}0.63 \text{ M}_\odot$  with a median  $M_s = 0.47 \text{ M}_\odot$ . Owing to the increased distances of many TIC members in our final sample, the distributions of  $R_s$  and  $M_s$  are both translated to slightly larger radii and masses compared to the values listed in the TIC-7.

## 7.3 Overview of the ORION transit detection pipeline

Here we present our independent transit detection pipeline **ORION** that borrows many of the strategies and vetting procedures from established methods focused on transit detections primarily with *Kepler* and *K2* (see references herein). Our pipeline can be thought of as having six sequential steps that take as input the TIC identifier and stellar parameters to produce a set of transiting planet candidates (PCs) with measured orbital periods  $P$ , times of mid-transit  $T_0$ , scaled semi-major axes  $a/R_s$ , scaled planetary radii  $r_p/R_s$ , and orbital inclinations  $i$ . The six steps within **ORION** are 1) to obtain the extracted TESS light curves and ancillary data for each input TIC member, 2) to derive an initial systematic model for light curve de-trending, 3) to perform a linear search for transit-like events, 4) to perform a periodic search for repeating transit-like events, 5) to subject putative PCs to a set of vetting criteria in an automated way, and 6) to re-model the light curve with a joint systematics plus transit model for all vetted transit-like events. These stages are described in detail in the subsequent sections.

### 7.3.1 TESS light curve acquisition

The execution of `ORION` on a TIC member begins with downloading the star’s publicly available 2 minute TESS extracted light curves and target pixel files for all available sectors. The TESS data is downloaded from the MAST data service<sup>2</sup>. Only TIC members observed at 2 minute cadence are considered at this time with their extracted light curve made available following its processing by the TESS SPOC. Efforts to extract 30 minute light curve data from the TESS Full Frame Images and significantly expand the list of TESS targets accessible to `ORION` are underway but are reserved for future work. Target pixels files are principally used to quickly assess the data quality and will be used to infer the TIC member’s point spread function during the statistical validation of putative PCs in Sect. 7.4.2.

For each available sector of data the chronologically sorted vectors of observing times  $\mathbf{t}$ , measured fluxes  $\mathbf{f}$ , and the associated  $1\sigma$  flux uncertainties  $\sigma_f$  are constructed. Fluxes are obtained from the Simple Aperture Photometry Pre-search Data Conditioning extraction which includes artifact mitigation (Smith et al., 2012). These vectors are attributed to the following fields: `TIME` [BJD], `PDCSAP_FLUX` [ $e^-/s$ ], and `PDCSAP_FLUX_ERR` [ $e^-/s$ ]. The flux and flux uncertainty vectors are converted into normalized flux units via division by  $\text{median}(\mathbf{f})$ .

### 7.3.2 Initial light curve de-trending

Residual systematic effects are clearly visible in the many of the extracted light curves. Due to the inherent photometric and pointing precision of the first TESS sectors, these systematic effects are often largely attributable to astrophysical noise sources such as flicker (Bastien et al., 2013) in Sun-like stars but more commonly for low mass stars from large-scale variability caused by active regions on the rotating stellar surface. As an initial de-trending step to correct for temporally-correlated noise sources from either systematics or intrinsic stellar phenomena, a semi-parametric Gaussian process (GP) regression model is fit to the extracted TESS photometry.

GP regression models provide a flexible and probabilistic framework to model the temporal covariances between photometric measurements following the removal of a mean model  $\mu(\theta)$  which is parameterized by the set of observable parameters  $\theta$  (e.g.  $P$ ,  $T_0$ ,  $a/R_s$ , etc.). The posterior PDFs of the  $\theta$  elements can be sampled simultaneously with the GP hyperparameters  $\Theta$  which parameterize the residual covariances through the covariance matrix  $K(\Theta)$  and are fit by optimizing the ln likelihood function

$$\ln \mathcal{L}(\theta, \Theta) = -\frac{1}{2} [(\mathbf{f} - \mu(\theta))^T \cdot K(\Theta)^{-1} \cdot (\mathbf{f} - \mu(\theta)) + \ln \det K + N \ln 2\pi] \quad (7.2)$$

along with appropriately chosen priors on the parameters in  $\theta$  and  $\Theta$ . Here,  $N$  is the number of photometric measurements in the light curve. Similar routines based on GP regression have been adopted for  $K2$  systematic corrections (e.g. Aigrain et al., 2015; Crossfield et al., 2015; Aigrain et al., 2016) and can also be used to infer accurate photometric stellar rotation periods (Angus et al., 2018). For cases in which the origin of the apparent photometric variations are likely dominated by active regions on rotating spotted stars, the resulting photometric signal will vary non-sinusoidally as the active regions evolve in size, brightness, and location over the observational baseline. This physically motivates the

---

<sup>2</sup><https://archive.stsci.edu/tess/index.html>

use of a quasi-periodic covariance matrix  $K_{i,j} = \delta_{i,j}\sigma_{f,i} + k_{i,j}$  where  $\delta_{i,j}$  is the Kronecker delta and  $k_{i,j}$  is the covariance kernel of the form

$$k_{i,j} = a_{\text{GP}}^2 \exp \left[ -\frac{|t_i - t_j|^2}{2\lambda^2} - \Gamma^2 \sin^2 \left( \frac{\pi|t_i - t_j|}{P_{\text{GP}}} \right) \right]. \quad (7.3)$$

The covariance kernel is parameterized by the four hyperparameters  $\Theta = \{a_{\text{GP}}, \lambda, \Gamma, P_{\text{GP}}\}$  where  $a$  is the correlation amplitude,  $\lambda$  the exponential timescale,  $\Gamma$  the coherence scale, and  $P_{\text{GP}}$  the periodic timescale of the correlations. Moreover, a quasi-periodic covariance kernel is favorable for cases in which the origin of apparent photometric variations are dominated by systematics which need not have a strong periodic component. In this limit, the coherence parameter  $\Gamma$  approaches small values such that the covariance kernel becomes well-approximated by a squared exponential kernel with effectively a single timescale  $\lambda$ . Note that because systematic and astrophysical noise sources within the GP noise model are not distinguished, the fitted hyperparameter values are unable to be used to interpret the origin of the photometric variability. Furthermore, during the remainder of this paper the covariance structures modelled by the GP will be solely referred to as ‘systematics’ despite the possibility that their (partial) origin may be astrophysical.

The logarithmic hyperparameters are initialized and subsequently optimized in an iterative manner and are performed on each TESS sector independently assuming a null mean function (i.e.  $\mu = \mathbf{0}$ ). The periodic GP timescale  $P_{\text{GP}}$  is initialized by peaks in the Lomb-Scargle periodogram (LSP; Scargle, 1982) of the extracted light curve whereby in each of the iterations performed,  $P_{\text{GP}}$  is initialized to the  $i^{\text{th}}$  most significant peak in the LSP where  $i$  is the iteration index  $\in [1, 10]$ . Following the use of GP regression modelling for radial velocity (RV) activity mitigation in Dittmann et al. (2017b), where the physical source of activity is largely common between the optical RVs and broadband TESS photometry,  $\ln \lambda$  is initialized to  $\ln 3P_{\text{GP}}$ . In each iteration,  $\ln \Gamma$  is initialized to 0 and  $\ln a_{\text{GP}} = \ln \max(|\mathbf{f}_{\text{bin}} - \text{median}(\mathbf{f}_{\text{bin}})|)$  where  $\mathbf{f}_{\text{bin}}$  is the vector of binned photometric points whose temporal bin width is set such that a single periodic GP timescale is sampled by at least eight measurements.

For each iteration in each TESS sector (if multiple sectors are available), the uniquely initialized GP hyperparameters are optimized using the `scipy.optimize.minimize` python function to minimize the negative  $\ln \mathcal{L}(\Theta)$  from Eq. 7.2 given the Jacobian of  $\ln \mathcal{L}(\Theta)$  with respect to the hyperparameters in  $\Theta$ . During optimization the  $\ln$  GP hyperparameters are bounded by broad uninformative priors which are explicitly reported in Table 7.1. Broad  $\ln$  uniform priors enable the generalization of the ORION de-trending method across all of the input TESS light curves which greatly benefit from semi-parametric modelling given the wide range of covariance timescales exhibited by TIC members in photometry. Given an optimized set of hyperparameters, the resulting GP posterior PDF is a  $N$ -dimensional multi-variate Gaussian distribution whose mean function is taken to be a potential systematic correction. The mean function of the GP from the iteration whose optimized hyperparameters maximize  $\ln \mathcal{L}(\Theta)$  is assigned as the initial systematic correction and is used to de-trend the photometry prior to the search for periodic transit events.

Fig. 7.3 depicts two examples of the results of this iterative de-trending procedure over individual TESS sectors for TIC members 235037759 and 262530407. The accuracy of each mean GP regression

Table 7.1: Model parameter priors

Parameter	Prior
<i>GP hyperparameters<sup>a</sup></i>	
Covariance amplitude, $\ln a_{\text{GP}}$	$\mathcal{U}(-20, 0)$
Exponential timescale, $\ln \lambda/\text{days}$	$\mathcal{U}(-3, 10)$
Coherence, $\ln \Gamma$	$\mathcal{U}(-5, 5)$
Periodic timescale, $\ln P_{\text{GP}}/\text{days}$	$\mathcal{U}(-3, 10)$
<i>Transit model parameters</i>	
Orbital period, $P$ [days]	$\mathcal{U}(0.9, 1.11) \cdot P_{\text{opt}}^{\text{b}}$
Time of mid-transit, $T_0$ [BJD-2,457,000]	$\mathcal{U}(-1.11, 1.11) \cdot P_{\text{opt}} + T_{0,\text{opt}}$
Scaled semimajor axis, $a/R_s$	$\mathcal{U}(0.58, 1.70) \cdot (a/R_s)_{\text{opt}}$
Planet-star radius ratio, $r_p/R_s$	$\mathcal{U}(0, 1)$
Orbital inclination, $i$	$\mathcal{U}(-1, 1) \cdot i((a/R_s)_{\text{opt}}, b = 1)^{\text{c}}$
<i>Single transit model parameters</i>	
Orbital period, $P$ [days]	$\mathcal{J}(1, 100) \cdot P_{\text{inner}}^{\text{d}}$
Time of mid-transit, $T_0$ [BJD-2,457,000]	$\mathcal{U}(-3, 3) \cdot D$
Scaled semimajor axis, $a/R_s$	$\mathcal{J}(1, 100) \cdot (a/R_s)_{\text{inner}}$
Planet-star radius ratio, $r_p/R_s$	$\mathcal{U}(0, 1)$
Orbital inclination, $i$	$\mathcal{U}(-1, 1) \cdot i((a/R_s)_i, b = 1)$

**Notes.**

<sup>a</sup> GP hyperparameter priors used during de-trending (i.e. with zero mean model) and during the simultaneous systematics plus transit modeling.

<sup>b</sup> The designation ‘opt’ is indicative of the optimized parameter values from the maximum likelihood model used for parameter initialization.

<sup>c</sup> The function  $i(a/R_s, b) = a \cos i/R_s$  returns the orbital inclination given  $a/R_s$  and the impact parameter  $b$  which is constrained to  $|b| < 1$  in our transit models.

<sup>d</sup> The designation ‘inner’ is indicative of the inner-most orbital period permissible for a single transit event over the TESS baseline.

model is clearly demonstrated. The systematics model for TIC 235037759 is required to be much more aggressive than that for TIC 262530407 given the star's large photometric variability with a peak-to-peak amplitude in the binned light curve of 280,000 ppm and a 85,000 ppm rms. Unlike in the raw light curve, the de-trended light curve lacks any low frequency variations and exhibits a significantly reduced rms of 12,000 ppm.

The TIC 262530407 light curve also exhibits photometric variability albeit with a much lower peak-to-peak amplitude and rms of 1700 and 1000 ppm respectively. After de-trending, the rms is slightly reduced to 810 ppm. The most important residual feature of the de-trended light curves is that they appear free of the majority of large-scale systematic effects. This fact will facilitate the linear search for transit-like events with minimal contamination from residual systematic features. Indeed a transiting planet candidate is detected around each of these systems although the putative PC around TIC 235037759 is ultimately favored by an astrophysical false positive interpretation as presented in Sect. 7.4.3.

One notable limitation to the systematics correction occurs during a period of a prevalent increase in rms for TIC 235037759 between BJD - 2,4570,00  $\sim$  1347 – 1349 during a brief period of loss of TESS pointing precision. This is a common feature to the TIC members observed during TESS sector 1. During this time the extracted light curve is only partially corrected for the pointing precision loss while the systematic GP model provides only marginal improvements if any at all. To ensure that the GP hyperparameters were not being strongly affected by the anomalous systematics structure during this period, those measurements were masked and the GP hyperparameters re-optimized with the remaining data. The resulting GP systematics only varies marginally from that shown in Fig. 7.3 for TIC 235037759 such that we are confident that the **ORION** de-trending is largely robust to the loss in pointing precision for TIC members observed in TESS sector 1.

Recall that in the initial de-trending step discussed in this section that the methodology assumes a null mean function which implies that any transient events such as flares or transits are still present during the optimization of the de-trending model. The principal caveat with this methodology is that one cannot guarantee that the GP model does not (partially) capture any of the in-transit light curve deprecations that **ORION** is searching for. If partially suppressed by the initial GP model, planets will be more difficult to detect due to the reduced signal-to-noise (S/N) of individual transit events. Furthermore, transit events that remain detectable within the de-trended light curve could result in underestimated transit depths and correspondingly smaller planet sizes. To ensure a self-consistent planet+systematics model for putative PCs from **ORION**, the light model is revisited in Sect. 7.3.6 with the inclusion of a transiting planet mean function in place of the null mean function used during the de-trending step.

### 7.3.3 Linear transit search

Next a linear search for individual transit-like events is conducted on the de-trended light curves over their full duration. The following methodology is reminiscent of a number of individual transit event search algorithms (e.g. Box Least Squares, **BLS**; Kovács et al. 2002, Transiting Planet Search, **TPS**; Jenkins et al. 2010; Christiansen et al. 2013, 2015, 2016, **TERRA**; Petigura et al. 2013b, Foreman-Mackey et al. 2015). The aim here is to identify high S/N transit-like events along with their associated mid-transit times  $T_0$ , durations  $D$ , and depths  $Z$  which will feed into the search for repeating transit-like events and

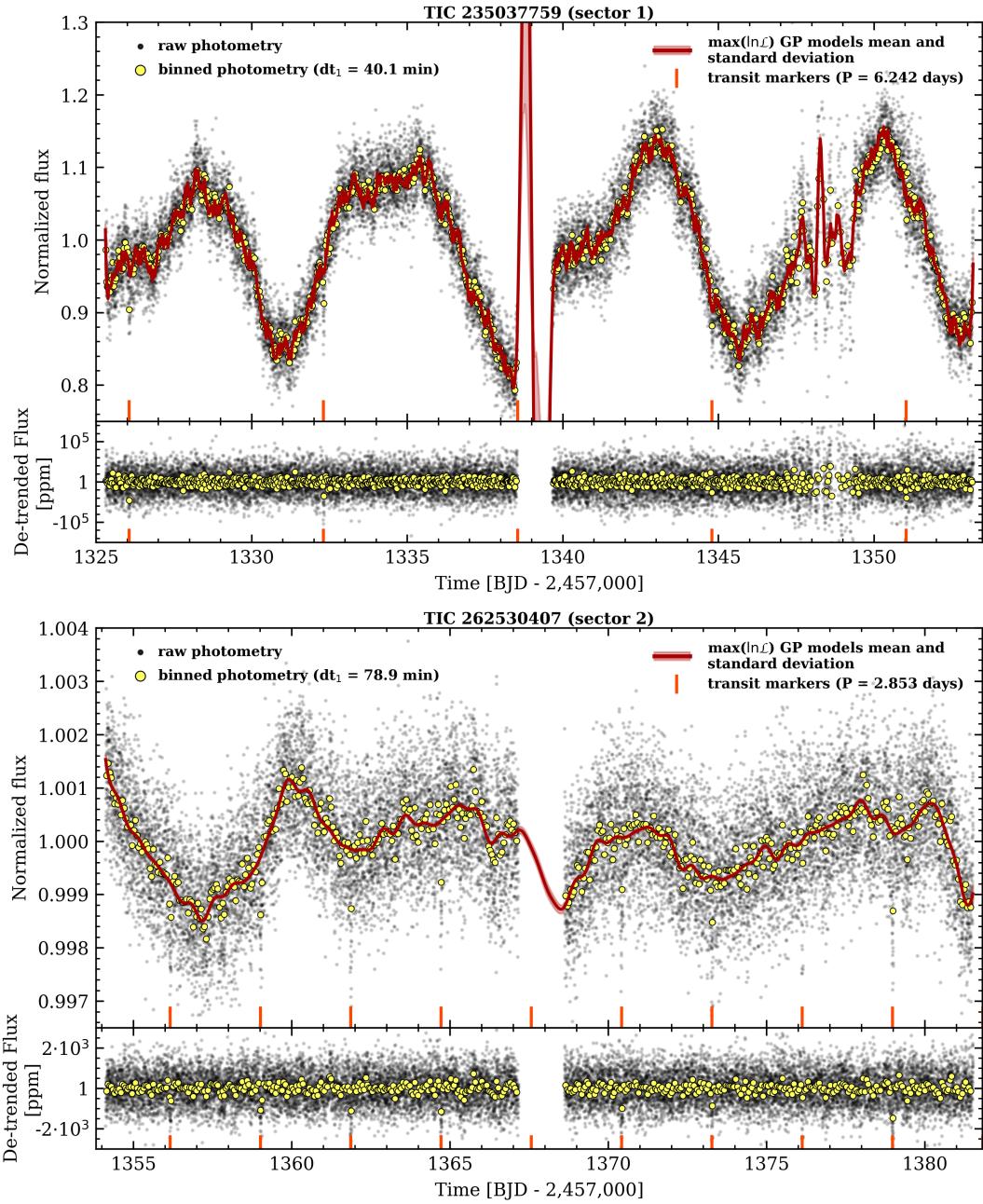


Figure 7.3: Two demonstrations of the ORION initial de-trending stage for TIC 235037759 and TIC 262530407. *Large panels*: the 2 minute raw and binned TESS light curves with temporal binning chosen to sample eight GP periodic timescales (40.1 and 78.9 minute bins for TIC 235037759 and 262530407 respectively). Solid red curves and their surrounding shaded regions depict the mean GP model used for de-trending and its  $1\sigma$  confidence intervals which are often small and difficult to visualize everywhere except during the data transfer gap at the centers of each light curve. Vertical ticks along the abscissa axes are indicative of the transit times of planet candidates flagged by ORION at 6.2 and 2.9 days around TIC 235037759 and TIC 262530407 respectively. *Shallow panels*: the raw and binned light curves after de-trending.

ultimately the list of putative transiting PCs.

The linear search for transit-like events begins with stepping through a two-dimensional grid of  $T_0$  and  $D$ . At each  $(T_0, D)$  grid point a simple box model of the form

$$m(t) = \begin{cases} 1 - Z & \text{if } T_0 - \frac{D}{2} \leq t \leq T_0 + \frac{D}{2} \\ 1 & \text{otherwise,} \end{cases} \quad (7.4)$$

is constructed with fixed  $T_0$  and  $D$ . The box depth  $Z$  (or mock transit depth) is fit by ln likelihood maximization and saved along with the value of ln  $\mathcal{L}$  given the unique set of parameters  $\{T_0, D, Z\}$ . Computing ln  $\mathcal{L}$  with Eq. 7.2 implicitly assumes that the flux uncertainties are Gaussian distributed which allows for the construction of a diagonal covariance matrix  $K$  with elements  $K_{i,j} = \delta_{i,j}\sigma_{f,i}$ . The linear search along the  $T_0$  dimension proceeds by stepping through the observation epochs  $t$  in 30 minute bins and assigning  $t_{\text{bin},i}$  to  $T_0 \forall i = 1, \dots, N_{\text{bin}}$ . This fixed binning is the first of many ORION free parameters listed in Table 7.2 along with their default values and brief explanations of their effects. Initial ORION tests on synthetic light curves with injected transit models determined that finer temporal binning did not result in a significant variation in the number of detected high S/N transit-like events. This is likely due to 30 minute bins being more comparable to typical transit durations of the types of planets that can be detected in 27-54 day baselines. Although notably a phase dependence on the sensitivity to short period planets should persist unless the bin width is reduced (Kovács et al., 2002).

The  $D$  dimension is sampled on a much coarser grid given that the precision of the box model parameters are not yet required to infer planet properties but only to identify epochs at which transit-like events are likely to have occurred. Explicitly, the adopted linear search  $D$  grid contains three possible transit durations of either 1.2, 2.4, or 4.8 hours.

Table 7.2: Descriptions of the free parameters controlling the performance of ORION

Parameter	Definition	Default Value	Summary of Behavior
<i>Linear search parameters</i>			
$\Delta t$	Linear search temporal bin width	30 minutes	Decreasing $\Delta t$ will improve sensitivity to ultra-short period planets while increasing the ORION runtime number of signals for confusion in the periodic search stage.
<i>Periodic search parameters</i>			
$f_P$	Maximum relative difference between two periods to be flagged as a multiple	0.01	Increasing $f_P$ makes fewer period pairs consistent with being multiples thus increasing the sensitivity to resonant planet pairs while simultaneously increasing the number of single planets misidentified as a resonant pair.
<i>Automated light curve vetting parameters</i>			
$c_1$	Minimum transit S/N	8.4	Increasing $c_1$ prevents the detection of some small planets but will also significantly reduce the number false positives due to residual systematics.
$c_2$	Minimum number of MADs from the out-of-transit flux that the difference in median in and out-of-transit fluxes must exceed	2.4	Similar behavior to $c_1$ .
$c_3$	Minimum fraction of in-transit points	0.7	Increasing $c_3$ may result in more accurate determinations of

	below $Z + \sigma_Z$		
$c_4$	Minimum fraction of in-transit points prior to $T_0$	0.1	correct periods but will also cause some transits to be discarded if residual noise is also present during the transit. Increasing $c_4$ increases sensitivity to asymmetric transit shapes such as those from disintegrating planets.
$c_5$	Minimum number of MADs for a flare above the flux continuum	8	Increasing $c_5$ makes flare detection more robust but at the risk of missing some lower amplitude flares.
$c_6$	Minimum number of successive points within a flare duration	2	Increasing $c_6$ decreases sensitivity to flares of short relative to the light curve cadence.
$c_7$	Assumed M dwarf flare duration	30 minutes	Increasing $c_7$ decreases the sensitivity to long duration flares while increasing the assumed fraction light curve fraction that is contaminated by flares.
$c_8$	Number of transit durations from $T_0$ which cannot be affected by flares	4	Decreasing $c_8$ decreases the probability of a transit being contaminated by flares.
$c_9$	Maximum time from the light curve edges to not be rejected due to possible contamination at the edges.	4.8 hours	Increasing $c_9$ decreases the probability of a transit being affected by light curve systematics at its edges but also narrows the baseline over which transits can be found.
$c_{10}$	Minimum autocorrelation of flux residuals	0.6	Increasing $c_{10}$ improves the robustness of transit detections but decreases the detection sensitivity in light curves with imperfect systematics corrections.

---

*Eclipsing binary vetting parameters*

---

$c_{EB,1}$	Maximum $r_p/R_s$ of a transit-like event	0.5	-
$c_{EB,2}$	Maximum planet radius	$30 R_\oplus$	-
$c_{EB,3}$	Maximum transit duration, $D(r_p, P, a/R_s, Z, i)$	$D(30 R_\oplus,$ $P, a/R_s, Z, i)$	-
$c_{EB,4}^a$	Minimum occultation S/N of an EB (Eq. 7.7)	5	Decreasing $c_{EB,4}$ makes a larger fraction of occultations consistent with being due to an EB rather than a transiting planet.
$c_{EB,5}^a$	Minimum fraction of iterative occultation searches consistent with an EB	0.5	Increasing $c_{EB,5}$ makes transit-like events more robust as the probability of being flagged as an EB is reduced.

$c_{\text{EB},6}^{\text{a}}$	Minimum ingress plus egress time fraction of $D$ for a ‘V’-shaped transit	0.9	Increasing $c_{\text{EB},6}$ makes fewer transit-like events consistent with having a ‘V’-shaped transit.
------------------------------	--	-----	--

**Notes.**

<sup>a</sup> These eclipsing binary (EB) parameters are intended to identify favorable EBs rather than transiting planets. Planetary signals of interest are rejected if any of the EB criteria are satisfied.

