Astrostatistics

Lecture 02: Monday, 20 January 2019

Recommended Reading:

Feigelson & Babu: Chapters 1-4

Ivezic: Chapters 1, 3-5

C. Schafer article:

“A Framework for Statistical Inference in Astrophysics”

Intro to Statistics in Astronomy

Review of Probability & Statistics Foundations

Classical & Bayesian Statistical Inference

Office Hour: Tuesday @ 1pm D 1.07

kmandel@statslab.cam.ac.uk

https://github.com/CambridgeAstroStat/PartIII-Astrostatistics-2020

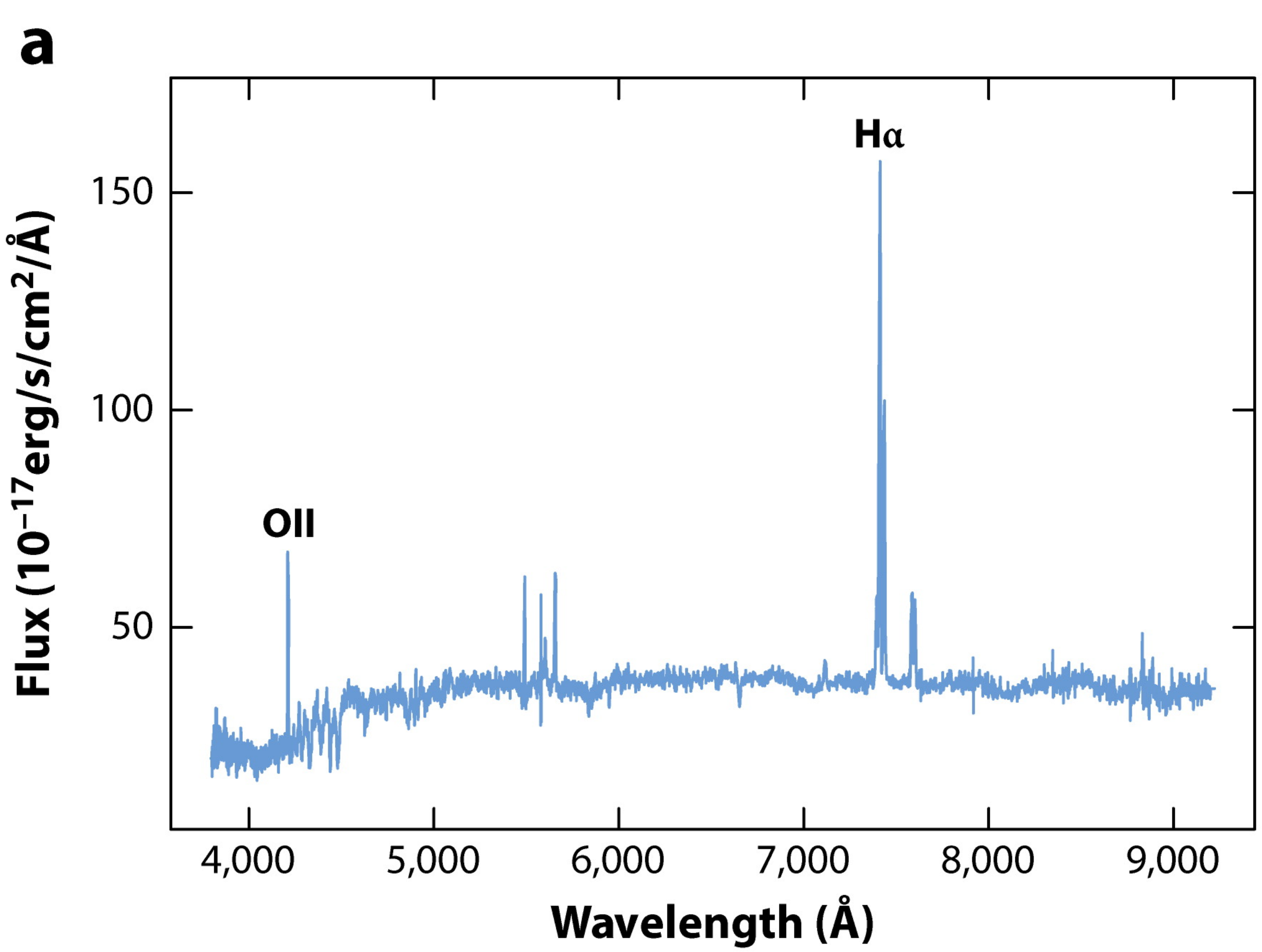
Today

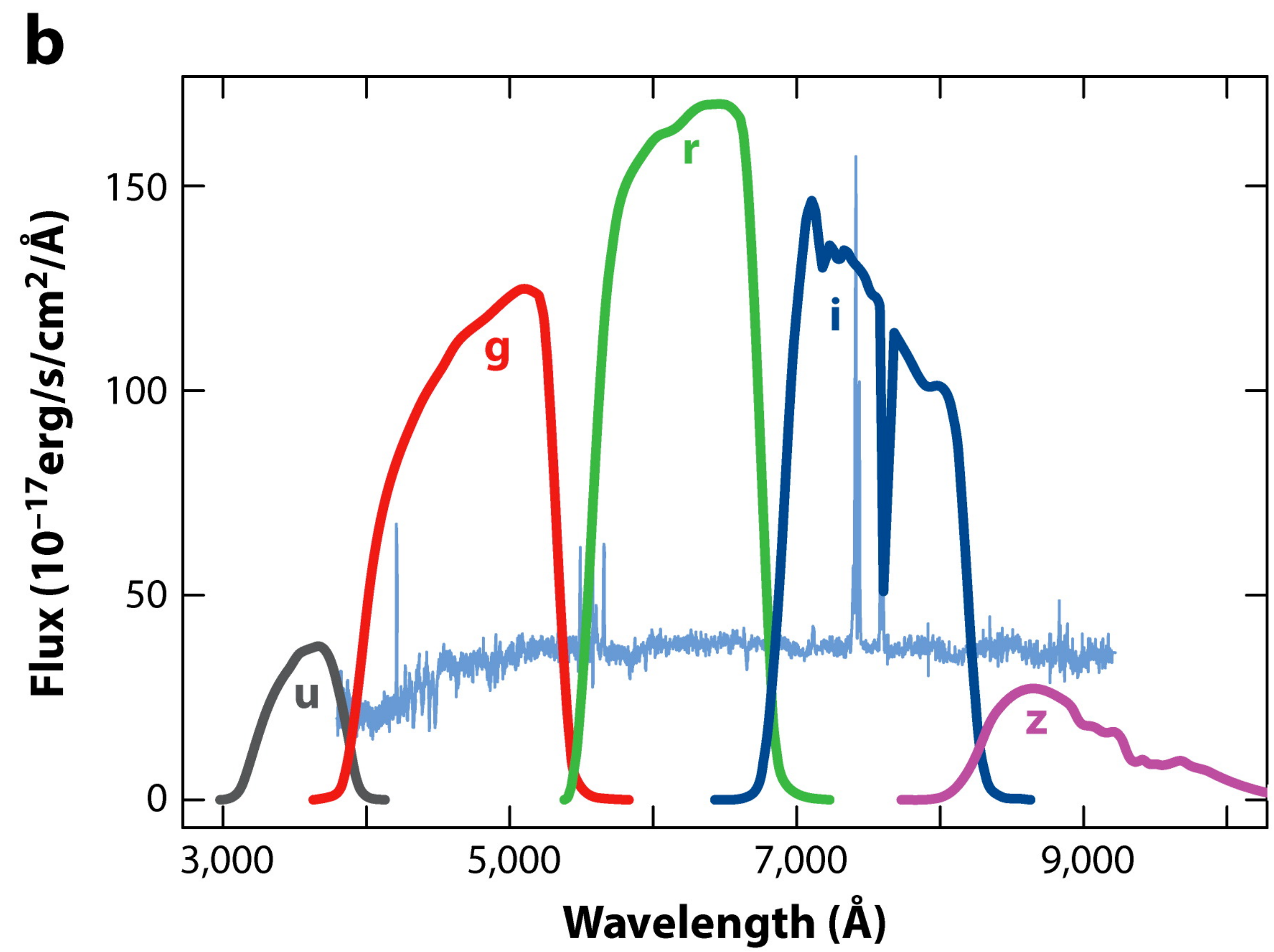
- Introduction to Astronomical Data Types
  - See also Schafer article “A Framework for Statistical Inference in Astrophhysics”
- Motivating Case Studies:
  - Gravitationally Lensed Time Delay Estimation
  - Bayesian Inference of the Milky Way Galactic Mass
  - Gaussian Processes for Spectral Time Series (Radial Velocity Analysis)
  - Hierarchical Bayes for Supernova Cosmology ?

