

ESSP Check-In 2

Nov. 13, 2020

Andrews

Scalpels

Andrew Collier Cameron



University of
St Andrews

Ancy Anna John



Eric Ford



PennState



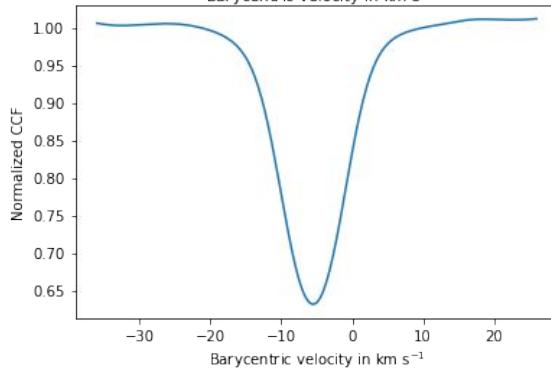
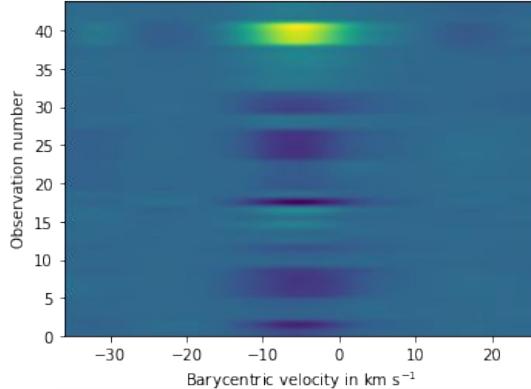
SCALPELS

Collier Cameron, Ford et al 2020arXiv201100018C

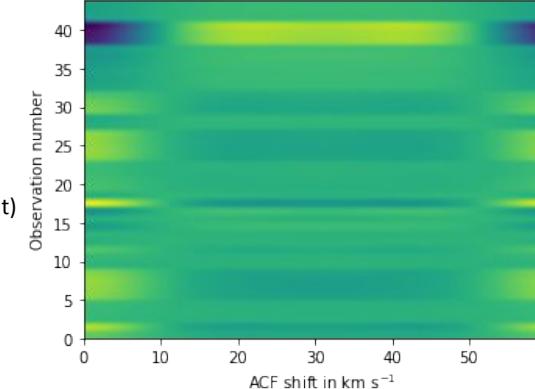
University of
St Andrews

Self-Correlation Analysis of Line Profiles for Extraction of Low-amplitude Shifts

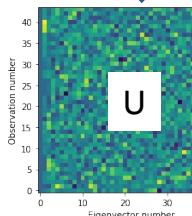
Residual CCF
timeseries



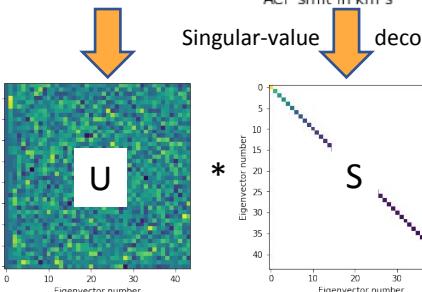
Autocorrelation
(Translation invariant)



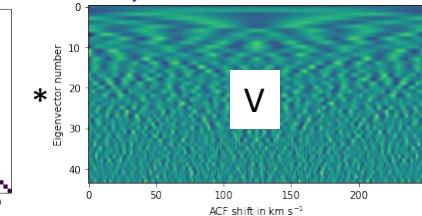
Residual ACF
timeseries



Time-domain
coefficients



Eigenvalues

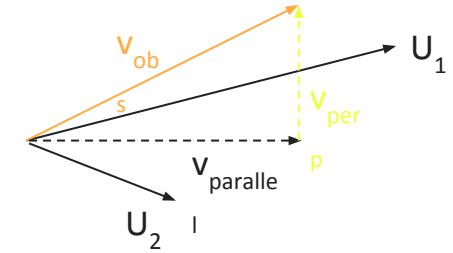
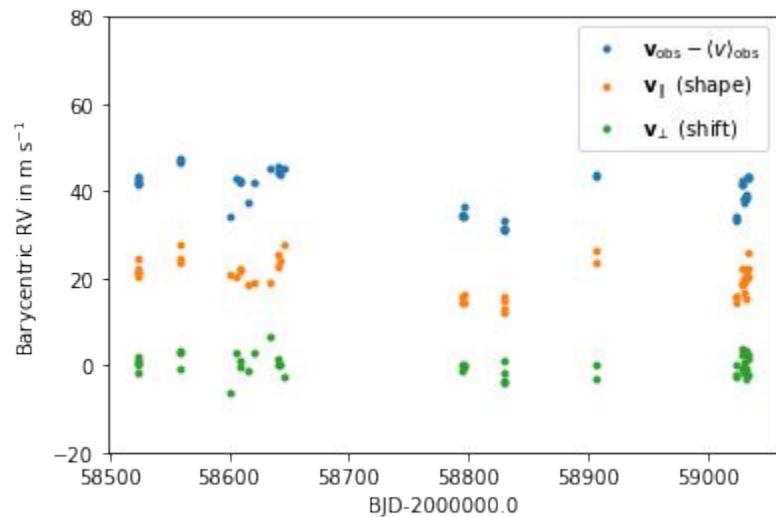


Eigenvectors
(shape-change modes)

Projection of RV timeseries on to U basis

$$\alpha = \mathbf{U}^T \cdot (\mathbf{v}_{\text{obs}} - \langle v_{\text{obs}} \rangle)$$

(Use reduced-rank basis to avoid overfitting)



Observed
RVs

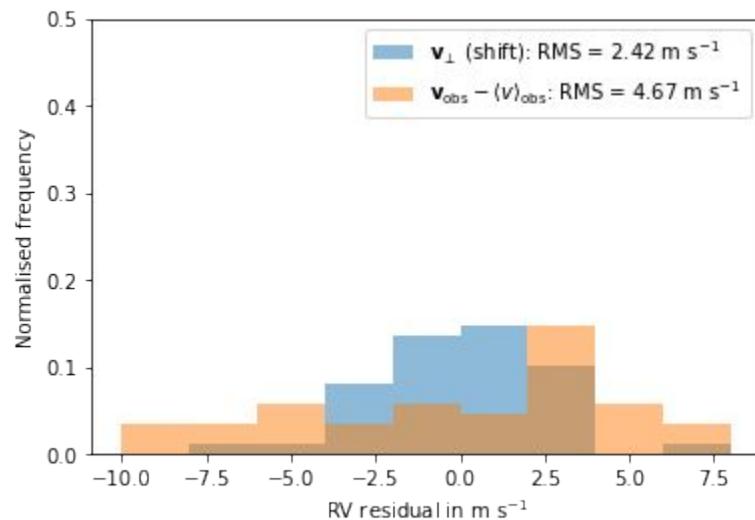
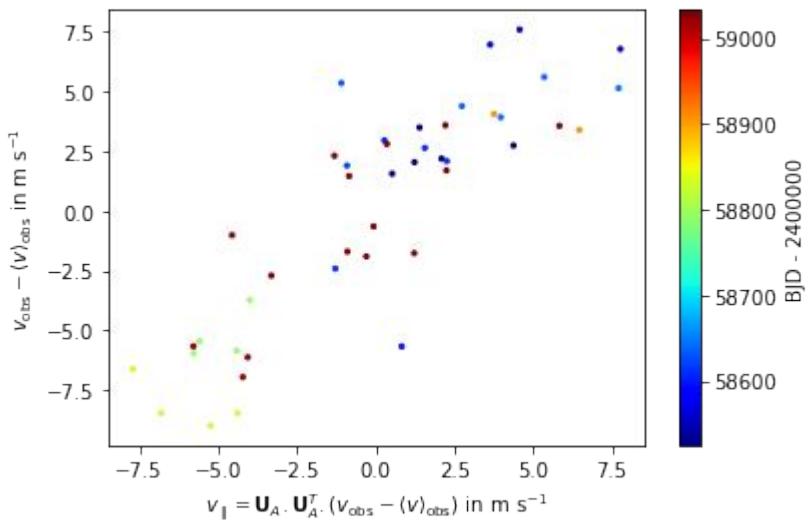
$$\mathbf{v}_{\parallel} = \alpha \cdot \mathbf{U}$$

Shape-driven
RVs

$$\mathbf{v}_{\perp} = \mathbf{v}_{\text{obs}} - \langle v_{\text{obs}} \rangle - \mathbf{v}_{\parallel}$$

Shift-driven
RVs

Most of the RV variation is shape-driven.