This procedure produces a  $N_{\text{bin}} \times 3$  matrix of transit times and durations with each entry having an associated ln likelihood and optimized depth  $Z$ . From the ln likelihoods a S/N spectrum as a function of transit times is computed for each  $D$  value considered. By translating the ln likelihoods by their median value and normalizing by their median absolute deviation (MAD), the aforementioned linear search S/N spectrum versus transit times is calculated. The conversion from ln likelihoods to the ad hoc S/N spectrum centered around zero aids in its interpretation as each TIC member's spectrum can be searched in absolute terms. In adopting the median and MAD ln likelihood values over the mean and standard deviation, the S/N is less sensitive to contamination by stochastic, short timescale photometric features and results in a S/N spectrum whose baseline is dominated by the light curve's inherent photometric precision. An example linear search S/N spectrum is shown in Fig. 7.4 for a fixed duration of 1.2 hours. We note that referring to the linear search S/N spectra as a S/N is a misnomer given that the spectrum values can be negative either in the presence of noise or due to an apparent brightening of the source. However we find this language to be a clear descriptor of the spectrum's aim and its interpretability.

The S/N spectra are then searched for high S/N transit-like events similarly to the individual significance events in the primary Kepler transit search which are combined into a multi-event statistic and flagged as a Threshold Crossing Event ([Jenkins et al., 2010](#)). Thus, the nature of the linear search is for single-event statistics and occurs prior to the periodic search step (see Sect. 7.3.4) for multi-event statistics in which the data are folding and repeating events are investigated. High S/N transit-like events are flagged as peaks in the linear search S/N spectrum when  $S/N_{\text{thresh}} \geq 5$ . All  $n_T$  transit times  $\mathbf{T}_0$  with S/N exceeding  $S/N_{\text{thresh}}$  for any value of  $D$  are compiled into a set of potential transit-like events. Because transit times in the linear search are sampled on a fixed grid (i.e.  $t$  binned to  $\Delta t$ ), each transit time is refined by Gaussian smoothing the light curve around  $\pm 2D$  of  $T_0$  using `scipy.ndimage.filters.gaussian_filter` and updating  $T_0$  to the central time of the box model minimum before proceeding to the search for periodic events that may be indicative of transiting PCs. In the example shown in Fig. 7.4, a PC exists with  $\sim 1.4$  day orbital period. Six out of the nineteen transit events that occur within the sector 1 baseline are detected above  $S/N_{\text{thres}}$ . This includes two consecutive transits between 1341 and 1343 BJD - 2,457,000 which are used in the subsequent section to infer its possible period equal to the time difference between the two events.

### 7.3.4 Periodic transit search

The chronologically sorted set of  $n_T$  high S/N transit times in  $\mathbf{T}_0$  are used to construct a matrix of differential transit times with elements  $P_{i,j} = T_{0,i} - T_{0,j} \forall i, j = 1, \dots, n_T$ . A separate matrix is populated for each unique value of  $D$ . An example of P is shown in Fig. 7.5 for TIC 234994474 for the fixed duration of 1.2 hours. The  $n_T \times n_T$  matrix P represents potential transit periods of PCs whose individual transit events may be separated in time by any of the off-diagonal elements of P or some multiple thereof. By its construction, the matrix P is skew-symmetric implying that only the positive non-zero matrix elements below the main diagonal are valid periods for consideration for a total of  $\sum_{i=1}^{n_T-1} i$  periods. Because the periodic search is for repeating transit-like events, it is required that  $n_T > 1$ . In this way, no more than two transit events are required to detect a repeating putative PC. This fact extends the ORION detection sensitivity to nearly the full observational baseline of  $\sim 27$  days for the majority of TIC members.

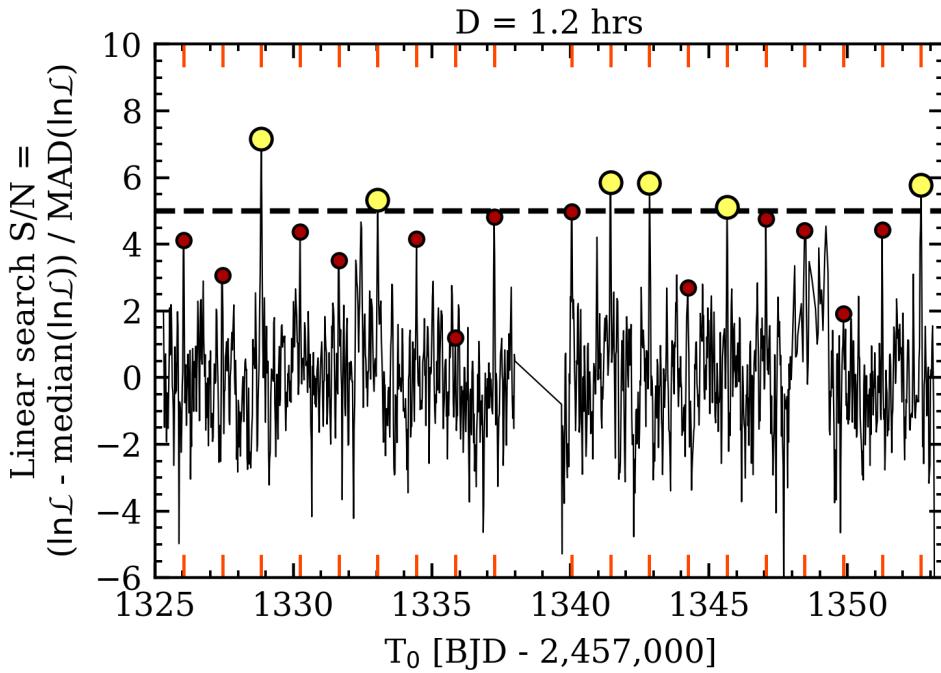


Figure 7.4: The linear search S/N spectrum versus transit times for TIC 234994474. The linear search S/N is calculated from the  $\ln$  likelihood of the de-trended light curve in 30 minute bins under a box model centered at each transit time and with a fixed transit duration of 1.2 hours in this example. Mid-transit times of the PC candidate hosted by TIC 234994474 (TOI-134.01,  $P = 1.4$  days) are depicted by the vertical ticks and highlighted on the S/N spectrum with circular markers. The larger yellow markers indicate transit times at which the S/N spectrum exceeds the imposed ORION detection threshold of  $S/N_{\text{thresh}} \geq 5\sigma$ .

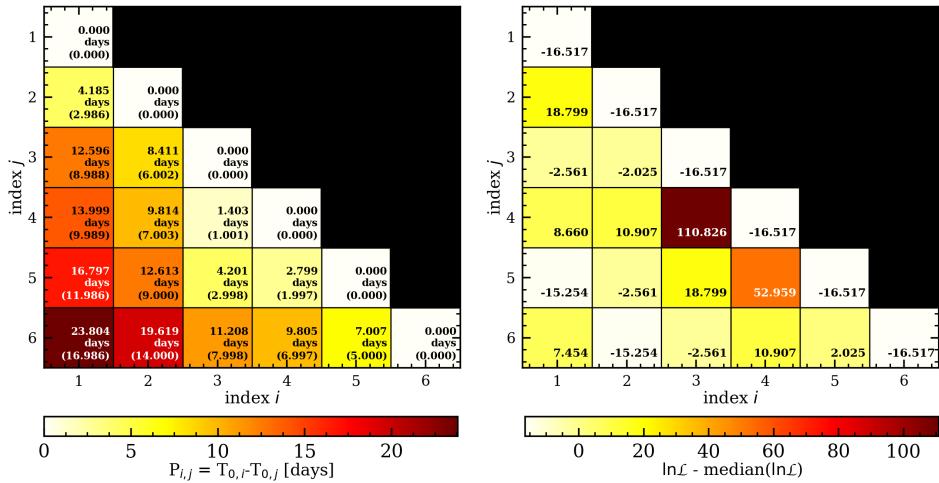


Figure 7.5: The results of the periodic search for repeating transit-like events in the light curve of TIC 234994474 which hosts TOI-134.01 at 1.40131 days. *Left panel:* the  $n_T \times n_T$   $P$  matrix of possible periods of repeating transit-like events from the  $n_T = 6$  high S/N transit times detected in the linear search stage (see Fig. 7.4).  $P$  is skew-symmetric such that only the periods  $> 0$  below the diagonal are valid potential periods. Each  $P$  element is annotated in each grid cell along with its ratio to the true orbital period. Because all of the transit times detected during the linear search are associated with a transit of the PC, the  $P$  matrix elements are all close to integer multiples of the true orbital period. *Right panel:* the  $n_T \times n_T$  matrix of data  $\ln L$  likelihoods under a box transit model with orbital periods from the matrix  $P$ , and with mid-transit times, depths, and fixed duration (i.e. 1.2 hours) from the linear search stage.  $\ln L$  likelihood values along the diagonal correspond to the null hypothesis: a transit model with zero period. The PC period at  $\sim 1.4$  days clearly exhibits the largest  $\ln L$  likelihood which then discards all other potential periods as multiples of  $\sim 1.4$  days.

Because each linear search with a unique  $D$  is independent of the others, the  $P$  matrices of differential transit times are considered together. In this way a single master set of periods is compiled whose elements are referred to as Periods-of-Interest (POIs). Recall that each POI has an associated time of mid-transit  $T_0$ , duration  $D$ , depth  $Z$ , and  $\ln \mathcal{L}$  from the linear search. At this stage, the computationally tractable box transit model is substituted in favor of a more physical transit model whose model parameters are initialized using the box model parameters before optimization. The [Mandel & Agol \(2002\)](#) transit model is used through its implementation within the `batman` python package ([Kreidberg, 2015](#)) to compute model realizations given the input parameters  $\theta = \{P, T_0, a/R_s, r_p/R_s, i, e, \omega, a_{\text{LDC}}, b_{\text{LDC}}\}$  where  $e$  is orbital eccentricity,  $\omega$  is the argument of periastron, and  $\{a_{\text{LDC}}, b_{\text{LDC}}\}$  are the quadratic limb darkening coefficients. In practice, only the parameters  $\theta = \{P, T_0, a/R_s, r_p/R_s, i\}$  are optimized by assuming circular orbits and fixing the quadratic limb darkening coefficients in the TESS bandpass to the values interpolated from the [Claret \(2017\)](#) grid over  $T_{\text{eff}}$ ,  $\log g$ , and assuming solar metallicity. The  $\theta$  parameters are optimized using the same routine which was used to optimize the GP hyperparameters during the initial de-trending stage (see Sect. 7.3.2). With each POI's optimized transit model, the  $\ln$  likelihood of the data is computed for use during the succeeding steps aimed at identifying repeating transit-like events from the initial set of POIs.

A series of cuts is performed on the set of optimized POIs to identify the most likely independent periods within the data. The first cut is to remove repeated period multiples. If POIs with integer period multiples (i.e.  $2P_i, 3P_i, \dots$ ) are indeed due to a transiting planet then those POIs are likely to be manifestations of the same object. The exact  $P$  value that is retained from a set of apparent period multiples is that with the largest  $\ln$  likelihood. Any arbitrary pair of POIs ( $P_i, P_j$ ) is flagged as a period multiple if  $P_i$  and  $n \cdot P_j$  are within  $f_P = 1\%$  for any  $n = 2, \dots, n_{\text{transit}}$ . One important caveat to the removal of integer multiple periods is that resonant multi-planet systems are undetectable within the periodic transit search because all but one of the POIs will always be rejected in favor of its maximum  $\ln$  likelihood multiple.

Similarly, because the set of POIs are derived from peaks in the linear search S/N spectrum, rational multiples of each POI must also be sampled (i.e.  $P_i/2, P_i/3, \dots$ ). Consider a S/N spectrum derived from a light curve containing a single transiting planet with  $n_{\text{transit}} > 2$  but whose individual transits are only marginally detectable due to their amplitude relative to the photometric precision. Consider in this case that only a fraction of transit events are detected during the linear search. The detection of only some transit events may result in a misidentified POI that is an integer multiple times greater than the underlying true period if one or more intermediate transit-events go undetected due to the effects of random noise. Therefore fractional multiples of each POI must be considered. These new periods are equal to  $P_i/n$  for all integers  $n$  resulting in a reduced period greater than or equal to the minimum orbital period considered by [ORION](#): 0.5 days. The  $\ln$  likelihood of the data under the box model with reduced period  $P_i/n$  and remaining parameters  $\{T_{0,i}, D_i, Z_i\}$  is calculated to be compared with the  $\ln \mathcal{L}$  value for the model with  $P_i$ . Here, the latter three parameters are fixed regardless of the input period. The search over period multiples retains the period with the largest  $\ln$  likelihood. The lower period limit of 0.5 days is imposed to limit the number of rational multiples of each POI that are investigated and because the temporal bin width used during the linear search stage already limits the sensitivity to short-period planets whose transit durations are comparable to the 30 minute bins

used therein. Furthermore, consideration of orbital periods  $< 0.5$  days in the periodic search stage is unlikely to result in a large number of missed transits owing to the intrinsically low occurrence rate of ultra-short-period planets ( $\lesssim 1\%$ ; Sanchis-Ojeda et al., 2014; Adams et al., 2016).

The result of these cuts to the initial set of POIs is a set of repeating transit-like events with a unique POI that is not deemed to be a multiple of another POI. The inclusion of multiple unique POIs allows ORION to search for multiple planets in a single light curve so long as those planets are not close to a resonant configuration. However this final set of POIs may or may not correspond to a transiting planet or some other form of periodic astrophysical source such as an eclipsing binary. Therefore the next steps in ORION are to vet the surviving POIs for systematic false positives given their distinctive light curve features that can be largely vetted in an automated way.

### 7.3.5 Automated planet vetting

#### Automated vetting based on light curve features

Here, POIs are automatically vetted using a set of eight vetting criteria which investigate the flagged transit-like features in the de-trended light curve (see Sect. 7.3.4). This automated vetting stage is intended to identify false or insignificant transit-like events and thus provides a preliminary list of putative PCs prior to more selective human vetting and statistical vetting for astrophysical false positives.

The automated vetting criteria are controlled by the set of free parameters  $\{c_i\}$  for  $i = 1, \dots, 10$  which are described below and are included in the summary Table 7.2. The adopted values of these parameters controls the performance of ORION in terms of its detection sensitivity and false positive rate and were derived from early ORION executions on both archival *Kepler* and simulated TESS light curves<sup>3</sup> prior to the first TESS data release. We do not however make any significant claims of their optimality.

The eight automated vetting criteria are defined as follows:

1. It is required that each POI’s transit depth  $Z$  from its optimized transit model be  $> 0$ .
2. The transit S/N is

$$\text{S/N}_{\text{transit}} = \frac{Z}{\sigma_{f,\text{transit}}} \sqrt{n_{\text{transit}}(P, T_0, \mathbf{t})} \quad (7.5)$$

where  $\sigma_{f,\text{transit}}$  is the photometric precision over the transit duration timescale and acts as a proxy for the Combined Differential Photometric Precision (CDPP<sub>transit</sub>; Christiansen et al., 2012). The number of observed transits is  $n_{\text{transit}}$  given the vector of observations  $\mathbf{t}$ , the POI’s orbital period  $P$ , and mid-transit time  $T_0$ . The S/N<sub>transit</sub> is required to be  $> c_1 = 8.4$ .

3. The transit parameters  $\{P, T_0, a/R_s, Z, i\}$  are used to phase-fold the light curve and compute the transit duration (Winn, 2010) such that the in-transit points, including those in ingress or egress, can be isolated. It is required that the difference in the median in and out-of-transit fluxes exceed  $c_2 = 2.4$  median absolute deviations of the out-of-transit flux.
4. If a misidentified POI happens to be less than the true period, then the phase-folded light curve will appear to contain out-of-transit points in-transit. This is combated by requiring that the

---

<sup>3</sup><https://archive.stsci.edu/tess/ete-6.html>

number of in-transit points lying below  $Z + \sigma_Z$  (where  $\sigma_Z$  is the  $1\sigma$  uncertainty on the transit depth) accounts for at least a  $c_3 = 0.7$  fraction of all in-transit points.

5. It is required that the in-transit sampling be approximately symmetric in time by insisting that the number of points between  $T_1$  and  $T_0$  be within  $50\% \pm c_4$  of the total number of in-transit points between  $T_1$  and  $T_4$ <sup>4</sup>.  $c_4$  is set to 10%.
6. Flare stars such as TIC 25200252 as shown in Fig. 7.6 are found to result in a number of misidentified transit-events. Flare events are therefore searched within each light curve by first flagging individual flux measurements which are  $> c_5 = 8$  median absolute deviations brighter than the median flux baseline. However, by the aforementioned criterion, individual stochastic flux jumps can also mimic flares. It is therefore required that any window over which a possible flare event occurs must contain  $> c_6 = 2$  successive bright measurements above the  $c_5$  threshold in order to identify a flare. Flux measurements occurring within a flare window are identified from the  $q^{\text{th}}$  percentile of the light curve flux distribution where  $q$  is the fraction of the observational baseline that occurs within a flare's duration. The total flare duration over the light curve is calculated from the number of detected flares multiplied by the characteristic M dwarf flare duration  $c_7 = 30$  minutes (Moffett, 1974; Walkowicz et al., 2011; Hawley et al., 2014). Transit-like events with an identified flare occurring within  $c_8 = 4$  transit durations from  $T_0$  of a POI are vetted as flares.
7. Visual inspection of a number of light curves observed during sector 1 frequently reveals sharp flux losses at the light curve edges. This signature is often falsely attributed to transit-like signal but is clearly a systematic effect that is not always well-modelled during de-trending stage. Because this edge effect appears to operate over the final  $\sim 4 - 5$  hours of the light curve, POIs with mid-transit times within  $c_9 = 4.8$  hours of either the first or final flux measurements around a sampling gap are automatically flagged as probable false positives.
8. The optimized transit models of the POIs that satisfy all of the seven aforementioned criteria are removed from the light curve. This produces a maximally clean light curve whose residuals should only arise from random noise in photometry or from inaccuracies in the independent systematic GP and transit models. The `numpy.correlate` python function is used to compute the autocorrelation of the residual light curve as a function of time delay as light curves demonstrating large autocorrelations due to imperfect systematic models, can often mimic transit-like events which satisfy the previous vetting criteria. This criterion is particularly important for the use of ORION on *K2* light curves which often exhibit significant temporal correlations due to the thrusts used for re-orientation of the spacecraft and its imperfect correction (Vanderburg & Johnson, 2014). It is required that the autocorrelation function for delays greater than zero be  $\leq c_{10} = 0.6$ .

### Automated vetting of eclipsing binaries

POIs that obey all eight automated vetting criteria from Sect. 7.3.5 are passed along to vetting for eclipsing binary (EB) astrophysical false positives. Six free parameters control the performance of the astrophysical vetting procedure:  $\{c_{\text{EB},i}\}$  for  $i = 1, \dots, 6$ . The EB vetting criteria are adopted from a variety of sources (Batalha et al., 2010; Bryson et al., 2013; Günther et al., 2017; Crossfield et al., 2018)

---

<sup>4</sup>The epochs  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$  represent the initial epoch in ingress, the final epoch in ingress, the initial epoch in egress, and the final epoch in egress respectively. Recall that  $T_0$  represents the mid-transit time between  $T_2$  and  $T_3$ .

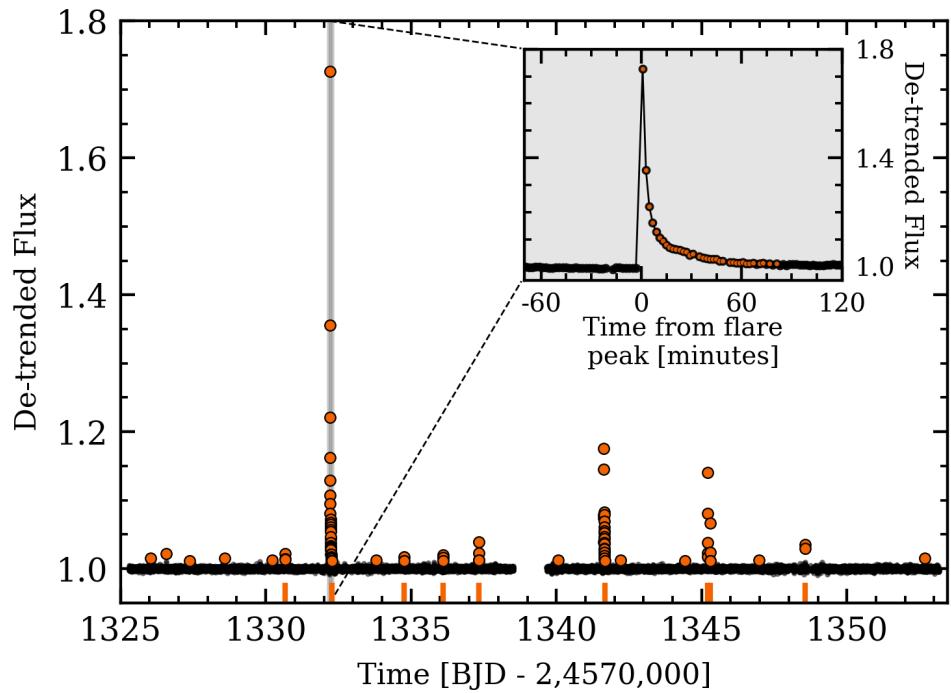


Figure 7.6: The de-trended light curve of the flare star TIC 25200252. Measurements initially flagged as being potentially associated with a flare event are highlighted by the large orange markers in the light curve. Only windows containing at least two successive bright measurements above this threshold are flagged as flares. TIC 25200252 shows nine flares during the TESS sector 1 baseline which are marked by the vertical ticks along the abscissa axis. This includes two flares in quick succession near  $\text{BJD}-2,457,000 = 1345.2$ . *Subpanel:* zoom-in of the event centered on  $1332.212 \text{ BJD}-2,457,000$  depicting the steep rise in flux and exponential decay which are characteristic of stellar flares.

and are used to flag light curve features consistent with being an EB rather than a transiting planet. The EB vetting criteria are defined as follows:

1. POIs are required to have  $r_p/R_s < c_{EB,1} = 0.5$ .
2. It is also required that the inferred companion radius  $r_p < c_{EB,2} = 30 R_\oplus$ .
3. The observed transit duration  $D$  is required to be less than the transit duration corresponding to a planet with radius  $c_{EB,2} = 30 R_\oplus$  at  $b = 0$ .
4. Searches for secondary eclipses are conducted by first sampling eclipse duty cycles (i.e. the fraction of the orbit during eclipse) from the [Shan et al. \(2015\)](#) duty cycle PDF (see their Fig. 4). This distribution was derived from a synthetic population of M dwarf EBs based on *Kepler* binary statistics and is dependent on the population of EB total radii, total masses, orbital periods, and eccentricities. In each POI’s light curve, phase-folded points occurring within a duty cycle centered on phase= 0.5 are considered for possible contamination by a secondary eclipse. With the in-eclipse points used to define the secondary eclipse depth  $Z_{occ}$  and photometric precision during the occultation  $\sigma_{occ}$ . It is required that EBs satisfy the following conditions:

$$\frac{Z_{occ}}{\sigma_{occ}} > c_{EB,4}, \quad (7.6)$$

$$\frac{Z - Z_{occ}}{\sqrt{\sigma_{f,transit}^2 + \sigma_{occ}^2}} > c_{EB,4}. \quad (7.7)$$

Recall that  $Z$  and  $\sigma_{f,transit}$  are the transit depth and in-transit photometric precision.  $c_{EB,4}$  is set to 5 ([Günther et al., 2017](#)) and EBs identified by this criterion are required to have  $> c_{EB,5} = 50\%$  duty cycle samples that satisfy the above conditions. Also note that eclipse searches here are restricted to circular orbits.

5. EBs also exhibit ‘V’-shaped transits due to the similarity of the companion sizes. To search for ‘V’-shaped transits the ingress time  $T_{12} = |T_1 - T_2|$  and egress time  $T_{34} = |T_3 - T_4|$  are calculated from the optimized transit model and compared to the total transit duration  $D$ . ‘V’-shaped transits are required to have  $T_{12} + T_{34}$  which are  $c_{EB,6} \geq 90\%$  of  $D$ . Notably, ‘V’-shaped transits may also be indicative of planetary transits at large impact parameters so ‘V’-shaped transits are not explicitly discarded but are instead assigned a non-definitive disposition based solely on this criterion.

