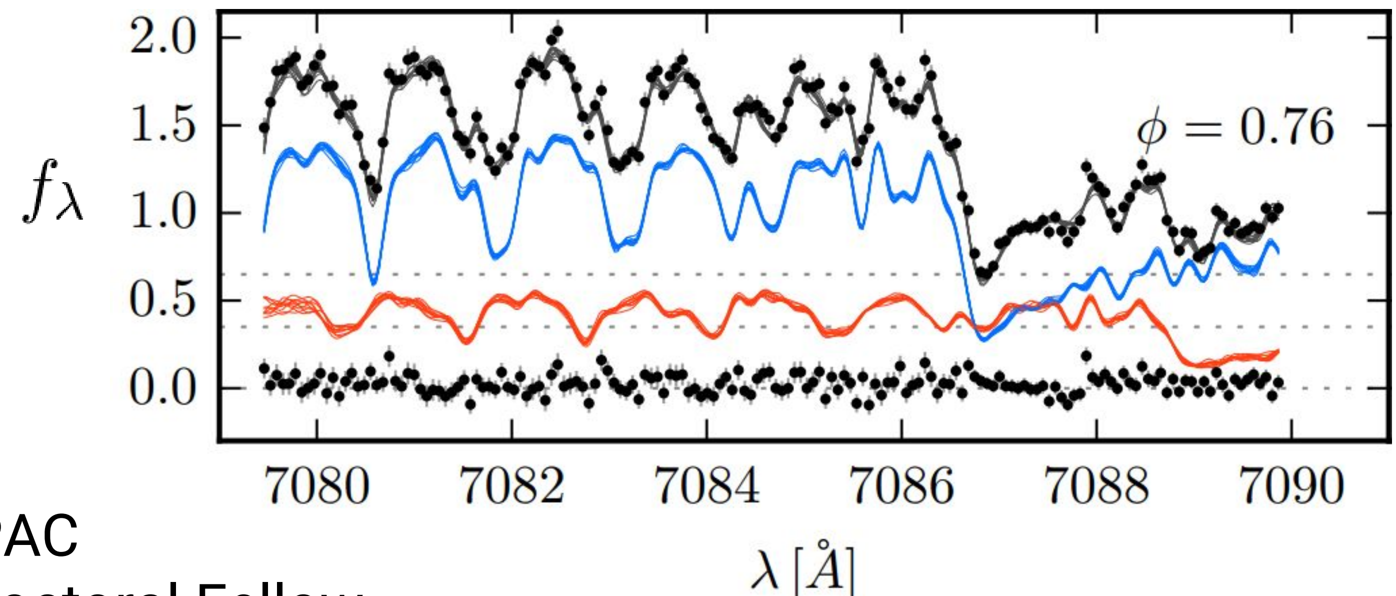
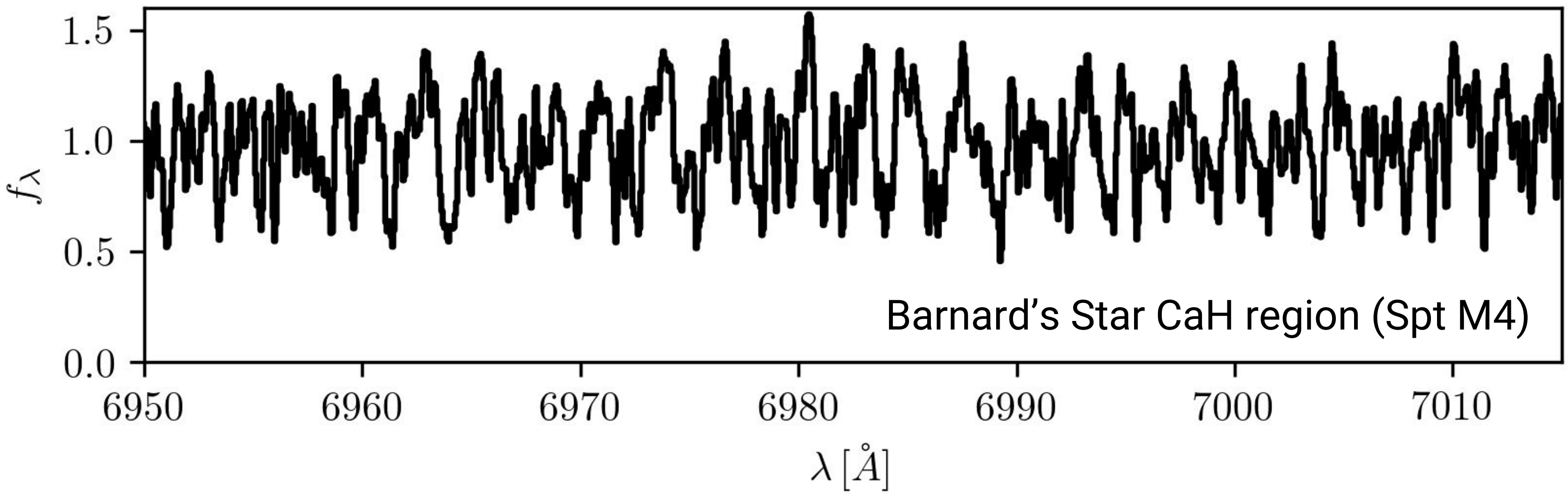
Using Gaussian Processes to Construct Flexible Models of Stellar Spectra



Ian Czekala
Stanford/KIPAC
KIPAC Postdoctoral Fellow
AAS January 10th, 2018

with Kaisey Mandel, Bruce Macintosh, Sean Andrews, Jason Dittmann, Sujit Ghosh, Ben Montet, Elisabeth Newton

Spectra are a fundamental data type of astronomy



- Flux density as a function of wavelength (or frequency)
- Analysis yields elemental abundances of stars, galaxies
- Accretion/explosion physics, mass flow rates
- Temperatures, ionization states, surface gravities
- Exoplanet discovery by radial velocity, stellar binaries, ...

Spectra are generated by complex astrophysical processes!

Physics-based spectral ***models*** are sometimes insufficient at high signal-to-noise and high resolution

- One approach is to continue iterating on the physical model, adding complexity, until models can fit the data
- ***This is hard, time-consuming, but also very important and should definitely be pursued!***

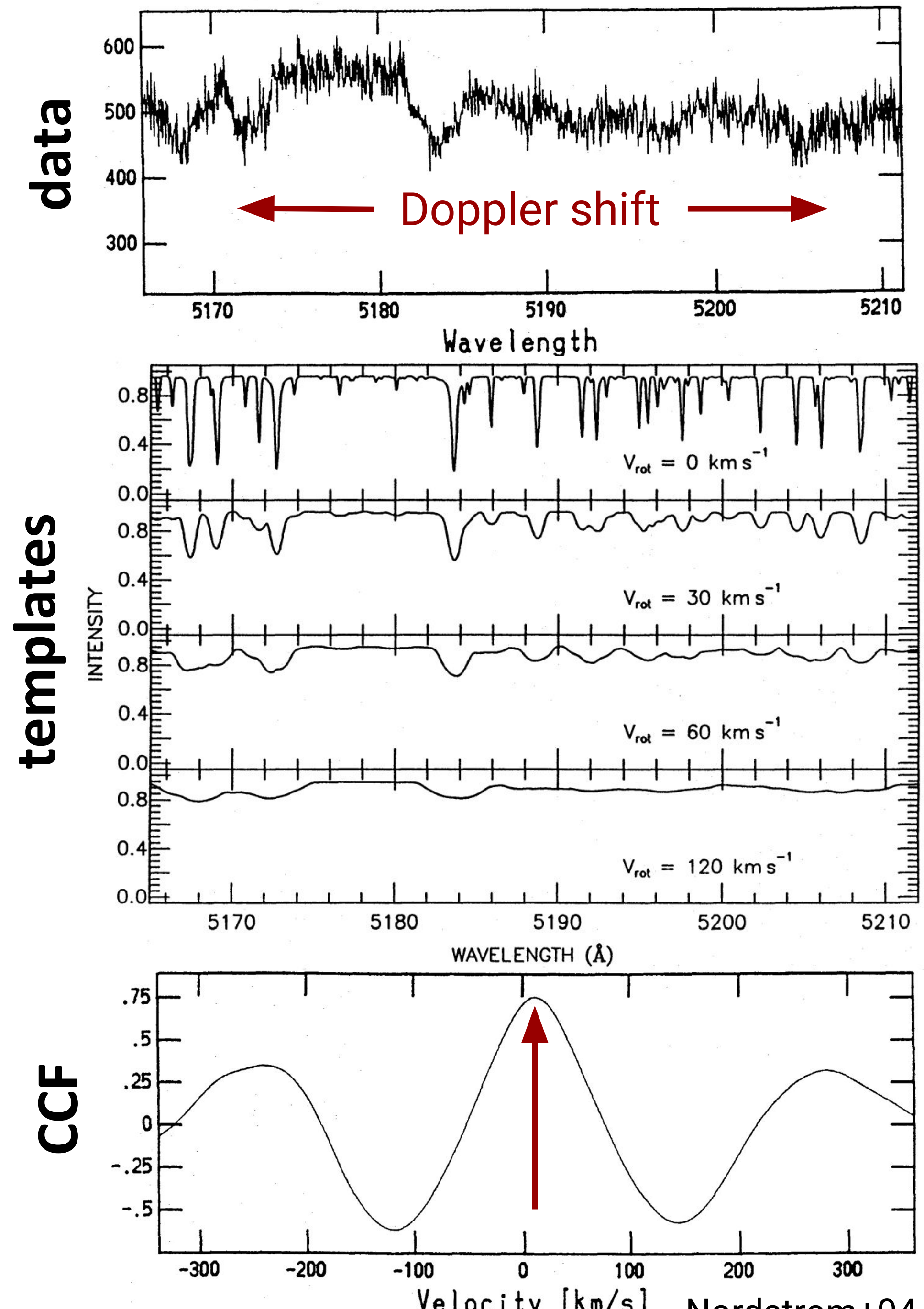
- In the meantime, what if the underlying physical spectral model is *peripheral* to our main science question? For example, determining redshift/radial velocity?