What astronomers measure

- Astrometry (angular position on sky, e.g. Gaia)
- Photometry (how bright is it?)
  - Flux = photons (or energy) per second per meter^2
  - (apparent) Magnitude = $-2.5 \times \mathrm{Log}_{10}$ [Flux] + const
    - Absolute Magnitude = $-2.5 \times \mathrm{Log}_{10}$ [Luminosity] + const
      = apparent magnitude at fixed distance of 10pc
- Spectroscopy (brightness versus wavelength)
- Time Series (light curves): Transients & Variables (e.g. stars, quasars, supernovae, exoplanets), Moving objects (e.g. asteroids)
- Spatial Variation (images, maps) - clustering, spatial correlation functions
- Combinations of the above, e.g.
  - Astrometry vs. Time = Proper Motion (e.g. stars, satellite galaxies)
  - Spectroscopic Time Series (Radial Velocity studies of stars)

Spectroscopy and Photometry



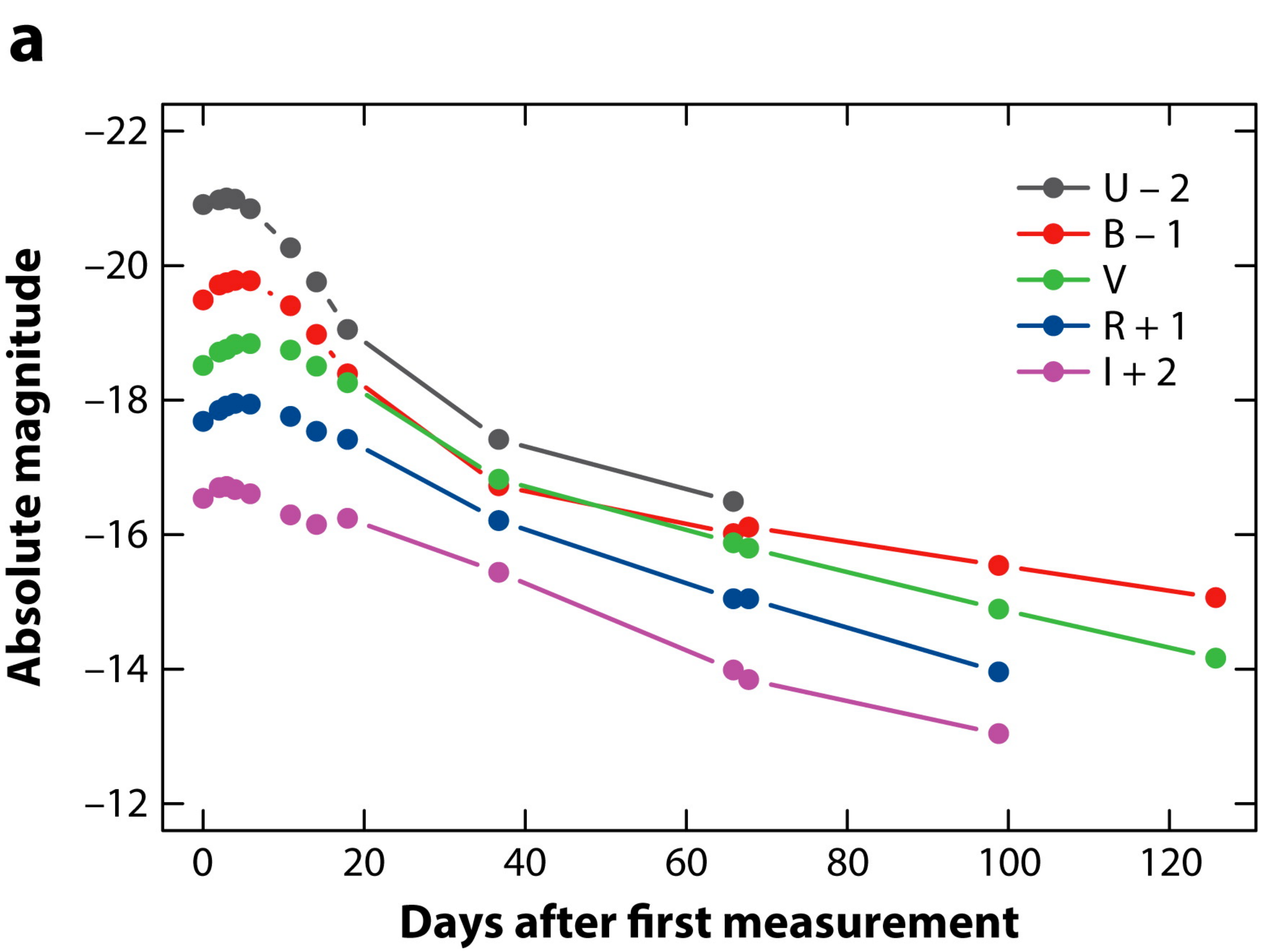


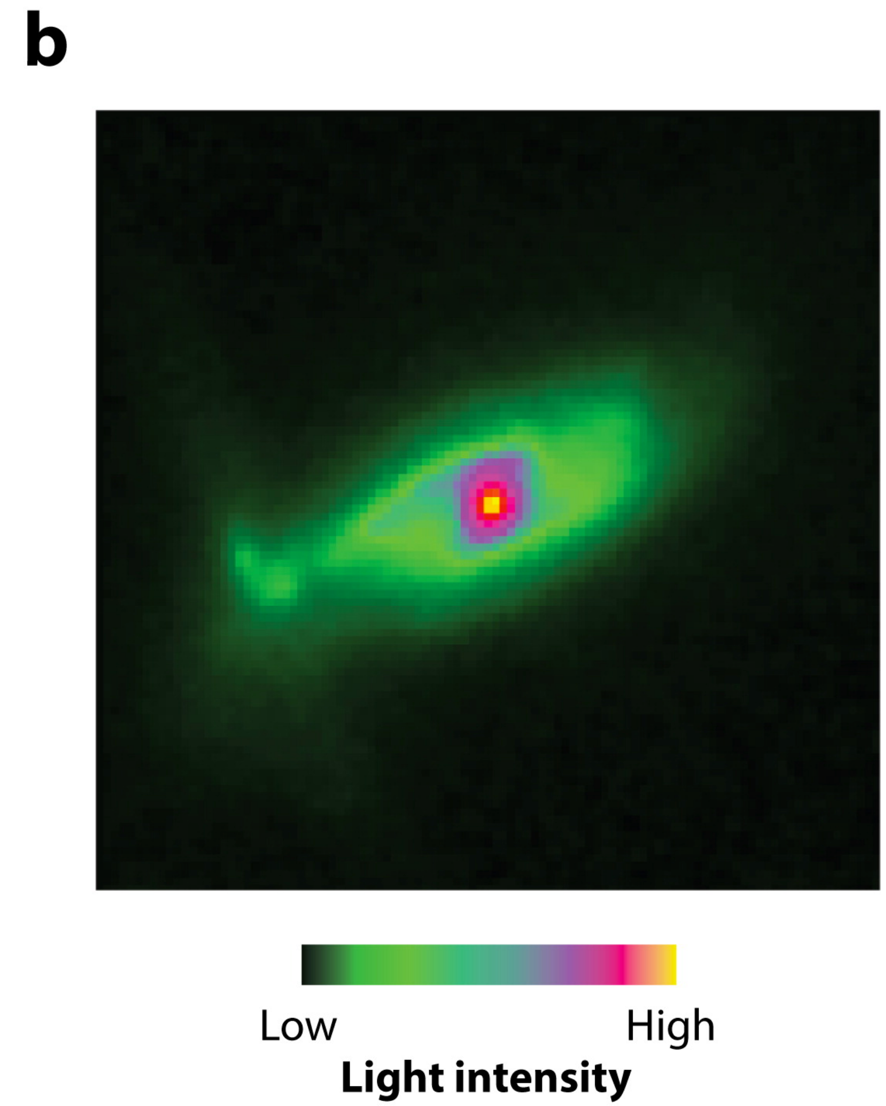
Schafer CM. 2015.
Annu. Rev. Stat. Appl. 2:141–62