### 7.3.6 Joint systematic plus transiting planet modelling

The set of transit-like events that satisfy all of the vetting criteria presented in Sects. 7.3.5 and 7.3.5 are treated as PCs in this, the final **ORION** stage. At this point the modelling of systematic light curve effects using a 1-dimensional GP regression model from Sect. 7.3.2 is revisited. The alternative is to simultaneously sample the joint GP plus transit light curve parameter posterior PDF using Markov Chain Monte-Carlo (MCMC) simulations. Explicitly, the light curve model is modified by replacing the previously null mean model  $\mu(t)$  with a full transit model containing all putative PCs. Overfitting by the systematic model, which can partially fill in planetary transits, is mitigated by simultaneously modelling systematics and PCs. The resulting joint systematic+planet model is therefore derived in a self-consistent manner with more robust solutions for the transiting PC parameters of interest. MCMC

sampling of the transit parameter marginalized posterior PDFs allows us to compute point estimates of their MAP values and uncertainties for later use.

For systems containing  $N_{\text{PC}}$  PCs,  $4 + 5N_{\text{PC}}$  parameter PDFs are sampled by continuing to insist on circular orbits and fixed limb-darkening coefficients. Explicitly, the GP hyperparameters  $\Theta = \{\ln a_{\text{GP}}, \ln \lambda, \ln \Gamma, \ln P_{\text{GP}}\}$  are fit along with the transiting planet parameters  $\theta = \{P_i, T_{0,i}, a_i/R_s, r_{p,i}/R_s, i_i\} \forall i = 1, \dots, N_{\text{PC}}$ . The GP hyperparameters are initialized to their maximum likelihood values from the de-trending stage and are continued to be bounded by the broad uniform priors listed in Table 7.1. Transit parameters are initialized by their maximum likelihood values assuming fixed GP hyperparameters from de-trending. The adopted prior PDFs on the transit model parameters are also reported in Table 7.1 for the most common case of fitting transiting PCs with multiple transits over the observational baseline. As we will see in Sect. 7.4.1, some priors will be modified when sampling transit parameters used to model only a single transit.

MCMC sampling is performed with the `emcee` ensemble-sampler (Foreman-Mackey et al., 2013). One hundred walkers are initialized in small Gaussian balls centered on each parameter’s initial value. Throughout the MCMC each walker’s acceptance fraction is monitored and warns the user when its mean value over all walkers fails to fall within the desired range of 20-60% in either of the burn-in or final sampling stages. The desired duration of each MCMC stage is  $\gtrsim 10$  autocorrelation times. Warnings are again produced if the MCMC chains fail to reach this length.

The main `ORION` output is a list of objects of interest (OIs) along with samples from the model parameter marginalized posterior PDFs and point estimates of each parameter’s MAP value and uncertainties derived from the 16<sup>th</sup> and 84<sup>th</sup> percentiles of their 1-dimensional marginalized PDF. The raw light curves and models sampled at observation times  $t$  are also saved to produce summary images such as the example shown in Fig. 7.7 for OI 234994474.01 whose PC is a known TESS Object of Interest: TOI-134.01, a close-in terrestrial planet currently being validated with HARPS and PFS RVs (Astudillo-Defru et al., 2019).

## 7.4 ORION planet search around low mass stars in TESS sectors 1 & 2

For its inaugural application we apply the `ORION` transit detection pipeline to the 2 minute extracted light curves from the first two TESS sectors. Overall, `ORION` produces a set of 121 OIs around 96 of the 1599 low mass TIC members in our stellar sample after automated vetting. It is expected that many of these postulated PCs will be false positives due to imperfect corrections of systematic effects or to astrophysical false positives other than eclipsing binaries as EBs have certain distinct photometric features which are flagged during the automated vetting stage. As such, we proceed with manual vetting of all `ORION` OIs via human inspection of the pipeline’s output. This step is particularly important for new transit search algorithms to develop an understanding of common sources of false positives and if or how they may be corrected in future versions.

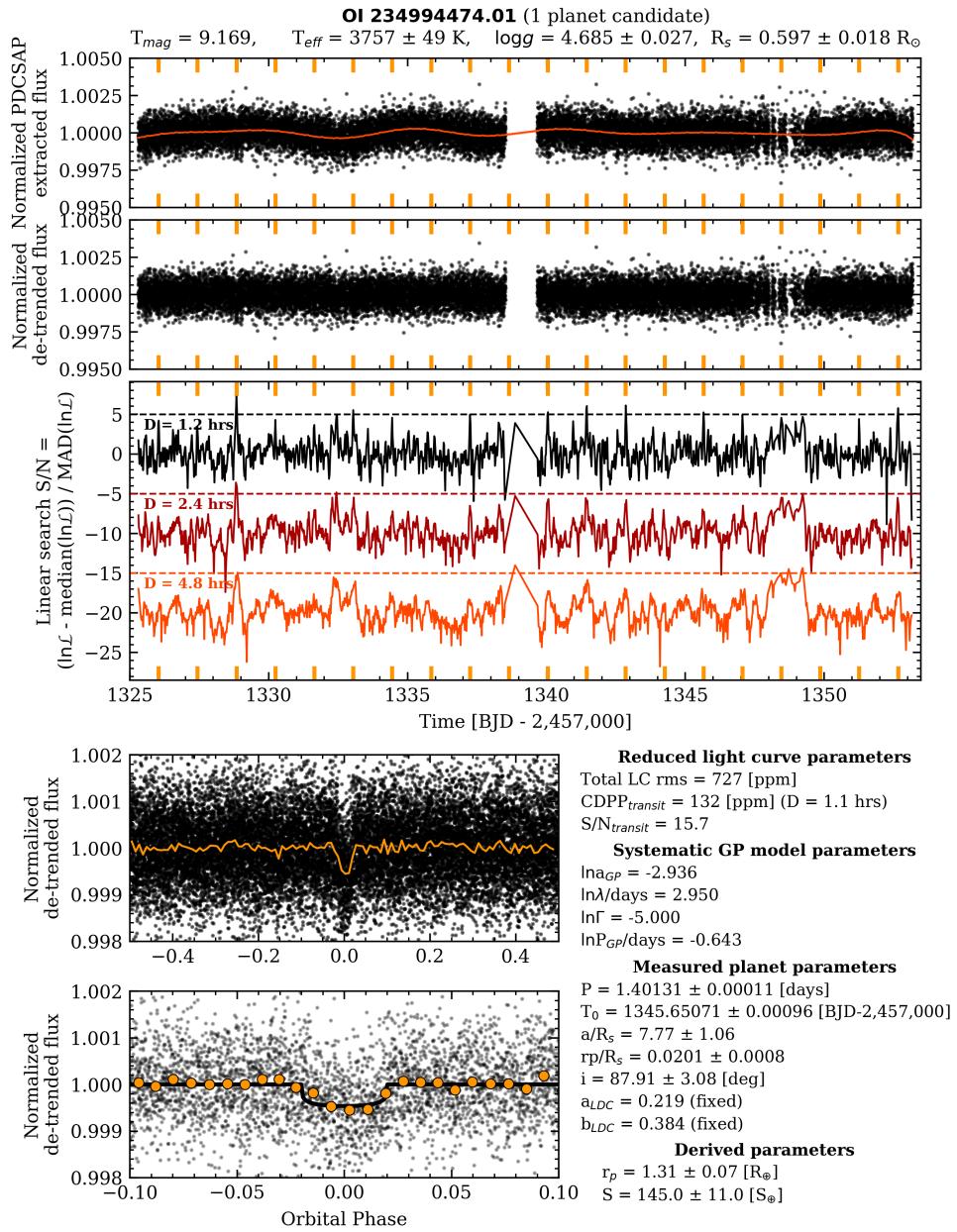


Figure 7.7: The summary image output from running ORION on TIC 234994474 and the resulting detection of OI 234994474.01 which is consistent with the known TOI-134.01. The TESS magnitude  $T$  and Gaia-derived physical stellar parameters are annotated at the top. *Top panel:* the 2 minute extracted light curve from the TESS SPOC along with the mean GP systematic model (orange line) and the times of TOI-134.01 planetary transits indicated by the vertical ticks. *Second panel:* the de-trended light curve. *Third panel:* the linear search S/N spectra calculated from the likelihood of the data given a box model with fixed mid-transit time  $T_0$  and for each of the three fixed transit durations  $D$  (i.e. 1.2, 2.4, and 4.8 hours). Each spectrum is offset for clarity along with the  $5\sigma$  S/N threshold. *Fourth panel:* the complete and binned ( $\Delta t = 0.2D/P$ ) de-trended light curve phase-folded to the MAP orbital period  $P$  and mid-transit time  $T_0$  of the planet candidate. *Bottom panel:* zoom-in on the transit in the de-trended and phase-folded light curve. Various diagnostic quantities are reported in the lower right corner along with measured and derived transit parameters.

The vast majority of the 1599 light curves investigated in this preliminary study exhibited multiple repeating transit-like events with independent periodicities. The median number of independent repeating transit-like signals per light curve was three or  $\sim 4800$  potential transit signals. The reduction of this number down to just 121 OIs demonstrates the effectiveness of the ORION automated vetting criteria (see Sect. 7.3.5) and its effect on limiting the manual vetting requirement.

#### 7.4.1 Manual vetting of ORION planet candidates

We conduct a visual inspection of each of the raw, de-trended, and phase-folded light curves for each OI detected by ORION. In this analysis we flag 97/121 PCs as being residual systematic effects which are misidentified as transits. Of those, 19 appear to have been directly affected by measurements obtained between  $\sim 1347$  and  $1349$  BJD - 2,4570,000 during sector 1 at times when TESS briefly lost much of its pointing precision. This effect is not perfectly corrected for in many of the sector 1 extracted light curves nor by ORION's own systematic modelling. Most of the remaining OIs flagged as false positives are attributable to residual systematics mimicking transits. Visual inspection indicates that these OIs are clearly inconsistent with being a planetary transit.

##### Single transit events

The interpretations of the remaining 24/121 OIs as transiting planets are deemed plausible by our manual vetting. These preliminary PC dispositions are based purely on the resemblance of the light curve features to periodic transit events or in some instances, to single transits (ST) events that do not show compelling evidence for periodicity over the observational baseline. We note that ORION is not optimized nor intended to be sensitive to the detection of ST events. However, some ST events with moderate to high S/N can be detected but incorrectly classified as periodic events if at least one other transit time in the linear search S/N spectrum exceeds  $S/N_{\text{thresh}}$ , even if that event is resulting from noise (see Sect. 7.3.3). In this case, the ST event is folded to the time difference between the ST and any of the other S/N events exceeding  $S/N_{\text{thresh}}$  with  $T_0$  being assigned the value of the mid-transit time of the maximum linear search S/N peak. If the ST event has a sufficiently high S/N on its own then the addition of noise by phase-folding to the incorrect period, but correct  $T_0$ , may result in a feature that still passes our automated vetting criteria due solely to the significance of the ST event. The inferred period of such a ST event by ORION will therefore always be less than its true period if the feature is indeed the result of a singly transiting planet.

Three OIs resembling ST events are identified during manual vetting around the TICs 49678165, 92444219, and 415969908. The latter TIC member already hosts the known TOI-233.01 at 11.7 days but the proposed ephemeris for TOI-233.01 is inconsistent with the  $T_0$  of our putative ST event. A more complete discussion of this and the other individual systems is reserved for Sect. 7.5.2.

For each of the three OIs classified as a putative ST, we refine their transit parameters by isolating the light curve around  $10D$  of  $T_0$  and use that reduced dataset and MCMC to sample the transit model parameters with just a single transit (Seager & Mallén-Ornelas, 2003). In these analyses, the  $P$ ,  $T_0$ , and  $a/R_s$  priors are modified as listed in Table 7.1. The orbital period of the ST is further restricted to

periods greater than the largest time difference between  $T_0$  and both edges of the light curve's baseline. The resulting periods are largely uncertain with their posterior PDFs showing extended tails out to long periods  $\gtrsim 80$  days as is expected for transit models which lack multiple events to constrain  $P$ . The refined transit parameters for our three putative ST events are used in place of the transit model solutions produced by ORION.

#### 7.4.2 Statistical validation of transiting planet candidates

We are currently not in a position to distinguish between confirmed planets and various astrophysical false positive scenarios in an absolute sense. This is because of the lack of follow-up observations in this study which are ultimately required to validate or disprove the planetary nature of our OIs. Despite the lack of such follow-up observations, it is still advisable to attempt to statistically validate OIs by inferring the relative probabilities of a variety of astrophysical false positive (AFP) scenarios which can be compared to the planetary interpretation. Such considerations are further motivated given that the rate of AFPs in the 2 minute TESS light curves is expected to be significant ( $\sim 60\%$ ; [Sullivan et al., 2015](#)).

We attempt to statistically validate our 24 OIs around 22 TIC members using the PyMultinest ([Buchner et al., 2014](#)) implementation of the probabilistic transit validation software `vespa` ([Morton, 2012, 2015](#)) for establishing the final dispositions of our OIs by computing the planetary false positive probabilities (FPP). `vespa` considers six AFP scenarios as potential explanations for transit-like signals. These include undiluted eclipsing binaries (EB), hierarchical triple EBs (HEB), and blended background or foreground EBs which are not physically associated with the target (BEB). Each of these scenarios then have two instances with the first assuming the input orbital period and the second assuming twice the input orbital period (i.e. EB2, HEB2, BEB2). We note however that the forthcoming statistical OI interpretations are not treated as absolutely definitive in lieu of the follow-up observations required to distinguish transiting planets from AFPs.

For `vespa` input we use the TIC member's celestial coordinates ( $\alpha, \delta$ ), stellar parameters  $T_{\text{eff}}$ ,  $\log g$ , and  $\varpi$ , along with the star's  $JHK_S$  photometry. `vespa` also requires the photometric band in which the putative transit is detected but the code cannot properly handle the TESS bandpass in its current version. Fortunately, the central wavelengths of the TESS bandpass and the Cousins  $I_C$ -band are similar but with the TESS band being much wider ([Sullivan et al., 2015](#)). Given the similarity of the  $I_C$  and SDSS  $i$ -band, and the compatibility of the latter passband within `vespa`, we use  $T$  and  $K_S$  to derive  $i$  using the color relation from [Muirhead et al. \(2018\)](#). We also pass to `vespa` the OI's MAP  $P$  and  $r_p/R_s$  along with its de-trended light curve following the removal of all candidate transit models that are not associated with the OI being statistically validated. The light curves are phase-folded and restricted to  $\pm 3D$  around  $T_0$  for comparison to light curve models generated under the transiting planet and all AFP scenarios.

`vespa` also requires constraints on the maximum angular separation (`maxrad`) from the target star that should be searched for potential blending sources. We limit this separation to be less than the median full width at half maximum (FWHM) of the target's approximate point spread function (PSF). The FWHM of the PSF is derived by fitting a 2-dimensional Gaussian profile to the target image in each target pixel file over time and adopting the median FWHM as `vespa` input. The median FWHM value

among the 22 TIC members is  $\sim 37$  arcsec or nearly two TESS pixels across. Over such a large field it is reasonable to expect that many of the OIs may be favored by either the BEB or BEB2 models.

Lastly, `vespa` requires the maximum permissible depth of a secondary eclipse of an EB (`secthresh`) to be specified. Recall that attempts within `ORION` were made to automatically vet EBs among our OIs in Sect. 7.3.5. We therefore expect that `vespa` is unlikely to detect any probable EBs. Nevertheless, the input `secthresh` value for each TIC member is derived from the box model depths fitted to each transit time in  $\mathbf{T}_0$  during the linear search stage (see Sect. 7.3.3). After masking measurements that occur within the PC’s transit window and extrapolating the fitted depths to the PC’s transit duration, we adopt the 95<sup>th</sup> percentile of the depth distribution depths as the value of `secthresh` (Crossfield et al., 2018). The median `secthresh` is  $\sim 2100$  ppm. The input `maxrad` and `secthresh` for each OI are reported in Table 7.3 along with the results of our `vespa` calculations.

Table 7.3: `vespa` input parameters and inferred probabilities of transiting planet and astrophysical false positive models for our 24 objects of interest

IDs		<code>vespa</code> Input				<code>vespa</code> Results					
OI	TOI	maxrad	secthresh	P <sub>EB</sub>	P <sub>EB2</sub>	P <sub>HEB</sub>	P <sub>HEB2</sub>	P <sub>BEB</sub>	P <sub>BEB2</sub>	FPP <sup>a</sup>	Disposition <sup>b</sup>
12421862.01	198.01	34.755	$9.9 \times 10^{-4}$	$2.0 \times 10^{-9}$	$4.9 \times 10^{-5}$	$1.1 \times 10^{-14}$	$9.8 \times 10^{-8}$	$5.8 \times 10^{-2}$	$3.7 \times 10^{-2}$	$9.6 \times 10^{-2}$	PC
47484268.01	226.01	34.147	$4.8 \times 10^{-3}$	$1.3 \times 10^{-7}$	$3.8 \times 10^{-3}$	$1.4 \times 10^{-9}$	$1.6 \times 10^{-4}$	$3.9 \times 10^{-1}$	$2.8 \times 10^{-1}$	$6.8 \times 10^{-1}$	pPC
49678165.01	-	75.094	$7.7 \times 10^{-3}$	$9.4 \times 10^{-27}$	$1.4 \times 10^{-14}$	$2.4 \times 10^{-26}$	$3.5 \times 10^{-14}$	$9.1 \times 10^{-3}$	$7.1 \times 10^{-3}$	$1.6 \times 10^{-2}$	ST
92444219.01	-	35.723	$3.8 \times 10^{-3}$	$4.1 \times 10^{-15}$	$6.7 \times 10^{-8}$	$4.4 \times 10^{-15}$	$1.6 \times 10^{-8}$	$2.6 \times 10^{-1}$	$1.6 \times 10^{-1}$	$4.1 \times 10^{-1}$	pST
100103200.01	-	47.832	$7.0 \times 10^{-4}$	$5.1 \times 10^{-19}$	$4.3 \times 10^{-12}$	$5.0 \times 10^{-33}$	$2.9 \times 10^{-20}$	$3.8 \times 10^{-3}$	$2.0 \times 10^{-3}$	$5.7 \times 10^{-3}$	pPC <sup>c</sup>
100103201.01	-	47.941	$8.1 \times 10^{-4}$	$2.2 \times 10^{-73}$	$1.7 \times 10^{-67}$	$5.1 \times 10^{-132}$	$3.7 \times 10^{-91}$	$6.2 \times 10^{-1}$	$3.5 \times 10^{-1}$	$9.7 \times 10^{-1}$	BEB
100103201.02	-	47.941	$8.1 \times 10^{-4}$	$2.2 \times 10^{-73}$	$1.7 \times 10^{-67}$	$5.1 \times 10^{-132}$	$3.7 \times 10^{-91}$	$6.2 \times 10^{-1}$	$3.5 \times 10^{-1}$	$9.7 \times 10^{-1}$	BEB
141708335.01	-	33.606	$3.6 \times 10^{-3}$	$6.7 \times 10^{-39}$	$2.9 \times 10^{-31}$	$3.4 \times 10^{-35}$	$8.6 \times 10^{-26}$	$6.3 \times 10^{-1}$	$3.3 \times 10^{-1}$	$9.5 \times 10^{-1}$	BEB
206660104.01	-	38.276	$8.4 \times 10^{-4}$	$1.7 \times 10^{-65}$	$8.6 \times 10^{-39}$	$2.5 \times 10^{-87}$	$1.2 \times 10^{-40}$	$1.3 \times 10^{-2}$	$9.3 \times 10^{-3}$	$2.2 \times 10^{-2}$	PC
231279823.01	-	36.346	$5.3 \times 10^{-4}$	$7.7 \times 10^{-43}$	$7.9 \times 10^{-47}$	$7.5 \times 10^{-55}$	$4.0 \times 10^{-78}$	$5.1 \times 10^{-1}$	$4.9 \times 10^{-1}$	$1.0 \times 10^0$	BEB
231702397.01	122.01	36.485	$6.9 \times 10^{-3}$	$7.5 \times 10^{-18}$	$8.7 \times 10^{-6}$	$1.8 \times 10^{-23}$	$5.1 \times 10^{-9}$	$2.3 \times 10^{-3}$	$2.7 \times 10^{-3}$	$5.0 \times 10^{-3}$	PC
234994474.01	134.01	37.794	$7.1 \times 10^{-4}$	$1.2 \times 10^{-51}$	$8.5 \times 10^{-34}$	$3.9 \times 10^{-107}$	$4.2 \times 10^{-61}$	$9.6 \times 10^{-4}$	$1.9 \times 10^{-4}$	$1.2 \times 10^{-3}$	PC
235037759.01	-	39.567	$1.9 \times 10^{-2}$	-	-	-	-	-	-	-	AFP <sup>d</sup>
238027971.01	-	37.085	$1.4 \times 10^{-3}$	$2.5 \times 10^{-67}$	$7.5 \times 10^{-92}$	$5.6 \times 10^{-76}$	$4.1 \times 10^{-79}$	$3.4 \times 10^{-1}$	$6.6 \times 10^{-1}$	$1.0 \times 10^0$	BEB2
260004324.01	-	35.803	$1.1 \times 10^{-3}$	$3.3 \times 10^{-44}$	$4.1 \times 10^{-17}$	$1.4 \times 10^{-61}$	$8.8 \times 10^{-20}$	$5.3 \times 10^{-2}$	$2.8 \times 10^{-2}$	$8.1 \times 10^{-2}$	PC
262530407.01	177.01	40.509	$8.0 \times 10^{-4}$	$6.5 \times 10^{-38}$	$3.2 \times 10^{-41}$	$7.6 \times 10^{-45}$	$5.5 \times 10^{-44}$	$1.8 \times 10^{-7}$	$5.7 \times 10^{-7}$	$7.5 \times 10^{-7}$	PC
278661431.01	-	48.894	$5.1 \times 10^{-3}$	$2.4 \times 10^{-31}$	$1.4 \times 10^{-37}$	$3.6 \times 10^{-48}$	$6.9 \times 10^{-55}$	$3.8 \times 10^{-1}$	$4.8 \times 10^{-1}$	$8.6 \times 10^{-1}$	pPC
279574462.01	-	146.950	$1.3 \times 10^{-2}$	-	$1.7 \times 10^{-86}$	$1.0 \times 10^{-304}$	$1.9 \times 10^{-77}$	$3.0 \times 10^{-1}$	$7.0 \times 10^{-1}$	$1.0 \times 10^0$	BEB2
303586421.01	-	32.084	$6.2 \times 10^{-3}$	$7.8 \times 10^{-53}$	$6.2 \times 10^{-47}$	$3.1 \times 10^{-67}$	$1.5 \times 10^{-31}$	$3.3 \times 10^{-1}$	$5.6 \times 10^{-1}$	$9.0 \times 10^{-1}$	BEB2
305040807.01	237.01	34.445	$8.7 \times 10^{-3}$	$2.6 \times 10^{-30}$	$1.4 \times 10^{-8}$	$4.8 \times 10^{-27}$	$8.0 \times 10^{-8}$	$1.8 \times 10^{-2}$	$1.4 \times 10^{-2}$	$3.2 \times 10^{-2}$	PC
307210830.01	175.01	38.817	$7.7 \times 10^{-4}$	$1.7 \times 10^{-5}$	$8.0 \times 10^{-3}$	$9.0 \times 10^{-10}$	$7.5 \times 10^{-5}$	$9.2 \times 10^{-3}$	$4.3 \times 10^{-3}$	$2.2 \times 10^{-2}$	PC
415969908.01	233.01	32.000	$2.3 \times 10^{-3}$	$1.5 \times 10^{-11}$	$1.1 \times 10^{-5}$	$1.4 \times 10^{-18}$	$2.0 \times 10^{-8}$	$6.0 \times 10^{-3}$	$5.5 \times 10^{-3}$	$1.2 \times 10^{-2}$	PC
415969908.02	-	32.000	$2.3 \times 10^{-3}$	$2.5 \times 10^{-19}$	$2.4 \times 10^{-10}$	$4.9 \times 10^{-28}$	$2.7 \times 10^{-13}$	$2.6 \times 10^{-3}$	$2.0 \times 10^{-3}$	$4.7 \times 10^{-3}$	ST
441056702.01	-	35.795	$1.9 \times 10^{-3}$	$5.5 \times 10^{-19}$	$5.3 \times 10^{-9}$	$1.8 \times 10^{-37}$	$1.8 \times 10^{-17}$	$6.0 \times 10^{-2}$	$2.5 \times 10^{-2}$	$1.7 \times 10^{-2}$	PC

**Notes.**

$P_i$  represents the relative probability (i.e. the product of the model's prior and likelihood relative to the probabilities of all other models) of the  $i^{th}$  astrophysical false positive scenario where  $i$  is one of six possible scenarios: blended eclipsing binaries (BEB), undiluted eclipsing binaries (EB), and hierarchical eclipsing binaries (HEB), each with either one or twice its input orbital period.