Traditional Radial Velocity Analysis

- The observed stellar spectrum is *cross-correlated* with a template
- Try different templates to find the best match
- The peak of the cross correlation function (CCF) is the radial velocity

$$\frac{\Delta v}{c} = \frac{\Delta \lambda}{\lambda}$$

- For binary stars, use a CCF with two stellar templates



Nordstrom+94

Mayor & Queloz 94

Traditional RV Analysis

1. Derive a list of velocities for each epoch of spectroscopy
2. Fit a Keplerian orbit and determine orbital parameters: ***amplitude*, *period*, *eccentricity*, *phase***

- If we want to measure relative radial velocity, do we need models of stars from **first principles**, or do we just need a **good model** of the stellar spectrum?
- *Gaussian processes provide a basis for modeling the stellar spectrum without physics-based models*

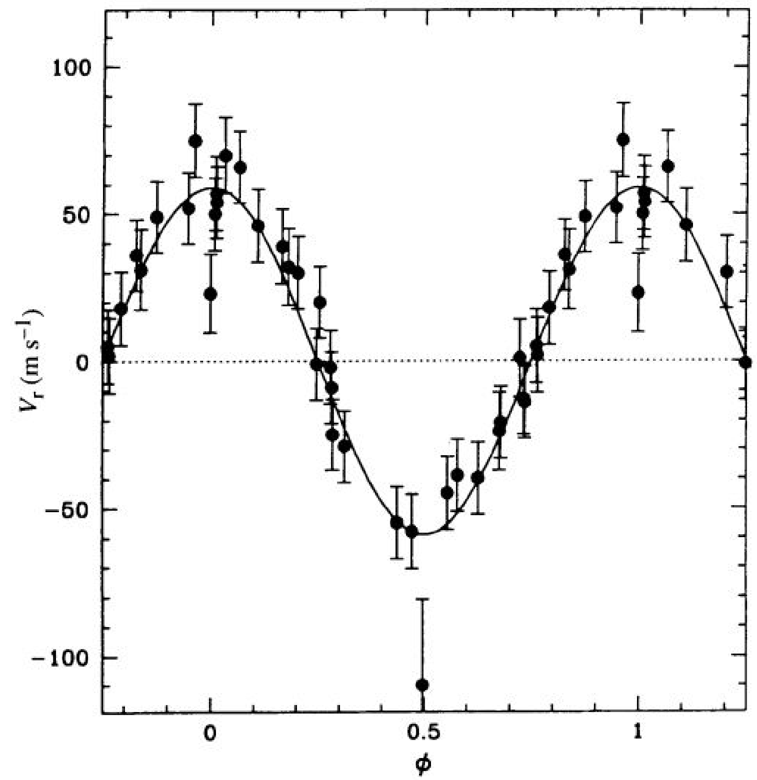


TABLE 1 Orbital parameters of 51 Peg

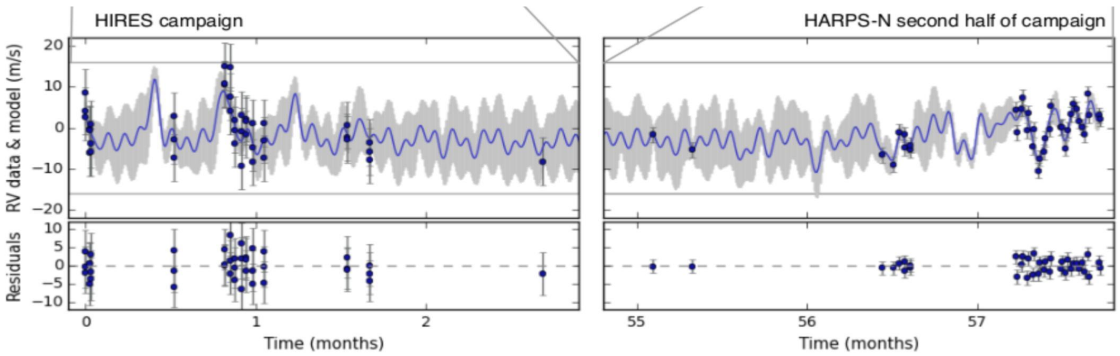
P	4.2293 ± 0.0011 d
T	$2{,}449{,}797.773 \pm 0.036$
e	0 (fixed)
K_1	0.059 ± 0.003 km s^{-1}
$a_1 \sin i$	$(34 \pm 2)\ 10^5$ m
$f_1(m)$	$(0.91 \pm 0.15)\ 10^{-10}\ M_\odot$
N	35 measurements
$(O-C)$	13 m s^{-1}

P, period; T, epoch of the maximum velocity; e, eccentricity; K_1, half-amplitude of the velocity variation; $a_1 \sin i$, where a_1 is the orbital radius; $f_1(m)$, mass function; N, number of observations; $(O-C)$, r.m.s. residual.

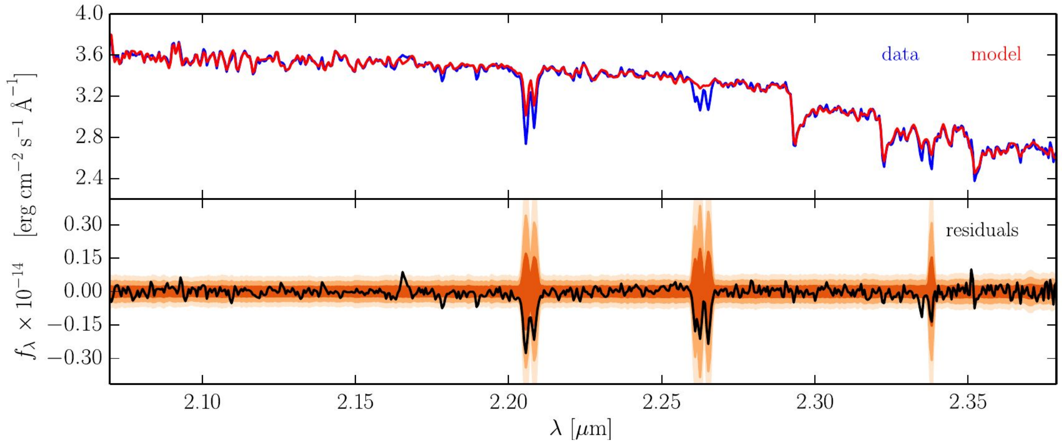
NATURE · VOL 378 · 23 NOVEMBER 1995

Two ways of using Gaussian Processes:

1) Model **Correlated Noise**

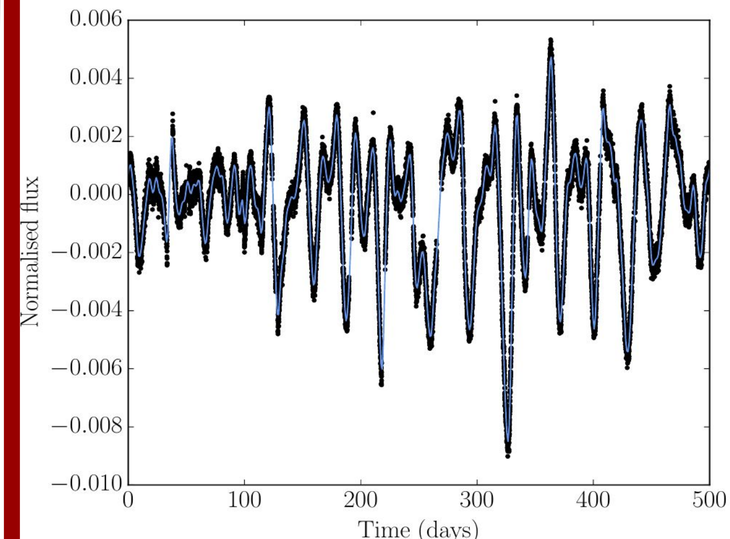


Fitting RV time series; Lopez-Morales, Heywood et al. 2016



Stellar properties with synthetic spectra; **Czekala** et al. 2015a; *Starfish*

2) Sophisticated **Interpolating Functions**



Kepler Rotation Periods; Angus et al. 2017

Measuring RVs without templates

This talk;