CCA

wobble

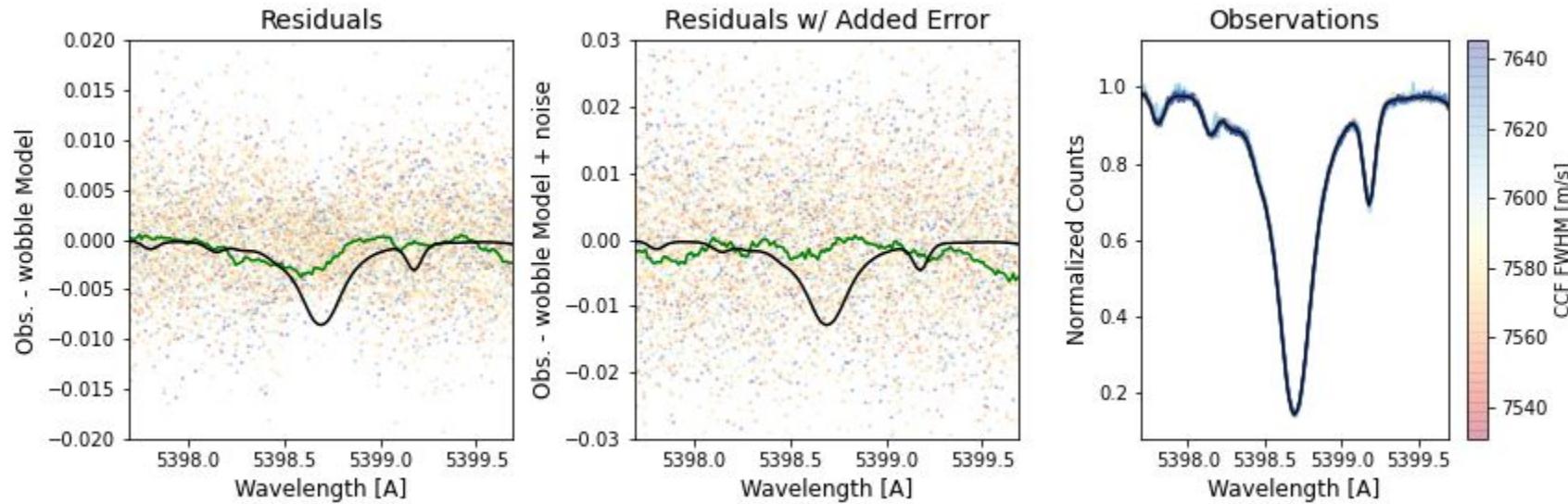
Megan Bedell

David Hogg

Lily Zhao

Rgress Wobble Residuals Against Meta Data

For a sliding window of a given resolution (0.01 Å) and width (0.03 Å),
find the slope of meta-data (i.e. CBC RV or BIS) w.r.t. wobble residuals.

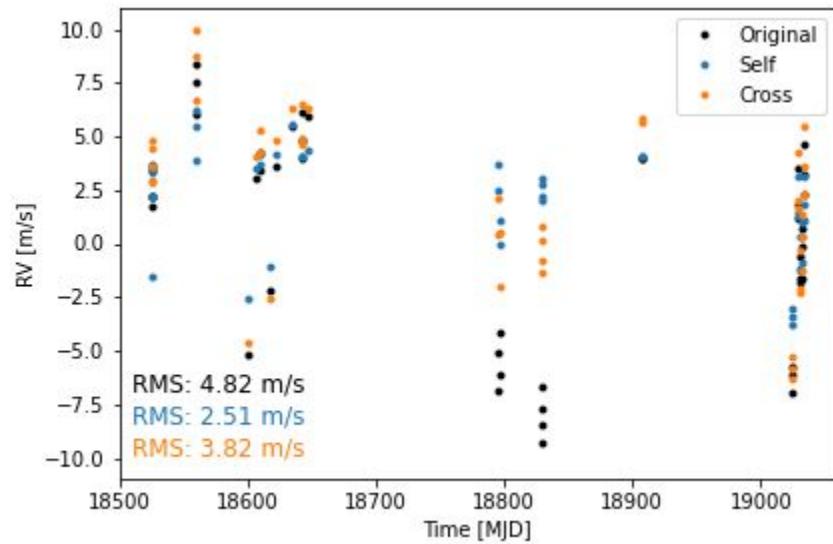


Framework

We use the slopes (df/dRV) to determine the amount of RV shift that is correlated with line shape variations (exposed by residuals) in the spectra.

$$\text{RV Estimate} \quad \text{Slopes}$$
$$\Delta \hat{v}_n = \Delta \vec{f}_n \cdot \frac{d\vec{f}_n}{dRV} \left[\frac{d\vec{f}_n}{dRV} \cdot \frac{d\vec{f}_n}{dRV} \right]^{-1}$$

↑ ↑
Residuals Information Content



Geneva

YARARA & Line-by-line RVs

Xavier Dumusque

Michael Cretignier

Reduction outlines

- Input product : 1d echelle order spectra by EXPRES DRS

https://github.com/MichaelCretignier/Rassine_public

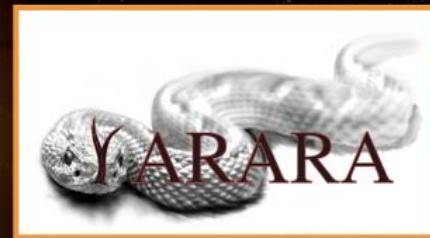
- Production of 1d merged spectra



- Spectra normalised by RASSINE (Cretignier et al. 2020)



- Tailored lines selection with Kit-Cat (Cretignier et al. 2020)

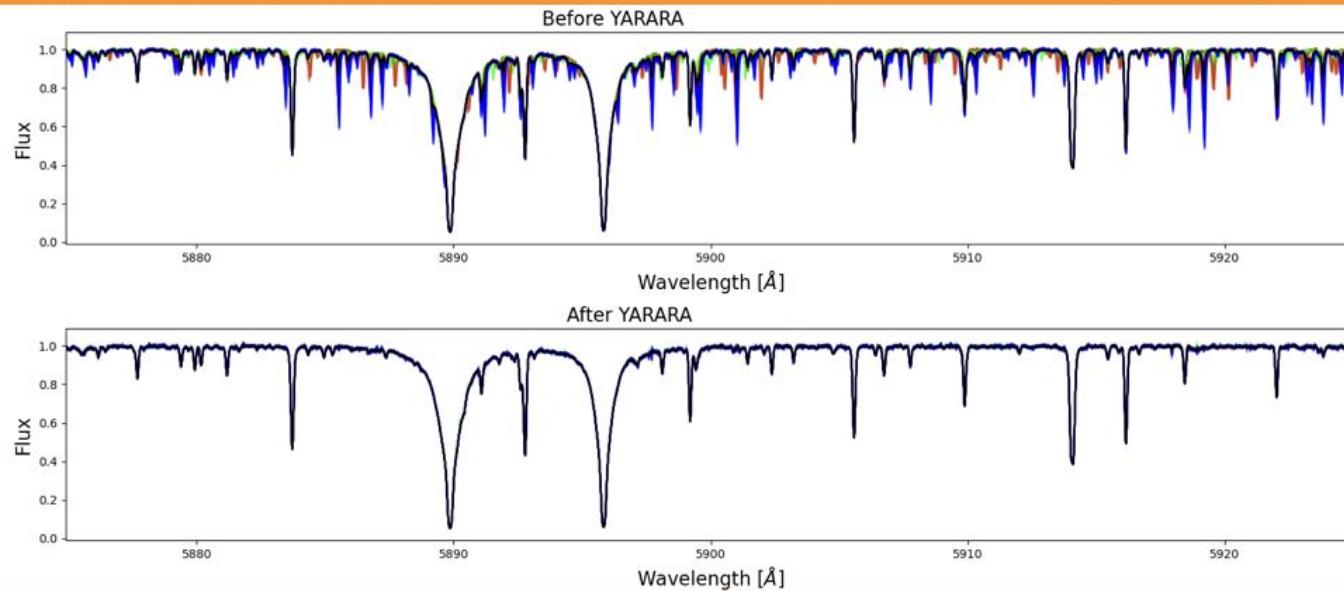


- Telluric correction by YARARA (Cretignier et al. in prep)

- LBL RVs derived and PCA correction by YARARA

YARARA

Telluric correction on HD101501



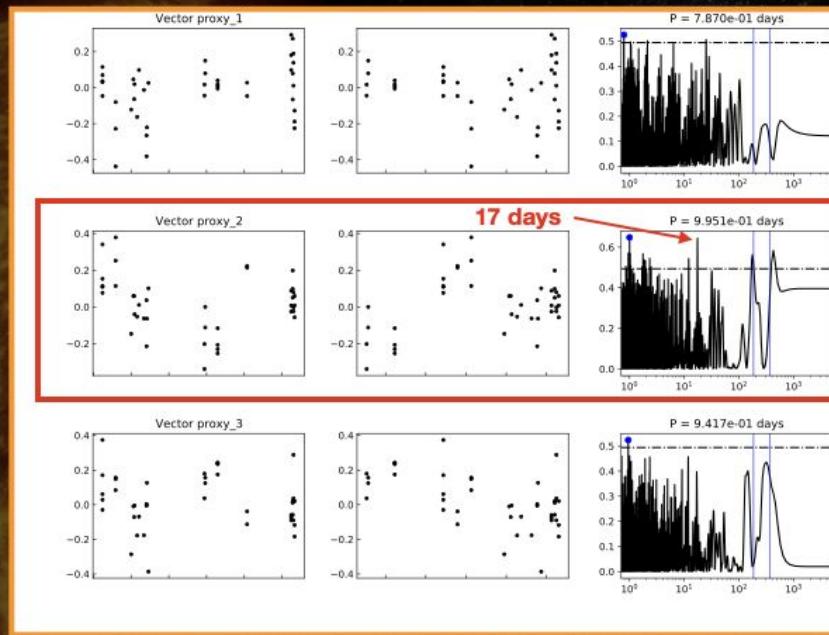
X

YARARA

PCA correction



- Three components fitted by PCA on LBL RVs
- Sampling unable any interpretation but 2nd component contains 17 days power