<sup>a</sup> The transiting planet false positive probability.

<sup>b</sup> Possible dispositions of objects of interest are a planet candidate (PC), a putative planet candidate (pPC), a single transit event (ST), a putative single transit event (pST), an unclassified astrophysical false positive (AFP), or any of the scenarios  $i$ . The putative dispositions have  $\text{FPP} \in [0.1, 0.9]$  whereas remaining candidates and astrophysical false positives have  $\text{FPP} < 0.1$  and  $> 0.9$  respectively.

<sup>c</sup> OI 100103200.01 is assigned the pPC disposition despite having a  $\text{FPP} < 0.1$  because of its proximity to the comparably bright TIC 100103201 (see Fig. 7.8).

<sup>d</sup> The MCMC for OI 235037759.01 failed to converge so we broadly classify it as an astrophysical false positive (AFP) based on the prevalence of nearby bright sources from Gaia (see panel ‘v’ in Fig. 7.8). We are unable to classify the object as a particular type of astrophysical false positive.

The true power of `vespa` is realized when additional follow-up observations such as contrast curves from AO-assisted imaging or photometric follow-up are used to inform the interpretation of transiting PCs. Given the lack of such data in this study, we adopt conservative limits on the interpretation of the resulting `vespa` probabilities. Similarly, we also do not claim to validate planets with ultra low FPP ( $< 0.01$ ; e.g. [Montet et al., 2015](#); [Crossfield et al., 2018](#); [Livingston et al., 2018](#)) as we will caution in Sect. 7.4.3 that `vespa` results should not be taken absolutely in the absence of follow-up observations. Our limiting values on interpreting FPPs are as follows: OIs with  $\text{FPP} < 0.1$  are classified as PCs. Similarly, OIs are classified as AFPs when  $\text{FPP} \geq 0.9$  and have their dispositions assigned to the specific AFP model with the highest probability. OIs with intermediate FPPs are classified as putative planet candidates (pPC).

The statistical validation calculations with `vespa` result in 13/24 of our OIs being classified as PCs plus 3/24 as pPCs. The OIs 49678165.01, 415969908.02, and 92444219.01 correspond to the three ST events detected during the manual vetting stage. We reclassify these objects as STs and a pST respectively. Seven of the remaining eight OIs are favored by either BEB model with 4/7 BEBs and 3/7 BEB2s. The MCMC during the `vespa` calculation of the lone remaining OI 235037759.01 failed to converge leaving its disposition as of yet undefined. We will show in the following subsection that despite the failure of the FPP calculation, the nature of OI 235037759.01 is likely to be an AFP although we are unable to distinguish between the different AFP scenarios. The derived rate of AFPs from this small sample of OIs is  $\sim 33 \pm 12\%$  which is somewhat lower than expected AFP rate of 60% from the TESS simulations by ([Sullivan et al., 2015](#)). A discrepancy which is likely explained by incompleteness in our vetting (see Sect. 7.5.1).

#### 7.4.3 Querying Gaia sources to supplement statistical validation calculations

`vespa` calculations are based on synthetic stellar populations from the TRILEGAL galaxy model ([Girardi et al., 2005](#)). These synthetic results can be supplemented by querying the Gaia DR2 in the vicinity of each TIC member to *empirically* investigate the number density and brightness distribution of nearby sources on the sky. In this way, we hope to find supporting empirical evidence for any of the BEB interpretations of OIs with high FPPs as those inferences should be expected if nearby bright sources fall within or near the TESS PSF of the targeted TIC member. The resulting maps of Gaia sources around the 22 TIC members with OIs in our sample are shown in Fig. 7.8. Querying the Gaia DR2 is performed identically to the method used in Sect. 7.2.2 to match TIC members with the Gaia DR2 catalog although here we conduct our searches with a fixed radius of 105 arcseconds or  $\sim 5$  TESS pixels.

From Fig. 7.8 it is clear that many of the statistically favored interpretations as either some form of planet candidate or a BEB are consistent with the lack or prevalence of nearby bright sources to the targeted TIC member respectively. All panels with planet candidate OIs in Fig. 7.8 show no or minor sources of comparable brightness within or very close to the target PSF edges as to significantly contaminate the measured TIC member photometry and consequently result in a probable FP. Similarly, all AFP panels other than ‘q’ and ‘s’, do have at least one neighbouring source of comparable brightness that may be responsible for the favorability of an AFP scenario by `vespa`. This includes the TIC 235037759 whose `vespa` calculation failed. Two bright sources are clearly seen to contribute to the flux within the target’s PSF thus supporting the probable interpretation of the OI 235037759.01 transit-like

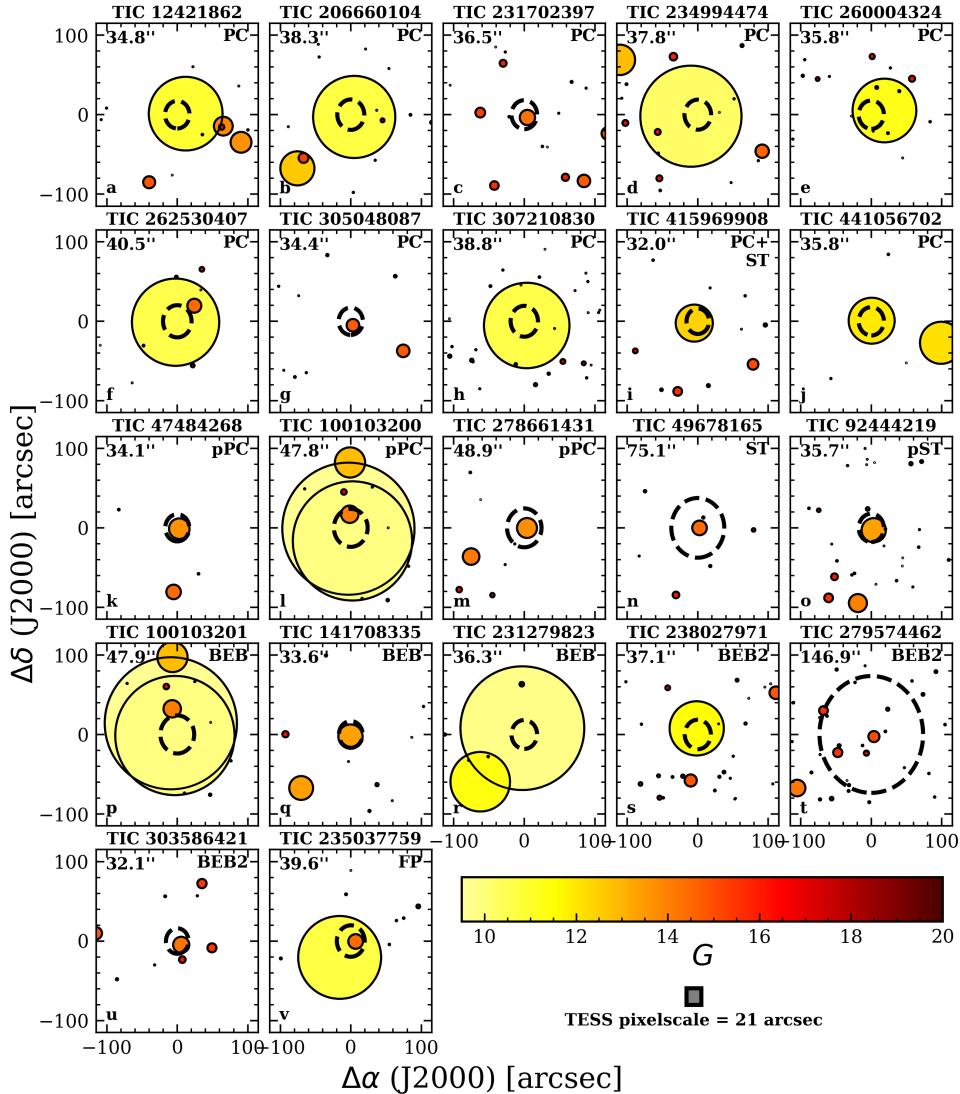


Figure 7.8: Star maps containing sources from the Gaia DR2 in the vicinity of each TIC member identified as hosting an object of interest during the manual vetting stage. Dispositions are annotated in the upper right of each panel (see Sect. 7.4.2 for definitions). The fitted FWHM of each targeted TIC member’s PSF is annotated in the upper left of its panel in arcsec as well as being depicted by the dashed black circle centered on the panel’s origin. The colorbar is indicative of  $G$ -band magnitudes while marker sizes are proportional to the source flux in that band. For reference, the size of a single TESS detector pixel is shown in the lower right corner of the figure.

event as being caused by an AFP.

We note however that some discrepancies between the distributions of Gaia sources and our `vespa` interpretations still persist. Particularly with regards to TIC 100103200.01 and TIC 100103201.01 (panels ‘l’ and ‘p’ in Fig. 7.8) which strongly favor the PC and BEB models respectively despite being located within 1 TESS pixel of one another on the sky, having very similar brightnesses (i.e.  $J = 7.50, 7.66$ ), and being located at effectively identical distances (i.e.  $d = 16.745$  pc). Perhaps naively, we might expect this architecture to favor a AFP scenario for both stars including blends, an EB, or an HEB. Indeed the apparent flux dips which appear qualitatively consistent with a transiting planet around either one of the TIC members, is also seen to have a clear manifestation in the light curve of the other as evidenced in Fig. 7.9. This is almost certainly caused by the overlap of each target’s PSF. However `vespa` results indicate FPPs that differ by over two orders of magnitude between the two TIC members. Perhaps it is feasible, although seemingly unlikely, for TIC 100103200 to host a detectable transiting PC while being blended with the nearby TIC 100103201 whose transit-like events are strongly favored by the BEB scenario. Indeed the transit times of each TIC member’s transit-like events flagged by `ORION` appear out of phase as they do not align nor do they overlap in Fig. 7.9 implying that transit-like events detected around each TIC member by `ORION` do not affect the transit-like events in the light curve of the other. Even if there are regions of the out-of-transit light curve that are mutually affected. We are therefore left with the questionable interpretation of these OIs based on the `vespa` dispositions alone and note that additional vetting criteria should be used in up-coming versions of `ORION` to flag other instances of probable blends. Here we opt to override the `vespa` dispositions and assign a BEB to each of these OIs (see Table 7.3).

#### 7.4.4 Population of planet candidates

After manual vetting and statistical validation we are left with fifteen candidate planets. These include ten PCs, two STs, two pPCs, and one pST. Half of our candidates are ‘new’ having not yet been released as TOIs<sup>5</sup>. Point estimates of observable and derived planetary parameters for these candidates are reported in Table 7.4. Fig. 7.10 also depicts their phase-folded light curves along with the transit models computed using the MAP parameter values from Table 7.4.

---

<sup>5</sup>As of December 19, 2018.

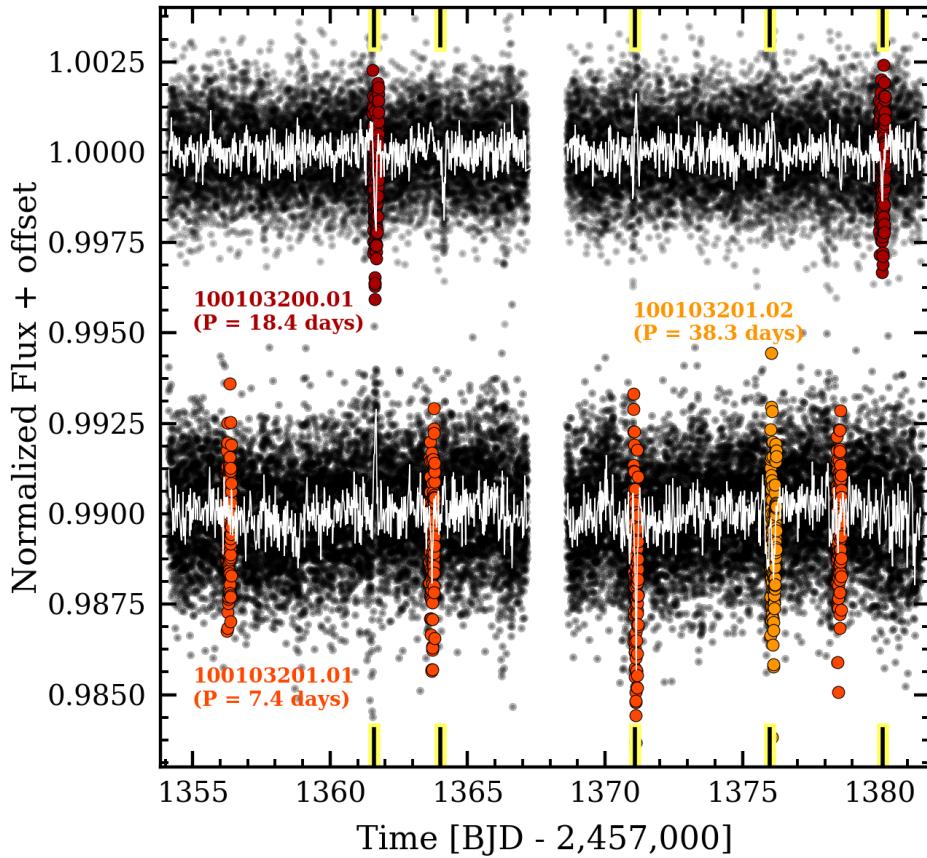


Figure 7.9: The de-trended light curves of TIC 100103200 (top) and TIC 100103201 (bottom). The in-transit points for the three OIs in these systems are highlighted by the colored points with each OI having a unique color. The OI identifications and periods are annotated next to one of its transit-like events. The white curves represent the light curves binned to 30 minutes. The binned light curves reveal  $\sim 5$  transit-like events (indicated by the vertical ticks along the abscissa axes) which appear to be due to transit-events in one TIC member's light curve but also having clear manifestations in the other TIC member's light curve despite the ephemerides of all three OIs being out of phase and non-commensurate.

Table 7.4: Planetary parameters for our 16 vetted candidates

OI	TOI	$P$ [days]	$T_0$ [BJD-2,457,000]	$a/R_s$	$r_p/R_s$	$i$ [deg]	$r_p$ [ $R_\oplus$ ]	$S$ [ $S_\oplus$ ]	Disposition <sup>a</sup>
12421862.01	198.01	$20.4282 \pm 0.0042$	$1376.8028 \pm 0.0023$	$68.05^{+6.21}_{-9.86}$	$0.033^{+0.001}_{-0.002}$	$90.24^{+0.44}_{-0.40}$	$1.59^{+0.09}_{-0.10}$	$2.7 \pm 0.2$	PC
47484268.01	226.01	$20.2833 \pm 0.0048$	$1378.7855 \pm 0.0042$	$56.34^{+9.13}_{-8.28}$	$0.090^{+0.004}_{-0.006}$	$90.62^{+0.32}_{-0.19}$	$3.82^{+0.20}_{-0.25}$	$1.4 \pm 0.1$	pPC
49678165.01	-	$35.64^{+49.78}_{-14.61}$	$1371.1945 \pm 0.0038$	$74.13^{+102.92}_{-26.55}$	$0.10^{+0.01}_{-0.00}$	$90.00^{+0.36}_{-0.36}$	$3.02^{+0.21}_{-0.19}$	$0.3^{+0.3}_{-0.2}$	ST
92444219.01	-	$39.52^{+45.64}_{-16.90}$	$1342.4729 \pm 0.0037$	$49.71^{+55.87}_{-19.54}$	$0.07^{+0.00}_{-0.00}$	$89.99^{+0.40}_{-0.41}$	$3.07^{+0.13}_{-0.14}$	$0.5^{+0.6}_{-0.3}$	pST
100103200.01	-	$18.4462 \pm 0.0024$	$1380.0894 \pm 0.0018$	$51.40^{+9.45}_{-5.97}$	$0.035^{+0.002}_{-0.001}$	$90.45^{+0.33}_{-0.23}$	$1.97^{+0.11}_{-0.09}$	$4.2 \pm 0.3$	pPC
206660104.01	-	$13.4507 \pm 0.0081$	$1379.7842 \pm 0.0079$	$27.84^{+1.82}_{-1.05}$	$0.021^{+0.001}_{-0.001}$	$90.00^{+0.44}_{-0.49}$	$1.21^{+0.08}_{-0.07}$	$5.2 \pm 0.4$	PC
231702397.01	122.01	$5.0789 \pm 0.0010$	$1349.4338 \pm 0.0031$	$31.23^{+3.05}_{-4.64}$	$0.077^{+0.004}_{-0.005}$	$90.55^{+0.36}_{-1.12}$	$2.80^{+0.18}_{-0.19}$	$7.5 \pm 0.7$	PC
234994474.01	134.01	$1.4013 \pm 0.0001$	$1345.6507 \pm 0.0010$	$8.68^{+0.73}_{-1.10}$	$0.021^{+0.001}_{-0.001}$	$88.84^{+0.67}_{-3.67}$	$1.39^{+0.07}_{-0.06}$	$145.1 \pm 11.0$	PC
260004324.01	-	$3.8157 \pm 0.0007$	$1357.9369 \pm 0.0036$	$17.89^{+1.98}_{-2.73}$	$0.021^{+0.001}_{-0.002}$	$91.12^{+1.23}_{-2.14}$	$1.15^{+0.07}_{-0.09}$	$24.6 \pm 1.9$	PC
262530407.01	177.01	$2.8540 \pm 0.0001$	$1364.7070 \pm 0.0004$	$16.69^{+1.86}_{-2.17}$	$0.033^{+0.001}_{-0.001}$	$90.95^{+1.45}_{-1.11}$	$1.87^{+0.08}_{-0.08}$	$39.7 \pm 3.3$	PC
278661431.01	-	$17.6317 \pm 0.0048$	$1343.9299 \pm 0.0025$	$46.41^{+0.26}_{-0.38}$	$0.094^{+0.002}_{-0.002}$	$89.98^{+0.12}_{-0.08}$	$2.82^{+0.11}_{-0.10}$	$1.1 \pm 0.1$	pPC
305048087.01	237.01	$5.4310 \pm 0.0014$	$1376.9753 \pm 0.0037$	$35.04^{+3.47}_{-5.12}$	$0.073^{+0.004}_{-0.006}$	$90.33^{+0.80}_{-1.14}$	$1.67^{+0.12}_{-0.14}$	$3.6 \pm 0.4$	PC
307210830.01	175.01	$3.6893 \pm 0.0001$	$1374.6508 \pm 0.0005$	$22.34^{+1.64}_{-2.52}$	$0.039^{+0.001}_{-0.001}$	$90.35^{+0.88}_{-0.58}$	$1.34^{+0.05}_{-0.04}$	$12.0 \pm 0.9$	PC
415969908.01	233.01	$11.6658 \pm 0.0056$	$1376.9247 \pm 0.0040$	$45.23^{+4.77}_{-7.04}$	$0.046^{+0.003}_{-0.003}$	$90.11^{+0.67}_{-0.73}$	$1.90^{+0.13}_{-0.13}$	$3.8 \pm 0.3$	PC
415969908.02	-	$52.74^{+56.33}_{-20.31}$	$1381.0703 \pm 0.0042$	$98.07^{+90.89}_{-33.09}$	$0.05^{+0.00}_{-0.00}$	$90.01^{+0.36}_{-0.37}$	$2.02^{+0.22}_{-0.20}$	$0.5^{+0.5}_{-0.3}$	ST
441056702.01	-	$6.3424 \pm 0.0093$	$1371.8111 \pm 0.0055$	$20.88^{+3.80}_{-3.40}$	$0.031^{+0.003}_{-0.003}$	$89.81^{+1.67}_{-1.24}$	$1.97^{+0.19}_{-0.19}$	$25.0 \pm 2.2$	PC

**Notes.**

<sup>a</sup> Possible dispositions of objects of interest are a planet candidate (PC), a putative planet candidate (pPC), a single transit event (ST), or a putative single transit event (pST). The putative dispositions have FPP  $\in [0.1, 0.9]$  whereas the remaining candidates have FPP  $< 0.1$ .

Fig. 7.11 depicts our candidate population and compares it to the twelve TOIs whose TIC hosts are included in our stellar sample. Our candidates have orbital periods from 1.4-20 days for PCs and MAP orbital periods from 35-50 days for ST events, although ST periods exhibit large uncertainties of order the MAP period value. Our candidates have radii between  $1.1\text{-}3.8 R_{\oplus}$  potentially making them targets of interest for the completion of the TESS level one science requirement of delivering 50 planets with  $r_p < 4 R_{\oplus}$  with measured masses. In Sect. 7.5.4 we will discuss the prospects that our candidates have for contributing to the realization of the TESS level one science requirement. We do not detect any hot sub-Neptunes in the photoevaporation desert (Lundkvist et al., 2016) nor any small planets ( $\lesssim 1.5 R_{\oplus}$ ) on orbits longer than  $\sim 20$  days. This is largely attributable to our poor detection sensitivity in that regime due to the limited TESS baselines of just 27 days.

Fig. 7.11 also depicts our candidates as a function of insolation. The majority of candidates (10/15) experience incident insolation levels  $S$  in excess of twice that of the Earth. However five candidates, including all three ST events and two pPCs, are likely more temperate and experience insolations  $\lesssim 1.5 S_{\oplus}$ . This  $S$  limit marks the ‘recent-Venus’ inner edge of the low mass star habitable zone (HZ; Kopparapu et al., 2013). Our STs also lie within the more conservative HZ definition bounded by the ‘water-loss’ inner edge, where an increase in insolation results in the photolysis of stratospheric water vapor causing the atmosphere to experience rapid hydrogen escape, and the ‘maximum-greenhouse’ outer edge where an increase in CO<sub>2</sub> no longer results in a net surface heating due to the increased albedo. Our five temperate candidates may represent attractive targets for the characterization of HZ exoplanets around nearby low mass stars. We will address the prospect of atmospheric characterization of these planets in Sect. 7.5.4.

The distribution of our candidates versus stellar parameters of interest are shown in Fig. 7.12. All candidates are detected around stars hotter than 3000 K which approximately corresponds to stars earlier than M5.5V (Pecaut & Mamajek, 2013). Most candidates are detected around stars with  $T_{\text{eff}} \in [3200, 3900]$  K which is largely consistent with the population of confirmed transiting planets recovered from the NASA Exoplanet Archive on December 13, 2018 (Akeson et al., 2013), modulo the TRAPPIST-1 planets (Gillon et al., 2017), GJ 1214b (Charbonneau et al., 2009), and the Kepler-42 planets (Muirhead et al., 2012b). The median effective temperature of the candidate-hosting TIC members in our sample is 3560 K. A notable dearth of PCs with  $r_p \gtrsim 2 R_{\oplus}$  exists around stars hotter than  $\sim 3500$  K. The cause of this is unlikely to be due to sensitivity losses around these relatively hot (and correspondingly bright) stars and is instead likely attributable to the sharp decrease in the number of stars within our sample at  $T_{\text{eff}} \sim 3550$  K (see Fig. 7.1).