Czekala et al. 2017a, *PSOAP*

Gaussian processes

We will model the latent stellar spectrum f_λ as a Gaussian process

$$f_\lambda \sim \mathrm{GP}(\mu(\lambda),\, k(\lambda, \lambda'))$$

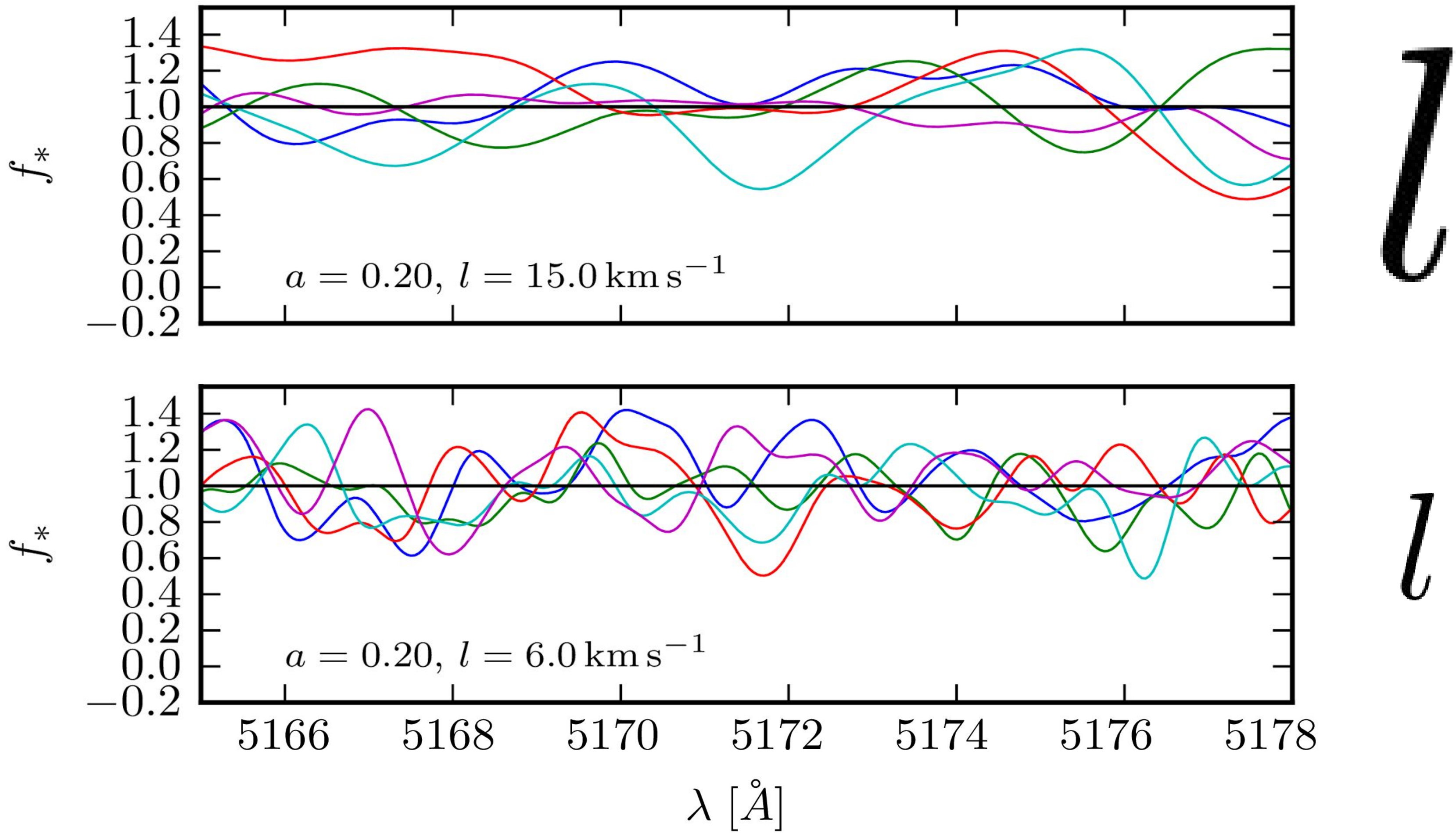
A function is said to have a Gaussian process if for any collection of inputs the random vector ***f*** has a multivariate Gaussian distribution with mean ***mu*** and covariance matrix given by *k* evaluated over ***lambda***

For a covariance kernel, we will use the commonly used squared exponential kernel, which relates pixels in the spectrum based upon their distance in log-wavelength (∝ velocity)

$$k_{ij}(r_{ij} \,|\, a, \boxed{l}) = a^2 \exp\left(-\frac{r_{ij}^2}{2l^2}\right)$$