OxBridGen

Pairwise-GP RV Extraction
Multi-dimensional GP

Suzanne Aigrain

Vinesh Maguire-Rajpaul

Lars Buchhave

Oscar Barragan Villanueva

Belinda Nicholson

Norbert Zicher

Pairwise spectral GP modelling

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY



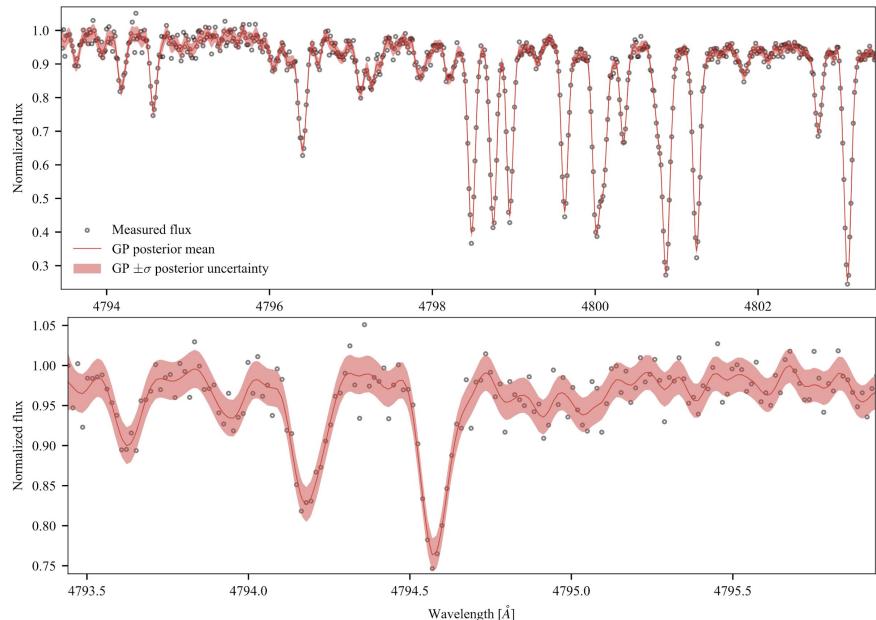
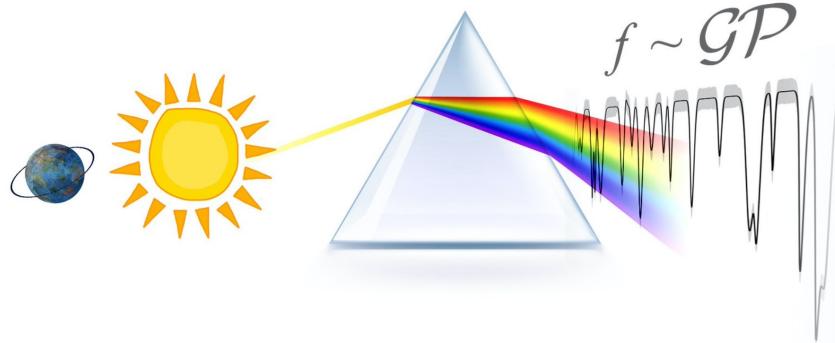
MNRAS 492, 3960–3983 (2020)

Advance Access publication 2020 January 3

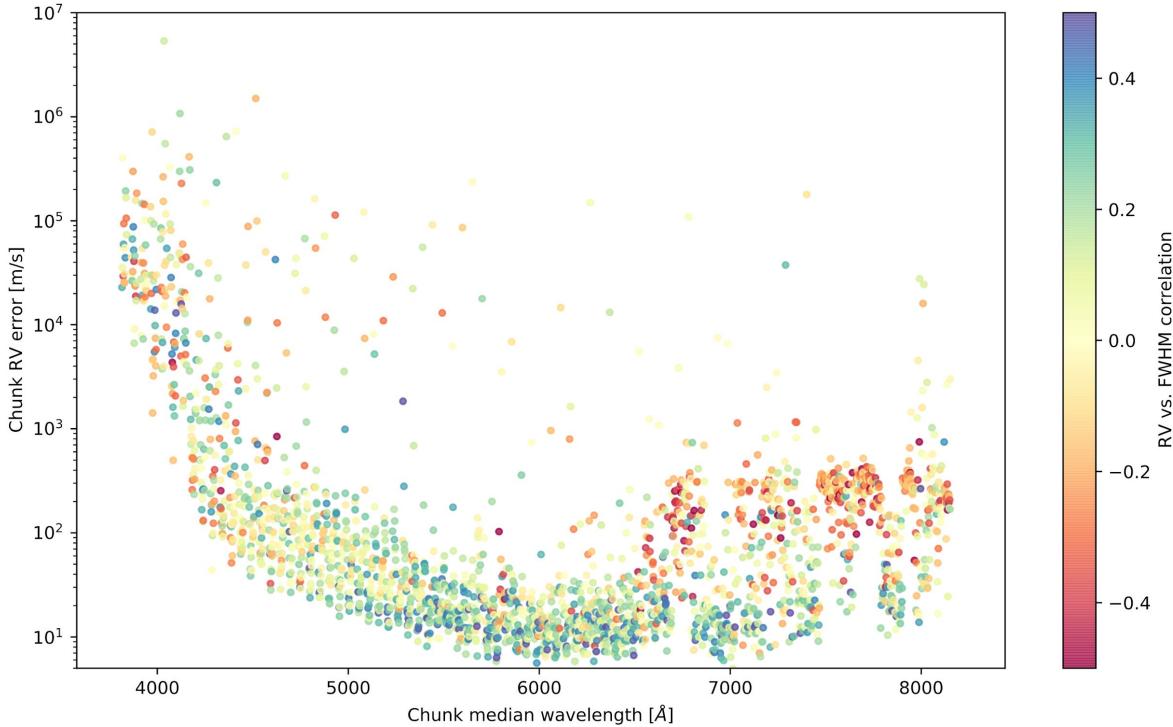
doi:10.1093/mnras/stz3599

A robust, template-free approach to precise radial velocity extraction

V. M. Rajpaul ¹★, S. Aigrain ^{1,2} and L. A. Buchhave ^{1,3}



Pairwise spectral GP modelling



Crude activity (etc.) removal

- Extract local chunk RVs
- Look for local correlations with activity indicators; unusually large RV errors; etc.
- Clustering analyses
- Exclude contaminated wavelengths

More sophisticated approach?

- Add extra covariance terms to account for changes between spectra that are not RV shifts
- Extract RVs jointly with spectral shape changes

Multi-dimensional Gaussian Processes method

OxBridGen

Rajpaul et al., (2015) created a multi-dimensional GP approach to model RVs with activity indicators A_i to constrain the induced stellar signals.

$$\Delta \mathcal{A}_1 = A_1 G(t) + B_1 \dot{G}(t)$$

⋮

$$\Delta \mathcal{A}_N = A_N G(t) + B_N \dot{G}(t)$$

$$\Delta\mathcal{A}_1 = A_1 G(t) + B_1 \dot{G}(t)$$

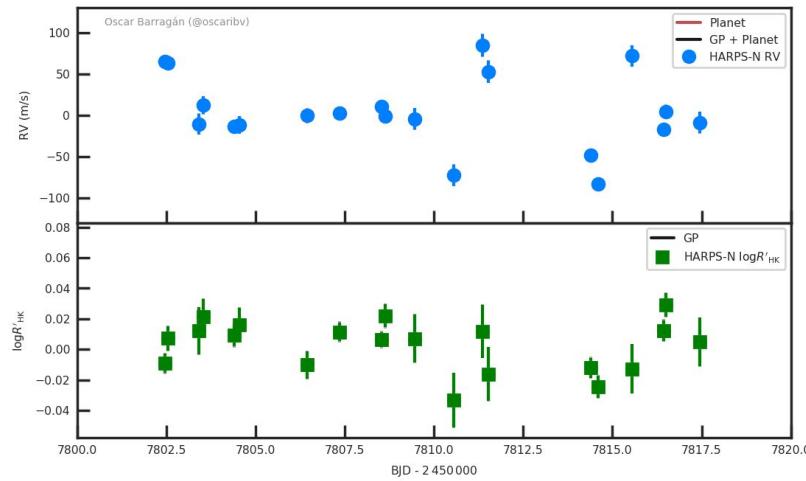
⋮

$$\Delta\mathcal{A}_N = A_N G(t) + B_N \dot{G}(t)$$

The variables $A_1, B_1, \dots, A_N, B_N$, are free parameters which relate the individual time series to an underlying Gaussian Process $G(t)$. Each time-series can be a different combination between RVs and activity indicators.