The distributions of previously confirmed planets and our candidates versus  $J$ -band magnitude reveals that many ( $\sim 9/15$ ) of our candidates orbit systematically brighter stars than the majority of confirmed transiting planets with  $J < 10$ . This directly demonstrates the power of a survey mission like TESS at discovering transiting planets orbiting nearby bright stars which are amenable to forthcoming detailed characterization efforts. The remaining six candidates still orbit moderately bright stars with  $J \in [10, 12]$ . The median  $J$  of our candidate-hosting TIC members is 9.9. From the number of ORION candidates as a function of  $J$ , it is clear that the ORION sensitivity to planets orbiting low mass TIC members starts to drop off around  $J \gtrsim 12$ . The large photometric uncertainties in this regime are largely

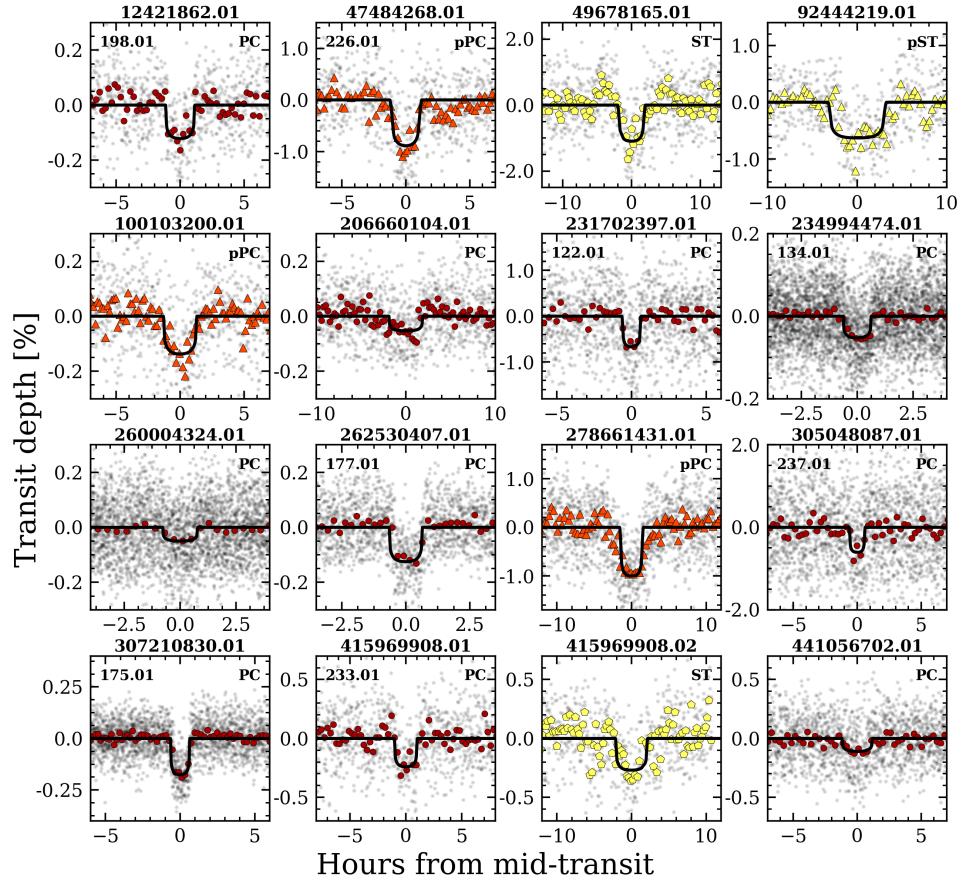


Figure 7.10: Phase-folded transit light curves for our set of 15 planet candidates. The temporal binning and axes ranges for each candidate are chosen to optimize visual clarity. The marker colors for each candidate's binned light curve is indicative of its disposition which also is annotated in the upper right of its panel. Candidates detected by ORION which are also TESS Objects of Interest have their TOI ID annotated in the upper left.

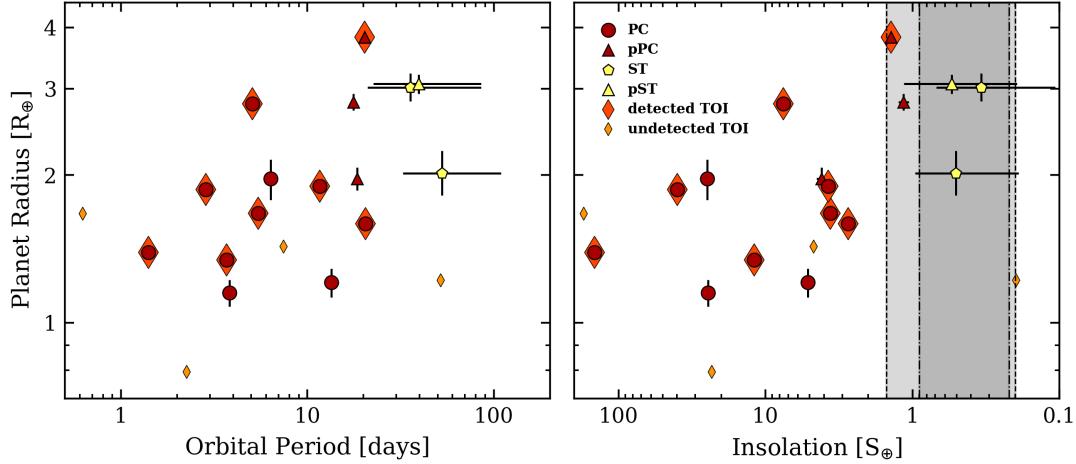


Figure 7.11: The resulting planet candidates from running ORION on the 2 minute extracted light curves from the first two TESS sectors in the period/radius and insolation/radius parameter spaces. The legend labels are planet candidates (PC), putative planet candidates (pPC), single transit events (ST), and putative single transit events (pST). TOIs which are also detected by ORION are highlighted with orange diamonds surrounding the associated candidate's marker. TOIs which remain undetected by ORION are depicted as small orange diamonds. The outer shaded region in the insolation panel marks the ‘recent-Venus’ and ‘early-Mars’ limits of the habitable zone around low mass stars from Kopparapu et al. (2013). The inner shaded region marks the more conservative ‘water-loss’ and ‘maximum-greenhouse’ habitable zone limits.

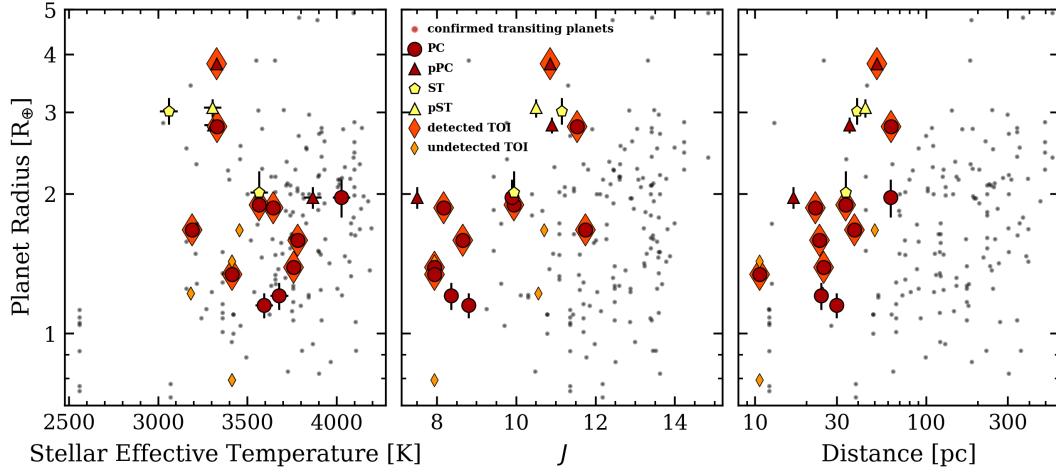


Figure 7.12: The planetary radii of our 15 ORION candidates as functions of their host stellar effective temperatures,  $J$ -band magnitudes, and Gaia distances. Our candidates are compared to the population of confirmed transiting planets around cool stars ( $T_{\text{eff}} < 4200$  K) from the NASA Exoplanet Archive which are depicted with small black circles. The legend labels are planet candidates (PC), putative planet candidates (pPC), single transit events (ST), and putative single transits (pST). TOIs which are also detected by ORION are highlighted with orange diamonds surrounding the associated PC marker. TOIs remaining undetected by ORION are depicted as small orange diamonds.

dominated by photon-noise from the target star, zodiacal light, and unresolved background stars ([Ricker et al., 2015](#)).

The distance distribution of our candidates is also included in Fig. 7.12. All of our candidates are found between 10-65 pc with a median distance of 34 pc. Our sample includes five candidates within  $\sim 25$  pc as well as the closest transiting PC around a low mass star to date: TIC 307210830.01 (TOI-175.01) at 10.6 pc which also contains two additional TOIs ([Kostov et al., 2019](#)) not detected by ORION (see Sect. 7.5.3). Among our seven new candidates not released as TOIs, we detect one candidate within 25 pc: the PC 206660104.01. Barring the rejection of this candidate and the TOIs around TIC 307210830, these two systems of four planet candidates represent 2/7 of the closest transiting planetary systems around low mass stars along with GJ 1132 ([Berta-Thompson et al., 2015; Bonfils et al., 2018](#)), TRAPPIST-1 ([Gillon et al., 2017; Luger et al., 2017](#)), LHS 1140 ([Dittmann et al., 2017b; Ment et al., 2019](#)), GJ 1214 ([Charbonneau et al., 2009](#)), and GJ 3470 ([Bonfils et al., 2012](#)).

## 7.5 Discussion

### 7.5.1 Outstanding vetting criteria

A number of potentially useful diagnostics of the validity of transit-like events should be implemented to improve the sensitivity and systematic false positive rate of future ORION versions. These include the characterization of internal scattered light from centroid offsets between the average in and out-of-transit pixel response functions, consideration of each epoch’s quality flag from the SPOC ([Jenkins et al., 2016](#)), and other stochastic effects on photometry such as imperfect systematics corrections at the edges of the light curves. These and other vetting criteria are expected to be realized through the analysis of additional TESS data from its many up-coming observing sectors and will help to make the ORION pipeline more robust.

### 7.5.2 Discussions of individual systems with planet candidates

*OI 12421862.01.* This PC has already been reported as TOI-198.01. The ORION planet parameters are consistent with those of TOI-198.01 with a 20.4 day orbital period and measured radius of  $1.6 R_{\oplus}$ , making it an interesting target for probing the photoevaporation valley around M dwarfs.

*OI 47484268.01.* This pPC based on its moderate FPP, has already been reported as TOI-226.01. The ORION orbital period is consistent with that of TOI-226.01 but we derive a 10% smaller planet radius of  $3.8 R_{\oplus}$  despite the star’s refined radius being 15% larger. At  $1.4 S_{\oplus}$ , this pPC orbits within the ‘recent-Venus’ habitable zone but remains an attractive target for rapid RV characterization owing to its expected large mass compared to most of the smaller candidate planets (see Sect. 7.5.4).

*OI 49678165.01.* A new candidate ST event with an estimated period between 21-85 days and a radius of  $3 R_{\oplus}$ . This candidate is likely the coldest object in our sample with a MAP  $S = 0.3 S_{\oplus}$  and a corresponding equilibrium temperature of  $T_{eq} = 212$  K assuming zero albedo and perfect heat redistribution.

*OI 92444219.01.* A pST based on its moderate FPP. This cool  $3 R_{\oplus}$  candidate is has an estimated period between 22-95 days placing it at  $S = 0.5 S_{\oplus}$  with an equilibrium temperature  $T_{\text{eq}} = 238$  K.

*OI 206660104.01.* A new terrestrial-sized ( $1.2 R_{\oplus}$ ) PC at 13.4 days. This PC is the second smallest candidate recovered by ORION and is smaller than most (3/4) of the TOIs missed by ORION (see Sect. 7.5.3).

*OI 231702397.01.* This PC has already been reported as TOI-122.01. The ORION orbital period is consistent with that of TOI-122.01 although we derive a 15% larger radius ( $2.8 R_{\oplus}$ ) corresponding to the 13% larger stellar radius in our sample. Given its large size, the expected mass from the [Chen & Kipping \(2017\)](#) mass-radius relation is  $8.26 M_{\oplus}$  making it an attractive target for rapid RV follow-up despite being relatively dim with  $J = 11.5$ .

*OI 234994474.01:* This PC has been reported as TOI-134.01 and is being validated with RVs from HARPS and PFS ([Astudillo-Defru et al., 2019](#)). The ORION parameters are largely consistent with the TOI-134.01 parameters albeit with a slightly smaller radius of  $1.38 R_{\oplus}$  (12% reduction) for an unchanged stellar radius. Being by-far the hottest target in candidate sample ( $T_{\text{eq}} = 965$  K) and orbiting an early M dwarf with  $J = 7.9$ , this PC is one of the best targets for any further RV characterization to search for longer period companions, transmission spectroscopy, and even thermal emission spectroscopy. Regarding the latter, TOI-134.01 is even more favorable than the previously most attractive terrestrial-sized planet for such observations: GJ 1132b ([Morley et al., 2017](#)).

*OI 260004324.01.* A new terrestrial-sized PC which is the smallest in our sample at  $1.15 R_{\oplus}$  with a 3.8 day period. Its small size and small expected mass of  $1.6 M_{\oplus}$  around a  $0.56 M_{\oplus}$  early M dwarf ( $J = 8.8$ ) will make this a slightly more challenging target for RV follow-up but may still be of interest for probing the low mass end of the 50 planets smaller than  $4 R_{\oplus}$  targeted for completion of the TESS level one science requirement.

*OI 262530407.01.* This PC has already been reported as TOI-177.01. The ORION orbital period is consistent with that of TOI-177.01 although we find a 12% smaller radius of  $1.87 R_{\oplus}$  for an unchanged stellar radius. The short period and host star brightness ( $J = 8.17$ ) make this PC an exceptional candidate for transmission spectroscopy observations and is one that is less affected by the *NIRISS* bright limit of its SOSS mode compared to other close-in super-Earth-sized planets in this catalog.

*OI 278661431.01.* A new pPC based on its moderate FFP. If validated, the candidate would have an orbital period of 17 days and a radius of  $2.8 R_{\oplus}$ . This pPC is temperate at  $S = 1.1 S_{\oplus}$  and near to the ‘water-loss’ HZ inner edge. This pPC orbits one of the cooler TIC members in our candidate-hosting sample ( $T_{\text{eff}} = 3300$  K) with a correspondingly small radius of  $0.28 R_{\odot}$ , thus making it an attractive target for transmission spectroscopy observations.

*OI 305048087.01.* This PC has already been reported as TOI-237.01. The ORION orbital period is consistent with the 5.4 day period and  $1.7 R_{\oplus}$  radius of TOI-177.01.

*OI 307210830.01.* This PC is one of the three reported TOIs around TIC 307210830 (i.e. 175-01, 02, 03). Only TOI-175.01 is detected by ORION for the reasons discussed in Sect. 7.5.3. The ORION planet parameters for this, the middle planet in this candidate three planet system, are all consistent with those for TOI-175.01. This PC is in the closest planetary system in our sample and is correspondingly an attractive target for RV characterization of individual masses and the RV analysis of the possible resonant pair 175.01/175.02 for comparisons to TTV analyzes from follow-up photometry. This PC is also a viable target for the atmospheric characterization of a terrestrial-sized planet ( $1.3 R_{\oplus}$ ).

*OI 415969908.01.* This PC has already been reported as TOI-233.01. The ORION orbital period is consistent with that of TOI-233.01 but finds a 26% smaller radius ( $1.9 R_{\oplus}$ ) in part because of the refined stellar radius being reduced by 21% in our sample.

*OI 415969908.02.* The ST event detected in the light of TIC 415969908 which already hosts the aforementioned candidate TOI-233.01 at 11.7 days. The estimated period of OI 415969908.02 is 32-108 days which effectively spans the conservative HZ limits from Kopparapu et al. (2013) and whose MAP value is 53 days or  $S = 0.5 S_{\oplus}$ . The measured radius is  $2 R_{\oplus}$  ( $r_p/R_s = 0.05$ ) making this ST event the most difficult to follow-up from the ground among the three ST events in our candidate catalog.

*OI 441056702.01.* A new PC with an orbital period of 6.3 days and radius of  $2 R_{\oplus}$  thus potentially being located within the photoevaporation valley around M dwarfs. Orbiting a moderately bright early-M to late-K dwarf ( $T_{\text{eff}} = 4030$  K) and with an expected mass of  $4.5 M_{\oplus}$ , this PC represents another attractive target for RV characterization aimed at addressing the TESS level one science requirement.

### 7.5.3 TOIs undetected by ORION

Our stellar sample using refined stellar parameters based on Gaia distances, contains ten TIC members with TOIs listed as part of the TESS alerts<sup>6</sup>. These ten systems host a total of twelve TOIs in nine single PC systems plus TIC 307210830 which hosts three TOIs. The ORION results presented in this paper include the independent detection of 8/12 TOIs with an additional ST event around TIC 415969908 which already hosts TOI-233.01. The discrepancies between the TESS alerts and our ORION results are described in what follows.

*TIC 259962054.* This star was observed in the consecutive TESS sectors 1 and 2. The TOI-203.01 has an orbital period of 52 days, longer than any repeating candidate in our catalog. A signal at  $\sim 52.1$  days is found in the ORION linear and periodic search stages with an orbital phase that is consistent with that reported for TOI-203.01. This suggests that the TOI-203.01 transit-like signal at 52 days does exist in the light curve. The phase-folded light curve satisfies all but the second criterion from the automated vetting stage (see Sect. 7.3.5) with  $S/N_{\text{transit}} = 5.6 < c_1 = 8.4$ . This is principally because the fitted transit depth within ORION is  $Z = 1839$  ppm which is just  $\sim 73\%$  of the TOI's reported depth thus making it difficult to confirm the nature of the repeating signal as due to a transiting planet with just two transits observed.

---

<sup>6</sup>As of December 19, 2018.

*TIC 307210830.* This system contains three TOIs (175.01, 175.02, 175.03) at 3.69, 7.45, and 2.25 days respectively. The innermost planet candidate is not found during the linear and periodic search stages. This is likely caused by the candidate's depth of 571 ppm (as reported by the SPOC) being less than the median photometric uncertainty of its light curve (i.e. median( $\sigma_f$ )  $\sim$  770 ppm). The two remaining planet candidates were each seen to be detected in the ORION linear search stage owing to their  $\sim 3$  times larger transit depths. However, this candidate pair has an orbital period ratio that is within 1% of a 2:1 resonant configuration. Recall that pairs of periods of interest which are that close to an integer period ratio will have one of those periods automatically discarded in the periodic search stage due to the detected periodic signals likely being aliases of the each other rather than being due to two separate planetary candidates.

*TIC 316937670.* TOI-221.01 has an orbital period of 0.624 days and a low transit depth of 954 ppm. By adopting the reported TOI-221.01 transit depth and duration, we estimate CDPP<sub>transit</sub> and find that  $Z/\text{CDPP}_{\text{transit}} \sim 1.1$ . Because of this, the results of the ORION linear search only detect a single transit-event above the S/N threshold such that no periodic events can be found. If the linear search S/N threshold is lowered such that multiple transits from TOI-221.01 are detected in the linear and periodic search stages, then the expected S/N<sub>transit</sub> is  $\sim 7.2 < c_1 = 8.4$  which would still be insufficient to detect the PC in a single TESS sector. We are further discouraged by the prospect of lowering the linear search S/N threshold as this would drastically increase the number of FPs purely from the noise.

### 7.5.4 Prospects for follow-up observations

#### Mass characterization via precision radial velocities

The TESS level one science requirement is to deliver at least 50 planets smaller than  $4 R_{\oplus}$  with measured masses via precision radial velocity (RV) follow-up. All fifteen of our candidates have a measured radius consistent with being  $< 4 R_{\oplus}$  (see Table 7.4). Using the empirical mass-radius relation for small planets  $< 14.26 R_{\oplus}$  from [Chen & Kipping \(2017\)](#), we compute the maximum likelihood masses  $m_p$  of our planet candidates and single transit events to then infer their expected RV semi-amplitudes using

$$K_{\text{RV}} = 2.4 \text{ m/s} \left( \frac{m_p}{5 M_{\oplus}} \right) \left( \frac{M_s}{0.5 M_{\odot}} \right)^{-2/3} \left( \frac{P}{10 \text{ days}} \right)^{-1/3}. \quad (7.8)$$

These values are shown in the first panel of Fig 7.13 versus *J*-band magnitude and are accompanied by the simulated TESS yield from [Barclay et al. \(2018\)](#) and the set of confirmed transiting planets from the NASA Exoplanet Archive. Both of the latter samples are restricted to cool host stars (i.e.  $T_{\text{eff}} < 4200$  K). Many existing high performance RV spectrographs are stable at the level of a few cm/s but RV observations are often limited by photon noise and intrinsic stellar activity at the level of one to a few m/s, even for relatively inactive stars ([Fischer et al., 2016](#)). The majority of our objects of interest have expected  $K_{\text{RV}}$  values in excess of this typical RV sensitivity limit and are reported in Table 7.5.

Table 7.5: Metric values indicating the feasibility of a variety of follow-up programs for our 16 vetted candidates

OI	TOI	$J$	$P$ [days]	$r_p$ [ $R_\oplus$ ]	$m_p^a$ [ $M_\oplus$ ]	$K$ [m/s]	$\Omega^b$	$T_{\text{eq}}^c$ [K]	TSM <sup>d</sup>	ESM <sup>e</sup>
12421862.01	198.01	8.65	20.43	1.59	3.16	1.19	0.58	358	53.4	0.6
47484268.01	226.01	10.85	20.28	3.82	14.03	5.93	<b>1.40</b>	302	69.0	0.9
49678165.01	-	11.14	35.64	3.02	9.40	4.40	0.92	212	57.7	0.1
92444219.01	-	10.49	39.52	3.07	9.68	3.35	0.90	238	51.1	0.1
100103200.01	-	7.50	18.45	1.97	4.54	1.61	<b>0.74</b>	397	<b>99.6</b>	2.0
206660104.01	-	8.36	13.45	1.21	1.92	0.76	0.51	419	5.9	0.7
231702397.01	122.01	11.53	5.08	2.80	8.26	6.35	<b>1.63</b>	461	69.3	3.2
234994474.01	134.01	7.94	1.40	1.39	2.52	1.98	<b>1.24</b>	965	<b>14.2</b>	<b>9.8</b>
260004324.01	-	8.80	3.82	1.15	1.62	0.98	0.74	619	7.6	2.1
262530407.01	177.01	8.17	2.85	1.87	4.16	2.74	<b>1.32</b>	698	<b>119.8</b>	<b>10.1</b>
278661431.01	-	10.89	17.63	2.82	8.36	5.14	1.08	288	<b>86.2</b>	0.7
305048087.01	237.01	11.74	5.43	1.67	3.44	4.06	0.95	384	66.2	1.3
307210830.01	175.01	7.93	3.69	1.34	2.37	2.15	<b>0.87</b>	517	<b>26.4</b>	6.7
415969908.01	233.01	9.94	11.67	1.90	4.26	2.21	0.84	387	55.5	1.1
415969908.02	-	9.94	52.74	2.02	4.74	1.48	0.54	234	36.4	0.1
441056702.01	-	9.89	6.34	1.97	4.53	2.17	<b>1.06</b>	622	39.8	2.4

**Notes.**

Bolded values are indicative of candidates that exceed threshold values of that parameter (see Cloutier et al. 2018a for  $\Omega$  and Kempton et al. 2018 for the TSM and ESM) and should be strongly considered for rapid confirmation and follow-up.

<sup>a</sup> Planet masses are estimated from the planet radius using the deterministic version of the mass-radius relation from Chen & Kipping (2017).

<sup>b</sup>  $\Omega$  is a diagnostic metric that is indicative of the observing time required to characterize a planet’s RV mass (Cloutier et al., 2018a).  $\Omega = r_p/P^{1/3}$  where  $r_p$  is given in Earth radii and  $P$  in days.

<sup>c</sup> Planetary equilibrium temperature is calculated assuming zero albedo and full heat redistribution via  $T_{\text{eq}} = T_{\text{eff}}\sqrt{R_s/2a}$ .

<sup>d</sup> The transmission spectroscopy metric from (Kempton et al., 2018). See Sect. 7.5.4 for the definition.

<sup>e</sup> The emission spectroscopy metric from (Kempton et al., 2018). See Sect. 7.5.4 for the definition.