Gaussian Process model for a single, stationary star

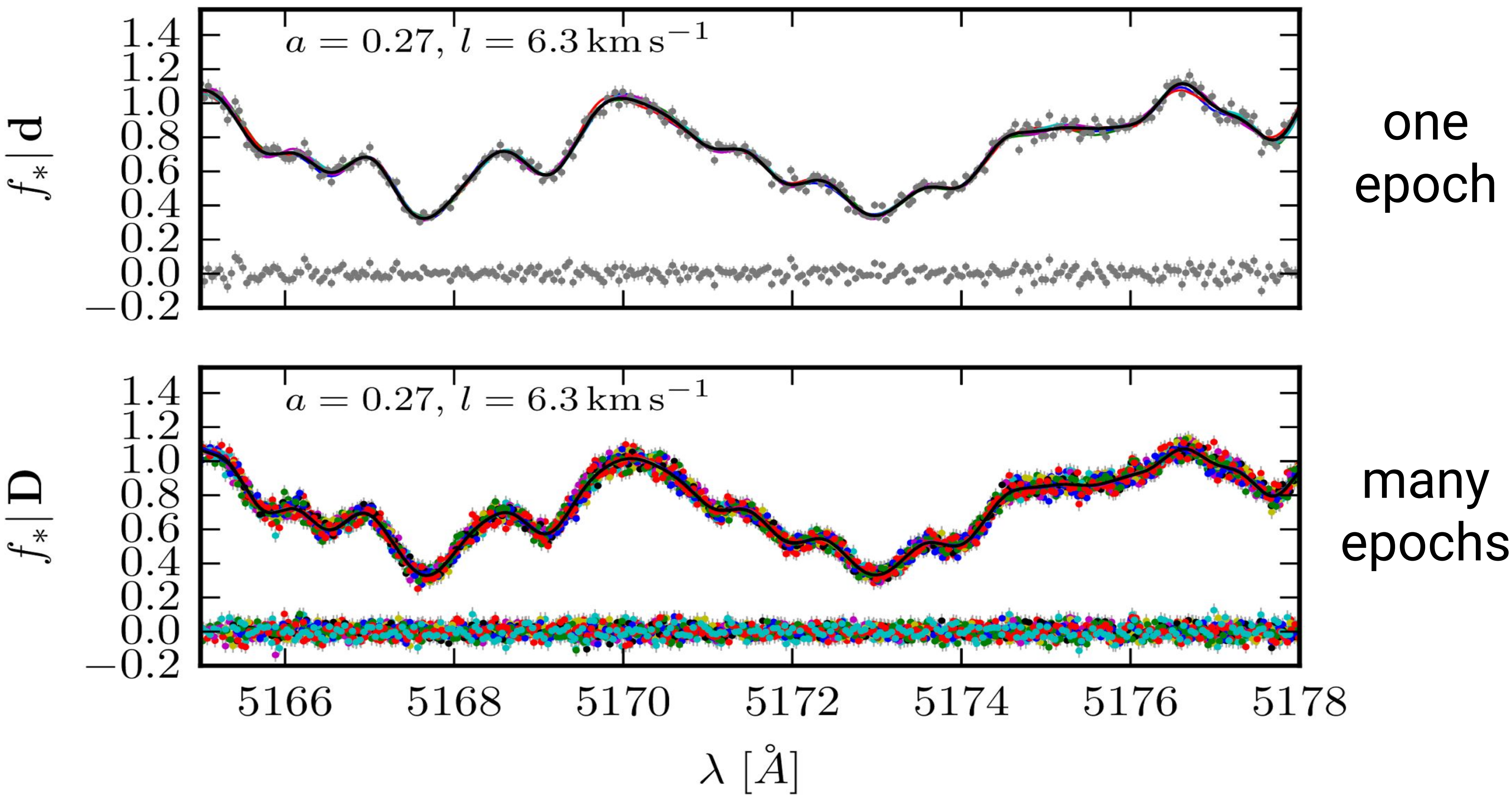
(Zoomed) draws from the prior



Czekala et al. 17a

Gaussian Process model for a single, stationary star

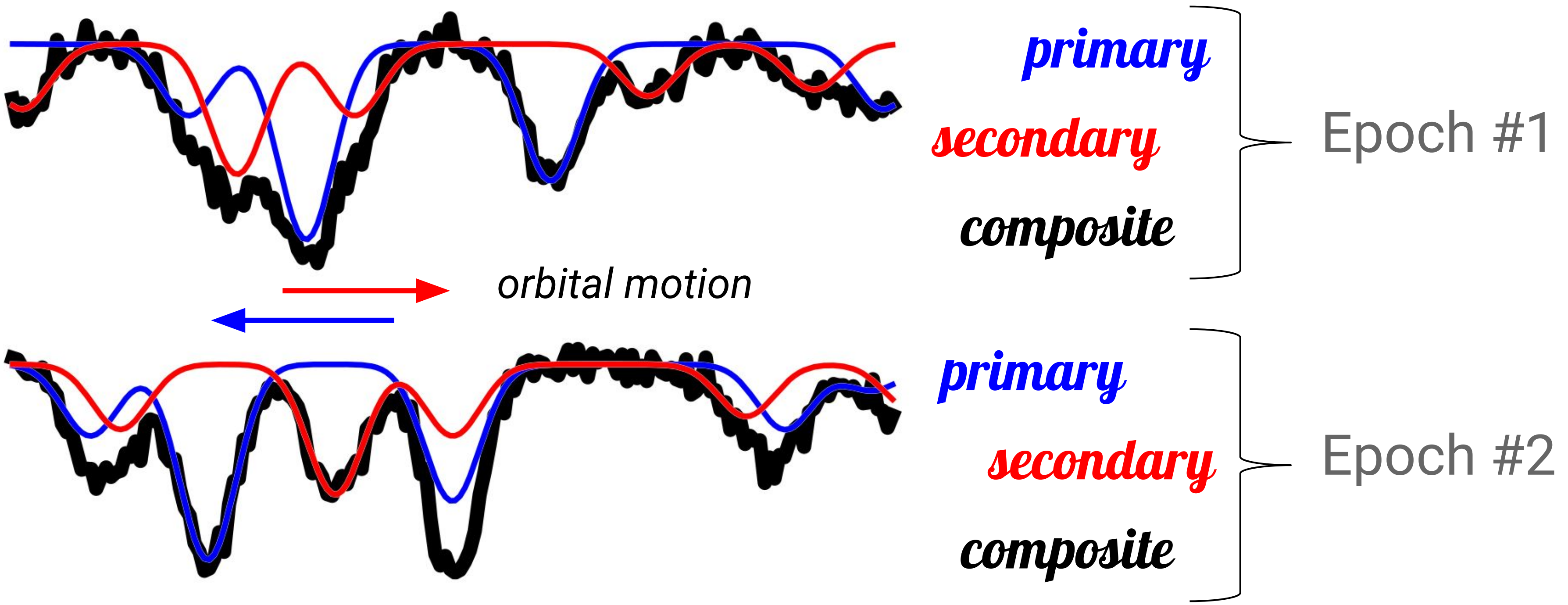
Conditioned on ***data***



a, ***ℓ*** optimized using Gaussian process likelihood

Czekala et al. 17a

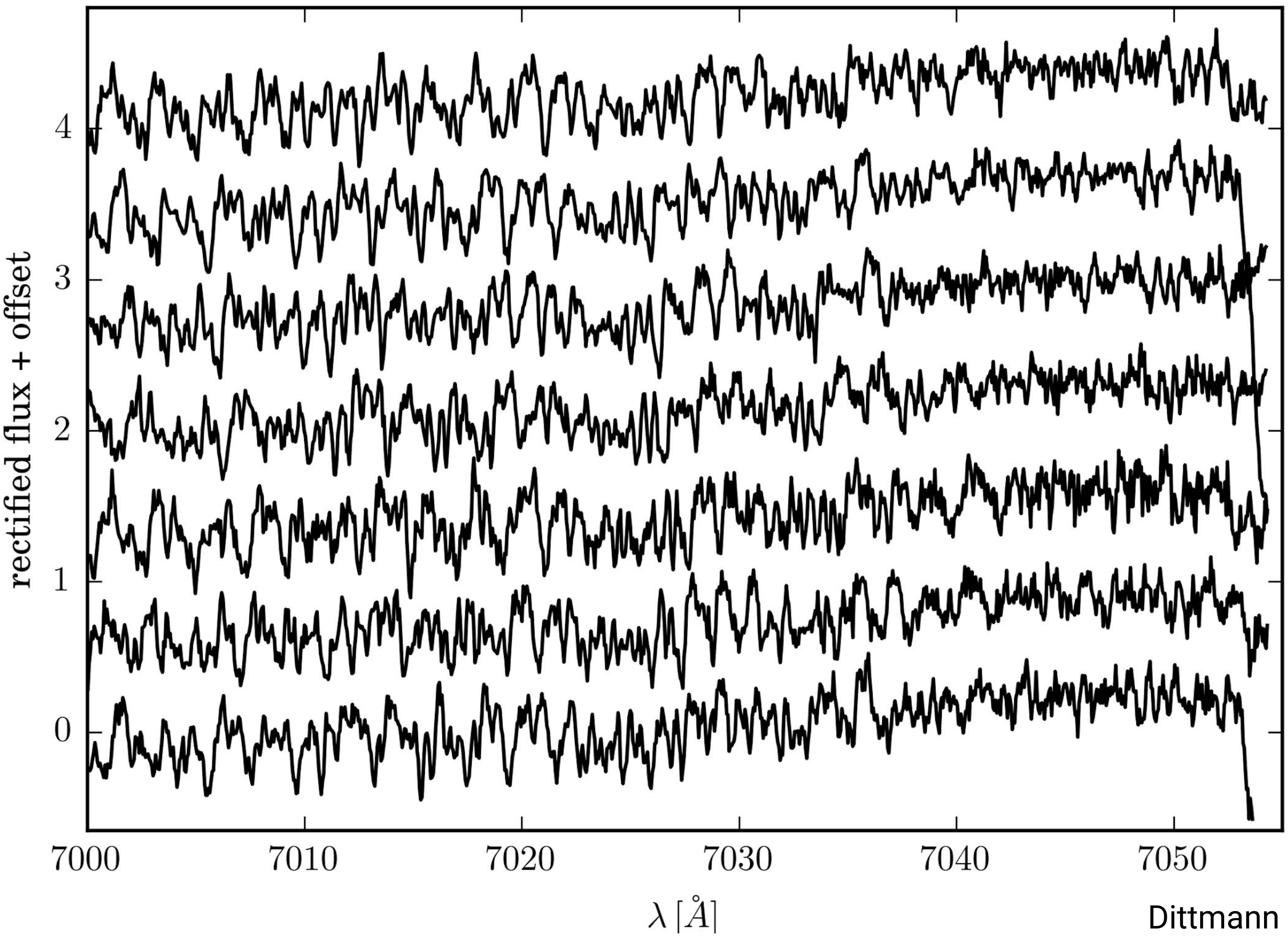
Spectroscopic Binary Stars



- Model the rest-frame spectra while only having access to **datasets** which represent the **sum** of the two spectra at different orbital phases
- The sum of two independent Gaussian processes is also a Gaussian process
- Joint GP likelihood maximized when the correct orbit is specified!