Galaxy Spectrum

Galaxy Photometry =
Integral of Spectrum x Filter
(Brightness: Flux / Magnitude)

Temporal & Spatial Variation



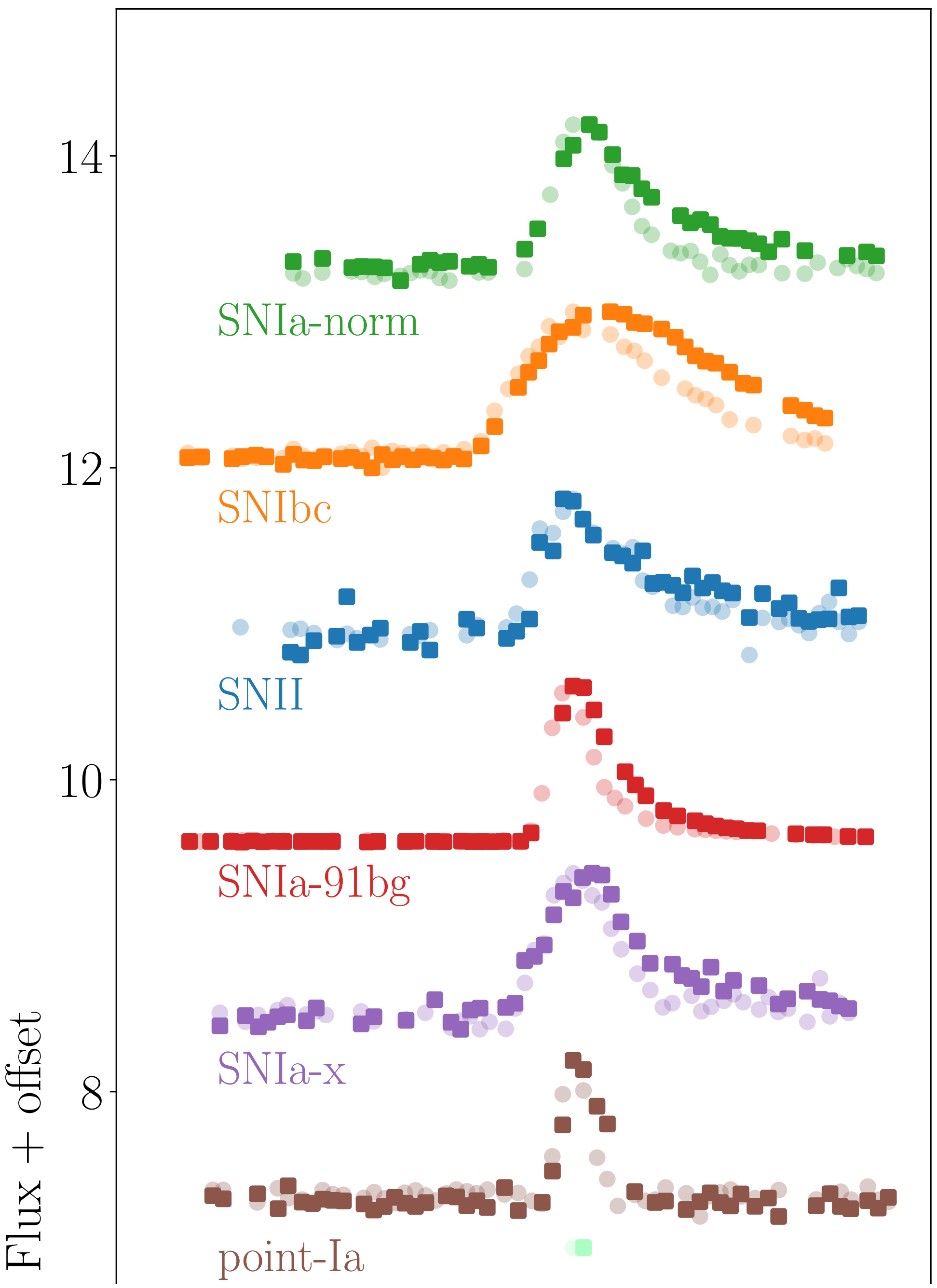


Schafer CM. 2015.
Annu. Rev. Stat. Appl. 2:141–62

Time Series (Light Curve)
Supernova

Galaxy Image
(Intensity Map)

Time-Series (Time Domain Astronomy)