[Cloutier et al. \(2018a\)](#) calculated the observing time required to complete the TESS level one science requirement based on the expected TESS yield from [Sullivan et al. \(2015\)](#). The 50 TOIs requiring the shortest time commitment to characterizing their planet masses at  $5\sigma$  with RVs satisfy the following empirically-derived conditions:

$$J < 11.7 \text{ and} \quad (7.9)$$

$$\Omega > 0.14J - 0.35, \quad (7.10)$$

where

$$\Omega = \left( \frac{r_p}{R_\oplus} \right) \left( \frac{P}{\text{day}} \right)^{-1/3} \quad (7.11)$$

is a proxy for  $K_{\text{RV}}$  that can be computed from transit-derived parameters. Six out of fifteen of our candidates satisfy Eqs. 7.9 and 7.10 and are highlighted in Table 7.5 and in Fig. 7.13. If astrophysical false positive scenarios can be ruled out, then these OIs represent highly favorable targets for RV follow-up observations and the rapid completion of the TESS level one science requirement. This is assuming that the host stars themselves are not active which is an important characteristic to consider for RV follow-up ([Moutou et al., 2017](#)) and is one that is not taken into account in Eqs. 7.9 and 7.10.

### Atmospheric characterization

TESS will provide many of the best transiting exoplanets for atmospheric characterization in the near future. [Kempton et al. \(2018\)](#) presented a framework to prioritize transiting planets for either transmission spectroscopy or emission spectroscopy observations with dedicated missions like *JWST* and *ARIEL*. This framework consists of analytical metrics which quantify the expected S/N ratio of transmission and emission signals from planetary atmospheres.

The transmission spectroscopy metric from [Kempton et al. \(2018\)](#) is

$$\text{TSM} = f \cdot 10^{-0.2J} \cdot \left( \frac{r_p^3 T_{\text{eq}}}{m_p R_s^2} \right) \quad (7.12)$$

The TSM represents the expected S/N of 10 hour observing programs with *JWST/NIRISS* assuming fixed atmospheric compositions for different planet types, cloud-free atmospheres, and a deterministic planet mass-radius relation. The planetary equilibrium temperature in Eq. 7.12 is  $T_{\text{eq}} = T_{\text{eff}} \sqrt{R_s/2a}$  where  $a$  is the planet's semimajor axis and is calculated assuming zero albedo and full heat redistribution over the planetary surface. The scale factor  $f$  is used to make the TSM non-dimensional and is used to correct discrepancies between the analytical TSM and the detailed simulations from [Louie et al. \(2018\)](#) using the *NIRISS* simulator for Single Object Slitless Spectroscopy (SOSS) observations. Values of  $f$  are reported in [Kempton et al. \(2018\)](#) each of four planet types separately: terrestrials ( $r_p < 1.5 R_\oplus$ ), super-Earths ( $1.5 < r_p/R_\oplus < 2.75$ ) sub-Neptunes ( $2.75 < r_p/R_\oplus < 4$ ), and giants ( $4 < r_p/R_\oplus < 10$ ). We calculate the TSM for our fifteen candidates and report those values in Table 7.5 and in Fig. 7.13. [Kempton et al. \(2018\)](#) highly recommends planets for atmospheric characterization (and therefore a-priori RV characterization) based on their TSM values relative to their derived cutoffs (see their Table

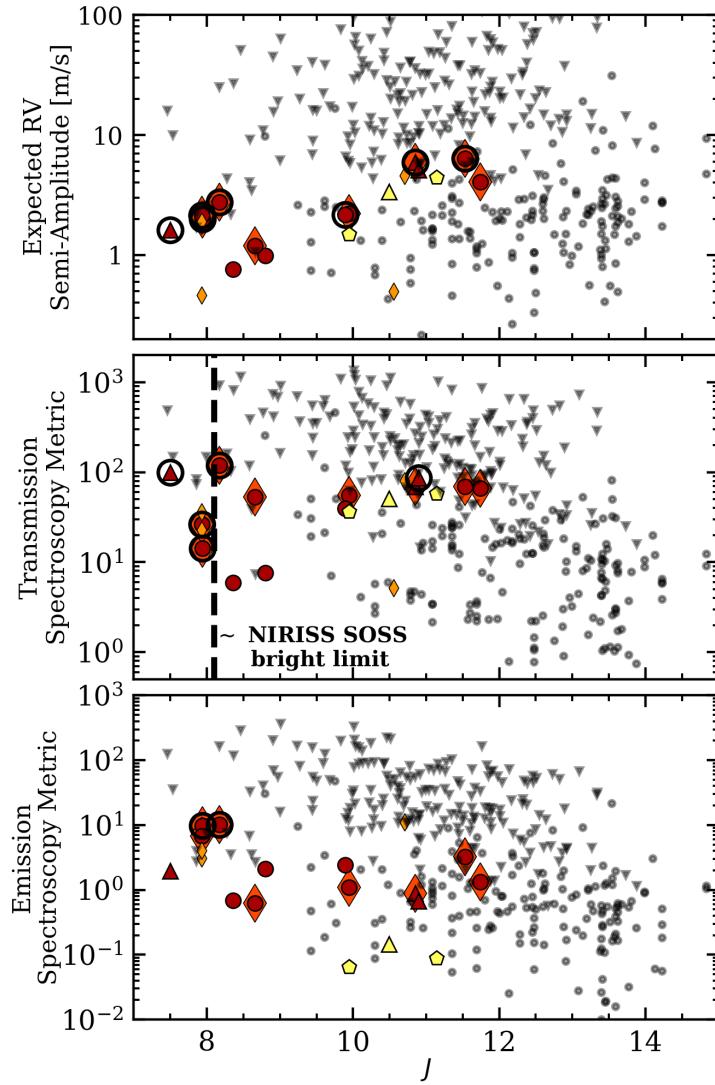


Figure 7.13: Expected values of the RV semi-amplitude (upper panel), transmission spectroscopy metric (middle panel), and emission spectroscopy metric (lower panel) for our set of 15 candidates as a function of  $J$ -band magnitude. The marker legend for our candidates is identical to those in Figs. 7.11 and 7.12 and are not included here for visual clarity. Candidate planets exceeding parameter cutoffs (Eqs. 7.9 and 7.10 for  $\Omega$  and see Kempton et al. 2018 for the TSM and ESM cutoffs) are highly favorable targets for follow-up observations and are highlighted by black rings. Also plotted are the expected values for simulated TESS planets around cool stars ( $T_{\text{eff}} < 4200$  K) from Barclay et al. (2018) (black inverted triangles) and for confirmed transiting planets around cool stars from the NASA Exoplanet Archive (black circles). To first order, our sample of candidate planets straddles the boundary between the expected TESS population and the population of known transiting planets.

1). Four out of fifteen candidates exceed the threshold TSM cutoff as highlighted in Table 7.5 and should be considered for confirmation as they represent highly attractive targets for transmission spectroscopy observations. We note however that host stars with  $J \lesssim 8.1$  begin to approach the *NIRISS* bright limit (depending on its spectral type) and will require specialized fast readout modes and/or the use of detector subarrays to be observed with *NIRISS* in its SOSS mode (Beichman et al., 2014).

Similarly to the TSM, Kempton et al. (2018) defined the thermal emission spectroscopy metric as

$$\text{ESM} = c \cdot 10^{-0.2K_s} \cdot \frac{B_{7.5}(T_{\text{day}})}{B_{7.5}(T_{\text{eff}})} \cdot \left( \frac{r_p}{R_s} \right)^2 \quad (7.13)$$

where  $B_{7.5}(T)$  is the Planck function of spectral irradiance evaluated for a given temperature  $T$  at  $7.5 \mu\text{m}$  and  $T_{\text{day}}$  is the planet's day-side temperature assumed to be  $1.1T_{\text{eq}}$ . The constant  $c = 4.29 \times 10^6$  is used to scale the ESM to yield the S/N of the reference planet GJ 1132b (Berta-Thompson et al., 2015; Dittmann et al., 2017a) in the center of the *MIRI* low resolution spectroscopy bandpass at  $7.5 \mu\text{m}$ . We calculate the ESM for our fifteen candidates and report those values in Table 7.5 and in Fig. 7.13. Kempton et al. (2018) advocate that planets with  $\text{ESM} > \text{ESM}_{\text{GJ } 1132} = 7.5$  should be considered favorable targets for thermal emission spectroscopy observations with *MIRI*. Within our candidate sample, 2/15 candidates exceed this threshold ESM and represent some of the best targets to-date for the characterization of terrestrial and super-Earth atmospheres with emission spectroscopy for the first time.

### 7.5.5 Comparison to yield simulations

Ballard (2018) performed a set of yield simulations focusing on M dwarfs not unlike the stellar population considered in this study. Ballard (2018) derive an ensemble completeness function for M dwarfs observed by TESS based on the simulated TESS yield from Sullivan et al. (2015) which includes details of the TESS footprint, systematics, the photometric error budget, and FP likelihoods. The expected TESS yield around M dwarfs is then derived by applying the completeness as a function of  $P$  and  $r_p$  to M dwarf planet occurrence rates. Said occurrence rate are derived from Dressing & Charbonneau (2015) and corrected for the eccentricity distribution (Limbach & Turner, 2015), dynamical stability (Fabrycky et al., 2012), and multiplicity effects according to the ‘Kepler dichotomy’ (Ballard & Johnson, 2016) of M dwarf planet populations: either high multiplicity systems ( $N > 5$ ) with low mutual inclinations or systems with lower multiplicity ( $N \sim 1 - 2$ ) and high mutual inclinations. The resulting TESS yield around M1-M4 dwarfs is predicted to be  $\sim 1100 \pm 220$  planets.

The following back-of-the-envelope calculation reveals how the expected M dwarf planet population to be discovered with TESS compares to our ORION results from the first two TESS sectors. First we note that the Sullivan et al. (2015) stellar population is a synthetic one derived from TRILEGAL galaxy model (Girardi et al., 2005). It contains 200,000 stars targeted by TESS which is effectively the size of the TIC (Stassun et al., 2017). Using the stellar parameters from the TIC-7 we find 53204 M1-M4 dwarfs in the TIC ( $T_{\text{eff}} \in [3200 - 3700] \text{ K}$ ; Pecaut & Mamajek 2013). Of these, 1624 and 1869 are targeted in sectors 1 and 2 respectively with 2849 being unique TIC members. However unlike in the TIC-7, our stellar sample is derived using parallaxes from the Gaia DR2 which results in a distinct population of just 1149 M1-M4 dwarfs observed in either or both of TESS sectors 1 and 2. We use a simple correction factor to the expected TESS yield to account for the fractionally fewer M1-M4 stars that we target for transit

searches compared to the TIC-7. This factor is  $f = 1149/2849 = 0.40 \pm 0.01$  where the  $f$  uncertainty is propagated from Poisson statistics.

The predicted number of M dwarf TESS planets discovered in sectors 1 and 2 is  $1100 \cdot w \sim 59 \pm 12$  where  $w = 2849/53204$  is the fraction of all M1-M4 dwarfs targeted in those sectors. Correcting this expected number of planets from the first two TESS sectors by the  $f$  times fewer M1-M4 dwarfs in our stellar sample compared to in the TIC-7, we find that we are expected to detect  $\sim 23 \pm 5$  planets in the first two TESS sectors. Note that this calculation inherently assumes that the detection completeness of ORION is equivalent to the ensemble completeness derived in [Ballard \(2018\)](#) which cannot be confirmed without performing a detailed characterization of the ORION completeness.

The ORION yield of fifteen planets is somewhat on the lower end of what is expected based on the TESS yield predictions. If the TESS alert TOIs around M1-M4 dwarfs that remain undetected by ORION in this study are also included in the cumulative number of planet detections then TESS has discovered  $\sim 19$  M1-M4 PCs in its first two sectors. The consistency between the ORION yield and the expected TESS yield speaks highly to TESS’s overall performance compared to its expected completeness from [Sullivan et al. \(2015\)](#) and [Ballard \(2018\)](#), as well as to the outstanding performance that the TESS mission has already achieved so early-on in its lifetime.

# Chapter 8

## Conclusions and Future Work

This thesis has focused on the development and implementation of semi-parametric Gaussian process (GP) regression modelling of stellar activity and other nuisance signals, with a focus on exoplanetary systems around low mass stars in particular. The GP formalism helps enable the detection as well as the precise and accurate characterization of planetary parameters from radial velocity and transit observations.

In chapter 3 I simulated the now ongoing SPIRou Legacy Survey-Planet Search (SLS-PS) in which I search for planets in synthetic RV time series using a GP treatment of activity to detect those planets and calculate the expected planet yield of the survey. In chapter 4 I derive an analytical framework to calculate the RV observational requirement to measure the masses of transiting planets at a desired level of precision in the presence of correlated RV residuals stemming from stellar activity and treated as a quasi-periodic GP. In chapters 5 and 6 the GP formalism is applied to empirical time series from K2, HARPS, and CARMENES of the transiting planetary system around the nearby mid-M dwarf K2-18.<sup>1</sup> Lastly, chapter 7 extends the one-dimensional GP regression modelling to the search for transiting planets in high cadence light curves from TESS. Here, GP modelling was used to treat stellar variability and residual systematic artifacts while searching for repeating transit-like events in the TESS 2 minute extracted light curves of low mass stars.

Each of the applications of the GP formalism in the aforementioned chapters has led to a number of potential improvements or continued paths of investigation that are discussed in the following sections.

### 8.1 On the Accuracy of the SLS-PS Planet Yield Predictions

Although not a critical assessment, in four years time it would be interesting to compare the actual yield of the SLS-PS to the predictions presented in chapter 3. The details of making this comparison one-to-one will need to be worked out as the actual time series obtained during the SLS-PS will certainly differ from those simulated in terms of their sizes, sampling, and measurement uncertainties. This may or may not prove to be a major hindrance however because it was determined through comparison of many

---

<sup>1</sup>I also applied this formalism to a pair of studies not presented in this thesis. Namely, [Bonfils et al. \(2018\)](#) and [Ment et al. \(2019\)](#).

flavours of the simulated SLS-PS that for  $N_{\text{RV}} \gtrsim 50$ , that the GP treatment of stellar activity results in nearly white RV residuals such that the detection sensitivity within the survey scales approximately as  $\sqrt{N_{\text{RV}}}$  and as  $\sigma_{\text{RV}}^{-1}$ .

Discrepancies between the predicted and actual planet yields could stem from any of a number of inconsistencies. The assumed planet occurrence rates used throughout were derived from Kepler statistics (i.e. as a function of planet radius) and converted to planetary masses using an empirical mass-radius relation. Inaccuracies from either calculation in general or their applicability to the stars in the SLS-PS target list may lead to a misestimated planet yield prediction. For example, the input Kepler statistics were derived from late K to early M dwarfs (Dressing & Charbonneau, 2015) and may not be applicable to the full range of M dwarf spectral types targeted in the SLS-PS. Another danger with simulated time series is that injected signals may be too simplistic or that the number of injected signals is incomplete. For example, the inclusion of detailed systematic RV noise sources may have had a direct affect the survey's detection sensitivity. Another major source of potential inconsistency is that the nature of detecting planets in the SLS-PS was based on an automated detection algorithm intended to emulate the steps taken by a human when searching for RV planets. Indeed the data flow from RV surveys is sufficiently small that manual planet detection is manageable (unlike in large transit surveys) and variances between human performance and the automated detection algorithm from chapter 3 may result in a different planet yield than was initially predicted. Overall, it would be an interesting exercise to isolate which of these discrepancies contributes to the observed differences between the predicted planet yield and the true planet yield following the conclusion of the SLS-PS.

## 8.2 Foreseen Improvements to the RVFC

Recall that the Radial Velocity Follow-up Calculator (RVFC) is a tool used to calculate the number of RV observations required to measure a transiting planet's RV semi-amplitude  $K$  at a desired level of precision given the properties of the planet, the host star, and the RV spectrograph. The properties of the latter two are required, in part, to compute the photon noise limited RV precision using model spectra from which the RV information content can be quantified. However, late K to M dwarf model spectra are known to be incomplete largely due to the abundance of molecular transitions that are not well understood and thus lead to discrepancies between empirical spectra of cool stars and their model atmospheres (Passegger et al., 2016; Behmard et al., 2019). Updating the M dwarf model spectra within the RVFC will improve the accuracy of the calculated  $\sigma_\gamma$  values and consequently the accuracy of the RVFC results for cool stars.

Another limiting factor to the accuracy of computed values of  $\sigma_\gamma$  is the treatment of telluric contamination. Currently in the RVFC regions of the spectral domain for which the vertically integrated atmospheric transmission is below some user-defined threshold are omitted from the calculation. As noted in Sect. 1.5.2, these regions are often wide in the infrared such that calculations of  $\sigma_\gamma$  in near-IR bands are likely underestimated relative to the values that could be obtained if a more sophisticated treatment of telluric contamination was applied (e.g. Artigau et al., 2014; Bedell et al., 2019). Implementation of an accurate model of the telluric spectrum and its subsequent subtraction from the observed spectrum would significantly increase the usable amount of spectral information for stars observed with

near-IR spectrographs and would thus decrease values of  $\sigma_\gamma$  in the near-IR bringing them closer to typical values obtained in the optical (see Fig. 4.2).

Another point of potential improvement to the accuracy of the RVFC is to allow for custom window functions to be used. In the special case of planets plus white noise in RV time series, the exact window function does not affect the  $K$  measurement precision  $\sigma_K$  as long as the time series approximately samples the planet’s orbital phase uniformly. However when correlated noise persists, the form of the window function matters because of the factors  $t_i - t_j$  in the quasi-periodic covariance function (Eq. 2.16). Currently the RVFC adopts a uniformly spaced window function for correlated noise calculations but the option for users to define their own window functions that take into account target visibility and observatory restrictions would be desirable.

### 8.3 Improving our Physical Understanding of M dwarf Stellar Activity

The Sun is unique in that its surface can be resolved with a variety of observational diagnostics of its brightness, surface velocity, magnetic field strength, and orientation distributions. When those observables are obtained simultaneously with other time series that are obtainable for distant stars (e.g. disk integrated photometry and RVs), then resolved structures on the solar surface can be related to the other observables and inform models of solar activity ([Dumusque et al., 2015](#); [Haywood et al., 2016](#)). Numerous programs that I have recently become involved with aim to observe a sample of M dwarfs with a similar strategy. That is, to obtain contemporaneous optical and near-IR RVs along with polarimetric diagnostics of bright M dwarfs spanning a range of stellar masses and rotation periods. This large collaboration known as the RVxTESS program has proposed to obtain those observations of M dwarfs within a suite of TESS sectors as they are being observed with TESS. The goal of this and other similar programs is to provide as many contemporaneous diagnostics of M dwarf stellar activity in order to infer the physical nature of activity signals that are seen in the RVs and that hinder our ability to detect low mass planets.

Deepening our understanding of the physical sources of M dwarf stellar may also be useful for modifying our GP models of stellar activity in that additional covariance terms may be required or these correlated noise models may need to be supplemented with other physical models of RV activity (e.g. [Aigrain et al., 2012](#); [Haywood et al., 2014](#)). Another potential outcome of these studies is to establish which observational diagnostics are optimal in terms of informing RV activity models as I have discerned that in practice, stellar photometry often provides the strongest constraints on RV activity despite the photometric time series used throughout chapters 5 and 6 not being contemporaneous with the RV measurements.

## 8.4 Detailing the Accuracy of RV Planet Models in the Presence of a GP

Since some of the original implementations of GP models of correlated noise in RV time series, the results have been widely accepted due to the ability of the formalism to return precise planetary parameters in models that are favoured over models lacking such a treatment of stellar activity (Haywood et al., 2014; Rajpaul et al., 2015) and to reconcile discrepant results from different RV teams (Rajpaul et al., 2017; Cloutier et al., 2019b). What remains to be quantified is the accuracy of the derived planetary parameters from RV models containing a GP activity model. For example, the GP treatment of stellar activity is known to at times absorb some of the planetary signal (e.g. Ribas et al., 2018). Furthermore, GP activity models that are untrained are at times disfavoured over models with trained GP models as they lack the constraining power provided by ancillary activity-sensitive time series like photometry (Cloutier et al., 2017b).

One avenue worth considering to quantify the accuracy of planet parameters derived from RV models containing a GP activity model is to compute planet parameters from a suite of synthetic RV time series for which the injected planet parameters are known. Calculating the bias function as the difference between the maximum a-posteriori  $K$  (for example) and the injected  $K$  for models with and without a GP would inform the ways in which the GP biases our inference of observable planetary parameters and consequently their physical properties. Such a study would likely be complex as the nature of the bias would likely depend on the exact form of the GP, its mean model (i.e. the planetary solution), and sampling that are all unique for every system studied with RVs. Inherent biases in the planetary mass-radius relation resulting from only reporting high S/N planet mass detections has already been noted (Burt et al., 2018). But whether or not the incompleteness of GP activity models is resulting in biased planetary mass measurements, we will want known about it.

## 8.5 Characterizing the Performance of ORION

Although the TESS mission was not designed as a statistical mission to study the bulk properties of the exoplanet population and subsets therein, characterization of the sensitivity and false positive rate of transit detection algorithms will enable planet occurrence rates to be derived from the primary mission results. Particularly for close-in planets where TESS is most sensitive due its short, 27 day fields. However these planetary systems are still of particular interest given structure that was recently detected in the occurrence rate of close-in planets around Sun-like stars postulated to be the result of photoevaporation on the planetary atmospheres Owen & Wu (2013); Fulton et al. (2017); Van Eylen et al. (2018).

A thorough suite of injection-recovery simulations using the ORION transit detection code (see chapter 7) can be used to characterize the algorithm's detection sensitivity and systematic false positive rate as a function of planetary orbital period and  $r_p/R_s$ . Application to the TESS data will then allow for the occurrence rate of close-in planets around M dwarfs to be calculated as TESS features an enhanced sensitivity to early-to-mid M dwarfs compared to Kepler. Focusing on the detailed occurrence rates of these planetary systems from TESS will enable the resolution of the radius valley for M dwarf planetary systems and for its measured location and its slope with orbital period to be compared to the results

obtained for Sun-like stars in various mass bins ([Fulton & Petigura, 2018](#)).

## Chapter 9

# Appendix: summaries of additional papers not presented in this thesis

The following is a chronological list of papers that include direct contributions from myself and have not been included as chapters in this thesis. Lead author papers in the following subsections are accompanied by a brief paper summary whereas papers for which I was a contributing author are accompanied by descriptions of those contributions.

### 9.1 Prospects for Detecting the Rossiter-McLaughlin Effect of Earth-like Planets: the test case of TRAPPIST-1b and c ([Cloutier & Triaud, 2016](#))

The Rossiter-McLaughlin (RM) effect, or spectroscopic transit, is an observable effect that is sensitive to the sky-projected spin-orbit angle  $\lambda$  of a planet's orbital plane relative to the host star's spin axis (see Fig. 9.1). The measurement of  $\lambda$  may be used to inform formation models and planet dynamical histories. Notably, it has yet to be measured for any terrestrial-sized exoplanet ( $\lesssim 1.6 R_{\oplus}$ ) owing to their typically shallow transit depths. In this paper, which followed shortly after the discovery of at least three terrestrial-sized planets orbiting the ultracool dwarf TRAPPIST-1 ([Gillon et al., 2016](#)), we argued that the two innermost planets represent ideal targets for being the best terrestrial-sized exoplanets amenable to the detection of the RM effect discovered to date. Due to the small size of TRAPPIST-1, its short rotation period (i.e. large  $v \sin i_s$ ), and its relatively low level of activity compared to other ultracool dwarfs, we expect the semi-amplitudes of the TRAPPIST-1b and c RM effects to be  $\sim 40 - 50 \text{ m s}^{-1}$  which is an order of magnitude greater than the amplitude of the Doppler reflex motion induced by the planets on their host star. Simulations of the RM effect showed that if the planets are well-aligned, then  $\lambda$  can be measured with a precision of  $\lesssim 10^\circ$  in an RV time series with typical measurement uncertainties of  $2 \text{ m s}^{-1}$ , although multiple transits will be required due to the faintness of TRAPPIST-1 ( $J = 11.4$ ).