This approach has been proven useful to disentangle planet and stellar RV signals in active stars.

In this example we see how the $\log R_{HK}$ helps to guide the RVs to model the activity properly and it allows to find the signal of K2-100b (Barragán et al., 2019)



Light curve fitting

- Because of the non-optimal sampling of the RV data, we first model the light curve in order to obtain an estimate of the rotational period of the star to use as a prior in the RV fit.
- We use a Quasi-periodic Kernel

$$\gamma(t_i, t_j) = \exp \left[-\frac{\sin^2[\pi(t_i - t_j)/P_{\text{GP}}]}{2\lambda_p^2} - \frac{(t_i - t_j)^2}{2\lambda_e^2} \right]$$

- We got a Period constraint of 16.3 ± 0.5 days.
- Note that we did not use priors on `lambda_e` and `lambda_p`.

RV fitting

- We ran a multi-dimensional GP for different combinations of RV and one activity indicator using a quasi-periodic Kernel.
- We show the results when fitting RVs and H_alpha emission.

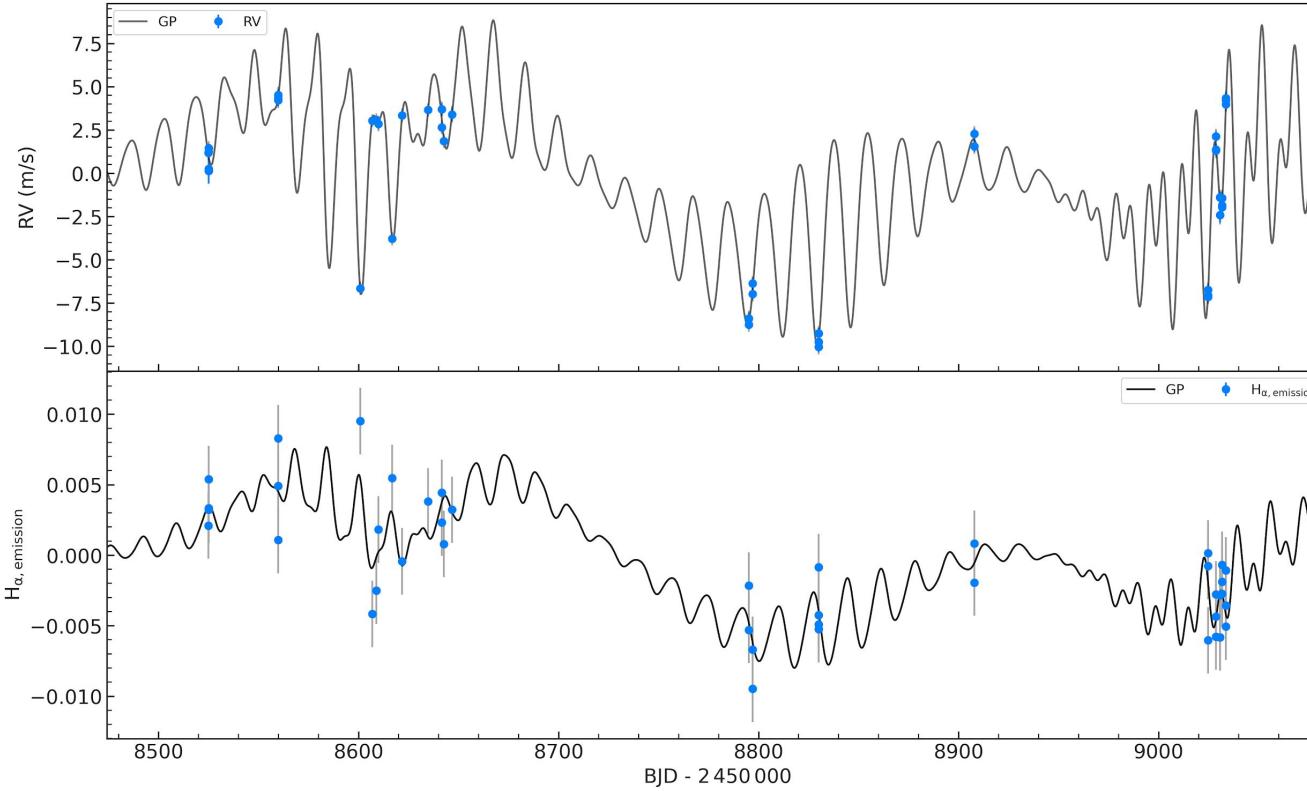
RVs and H_alpha emission

$$l_e = 38 - 20 + 191 d$$

$$l_p = 1.5 - 0.7 + 1.1$$

$$P_{GP} = 16.7 - 0.5 + 0.2 d$$

RMS residuals
0.33 m/s



Radial velocity (top), and $H_{\alpha, \text{emission}}$ (bottom) time series. Both time series have been corrected by the inferred offset. Inferred models are presented as solid continuous lines. Measurements are shown with filled symbols with error bars.

PennState

DCPA
Scalpels Some More
Fiesta

Eric Ford

Christian Gilbertson

Michael Palumbo

Alex Wise

Jinglin Zhao

Penn State: Oct 31 submission

Two algorithms for modeling stellar variability:

- Scalpels ([Collier-Cameron et al. 2020](#))
 - Applied to custom CCFs and RVs from those CCFs
 - Subtracted projection of RVs onto two basis vectors
- Doppler-constrained PCA ([Jones et al. 2017/20](#))
 - Applied to flux in same wavelength ranges as used for custom CCFs
 - Used linear regression on score for leading basis vector

Each algorithm used three line lists/masks:

- ESPRESSO G9 mask
 - Filtered to stay in LFC range
 - Avoid tellurics & order boundaries
- Two custom line lists
 - Coadd spectra to make template
 - Line finder based on 2nd deriv of GP fit
 - Fit each line at each time
 - Select lines based on width, RMS of measured properties
 - Reject lines with blend predicted by VALD
 - Filtered to stay in LFC range
 - Avoid tellurics & order boundaries
 - “1” is less aggressive in filtering than “2”

Penn State: Other things brewing

- Multivariate GP regression ([Jones et al. 2017/20](#) & [Gilbertson et al. 2020](#); generalization of Rajpaul et al. 2015)
 - Tried passing scores for basis vectors from both DCPCA and Scalpels as indicators
 - Overfit to RVs, so didn't submit that.
 - Christian robustifying code for Bayesian model comparison of no-planet versus planet model.
- Making use of multiple CCF masks based on line properties (Alex)
- Extending DCPCA to wavelengths with blends or modest telluric contamination (Christian)
- Fiesta (Fourier decomposition of CCF; Jinglin)

Porto

GPRN

João Camacho

João Faria

Pedro Viana

Gaussian process regression network (GPRN)

A general GP framework that can take into account multiple inputs and outputs

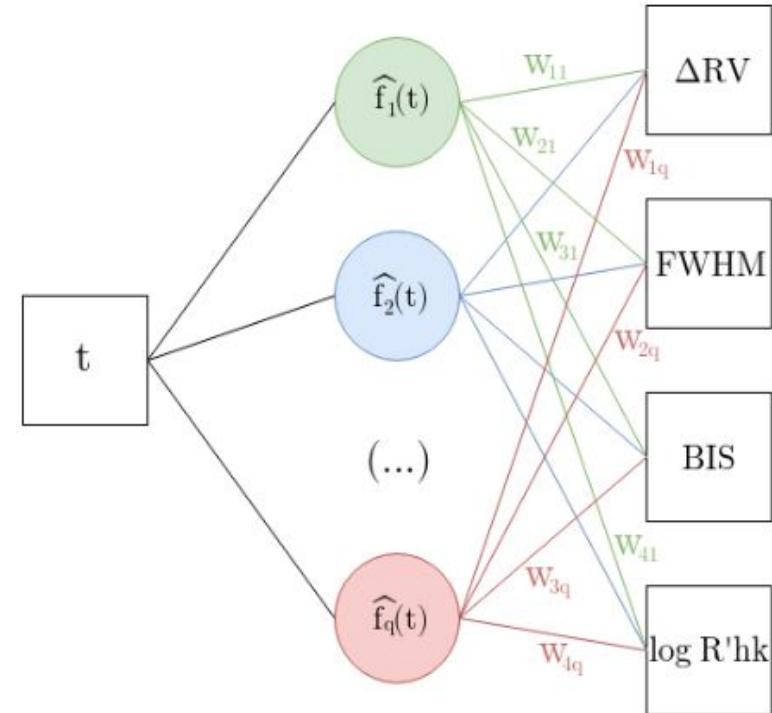
Wilson et al. (2012)