Raw Observations of the LP661-13 M4 Binary

time

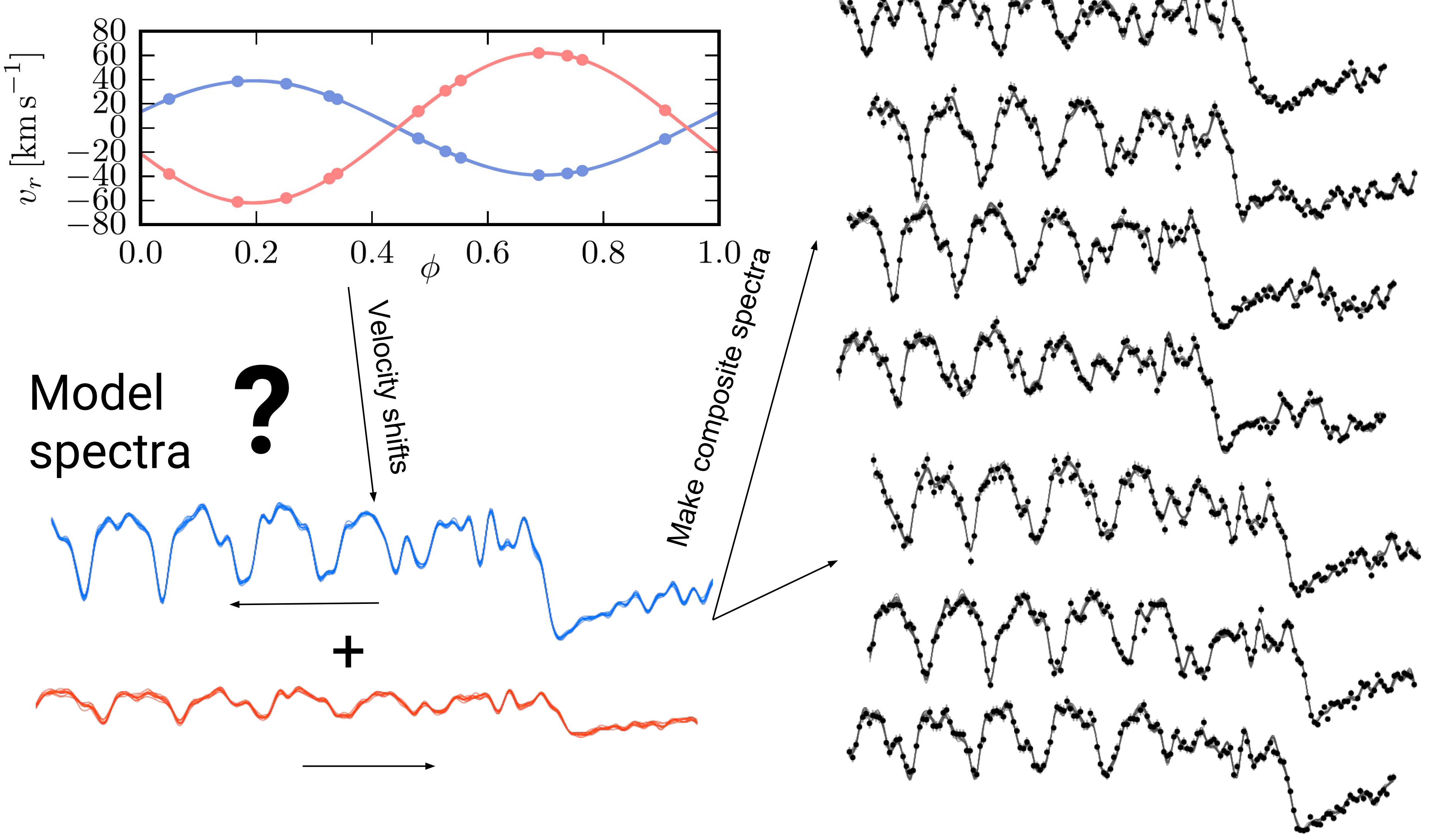


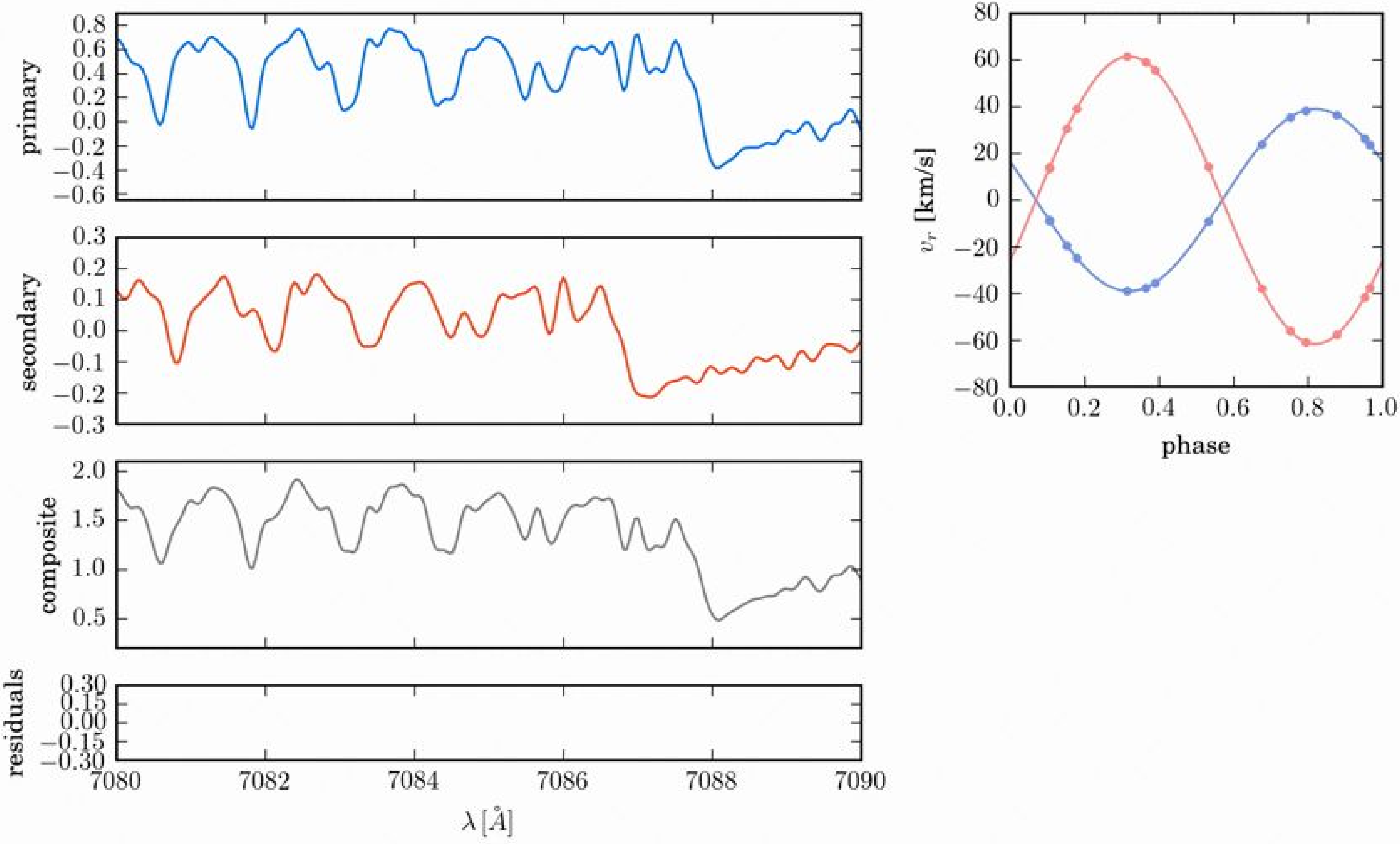
Dittmann et al. 17
Czekala et al. 17a

Problem setup

Orbit: period, eccentricity, phase, etc. ?

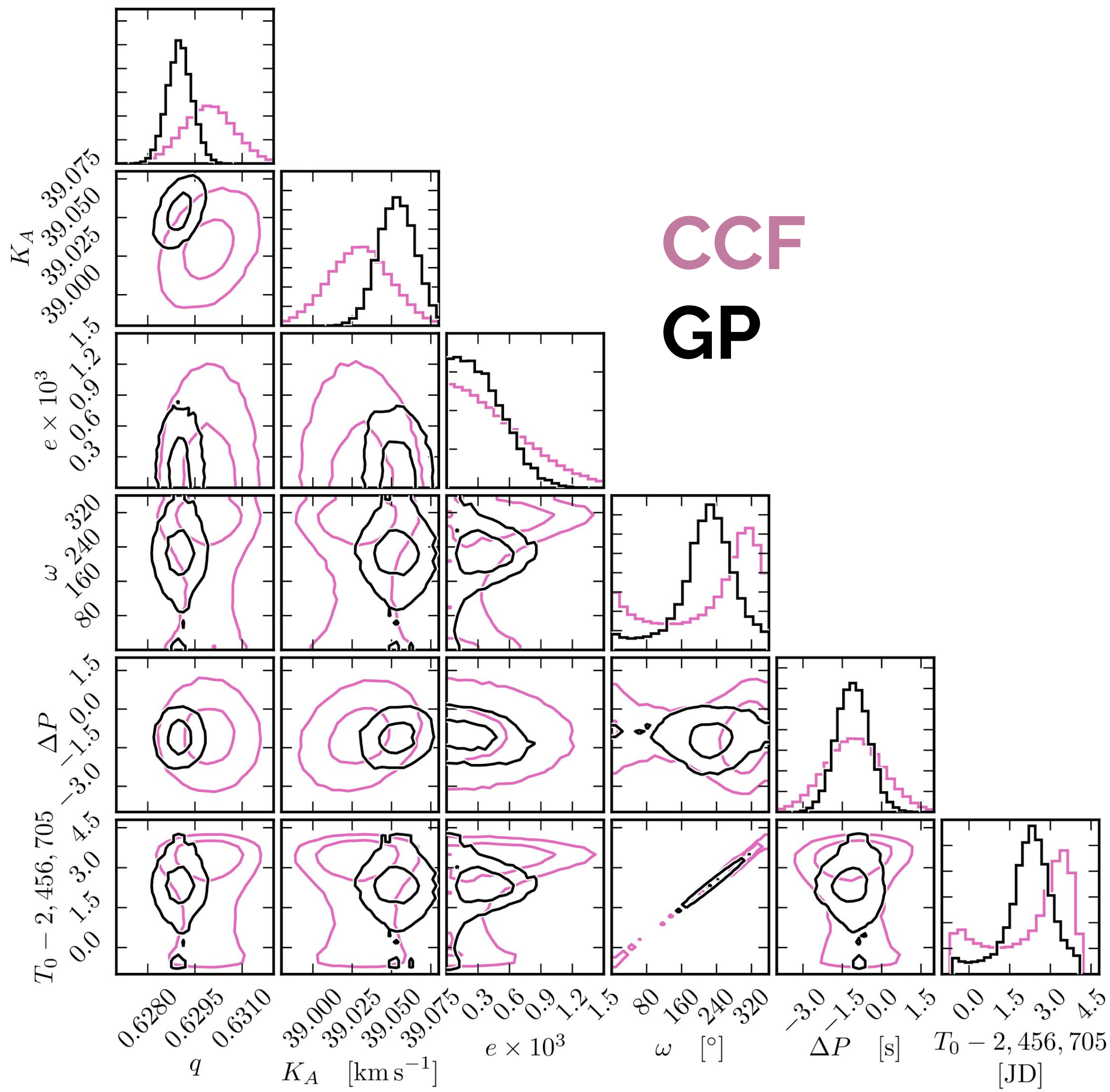
Data spectra





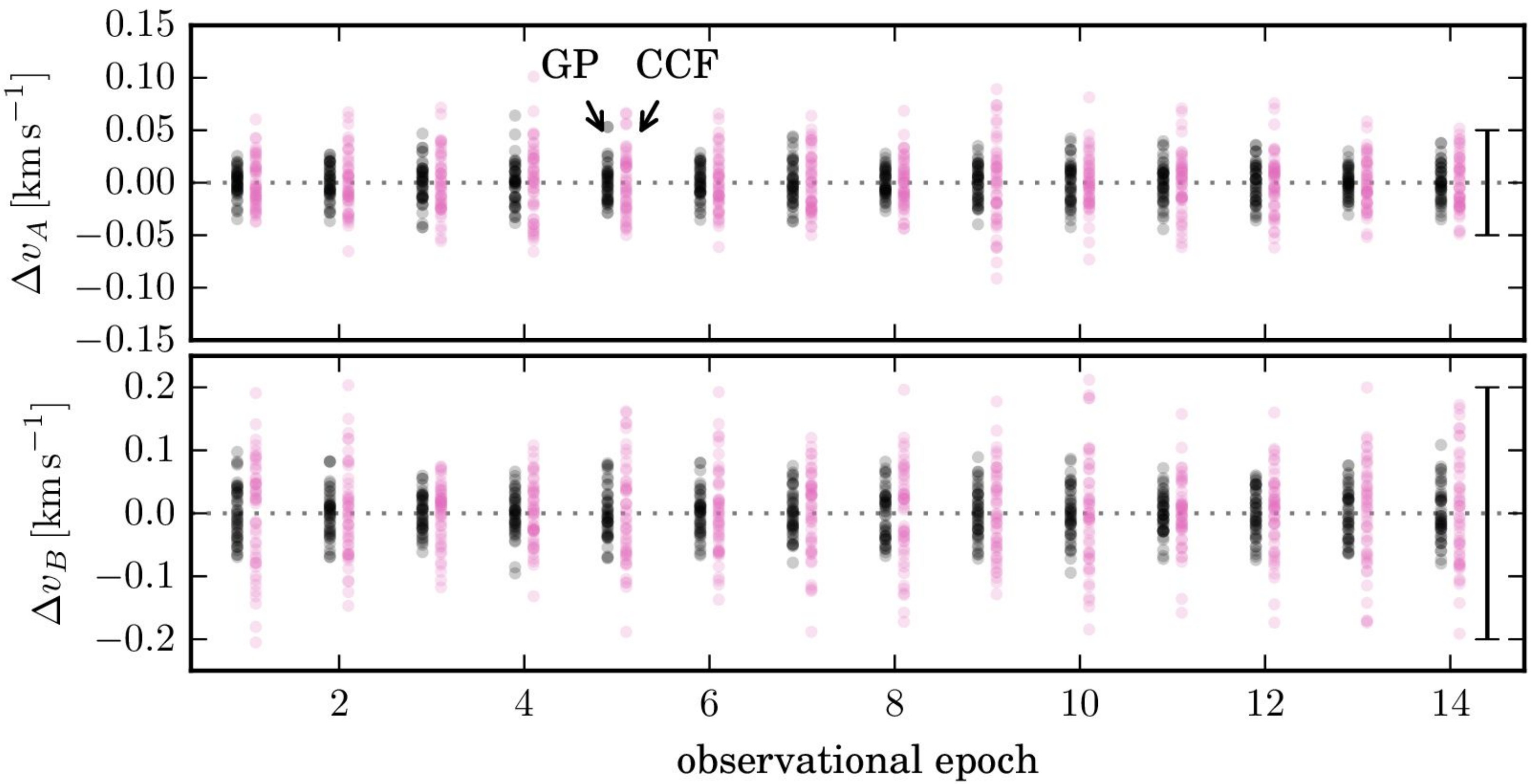
primary
secondary
composite
residuals
λ [Å]
7080
7082
7084
7086
7088
7090
v_r [km/s]
phase