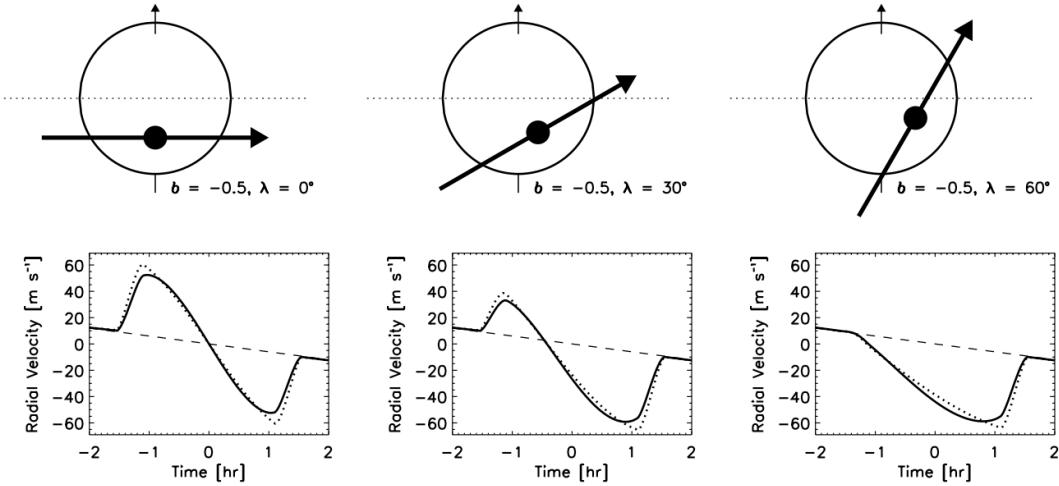


Figure 9.1: Dependence of the anomalous RV waveform from the Rossiter-McLaughlin effect on the sky-projected spin-orbit angle  $\lambda$ . All depicted trajectories during the planetary transit have the same impact parameter  $b = -0.5$ . The solid and dashed curves correspond to different assumed limb darkening coefficients. (Image credit: [Gaudi & Winn, 2007](#))

## 9.2 On the Radial Velocity Detection of Additional Planets in Transiting, Slowly Rotating M dwarf Systems: The Case of GJ 1132 ([Cloutier et al., 2017a](#))

M dwarfs are known to frequently host multi-planet systems such that we can reasonably expect to find additional planets with RVs around M dwarfs that are known to host at least one transiting planet. In this paper we focused on one such system, namely the nearby mid-M dwarf GJ 1132, which hosts a small transiting planet at  $P = 1.6$  days. This system is amenable to deep RV follow-up both for the precise mass measurement of GJ 1132b and for the search for non-transiting planets in the system. We created synthetic RV time series by injecting planets from the empirical M dwarf occurrence rates plus physical models of stellar activity. We then applied our GP formalism for modelling stellar activity (Sect. 2.2) to attempt to recover those planets. From these calculations we compute our sensitivity to the detection of additional planets in the GJ 1132 planetary system as a function of the number of RV measurements. We show that at  $1 \text{ m s}^{-1}$  precision,  $\sim 50\%$  of all non-transiting planets from the expected planet population can be detected with  $\sim 50$  measurements.

Indeed there turns out to be at least one additional planet in the GJ 1132 planetary system that is not transiting and is presented in Sect. 9.5.

## 9.3 Near-Infrared Planet Searcher to Join HARPS on the ESO 3.6-metre Telescope ([Bouchy et al., 2017](#))

This paper summarized the current status of the up-coming near-IR spectrograph NIRPS (*Near-Infrared Planet Searcher*) including an overview of its design and primary science goals. As an instrument paper

all science team members were added as co-authors including myself as a Canadian collaborator. I did however contribute directly to the discussion of the NIRPS science goal of searching for new RV planetary systems in the southern sky that may be amenable to direct imaging with ELTs. Those specific contributions were calculations of the expected NIRPS planet yield. These calculations were based on the simulations presented in Chapter 3 for the SPIRou Legacy Survey-Planet Search after modifications to the simulated observing strategy and instrument performance were made as they pertain to NIRPS.

## 9.4 Quantifying the Evidence for a Planet in Radial Velocity Data ([Nelson et al., 2018](#))

The paper presents the results of a collaborative study that was conceived in a breakout session during the *Extremely Precise Radial Velocities III* conference at Penn State. When searching for planets in an RV dataset  $\mathbf{y}$ , robustly quantifying the detection of a planet requires the Bayesian evidence  $\mathcal{Z}$  of an RV model  $M$ , that contains that planet, to be greater than the evidence of a competing model  $M'$  that does not. If the model is parameterized by a set of model parameters  $\theta$  then the evidence integral is written as

$$\mathcal{Z}(\mathbf{y}|M) = \int_{\theta} \mathcal{L}(\mathbf{y}|\theta, M) \Pi(\theta|M) d\theta \quad (9.1)$$

and requires one to integrate the product of the data likelihood  $\mathcal{L}(\mathbf{y}|\theta, M)$  and full model parameter prior  $\Pi(\theta|M)$  over the full model parameter space  $\theta$ . The evidence integral is expensive to compute accurately and many computational methods have been proposed to calculate or approximate  $\mathcal{Z}$ .

In this study, a number of astronomers from differing RV groups were given identical synthetic RV datasets with injected planets that we participants were agnostic to. The groups were tasked with using our favourite methods to calculate or approximate  $\mathcal{Z}$  using those datasets and a consistent set of priors. This paper was comparative study of those methods.

My contribution was to test two non-Bayesian methods that operate analogously to the Bayesian evidence ratio (i.e. ratio of two model evidences  $\mathcal{Z}_1$  and  $\mathcal{Z}_2$ ) for the purposes of model comparison. These methods were leave-one-out cross-validation and time series cross-validation. Instead of calculating  $\mathcal{Z}$  they are used to compute the predictive power of a given model such that if a model containing  $N + 1$  planets demonstrates more predictive power on previously unseen RV measurements (i.e. it is more accurate), then  $N + 1$  planets are favoured in the dataset than just  $N$  planets where  $N = 0, 1, 2$  in the study. The two methods were shown to perform comparably to the median Bayesian method when comparing their model comparison diagnostics to Bayesian evidence ratios.

## 9.5 Radial Velocity Follow-up of GJ 1132 with HARPS: a Precise Mass for Planet ‘b’ and the Discovery of a Second Planet ([Bonfils et al., 2018](#))

This paper presents the results of an intensive RV follow-up campaign with HARPS of the GJ 1132 mid-M dwarf planetary system. My contribution to this work was to conduct an independent analysis of the RV time series using my GP formalism. This analysis was complementary to those performed by the lead authors which were based on more traditional methods that lacked a treatment of temporally correlated stellar activity signals. With the results of my analysis we measured a precise mass of the known transiting planet GJ 1132b of  $1.66 \pm 0.23 M_{\oplus}$  (i.e. a  $7.2\sigma$  mass detection). We also detected an additional signal at  $\sim 8.9$  days that favours a second terrestrial mass planet as postulated by my work in Sect. 9.2. Yet another significant signal was seen in the RVs at  $\sim 177$  days although its interpretation as planetary or as an activity-induced signal remains an open question.

## 9.6 A Second Terrestrial Planet Orbiting the Nearby M Dwarf LHS 1140 ([Ment et al., 2019](#))

This paper presents the results of another intensive RV follow-up campaign with HARPS. This time of the LHS 1140 mid-M dwarf planetary system. LHS 1140 was known to host a small transiting HZ planet ( $P = 24.7$  days,  $r_{p,b} = 1.73 R_{\oplus}$ ) discovered from the ground with the MEarth telescope array ([Dittmann et al., 2017b](#)). Our RV follow-up campaign, which included an independent analysis from myself using my GP formalism, resulted in a significant improvement to the measurement precision of the planet’s mass and also revealed a second strong periodic signal at  $\sim 3.8$  days. Re-investigation of the MEarth photometry by the lead authors revealed a second small transiting planet that was missed in the initial light curve analysis, and with an orbital period that was consistent with the 3.8 day RV signal. Through my reanalysis of the RV data now including two planetary signals we recovered the planet masses and confirmed the presence of at least two terrestrial planets around LHS 1140.

## 9.7 Characterization of the L 98-59 Multi-Planetary System with HARPS: Two Confirmed Terrestrial Planets and a Mass Upper Limit on the Third ([Cloutier et al., 2019a](#))

Similarly to the announcement of the planetary candidate around GJ 4332, the TESS team also reported a compact system of three transiting planet candidates orbiting the mid-M dwarf L 98-59 from sector 1 (i.e. TOI-175.01,02,03). From follow-up observations including ground-based photometry, reconnaissance spectroscopy, high-resolution imaging, and dynamical stability arguments, [Kostov et al. \(2019\)](#) statistically validated each of the three planets. I led the HARPS RV follow-up of this system and applied my usual analysis techniques to measure precise planet masses for the two outermost planets ( $r_{p,c} = 1.35, r_{p,d} = 1.57 R_{\oplus}$ ) and place an upper mass limit on the smallest inner planet ( $r_{p,b} = 0.80 R_{\oplus}$ ).

Similarly to the dynamical analysis performed on the K2-18 two-planet system, I ran a suite of N-body simulations of the L 98-59 planetary system to constrain their eccentricities. These results supplemented the orbital eccentricity constraints from the RVs themselves and together placed upper limits on each planet's orbital eccentricity of  $< 0.1$  at 95% confidence to ensure a dynamically stable system given their newly measured masses.

# Bibliography

- Adams, E. R., Jackson, B., & Endl, M. 2016, AJ, 152, 47
- Agol, E., & Deck, K. 2016, ApJ, 818, 177
- Aigrain, S., Hodgkin, S. T., Irwin, M. J., Lewis, J. R., & Roberts, S. J. 2015, MNRAS, 447, 2880
- Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, MNRAS, 459, 2408
- Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
- Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, PASP, 125, 989
- Alibert, Y., & Benz, W. 2017, A&A, 598, L5
- Alibert, Y., Carron, F., Fortier, A., et al. 2013, A&A, 558, A109
- Allen, L. E., Wright, J., Rajagopal, J., et al. 2018, in American Astronomical Society Meeting Abstracts, Vol. 231, American Astronomical Society Meeting Abstracts, 246.08
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O’Neil, M. 2014, ArXiv e-prints, arXiv:1403.6015
- . 2015, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, arXiv:1403.6015
- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Nature, 536, 437
- Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018, MNRAS, 474, 2094
- Arlot, S., & Celisse, A. 2010, Statist. Surv., 4, 40
- Artigau, É., Malo, L., Doyon, R., et al. 2018, AJ, 155, 198
- Artigau, É., Kouach, D., Donati, J.-F., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 15
- Astraatmadja, T. L., & Bailer-Jones, C. A. L. 2016, ApJ, 832, 137
- Astudillo-Defru, N., Delfosse, X., Bonfils, X., et al. 2017a, A&A, 600, A13
- Astudillo-Defru, N., Bonfils, X., Delfosse, X., et al. 2015, A&A, 575, A119
- Astudillo-Defru, N., Forveille, T., Bonfils, X., et al. 2017b, A&A, 602, A88

- Astudillo-Defru, N., Wang, S., Cloutier, R., et al. 2019, ArXiv e-prints, arXiv:in prep.
- Bailer-Jones, C. A. L. 2015, PASP, 127, 994
- Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
- Baliunas, S. L., Donahue, R. A., Soon, W., & Henry, G. W. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 154, Cool Stars, Stellar Systems, and the Sun, ed. R. A. Donahue & J. A. Bookbinder, 153
- Baliunas, S. L., Donahue, R. A., Soon, W. H., et al. 1995, ApJ, 438, 269
- Ballard, S. 2018, ArXiv e-prints, arXiv:1801.04949
- Ballard, S., & Johnson, J. A. 2016, ApJ, 816, 66
- Ballard, S., Fabrycky, D., Fressin, F., et al. 2011, ApJ, 743, 200
- Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, A&A, 337, 403
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&As, 119, 373
- Barclay, T., Pepper, J., & Quintana, E. V. 2018, ArXiv e-prints, arXiv:1804.05050
- Barnes, J. R., & Collier Cameron, A. 2001, MNRAS, 326, 950
- Barnes, J. R., Collier Cameron, A., Donati, J.-F., et al. 2005, MNRAS, 357, L1
- Barnes, J. R., James, D. J., & Collier Cameron, A. 2004, MNRAS, 352, 589
- Barnes, R., & Greenberg, R. 2006, ApJL, 647, L163
- Barnes, S. A. 2003, ApJ, 586, 464
- Bastien, F. A., Stassun, K. G., Basri, G., & Pepper, J. 2013, Nature, 500, 427
- Bastien, F. A., Stassun, K. G., Pepper, J., et al. 2014, AJ, 147, 29
- Batalha, N. M., Rowe, J. F., Gilliland, R. L., et al. 2010, ApJL, 713, L103
- Beaulieu, J.-P., Bennett, D. P., Fouqué, P., et al. 2006, Nature, 439, 437
- Bedding, T. R., Butler, R. P., Kjeldsen, H., et al. 2001, ApJL, 549, L105
- Bedell, M., Hogg, D. W., Foreman-Mackey, D., Montet, B. T., & Luger, R. 2019, arXiv e-prints, arXiv:1901.00503
- Behmard, A., Petigura, E. A., & Howard, A. W. 2019, arXiv e-prints, arXiv:1904.00094
- Beichman, C., Benneke, B., Knutson, H., et al. 2014, PASP, 126, 1134
- Benedict, G. F., Henry, T. J., Franz, O. G., et al. 2016, AJ, 152, 141
- Benneke, B., Werner, M., Petigura, E., et al. 2017, ApJ, 834, 187
- Berdyugina, S. V. 2005, Living Reviews in Solar Physics, 2, doi:10.12942/lrsp-2005-8

- Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, *ApJ*, 866, 99
- Berta-Thompson, Z. K., Irwin, J., Charbonneau, D., et al. 2015, *Nature*, 527, 204
- Bertaux, J. L., Lallement, R., Ferron, S., Boonne, C., & Bodichon, R. 2014, *A&A*, 564, A46
- Birkby, J. L., de Kok, R. J., Brogi, M., et al. 2013, *MNRAS*, 436, L35
- Boisse, I., Bonfils, X., & Santos, N. C. 2012, *A&A*, 545, A109
- Boisse, I., Bouchy, F., Hébrard, G., et al. 2011, *A&A*, 528, A4
- Boisse, I., Moutou, C., Vidal-Madjar, A., et al. 2009, *A&A*, 495, 959
- Bonfils, X., Mayor, M., Delfosse, X., et al. 2007, *A&A*, 474, 293
- Bonfils, X., Gillon, M., Udry, S., et al. 2012, *A&A*, 546, A27
- Bonfils, X., Delfosse, X., Udry, S., et al. 2013a, *A&A*, 549, A109
- Bonfils, X., Lo Curto, G., Correia, A. C. M., et al. 2013b, *A&A*, 556, A110
- Bonfils, X., Almenara, J. M., Jocou, L., et al. 2015, in SPIE, Vol. 9605, Techniques and Instrumentation for Detection of Exoplanets VII, 96051L
- Bonfils, X., Astudillo-Defru, N., Díaz, R., et al. 2017, ArXiv e-prints, arXiv:1711.06177
- Bonfils, X., Almenara, J.-M., Cloutier, R., et al. 2018, *A&A*, 618, A142
- Bouchy, F., Pepe, F., & Queloz, D. 2001, *A&A*, 374, 733
- Bouchy, F., & Sophie Team. 2006, in Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies, ed. L. Arnold, F. Bouchy, & C. Moutou, 319–325
- Bouchy, F., Udry, S., Mayor, M., et al. 2005, *A&A*, 444, L15
- Bouchy, F., Doyon, R., Artigau, É., et al. 2017, *The Messenger*, 169, 21
- Bouma, L. G., Winn, J. N., Kosiarek, J., & McCullough, P. R. 2017, ArXiv e-prints, arXiv:1705.08891
- Bovy, J., Rix, H.-W., Green, G. M., Schlafly, E. F., & Finkbeiner, D. P. 2016, *ApJ*, 818, 130
- Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012, *ApJ*, 757, 112
- Broeg, C., Fortier, A., Ehrenreich, D., et al. 2013, in European Physical Journal Web of Conferences, Vol. 47, European Physical Journal Web of Conferences, 03005
- Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2012, *Nature*, 486, 502
- Browning, M. K. 2008, *ApJ*, 676, 1262
- Bryson, S. T., Jenkins, J. M., Gilliland, R. L., et al. 2013, *PASP*, 125, 889
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *A&A*, 564, A125
- Burke, C. J., McCullough, P. R., Valenti, J. A., et al. 2007, *ApJ*, 671, 2115

- Burt, J., Holden, B., Wolfgang, A., & Bouma, L. G. 2018, AJ, 156, 255
- Butler, R. P., & Marcy, G. W. 1996, ApJL, 464, L153
- Carson, J. C., Eikenberry, S. S., Smith, J. J., & Cordes, J. M. 2006, AJ, 132, 1146
- Carter, J. A., Yee, J. C., Eastman, J., Gaudi, B. S., & Winn, J. N. 2008, ApJ, 689, 499
- Cegla, H. M., Stassun, K. G., Watson, C. A., Bastien, F. A., & Pepper, J. 2014, ApJ, 780, 104
- Cegla, H. M., Watson, C. A., Marsh, T. R., et al. 2012, MNRAS, 421, L54
- Chakraborty, A., Richardson, E. H., & Mahadevan, S. 2008, in SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 70144G
- Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJL, 529, L45
- Charbonneau, D., Berta, Z. K., Irwin, J., et al. 2009, Nature, 462, 891
- Chen, J., & Kipping, D. 2017, ApJ, 834, 17
- Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, PASP, 124, 1279
- Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2013, ApJs, 207, 35
- . 2015, ApJ, 810, 95
- . 2016, ApJ, 828, 99
- Christiansen, J. L., Vanderburg, A., Burt, J., et al. 2017, AJ, 154, 122
- Ciardi, D. R., Fabrycky, D. C., Ford, E. B., et al. 2013, ApJ, 763, 41
- Claret, A. 2017, A&A, 600, A30
- Claret, A., & Bloemen, S. 2011, A&A, 529, A75
- Claudi, R., Benatti, S., Carleo, I., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99081A
- Clough, S. A., Shephard, M. W., Mlawer, E. J., et al. 2005, JQSRT, 91, 233
- Cloutier, R. 2018, arXiv e-prints, arXiv:1812.08145
- Cloutier, R., Doyon, R., Bouchy, F., & Hébrard, G. 2018a, AJ, 156, 82
- Cloutier, R., Doyon, R., Menou, K., et al. 2017a, AJ, 153, 9
- Cloutier, R., Tamayo, D., & Valencia, D. 2015, ApJ, 813, 8
- Cloutier, R., & Triaud, A. H. M. J. 2016, MNRAS, 462, 4018
- Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2017b, A&A, 608, A35
- Cloutier, R., Artigau, É., Delfosse, X., et al. 2018b, AJ, 155, 93

- Cloutier, R., Astudillo-Defru, N., Bonfils, X., et al. 2019a, arXiv e-prints, arXiv:1905.10669
- Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2019b, *A&A*, 621, A49
- Collier Cameron, A., Wilson, D. M., West, R. G., et al. 2007, *MNRAS*, 380, 1230
- Cooke, B. F., Pollacco, D., West, R., McCormac, J., & Wheatley, P. J. 2018, *A&A*, 619, A175
- Correia, A. C. M., Couetdic, J., Laskar, J., et al. 2010, *A&A*, 511, A21
- Cosentino, R., Lovis, C., Pepe, F., et al. 2012a, in SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461V
- Cosentino, R., Lovis, C., Pepe, F., et al. 2012b, in SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461V
- Cowan, N. B., & Strait, T. E. 2013, *ApJl*, 765, L17
- Cowan, N. B., Agol, E., Meadows, V. S., et al. 2009, *ApJ*, 700, 915
- Crane, J. D., Shectman, S. A., Butler, R. P., et al. 2010, in SPIE, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 773553
- Crause, L. A., Sharples, R. M., Bramall, D. G., et al. 2014, in SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91476T
- Crepp, J. R., Crass, J., King, D., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 990819
- Crossfield, I. J. M. 2013, *A&A*, 551, A99
- Crossfield, I. J. M., Petigura, E., Schlieder, J. E., et al. 2015, *ApJ*, 804, 10
- Crossfield, I. J. M., Guerrero, N., David, T., et al. 2018, *ApJs*, 239, 5
- Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, *PASP*, 120, 531
- Currie, T., Burrows, A., Girard, J. H., et al. 2014, *ApJ*, 795, 133
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of point sources.
- Davis, A. B., Cisewski, J., Dumusque, X., Fischer, D. A., & Ford, E. B. 2017, *ApJ*, 846, 59
- Dawson, R. I., & Fabrycky, D. C. 2010, *ApJ*, 722, 937
- de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, *A&A*, 554, A82
- Deck, K. M., Agol, E., Holman, M. J., & Nesvorný, D. 2014, *ApJ*, 787, 132
- Delfosse, X., Forveille, T., Perrier, C., & Mayor, M. 1998, *A&A*, 331, 581
- Delfosse, X., Forveille, T., Ségransan, D., et al. 2000, *A&A*, 364, 217
- Delfosse, X., Bonfils, X., Forveille, T., et al. 2013a, *A&A*, 553, A8

- Delfosse, X., Donati, J.-F., Kouach, D., et al. 2013b, in SF2A-2013: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. L. Cambresy, F. Martins, E. Nuss, & A. Palacios, 497–508
- Demory, B.-O., Queloz, D., Alibert, Y., Gillen, E., & Gillon, M. 2016, *ApJl*, 825, L25
- Desort, M., Lagrange, A.-M., Galland, F., Udry, S., & Mayor, M. 2007, *A&A*, 473, 983
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., Berta-Thompson, Z. K., & Newton, E. R. 2017a, *AJ*, 154, 142
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017b, *Nature*, 544, 333
- Donati, J.-F., Catala, C., Landstreet, J. D., & Petit, P. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 358, Astronomical Society of the Pacific Conference Series, ed. R. Casini & B. W. Lites, 362
- Donati, J.-F., Morin, J., Petit, P., et al. 2008, *MNRAS*, 390, 545
- Donati, J.-F., Yu, L., Moutou, C., et al. 2017, *MNRAS*, 465, 3343
- Doyle, L. R., Carter, J. A., Fabrycky, D. C., et al. 2011, *Science*, 333, 1602
- Dravins, D., Lindegren, L., & Nordlund, A. 1981, *A&A*, 96, 345
- Dressing, C. D., & Charbonneau, D. 2013, *ApJ*, 767, 95
- . 2015, *ApJ*, 807, 45
- Dressing, C. D., Charbonneau, D., Dumusque, X., et al. 2015, *ApJ*, 800, 135
- Drimmel, R., Cabrera-Lavers, A., & López-Corredoira, M. 2003, *A&A*, 409, 205
- Dumusque, X. 2016, *A&A*, 593, A5
- . 2018, *A&A*, 620, A47
- Dumusque, X., Boisse, I., & Santos, N. C. 2014a, *ApJ*, 796, 132
- Dumusque, X., Santos, N. C., Udry, S., Lovis, C., & Bonfils, X. 2011a, *A&A*, 527, A82
- Dumusque, X., Udry, S., Lovis, C., Santos, N. C., & Monteiro, M. J. P. F. G. 2011b, *A&A*, 525, A140
- Dumusque, X., Pepe, F., Lovis, C., et al. 2012, *Nature*, 491, 207
- Dumusque, X., Bonomo, A. S., Haywood, R. D., et al. 2014b, *ApJ*, 789, 154
- Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015, *ApJL*, 814, L21
- Dumusque, X., Borsa, F., Damasso, M., et al. 2017, *A&A*, 598, A133
- Duncan, D. K., Vaughan, A. H., Wilson, O. C., et al. 1991, *ApJS*, 76, 383
- Esposito, M., Armstrong, D. J., Gandolfi, D., et al. 2018, arXiv e-prints, arXiv:1812.05881

- Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4
- Fabrycky, D. C., Ford, E. B., Steffen, J. H., et al. 2012, ApJ, 750, 114
- Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
- Faria, J. P., Haywood, R. D., Brewer, B. J., et al. 2016, A&A, 588, A31
- Figueira, P., Marmier, M., Boué, G., et al. 2012, A&A, 541, A139
- Figueira, P., Adibekyan, V. Z., Oshagh, M., et al. 2016, A&A, 586, A101
- Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
- Ford, E. B. 2006, ApJ, 642, 505
- Ford, E. B., & Gregory, P. C. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 371, Statistical Challenges in Modern Astronomy IV, ed. G. J. Babu & E. D. Feigelson, 189
- Ford, E. B., Seager, S., & Turner, E. L. 2001, Nature, 412, 885
- Foreman-Mackey, D. 2015, George: Gaussian Process regression, Astrophysics Source Code Library, ascl:1511.015
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
- Foreman-Mackey, D., Montet, B. T., Hogg, D. W., et al. 2015, ApJ, 806, 215
- Forveille, T., Bonfils, X., Delfosse, X., et al. 2009, A&A, 493, 645
- Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
- Fujii, Y., Kawahara, H., Suto, Y., et al. 2011, ApJ, 738, 184
- . 2010, ApJ, 715, 866
- Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264
- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
- Fürész, G., Simcoe, R., Barnes, S. I., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 990814
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016a, A&A, 595, A2
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016b, A&A, 595, A1
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
- Gaidos, E., Kitzmann, D., & Heng, K. 2017, MNRAS, 468, 3418
- Gaidos, E., Mann, A. W., Kraus, A. L., & Ireland, M. 2016, MNRAS, 457, 2877
- Gandolfi, D., Barragán, O., Livingston, J. H., et al. 2018, A&A, 619, L10