Combines a Bayesian neural network with the flexibility of GPs

$$\mathbf{y}(\mathbf{x}) = \mathbf{W}(\mathbf{x}) [\mathbf{f}(\mathbf{x}) + \sigma_f \boldsymbol{\epsilon}] + \sigma_y \mathbf{z}$$

$$f_i(x) \sim \mathcal{GP}(0, k_f) \quad \leftarrow q \text{ nodes}$$

$$W_{ij} \sim \mathcal{GP}(0, k_w) \quad \leftarrow q \cdot p \text{ weights}$$

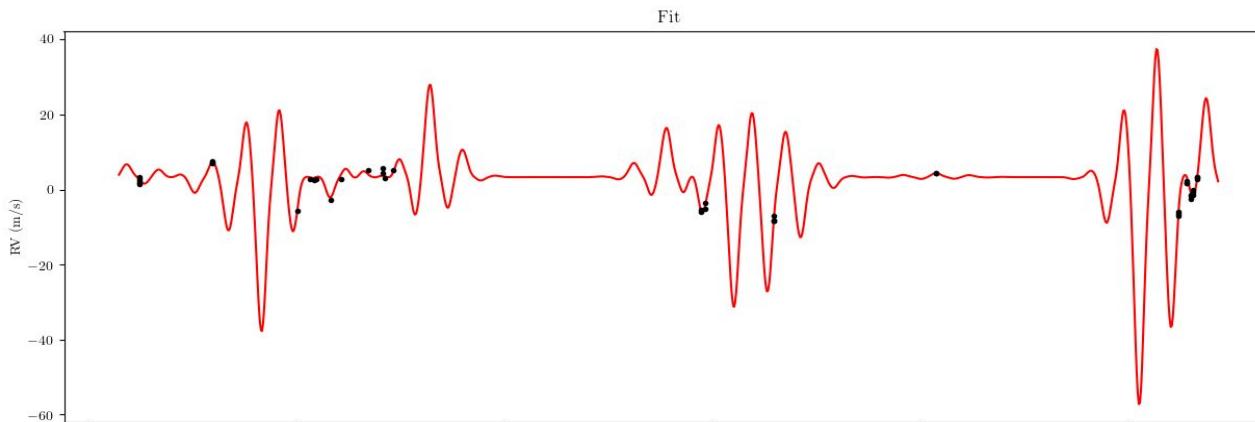


Outputs $\mathbf{y}(\mathbf{x})$ will be *non-stationary*

GPRN on the HD101501 RVs

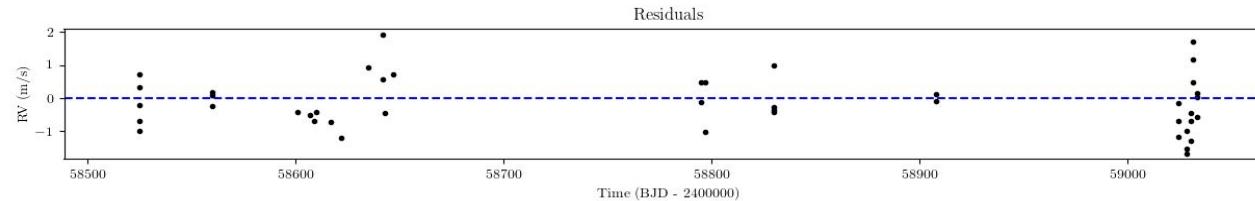
relatively few observations (lack of rotation phase coverage)

we considered only the RVs: $p=1$, $q=1$, Periodic node and SE weight
(this is close, but not equal, to the standard quasi-periodic GP)



weight amplitude: 6.5 m/s
node period: 16.2 days

jitter: 0.33 m/s



residual rms
0.82 m/s

we submitted the wrong residuals... 😳

Sidera

Coherence
(Phase)
Debiased periodograms

Sally Dodson-Robinson

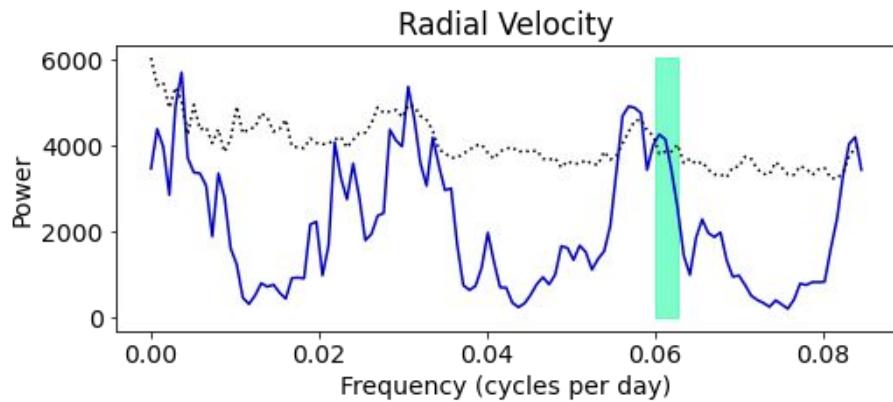
Charlotte Haley

Victor Ramirez Delgado

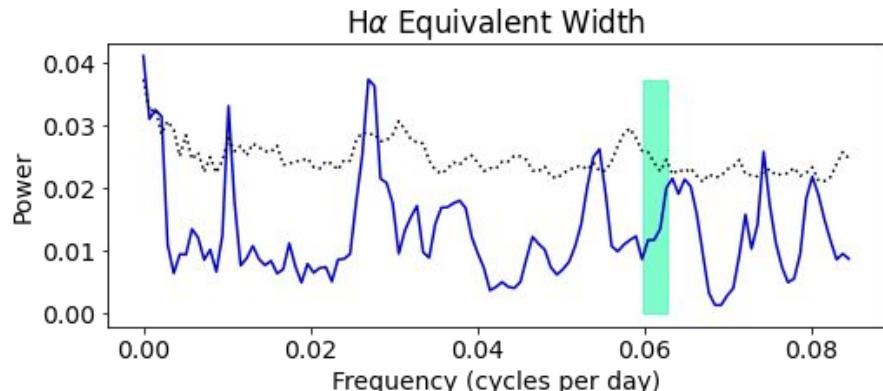
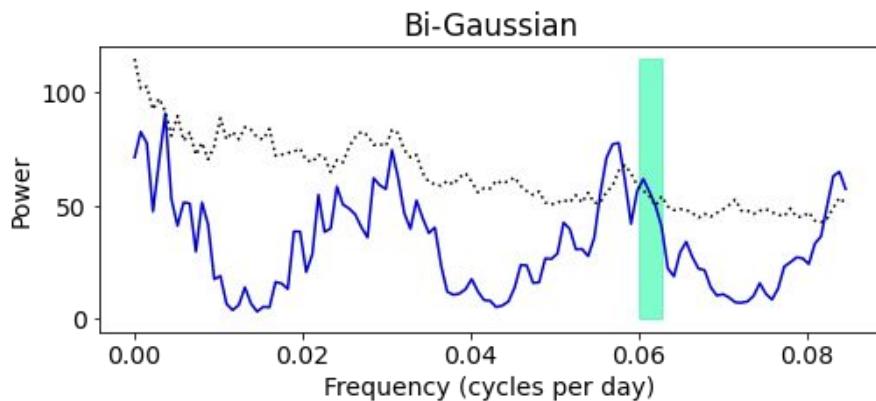
Justin Harrell

Catherine Lembo

Debiased periodograms



Background noise model: AR(1)
Dotted line: 5% FAP from Monte Carlo simulation
Green shaded region: star rotation

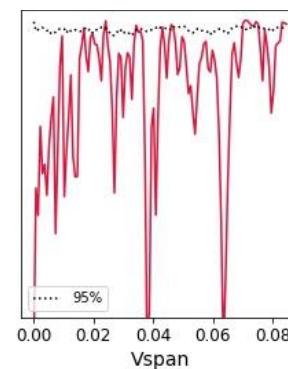
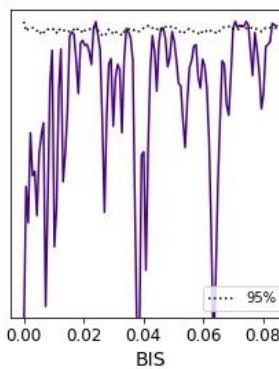
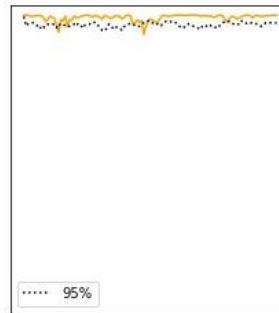
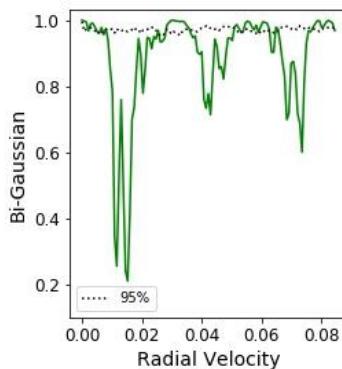
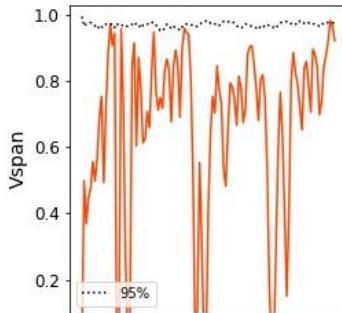
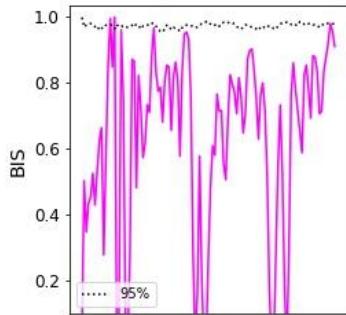


Coherence

Question: To what extent does an activity indicator predict RV?