<u>Goals:</u>

- Classification
- Astrophysics
- Cosmology

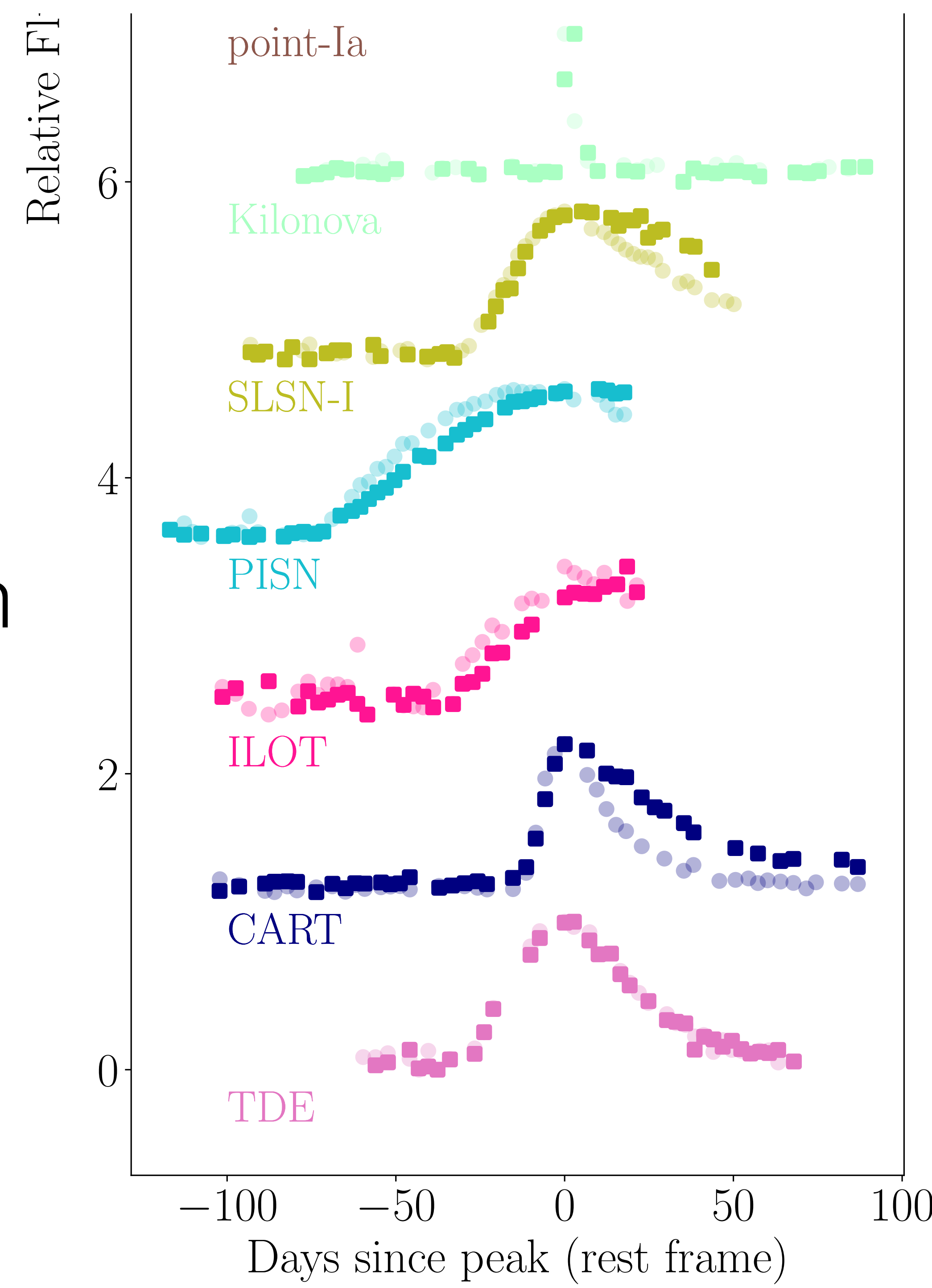


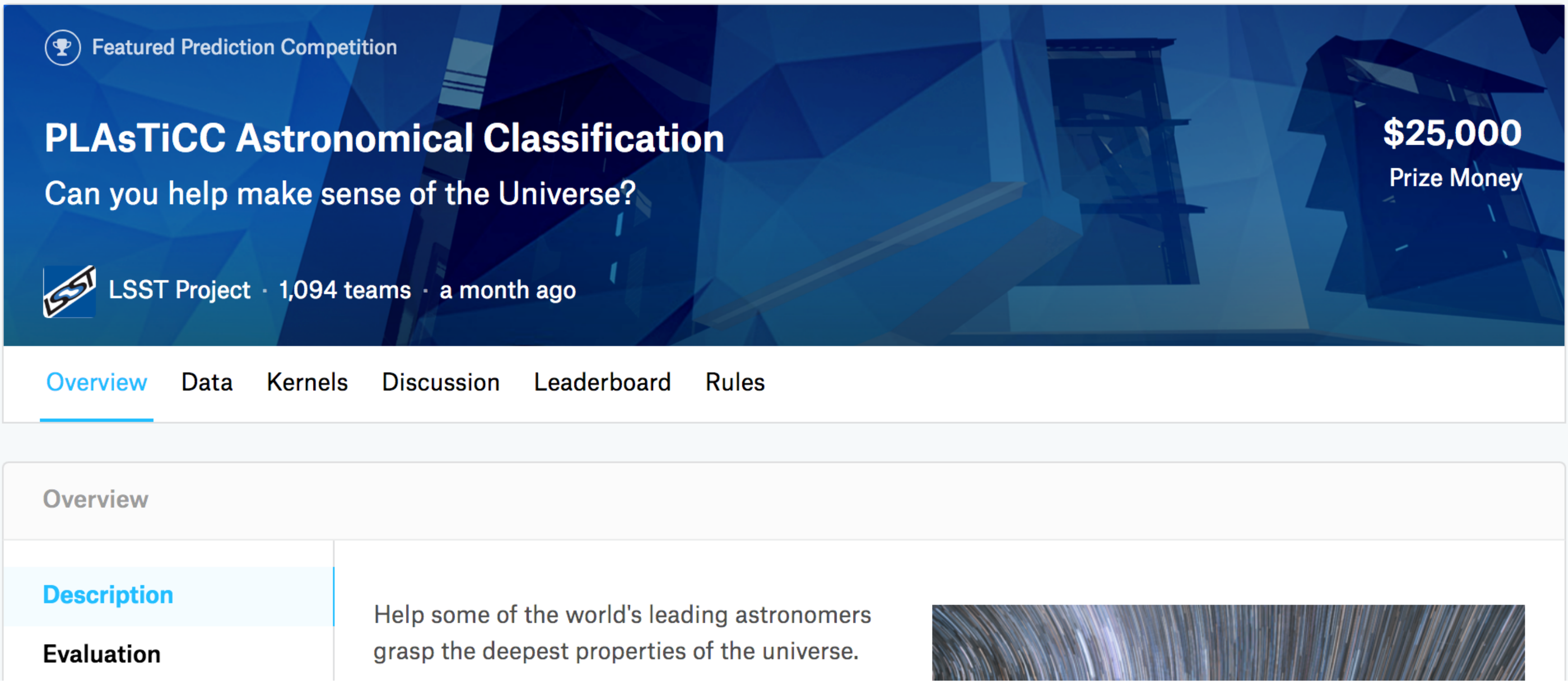
Figure 2. The light curves of one example transient from each of the 12 transient classes is plotted with an offset. We have only plotted transients with a high signal-to-noise and with a low simulated host redshift ($z < 0.2$) to facilitate comparison of light curve shape between the classes. The opaque square markers plots the r band light curves of each transient, while the transparent circle markers are the g band light curves of each transient.