- Gardner, J. R., Pleiss, G., Bindel, D., Weinberger, K. Q., & Wilson, A. G. 2018, arXiv e-prints, arXiv:1809.11165
- Gaudi, B. S., & Winn, J. N. 2007, ApJ, 655, 550
- Ge, J., Ma, B., Sithajan, S., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99086I
- Gibson, S. R., Howard, A. W., Marcy, G. W., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 990870
- Giguere, M. J., Fischer, D. A., Zhang, C. X. Y., et al. 2016, ApJ, 824, 150
- Giles, H. A. C., Collier Cameron, A., & Haywood, R. D. 2017, MNRAS, 472, 1618
- Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, ApJS, 197, 6
- Gillon, M., Jehin, E., Delrez, L., et al. 2013, in Protostars and Planets VI Posters
- Gillon, M., Jehin, E., Magain, P., et al. 2011, in European Physical Journal Web of Conferences, Vol. 11, European Physical Journal Web of Conferences, 06002
- Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, Nature, 533, 221
- Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Nature, 542, 456
- Ginsburg, A., Parikh, M., Woillez, J., et al. 2017, Astroquery: Access to online data resources, Astrophysics Source Code Library, ascl:1708.004
- Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895
- Gladman, B. 1993, Icarus, 106, 247
- Gomes da Silva, J., Santos, N. C., Bonfils, X., et al. 2012, A&A, 541, A9
- Gray, D. F. 1989, PASP, 101, 832
- . 2008, The Observation and Analysis of Stellar Photospheres (Cambridge University Press)
- Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, ApJ, 810, 25
- Green, G. M., Schlafly, E. F., Finkbeiner, D., et al. 2018, MNRAS, 478, 651
- Gregory, P. C. 2005, ApJ, 631, 1198
- Grunblatt, S. K., Howard, A. W., & Haywood, R. D. 2015, ApJ, 808, 127
- Günther, M. N., Queloz, D., Demory, B.-O., & Bouchy, F. 2017, MNRAS, 465, 3379
- Hale, G. E. 1908, PASP, 20, 287
- Hall, J. C. 2008, Living Reviews in Solar Physics, 5, doi:10.1007/lrsp-2008-2
- Hall, J. C., Lockwood, G. W., & Skiff, B. A. 2007, AJ, 133, 862

- Halverson, S., Terrien, R., Mahadevan, S., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99086P
- Hathaway, D. H. 2010, Living Reviews in Solar Physics, 7, doi:10.12942/lrsp-2010-1
- Hatzes, A. P. 1996, PASP, 108, 839
- . 2016, A&A, 585, A144
- Hawley, S. L., Davenport, J. R. A., Kowalski, A. F., et al. 2014, ApJ, 797, 121
- Haywood, R. D. 2015, PhD thesis, University of St Andrews, doi:10.5281/zenodo.35161
- Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517
- Haywood, R. D., Collier Cameron, A., Unruh, Y. C., et al. 2016, MNRAS, 457, 3637
- Hébrard, É. M., Donati, J.-F., Delfosse, X., et al. 2014, MNRAS, 443, 2599
- . 2016, MNRAS, 461, 1465
- Henden, A., & Munari, U. 2014, Contributions of the Astronomical Observatory Skalnate Pleso, 43, 518
- Henry, T. 2009, in APS Southeastern Section Meeting Abstracts, H2
- Henry, T. J., Soderblom, D. R., Donahue, R. A., & Baliunas, S. L. 1996, AJ, 111, doi:10.1086/117796
- Herrero, E., Ribas, I., Jordi, C., et al. 2016, A&A, 586, A131
- Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2011, ApJ, 730, 10
- Howard, A. W., Sanchis-Ojeda, R., Marcy, G. W., et al. 2013, Nature, 503, 381
- Huang, C. X., Shporer, A., Dragomir, D., et al. 2018a, arXiv e-prints, arXiv:1807.11129
- Huang, C. X., Burt, J., Vanderburg, A., et al. 2018b, arXiv e-prints, arXiv:1809.05967
- Huélamo, N., Figueira, P., Bonfils, X., et al. 2008, A&A, 489, L9
- Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
- Irwin, J., Berta, Z. K., Burke, C. J., et al. 2011, ApJ, 727, 56
- Irwin, J. M., Berta-Thompson, Z. K., Charbonneau, D., et al. 2015, in Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, Vol. 18, 18th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. G. T. van Belle & H. C. Harris, 767–772
- Isaacson, H., & Fischer, D. 2010, ApJ, 725, 875
- Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010, in SPIE, Vol. 7740, Software and Cyberinfrastructure for Astronomy, 77400D
- Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, in SPIE, Vol. 9913, Software and Cyberinfrastructure for Astronomy IV, 99133E

- Johnson, J. A., Gazak, J. Z., Apps, K., et al. 2012, AJ, 143, 111
- Jurgenson, C., Fischer, D., McCracken, T., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99086T
- Kaltenegger, L., Segura, A., & Mohanty, S. 2011, ApJ, 733, 35
- Kaltenegger, L., & Traub, W. A. 2009, ApJ, 698, 519
- Kasper, M., Beuzit, J.-L., Verinaud, C., et al. 2010, in SPIE, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 77352E–77352E–9
- Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108
- Kempton, E. M.-R., Bean, J. L., Louie, D. R., et al. 2018, ArXiv e-prints, arXiv:1805.03671
- Kipping, D. M. 2013, MNRAS, 434, L51
- Kitchatinov, L. L., & Olemskoy, S. V. 2011, MNRAS, 411, 1059
- Knutson, H. A., Fulton, B. J., Montet, B. T., et al. 2014, ApJ, 785, 126
- Kochukhov, O. 2016, Basic Principles of Doppler Imaging
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
- Kostov, V. B., Schlieder, J. E., Barclay, T., et al. 2019, arXiv e-prints, arXiv:1903.08017
- Kotani, T., Tamura, M., Suto, H., et al. 2014, in SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 914714
- Kovács, G., Zucker, S., & Mazeh, T. 2002, A&A, 391, 369
- Kreidberg, L. 2015, PASP, 127, 1161
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, Nature, 505, 69
- Kroupa, P. 2001, MNRAS, 322, 231
- Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007a, ApJ, 660, 770
- Lafrenière, D., Doyon, R., Marois, C., et al. 2007b, ApJ, 670, 1367
- Lanotte, A. A., Gillon, M., Demory, B.-O., et al. 2014, A&A, 572, A73
- Leggett, S. K. 1992, ApJs, 82, 351
- Leung, H. W., & Bovy, J. 2019, MNRAS, 483, 3255
- Li, J., Tenenbaum, P., Twicken, J. D., et al. 2018, arXiv e-prints, arXiv:1812.00103
- Limbach, M. A., & Turner, E. L. 2015, Proceedings of the National Academy of Science, 112, 20
- Lindegren, L., & Dravins, D. 2003, A&A, 401, 1185
- Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2

- Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, *ApJs*, 197, 8
- Lithwick, Y., Xie, J., & Wu, Y. 2012, *ApJ*, 761, 122
- Livingston, J. H., Endl, M., Dai, F., et al. 2018, *AJ*, 156, 78
- Lockwood, G. W., Skiff, B. A., & Radick, R. R. 1997, *ApJ*, 485, 789
- Lopez, E. D., & Fortney, J. J. 2014, *ApJ*, 792, 1
- López-Morales, M., Haywood, R. D., Coughlin, J. L., et al. 2016, *AJ*, 152, 204
- Louie, D. R., Deming, D., Albert, L., et al. 2018, *PASP*, 130, 044401
- Lovis, C., & Pepe, F. 2007, *A&A*, 468, 1115
- Lovis, C., Mayor, M., Bouchy, F., et al. 2005, *A&A*, 437, 1121
- Lovis, C., Dumusque, X., Santos, N. C., et al. 2011, ArXiv e-prints, arXiv:1107.5325
- Lovis, C., Snellen, I., Mouillet, D., et al. 2017, *A&A*, 599, A16
- Luger, R., Agol, E., Kruse, E., et al. 2016, *AJ*, 152, 100
- Luger, R., Sestovic, M., Kruse, E., et al. 2017, *Nature Astronomy*, 1, 0129
- Luhman, K. L., Stauffer, J. R., Muench, A. A., et al. 2003, *ApJ*, 593, 1093
- Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, *Nature Communications*, 7, 11201
- Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, *A&A*, 616, A9
- Macintosh, B., Troy, M., Doyon, R., et al. 2006, in SPIE, Vol. 6272, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 62720N
- Mahadevan, S., Ramsey, L., Bender, C., et al. 2012, in SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461S
- Mahmud, N. I., Crockett, C. J., Johns-Krull, C. M., et al. 2011, *ApJ*, 736, 123
- Mandel, K., & Agol, E. 2002, *ApJL*, 580, L171
- Mann, A. W., Brewer, J. M., Gaidos, E., Lépine, S., & Hilton, E. J. 2013, *AJ*, 145, 52
- Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, *ApJ*, 804, 64
- Mann, A. W., Dupuy, T., Kraus, A. L., et al. 2019, *ApJ*, 871, 63
- Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007, *ApJ*, 655, 541
- Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, *ApJ*, 641, 556
- Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, *A&A*, 453, 635
- Martín, E. L., Guenther, E., Zapatero Osorio, M. R., Bouy, H., & Wainscoat, R. 2006, *ApJL*, 644, L75

- Maunder, E. W. 1904, MNRAS, 64, 747
- Mayor, M., & Queloz, D. 1995, Nature, 378, 355
- Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20
- Mazeh, T., Zucker, S., Dalla Torre, A., & van Leeuwen, F. 1999, ApJL, 522, L149
- McQuillan, A., Aigrain, S., & Mazeh, T. 2013, MNRAS, 432, 1203
- Ment, K., Dittmann, J. A., Astudillo-Defru, N., et al. 2019, AJ, 157, 32
- Meunier, N., Desort, M., & Lagrange, A.-M. 2010, A&A, 512, A39
- Meunier, N., Lagrange, A.-M., Mbemba Kabuiku, L., et al. 2017, A&A, 597, A52
- Miller-Ricci, E., Seager, S., & Sasselov, D. 2009, ApJ, 690, 1056
- Moffett, T. J. 1974, ApJs, 29, 1
- Montet, B. T., Morton, T. D., Foreman-Mackey, D., et al. 2015, ApJ, 809, 25
- Morgan, J., Kerins, E., Awiphan, S., et al. 2018, ArXiv e-prints, arXiv:1802.05645
- Morin, J., Donati, J.-F., Petit, P., et al. 2008, MNRAS, 390, 567
- Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., & Fortney, J. J. 2017, ApJ, 850, 121
- Mortier, A., Faria, J. P., Correia, C. M., Santerne, A., & Santos, N. C. 2015, A&A, 573, A101
- Morton, T. D. 2012, ApJ, 761, 6
- . 2015, VESPA: False positive probabilities calculator, Astrophysics Source Code Library, ascl:1503.011
- Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86
- Morton, T. D., & Swift, J. 2014, ApJ, 791, 10
- Moutou, C., Hébrard, E. M., Morin, J., et al. 2017, MNRAS, 472, 4563
- Muirhead, P. S., Dressing, C. D., Mann, A. W., et al. 2018, AJ, 155, 180
- Muirhead, P. S., Hamren, K., Schlawin, E., et al. 2012a, ApJL, 750, L37
- Muirhead, P. S., Johnson, J. A., Apps, K., et al. 2012b, ApJ, 747, 144
- Mulders, G. D., Pascucci, I., & Apai, D. 2015, ApJ, 798, 112
- Nelson, B. E., Ford, E. B., Buchner, J., et al. 2018, ArXiv e-prints, arXiv:1806.04683
- Newton, E. R., Charbonneau, D., Irwin, J., et al. 2014, AJ, 147, 20
- Newton, E. R., Irwin, J., Charbonneau, D., et al. 2016, ApJ, 821, 93
- Noguchi, K., Ando, H., Izumiura, H., et al. 1998, in SPIE, Vol. 3355, Optical Astronomical Instrumentation, ed. S. D'Odorico, 354–362

- Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, *ApJ*, 279, 763
- Nutzman, P., & Charbonneau, D. 2008, *PASP*, 120, 317
- Oliva, E., Origlia, L., Baffa, C., et al. 2006, in SPIE, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 626919
- O'Malley-James, J. T., & Kaltenegger, L. 2017, *MNRAS*, 469, L26
- O'Neal, D., Saar, S. H., Neff, J. E., & Cuntz, M. 2005, in ESA Special Publication, Vol. 560, 13th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun, ed. F. Favata, G. A. J. Hussain, & B. Battrick, 853
- Oshagh, M., Santos, N. C., Figueira, P., et al. 2017, *A&A*, 606, A107
- Owen, J. E., & Mohanty, S. 2016, *MNRAS*, 459, 4088
- Owen, J. E., & Wu, Y. 2013, *ApJ*, 775, 105
- Pan, M., & Schlichting, H. E. 2017, ArXiv e-prints, arXiv:1704.07836
- Passegger, V. M., Wende-von Berg, S., & Reiners, A. 2016, *A&A*, 587, A19
- Pecaut, M. J., & Mamajek, E. E. 2013, *ApJs*, 208, 9
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *Journal of Machine Learning Research*, 12, 2825
- Pepe, F., Mayor, M., Galland, F., et al. 2002, *A&A*, 388, 632
- Pepe, F., Mayor, M., Queloz, D., et al. 2004, *A&A*, 423, 385
- Pepe, F. A., Cristiani, S., Rebolo Lopez, R., et al. 2010, in SPIE, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 77350F
- Perrakis, K., Ntzoufras, I., & Tsionas, E. G. 2013, ArXiv e-prints, arXiv:1311.0674
- Perruchot, S., Bouchy, F., Chazelas, B., et al. 2011, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8151, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 15
- Petigura, E. A., Marcy, G. W., & Howard, A. W. 2013a, *ApJ*, 770, 69
- . 2013b, *ApJ*, 770, 69
- Petrovay, K., & van Driel-Gesztelyi, L. 1997, *Solar Physics*, 176, 249
- Pizzolato, N., Maggio, A., Micela, G., Sciortino, S., & Ventura, P. 2003, *A&A*, 397, 147
- Podgorski, W., Bean, J., Bergner, H., et al. 2014, in SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91478W
- Pont, F., Sing, D. K., Gibson, N. P., et al. 2013, *MNRAS*, 432, 2917

- Prato, L., Huerta, M., Johns-Krull, C. M., et al. 2008, ApJL, 687, L103
- Queloz, D., Henry, G. W., Sivan, J. P., et al. 2001, A&A, 379, 279
- Queloz, D., Bouchy, F., Moutou, C., et al. 2009, A&A, 506, 303
- Quillen, A. C., Morbidelli, A., & Moore, A. 2007, MNRAS, 380, 1642
- Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2014, in SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91471F
- Rajpaul, V., Aigrain, S., Osborne, M. A., Reece, S., & Roberts, S. 2015, MNRAS, 452, 2269
- Rajpaul, V., Aigrain, S., & Roberts, S. 2016, MNRAS, 456, L6
- Rajpaul, V., Buchhave, L. A., & Aigrain, S. 2017, MNRAS, 471, L125
- Rasmussen, C. E., & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning) (The MIT Press)
- Rayner, J., Bond, T., Bonnet, M., et al. 2012, in SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84462C
- Rayner, J. T., Cushing, M. C., & Vacca, W. D. 2009, ApJS, 185, 289
- Rein, H., & Liu, S.-F. 2012, A&A, 537, A128
- Rein, H., & Tamayo, D. 2015, MNRAS, 452, 376
- Reiners, A. 2009, A&A, 498, 853
- Reiners, A., & Basri, G. 2007, ApJ, 656, 1121
- Reiners, A., Bean, J. L., Huber, K. F., et al. 2010, ApJ, 710, 432
- Reiners, A., Shulyak, D., Anglada-Escudé, G., et al. 2013, A&A, 552, A103
- Reiners, A., Zechmeister, M., Caballero, J. A., et al. 2018, A&A, 612, A49
- Ribas, I., Tuomi, M., Reiners, A., et al. 2018, Nature, 563, 365
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in SPIE, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, 914320
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
- Rincon, F., Roudier, T., Schekochihin, A. A., & Rieutord, M. 2017, A&A, 599, A69
- Robertson, P., Bender, C., Mahadevan, S., Roy, A., & Ramsey, L. W. 2016, ApJ, 832, 112
- Robertson, P., Mahadevan, S., Endl, M., & Roy, A. 2014, Science, 345, 440
- Rodler, F., & López-Morales, M. 2014, ApJ, 781, 54

- Rodríguez-López, C., Gizis, J. E., MacDonald, J., Amado, P. J., & Carosso, A. 2015, MNRAS, 446, 2613
- Rogers, L. A. 2015, ApJ, 801, 41
- Route, M. 2016, ApJl, 830, L27
- Saar, S. H., & Donahue, R. A. 1997, ApJ, 485, 319
- Saar, S. H., Hatzes, A., Cochran, W., & Paulson, D. 2003, in Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, Vol. 12, The Future of Cool-Star Astrophysics: 12th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. A. Brown, G. M. Harper, & T. R. Ayres, 694–698
- Sahlmann, J., Lazorenko, P. F., Ségransan, D., et al. 2013, A&A, 556, A133
- Sanchis-Ojeda, R., Rappaport, S., Winn, J. N., et al. 2014, ApJ, 787, 47
- Santos, N. C., Gomes da Silva, J., Lovis, C., & Melo, C. 2010, A&A, 511, A54
- Santos, N. C., Mayor, M., Naef, D., et al. 2000, A&A, 361, 265
- Sarkis, P., Henning, T., Kürster, M., et al. 2018, AJ, 155, 257
- Scargle, J. D. 1982, ApJ, 263, 835
- Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
- Schmidt, S. J., Kowalski, A. F., Hawley, S. L., et al. 2012, ApJ, 745, 14
- Seager, S., & Mallén-Ornelas, G. 2003, ApJ, 585, 1038
- Seifahrt, A., Stürmer, J., Bean, J. L., & Schwab, C. 2018, ArXiv e-prints, arXiv:1805.09276
- Shan, Y., Johnson, J. A., & Morton, T. D. 2015, ApJ, 813, 75
- Shulyak, D., Reiners, A., Engeln, A., et al. 2017, Nature Astronomy, 1, 0184
- Shulyak, D., Reiners, A., Seemann, U., Kochukhov, O., & Piskunov, N. 2014, A&A, 563, A35
- Siverd, R. J., Brown, T. M., Hygelund, J., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99086X
- Skumanich, A. 1972, ApJ, 171, 565
- Sliski, D., Blake, C., Johnson, J. A., et al. 2017, in American Astronomical Society Meeting Abstracts, Vol. 229, American Astronomical Society Meeting Abstracts #229, 146.09
- Sliski, D. H., & Kipping, D. M. 2014, ApJ, 788, 148
- Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, PASP, 124, 1000
- Snellen, I., de Kok, R., Birkby, J. L., et al. 2015, A&A, 576, A59
- Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049

- Soummer, R., Pueyo, L., & Larkin, J. 2012, *ApJl*, 755, L28
- Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2017, ArXiv e-prints, arXiv:1706.00495
- Stefansson, G., Mahadevan, S., Hebb, L., et al. 2017, *ApJ*, 848, 9
- Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, *Nature*, 464, 1161
- Stewart, G. R., & Ida, S. 2000, *Icarus*, 143, 28
- Strassmeier, K. G., Ilyin, I., Järvinen, A., et al. 2015, *Astronomische Nachrichten*, 336, 324
- Suárez Mascareño, A., Rebolo, R., & González Hernández, J. I. 2016, *A&A*, 595, A12
- Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, *ApJ*, 809, 77
- Swift, J. J., Bottom, M., Johnson, J. A., et al. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 027002
- Szentgyorgyi, A., Baldwin, D., Barnes, S., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 990822
- Tal-Or, L., Zechmeister, M., Reiners, A., et al. 2018, *A&A*, 614, A122
- Tamura, M., Suto, H., Nishikawa, J., et al. 2012, in SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461T
- Thompson, S. J., Queloz, D., Baraffe, I., et al. 2016, in SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 99086F
- Torres, G., Andersen, J., & Giménez, A. 2010, *A&A*, 18, 67
- Torres, G., Kipping, D. M., Fressin, F., et al. 2015, *ApJ*, 800, 99
- Trauger, J., Gordon, B., Krist, J., & Moody, D. 2015, in SPIE, Vol. 9605, Techniques and Instrumentation for Detection of Exoplanets VII, 96050N
- Trifonov, T., Rybizki, J., & Kürster, M. 2018, ArXiv e-prints, arXiv:1812.04501
- Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2018, *Publications of the Astronomical Society of the Pacific*, 130, 064502
- Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, *PASP*, 115, 389
- Valencia, D., Guillot, T., Parmentier, V., & Freedman, R. S. 2013, *ApJ*, 775, 10
- Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007, *ApJ*, 665, 1413
- Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, *MNRAS*, 479, 4786
- Van Eylen, V., & Albrecht, S. 2015, *ApJ*, 808, 126
- Van Grootel, V., Gillon, M., Valencia, D., et al. 2014, *ApJ*, 786, 2
- Vanderburg, A., & Johnson, J. A. 2014, *PASP*, 126, 948

- Vanderburg, A., Plavchan, P., Johnson, J. A., et al. 2016, ArXiv e-prints, arXiv:1604.03143
- Vanderburg, A., Montet, B. T., Johnson, J. A., et al. 2015, ApJ, 800, 59
- Vanderspek, R., Huang, C. X., Vanderburg, A., et al. 2018, arXiv e-prints, arXiv:1809.07242
- Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in SPIE, Vol. 2198, Instrumentation in Astronomy VIII, ed. D. L. Crawford & E. R. Craine, 362
- Vogt, S. S., Radovan, M., Kibrick, R., et al. 2014, PASP, 126, 359
- Walkowicz, L. M., Basri, G., Batalha, N., et al. 2011, AJ, 141, 50
- Wang, J. J., Graham, J. R., Dawson, R., et al. 2018, AJ, 156, 192
- Wang, K. A., Pleiss, G., Gardner, J. R., et al. 2019, arXiv e-prints, arXiv:1903.08114
- Wargelin, B. J., Saar, S. H., Pojmański, G., Drake, J. J., & Kashyap, V. L. 2017, MNRAS, 464, 3281
- Weiss, L. M., & Marcy, G. W. 2014, ApJL, 783, L6
- Weiss, L. M., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 768, 14
- Weiss, L. M., Rogers, L. A., Isaacson, H. T., et al. 2016, ApJ, 819, 83
- West, A. A., Weisenburger, K. L., Irwin, J., et al. 2015, ApJ, 812, 3
- Wilson, O. C. 1968, ApJ, 153, 221
- Winn, J. N. 2010, Exoplanet Transits and Occultations, ed. S. Seager (University of Arizona Press), 55–77
- Winters, J. G., Henry, T. J., Lurie, J. C., et al. 2015, AJ, 149, 5
- Wise, A. W., Dodson-Robinson, S. E., Bevenour, K., & Provini, A. 2018, AJ, 156, 180
- Wolfgang, A., Rogers, L. A., & Ford, E. B. 2016, ApJ, 825, 19
- Wright, J. T., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2004, ApJs, 152, 261
- Yu, L., Donati, J.-F., Hébrard, E. M., et al. 2017, MNRAS, 467, 1342
- Zacharias, N., Finch, C. T., Girard, T. M., et al. 2013, AJ, 145, 44
- Zeng, L., & Sasselov, D. 2013, PASP, 125, 227