Answer: $C_{xy}(f) = \frac{|\mathcal{F}\{x(t) \star y(t)\}|^2}{|\mathcal{F}\{x(t)\}|^2 |\mathcal{F}\{y(t)\}|^2}$

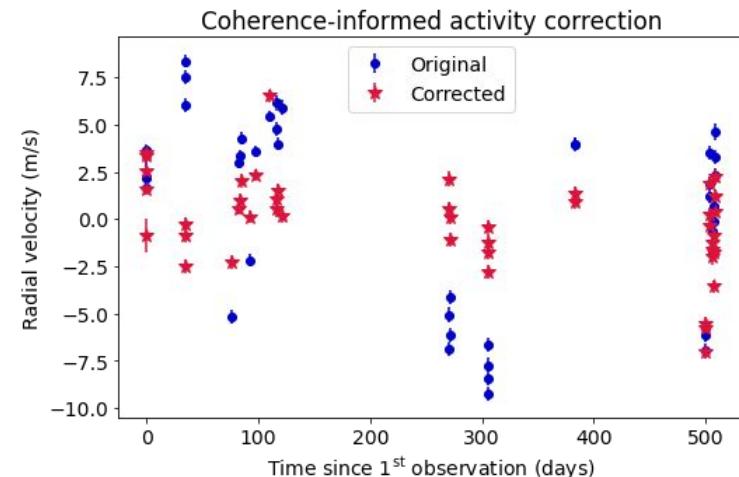
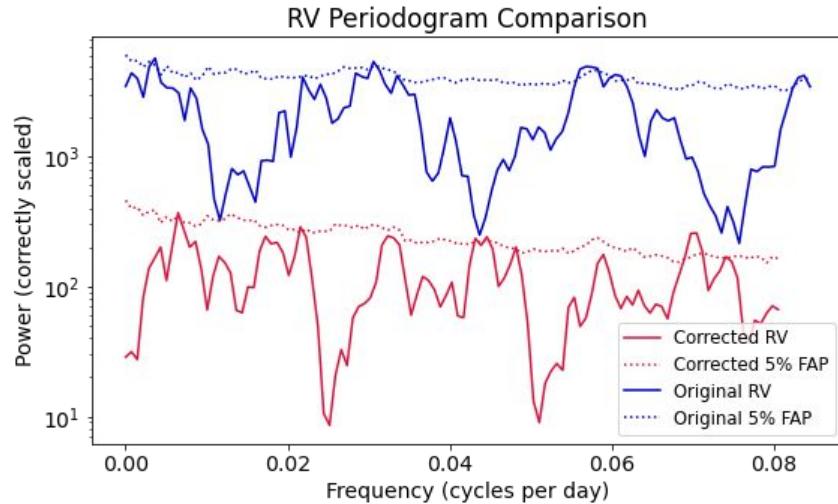
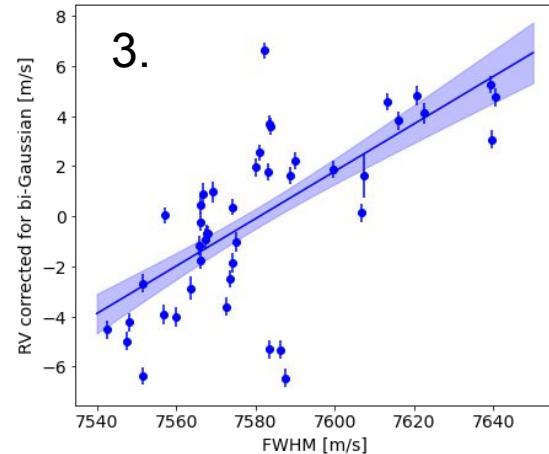
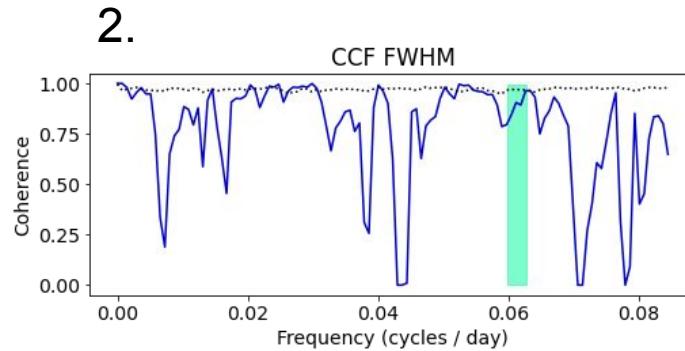
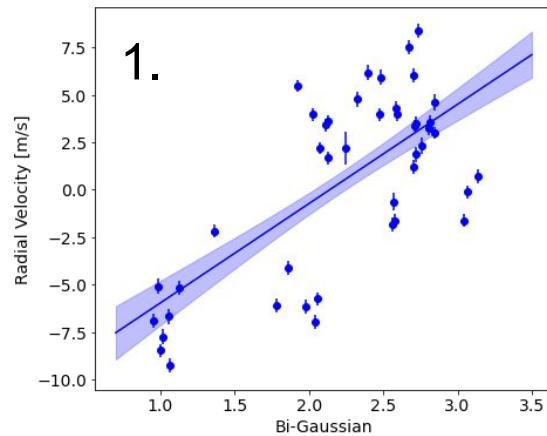
$x(t)$: RV \star : cross-correlation
 $y(t)$: activity indicator f : frequency
 $\mathcal{F}\{\}$: Fourier transform



$C_{xy}(f) = 0$: no stellar activity in RV at f

$C_{xy}(f) = 1$: entire RV signal at f is due to activity

Simple activity correction



YaleWI

HGRV
SAFE

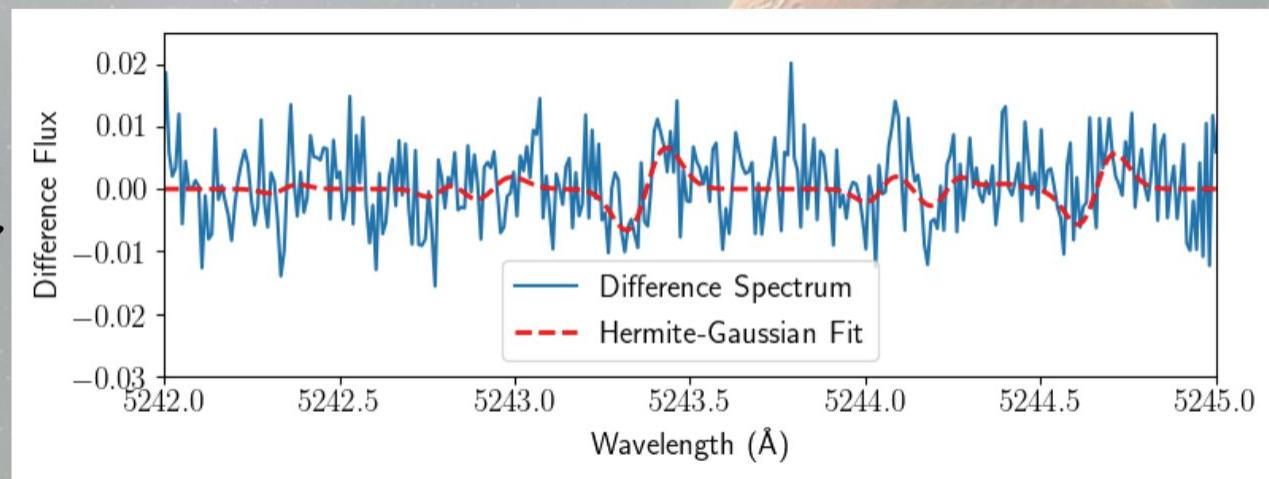
Parker Holzer

Jessi Cisewski-Kehe

HGRV Linear Model

$$y_i = v_r \sum_{j=1}^m \frac{\sqrt{\sqrt{\pi} d_j \mu_j}}{c \sqrt{2 \sigma_j}} \psi_1(x_i; \mu_j, \sigma_j) + \varepsilon_i \quad , \quad \varepsilon_i \sim N(0, Q_i)$$

Coefficient to estimate Sum across m absorption features Relative amplitude of absorption feature j First-degree Hermite-Gaussian function centered at μ_j



1. Stellar Activity F-statistic for Exoplanet surveys (SAFE) (Holzer et al. (in review))

Define $\Psi_k(x) := \sum_{i=0}^m \gamma_{k,i} \psi_k(x; \mu_i, \sigma_i)$, $\gamma_{k,i}$'s estimated using PCA on models of stellar activity on the Sun

$$y_i = v_r \sum_{j=1}^m \frac{\sqrt{\pi d_j} \mu_j}{c \sqrt{2\sigma_j}} \psi_1(x_i; \mu_j, \sigma_j) + \sum_{k=0, k \neq 1}^5 \beta_k \Psi_k(x_i) + \varepsilon_i , \quad \varepsilon_i \sim N(0, Q_i)$$