GP framework can deliver more precise orbital constraints than Cross-Correlation



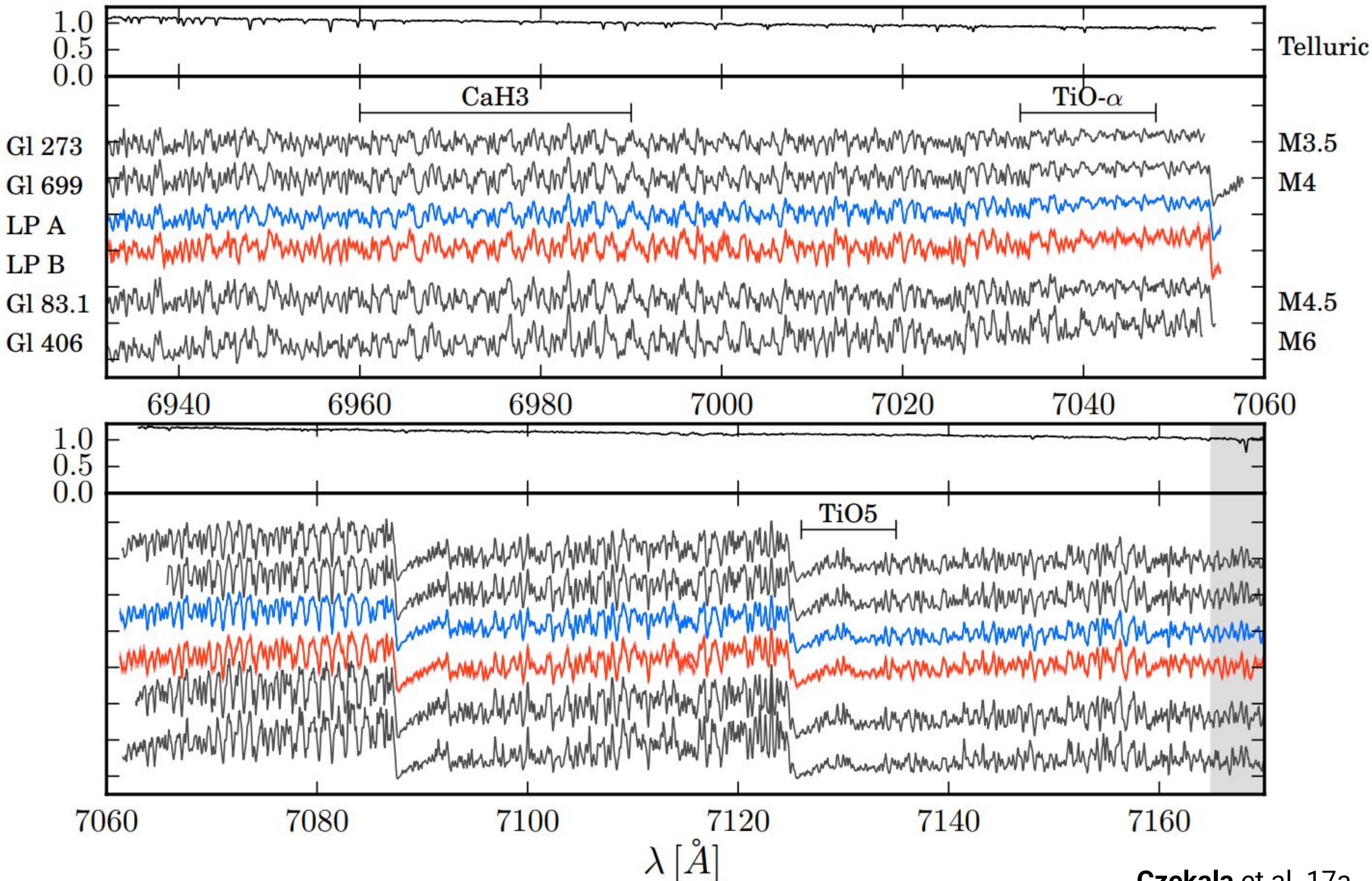
Czekala et al. 17a

GP reduces RV scatter compared to CCF



Czekala et al. 17a

Disentangled spectra match other single standard stars



Czekala et al. 17a

PSOAP

0.2.0

Search docs

psoap.readthedocs.io

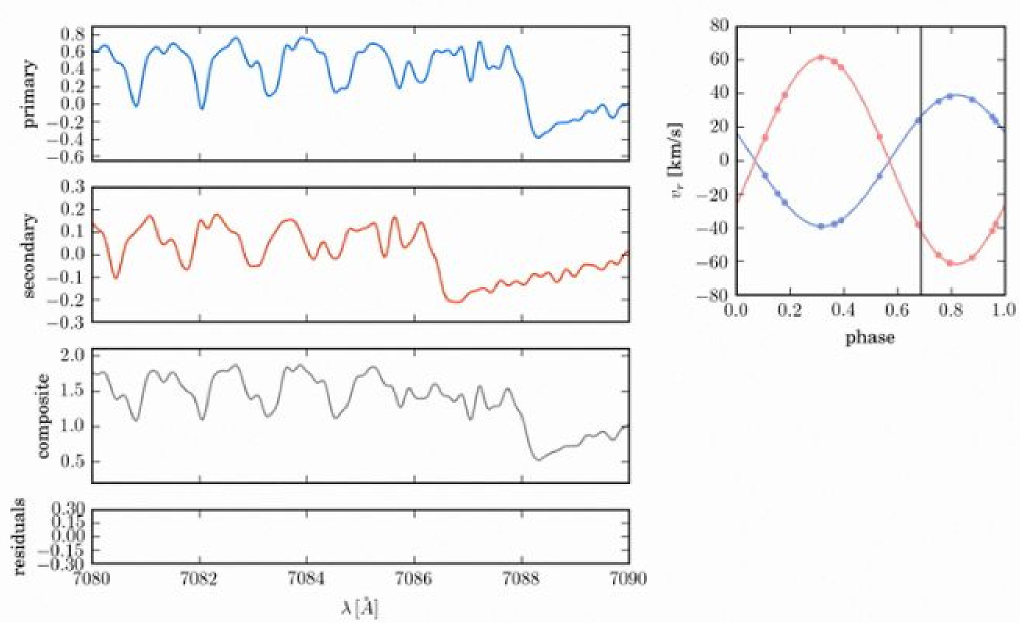
PSOAP

Precision Spectroscopic Orbits A-Parametrically (pronounced 'soap')

PSOAP is a package to model astronomical spectra nonparametrically, for the purposes of disentangling spectra (as in double-lined spectroscopic binaries) or simply determining orbits (in the case of (single-lined spectroscopic binaries or exoplanet hosts). For more information about the mathematical framework underlying *PSOAP*, please see our paper

Disentangling Time Series Spectra with Gaussian Processes: Applications to Radial Velocity Analysis, Czekala et al., 2017ApJ...840...49C

PSOAP is actively developed on github here.



Extensions: exploring variability

- **High variability**: some pre-main sequence stars are still actively accreting gas
  - Disentangling photosphere and accretion spectrum to measure the accretion rate from the protoplanetary disk
- **Low variability**: Starspots, chromospheric activity can bias precision radial velocity measurements

Joey Murphy (Stanford)

Poster
Murphy et al. **339.08**
Thursday 5:30pm

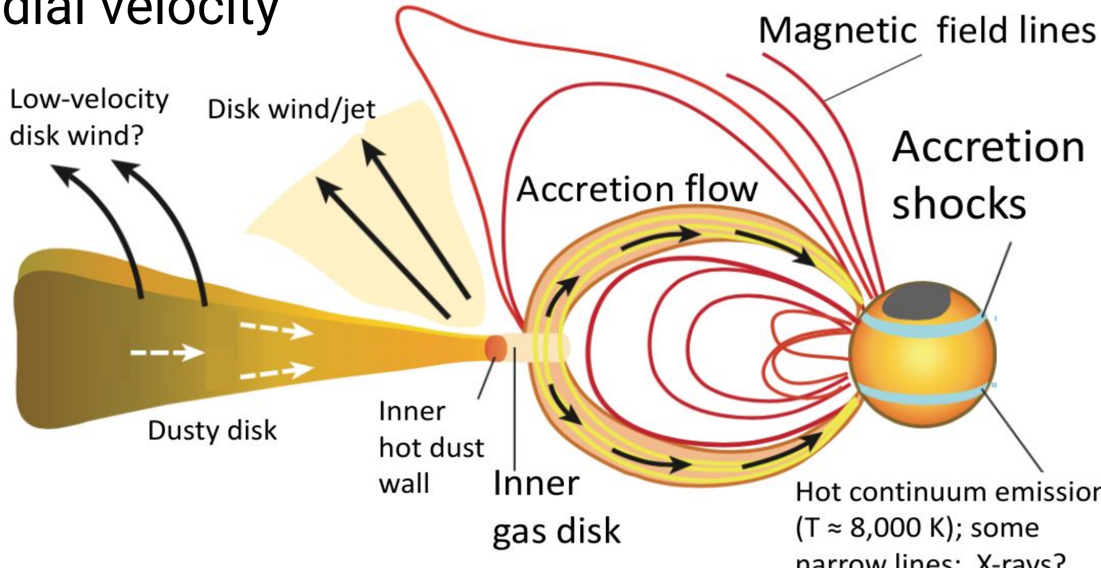


Figure adapted from Hartmann et al. 16

Murphy, **Czekala**, et. al. *in prep*

Conclusions

- Gaussian processes provide a flexible basis for modeling complex astrophysical spectra
- Data-driven models can sometimes circumvent limitations of physics-based models
- E.g., radial-velocity inference: recover intrinsic spectra to an accuracy far exceeding the average pixel noise in the dataset (see also **Czekala et al. 2017b**)
- No knowledge of spectral types, flux ratio, synthetic models needed!
- Many possible future applications: variability, micro-telluric modelling, non-stellar applications (e.g., high dispersion coronography)

Thank you!