Define SAFE : F-statistic for testing the null hypothesis of no stellar activity

$$H_0 : \beta_0 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

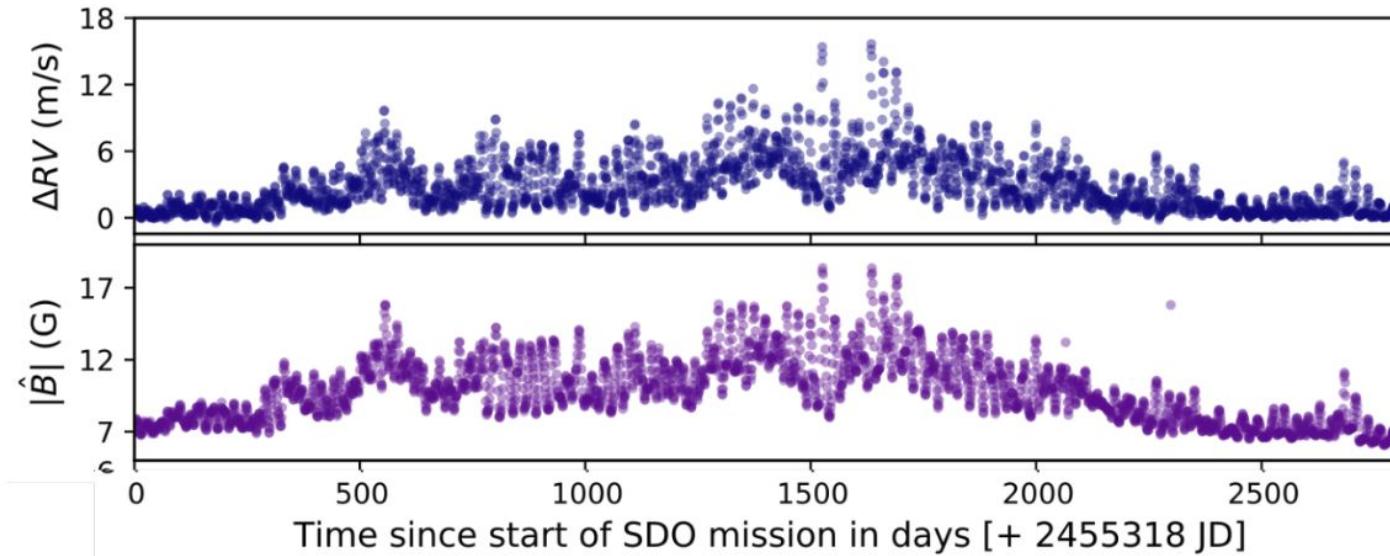
LienhardMortier

ZLSD

Annelies Mortier

Florian Lienhard

B-field as an activity indicator



Zeeman effect

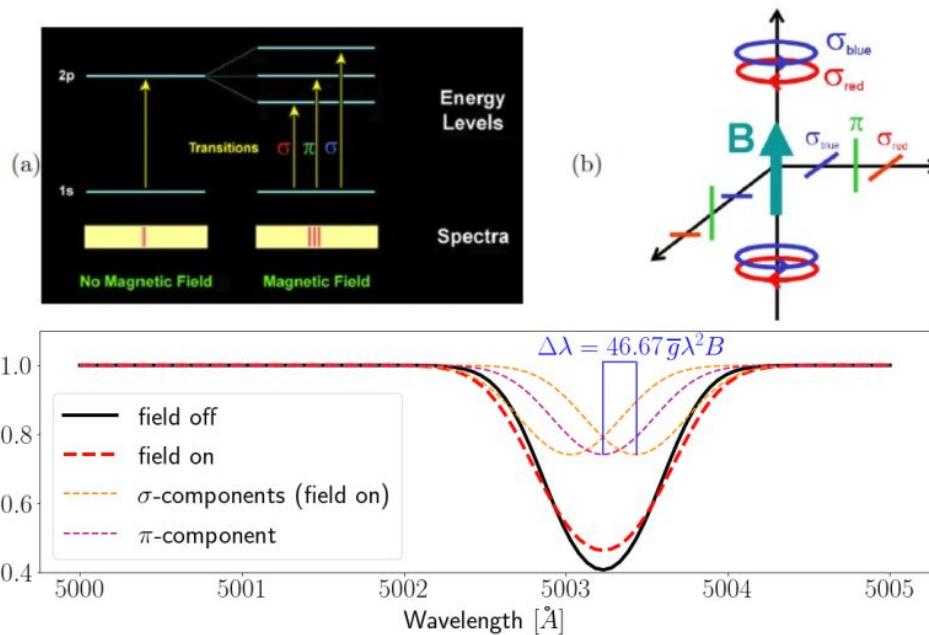


Fig. a & b: Reiners 2012.

Input & output

Input

- ▶ EXPRES:
 - ▶ bary-chrom wavelengths
 - ▶ intensities
 - ▶ uncertainties
 - ▶ tellurics
 - ▶ continuum
 - ▶ (blaze)
- ▶ VALD database:
 - ▶ line wavelength
 - ▶ line depth
 - ▶ line Landé factor

Output

- ▶ Active region filling factor
- ▶ Magnetic field strength averaged and in active regions
- ▶ Radial velocity estimates (with/without Zeeman effect)

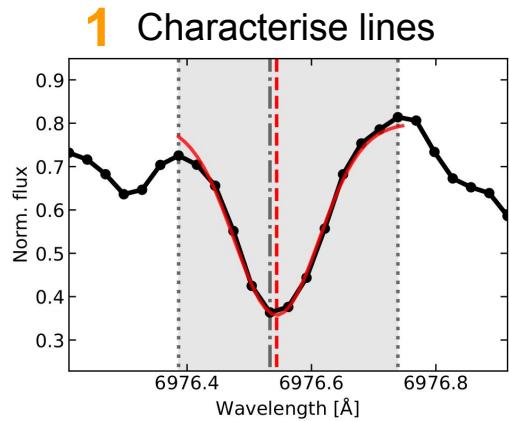
Warwick

Heather Cegla

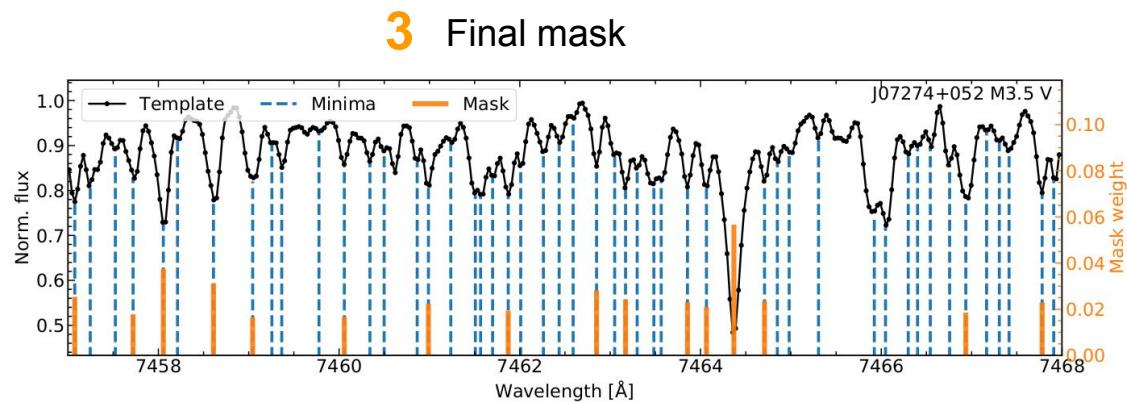
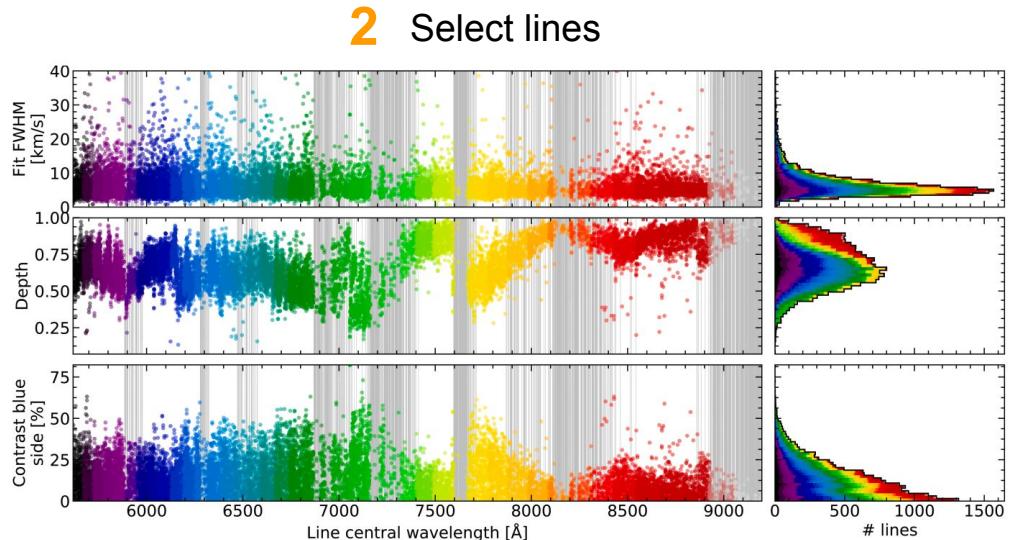
Lauren Doyle

Marina Lafarga Margo

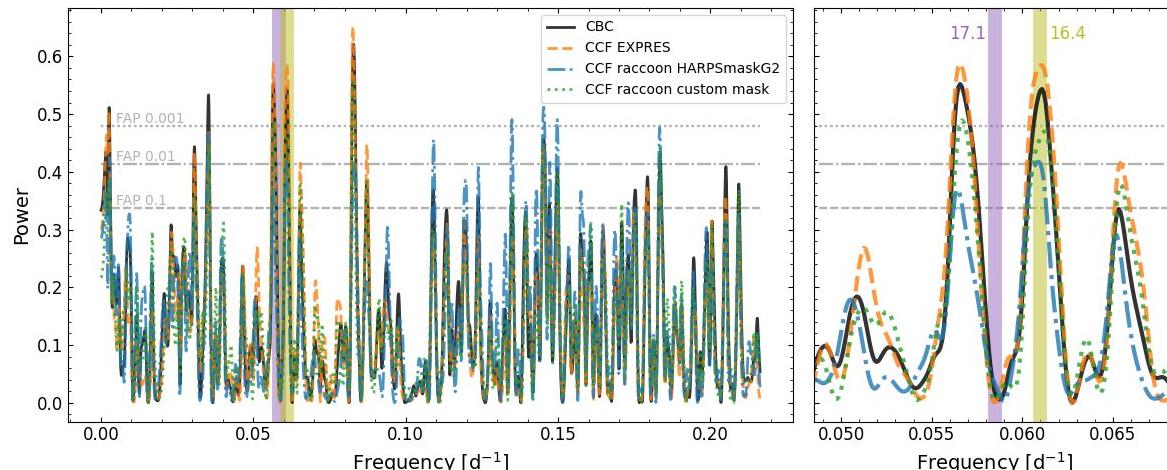
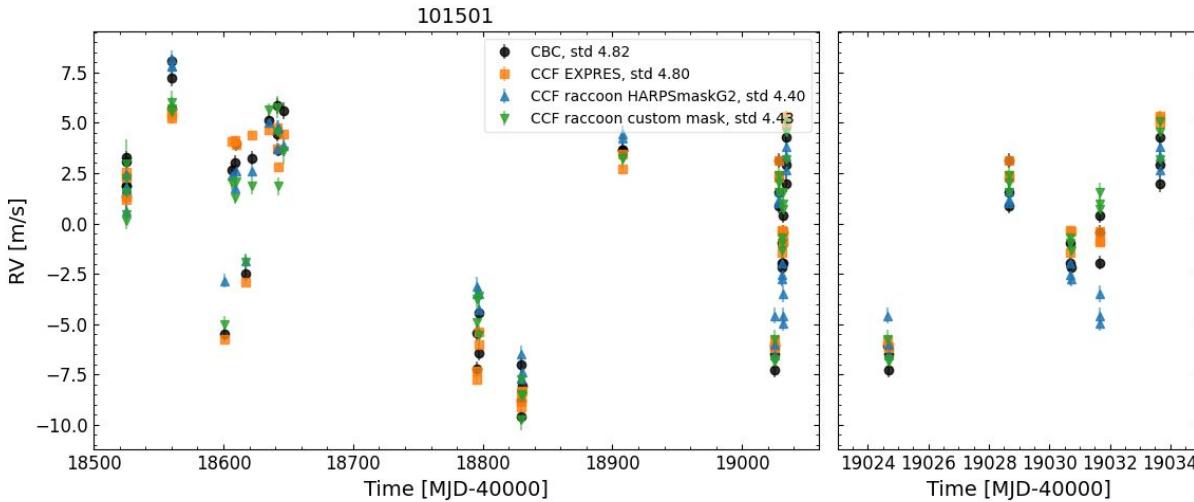
- Create custom mask from observations
- Compute new CCF



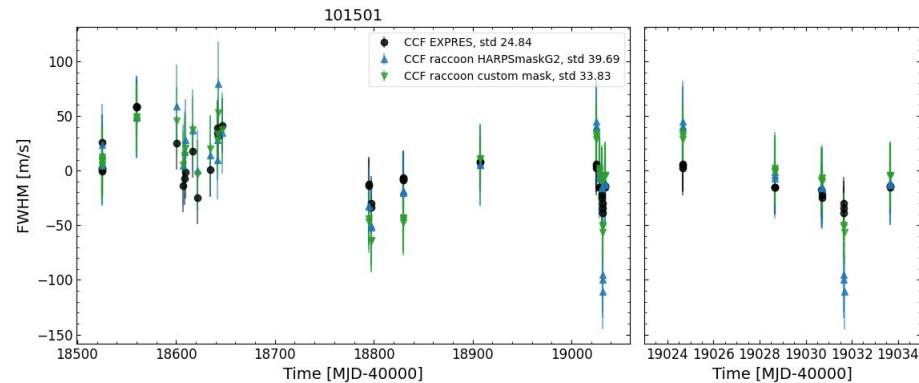
Example data from CARMENES observations of an M dwarf



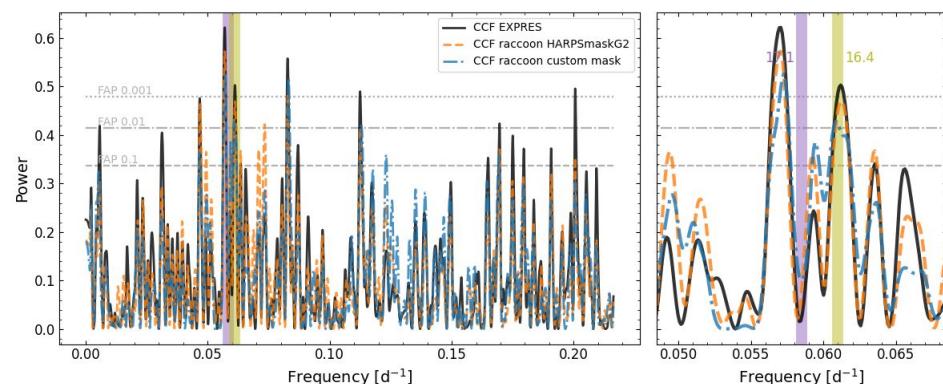
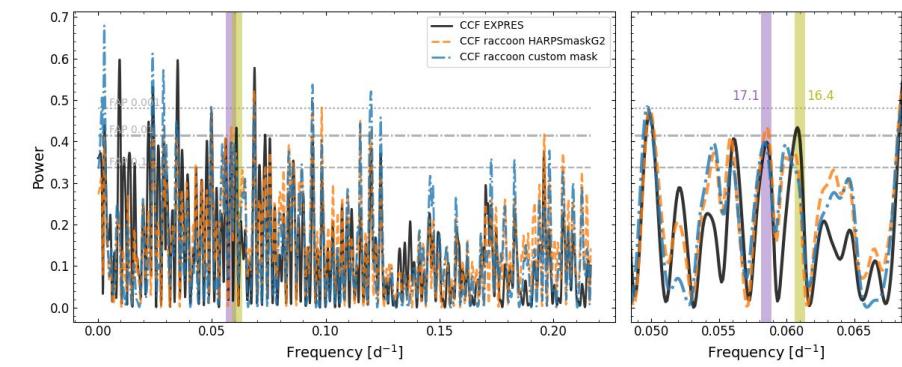
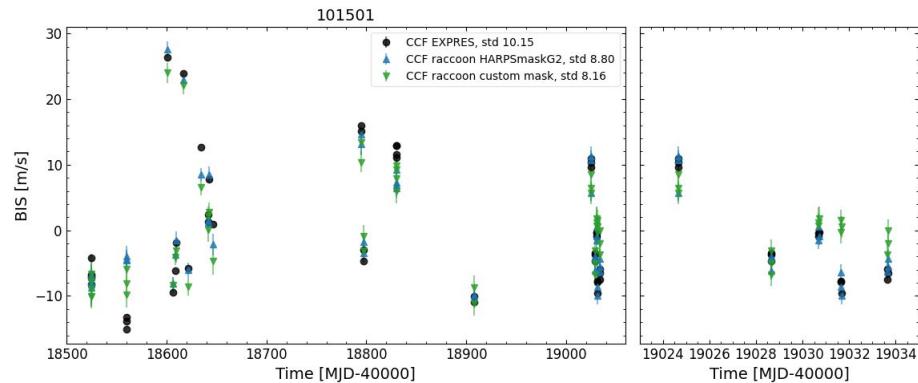
RVs



FWHM



BIS



Austin

Neural Networks

Zoe L De Beurs

Andrew Vanderburg

Christopher Shallue

(Very) Initial Results