PSOAP

psoap.readthedocs.io

samsi

NSF • Duke • NCSU • UNC • NISS

Thanks to SAMSI astrophysics workshop and funding for this project.

Extra slides

Multivariate Gaussians

x is a random vector drawn from the Normal distribution

$$\boldsymbol{x} \;\sim\; \mathcal{N}\left(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx}\right)$$

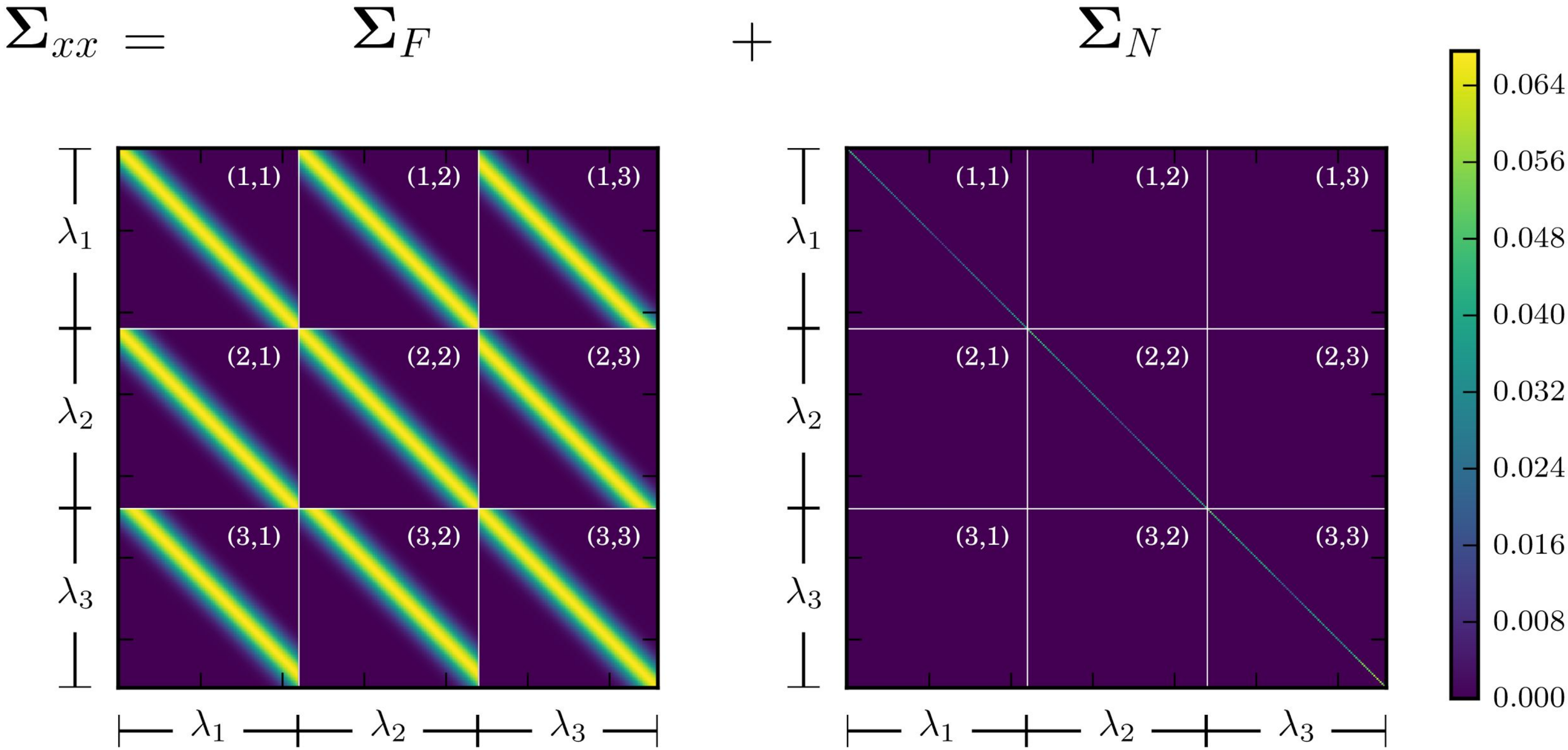
E.g., `numpy.random.multivariate_normal()` or `MRANDOMN()`

the likelihood function for **x** is given by

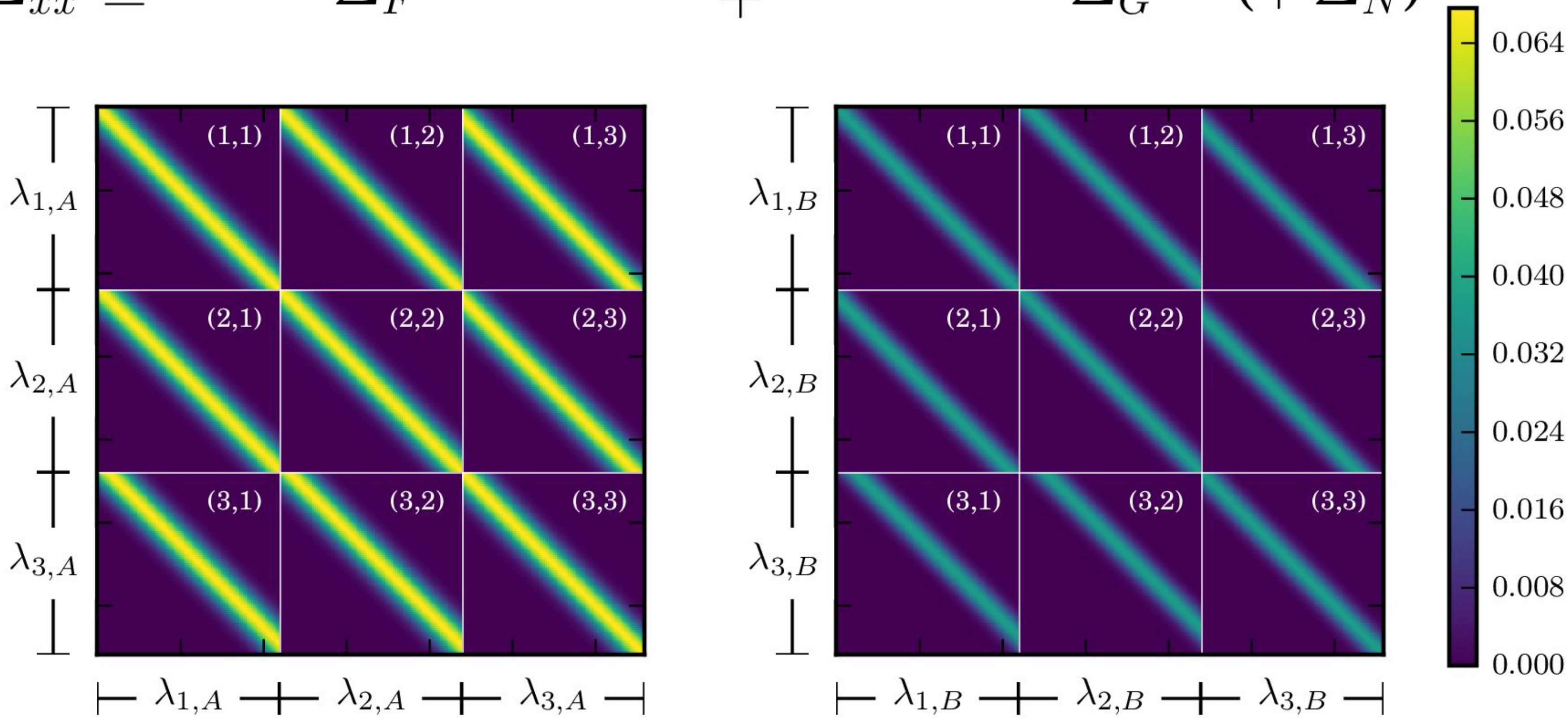
$$p(\boldsymbol{x}|\,\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx}) = \frac{1}{\left[(2\pi)^N \det \boldsymbol{\Sigma}_{xx}\right]^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_x)^{\mathrm{T}}\boldsymbol{\Sigma}_{xx}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_x)\right)$$

If the covariance matrix is diagonal, we arrive back at the well-known

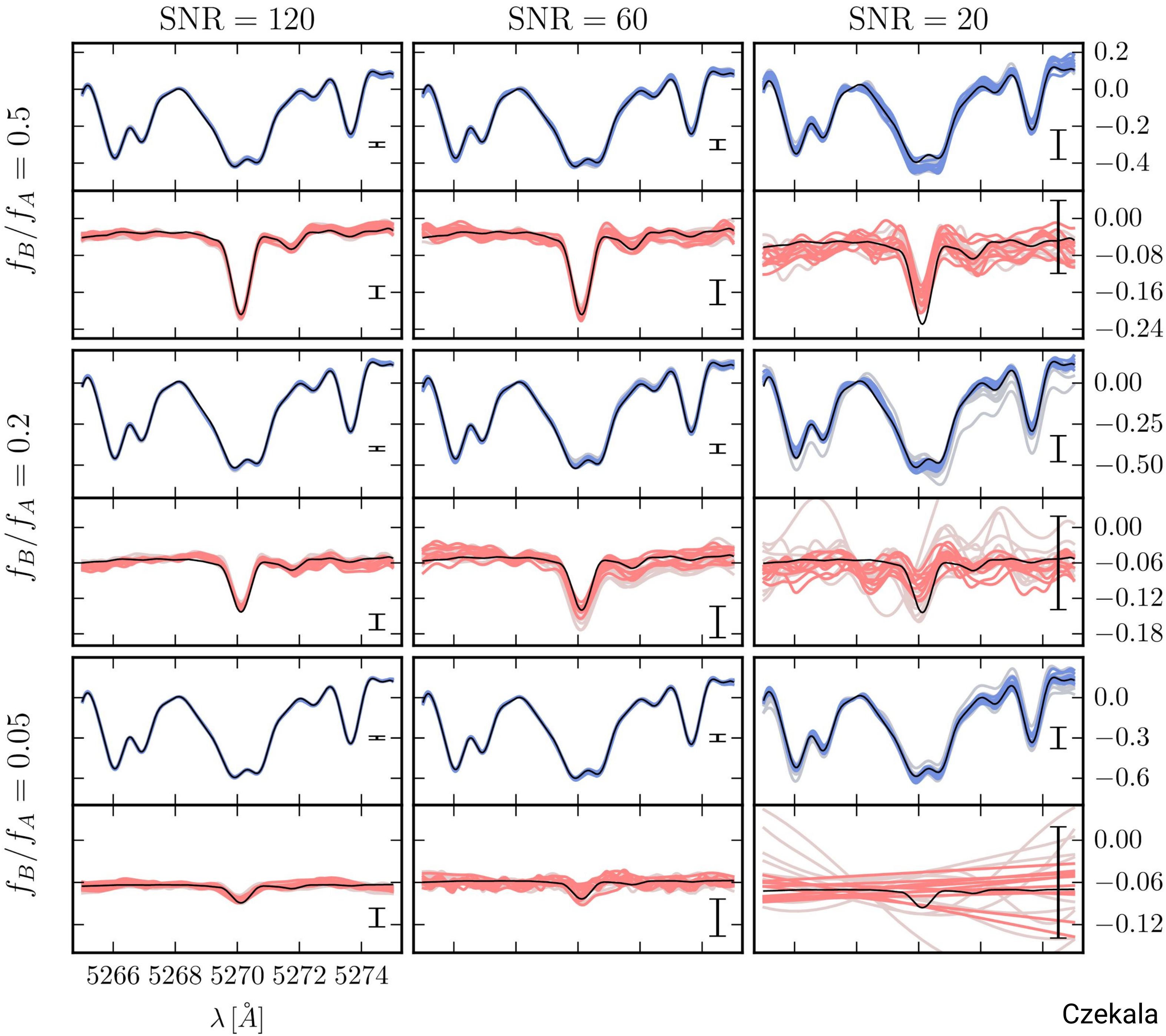
$$\ln p(\mathbf{x}|\,\mu_x, \boldsymbol{\Sigma}_{xx}) = \frac{-\chi^2}{2}$$


$\mathbf{\Sigma}_{xx} =$
$\mathbf{\Sigma}_F$
$+$
$\mathbf{\Sigma}_N$
(1,1)
(1,2)
(1,3)
(2,1)
(2,2)
(2,3)
(3,1)
(3,2)
(3,3)
λ_1
λ_2
λ_3
0.064
0.056
0.048
0.040
0.032
0.024
0.016
0.008
0.000

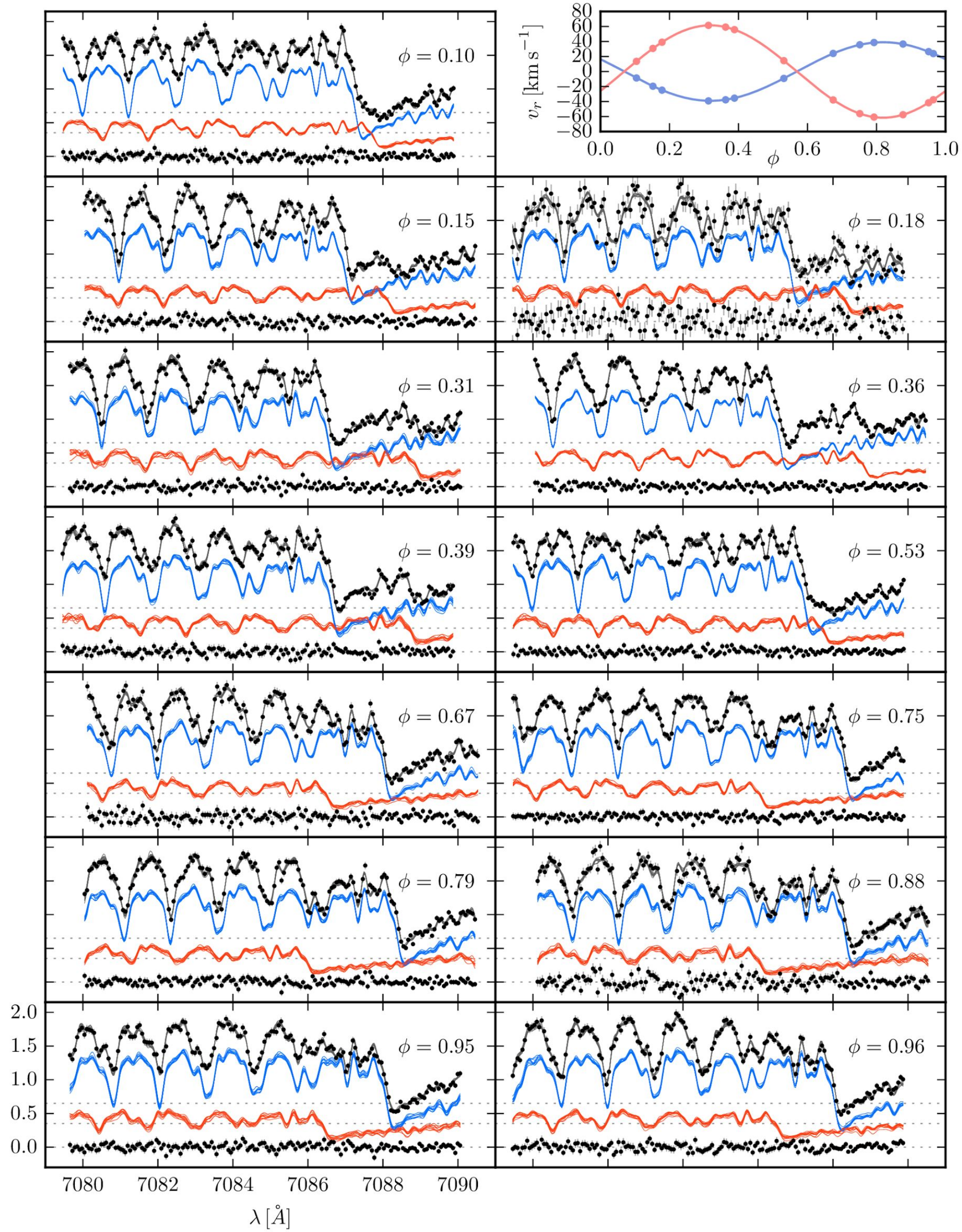
$$\mathbf{\Sigma}_{xx} = \mathbf{\Sigma}_F + \mathbf{\Sigma}_G \quad (+\mathbf{\Sigma}_N)$$



$$\begin{bmatrix} \boldsymbol{f}_* \\ \boldsymbol{g}_* \\ \boldsymbol{D} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_f \\ \boldsymbol{\mu}_g \\ \boldsymbol{\mu}_{FG} \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma}_f & \mathbf{0} & \mathbf{\Sigma}_{fF} \\ \mathbf{0} & \mathbf{\Sigma}_g & \mathbf{\Sigma}_{gG} \\ \mathbf{\Sigma}_{Ff} & \mathbf{\Sigma}_{Gg} & (\mathbf{\Sigma}_F + \mathbf{\Sigma}_G + \mathbf{\Sigma}_N) \end{bmatrix} \right)$$



Czekala et al. 17



$\phi = 0.10$
$\phi = 0.15$
$\phi = 0.18$
$\phi = 0.31$
$\phi = 0.36$
$\phi = 0.39$
$\phi = 0.53$
$\phi = 0.67$
$\phi = 0.75$
$\phi = 0.79$
$\phi = 0.88$
$\phi = 0.95$
$\phi = 0.96$
$v_r\,[\mathrm{km\,s^{-1}}]$
ϕ
0.0 0.2 0.4 0.6 0.8 1.0
80 60 40 20 0 −20 −40 −60 −80
2.0 1.5 1.0 0.5 0.0
7080 7082 7084 7086 7088 7090
$\lambda\,[\mathring{A}]$

Test	1	2	3	4	5	6	7	8	9
f_B/f_A	0.50	0.50	0.50	0.20	0.20	0.20	0.05	0.05	0.05
SNR	120	60	20	120	60	20	120	60	20
σ_A	0.219	0.219	0.219	0.191	0.191	0.191	0.153	0.153	0.153
σ_B	0.005	0.005	0.005	0.019	0.019	0.019	0.038	0.038	0.038