Muthukrishna et al. 2019
Deep Learning for
Transient Classification

Photometric LSST Astronomical Time Series Classification Challenge (PLAsTiCC):

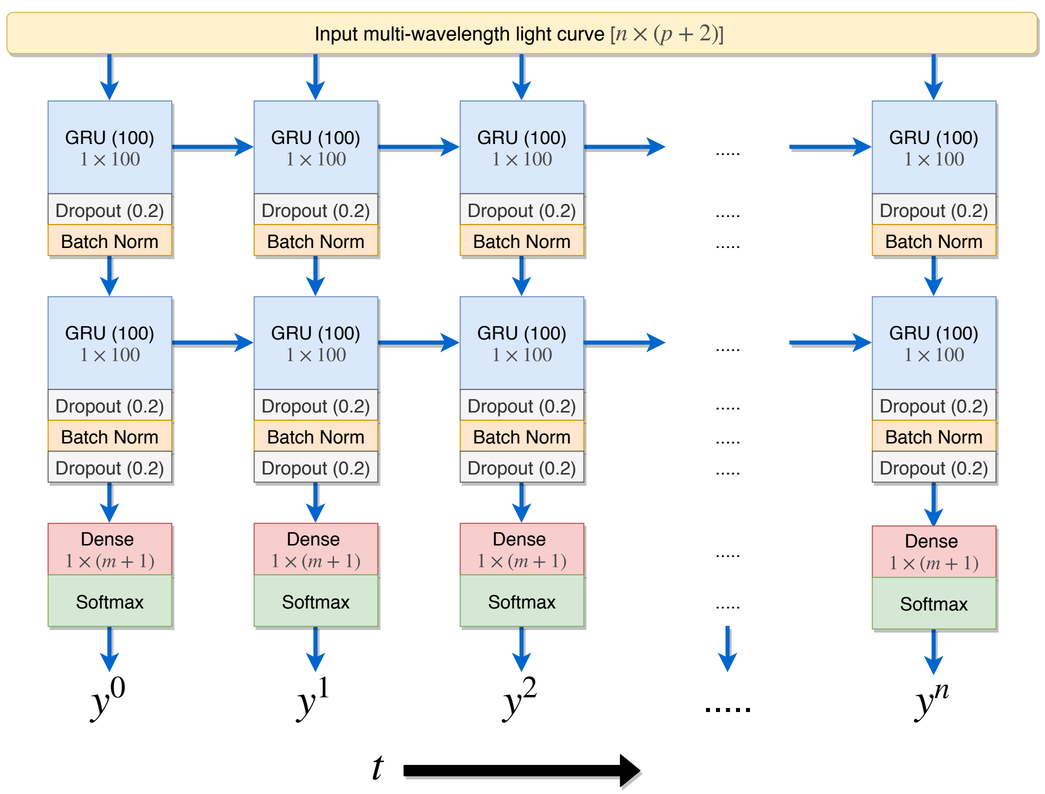
(arXiv:1810.00001)

https://www.kaggle.com/c/PLAsTiCC-2018

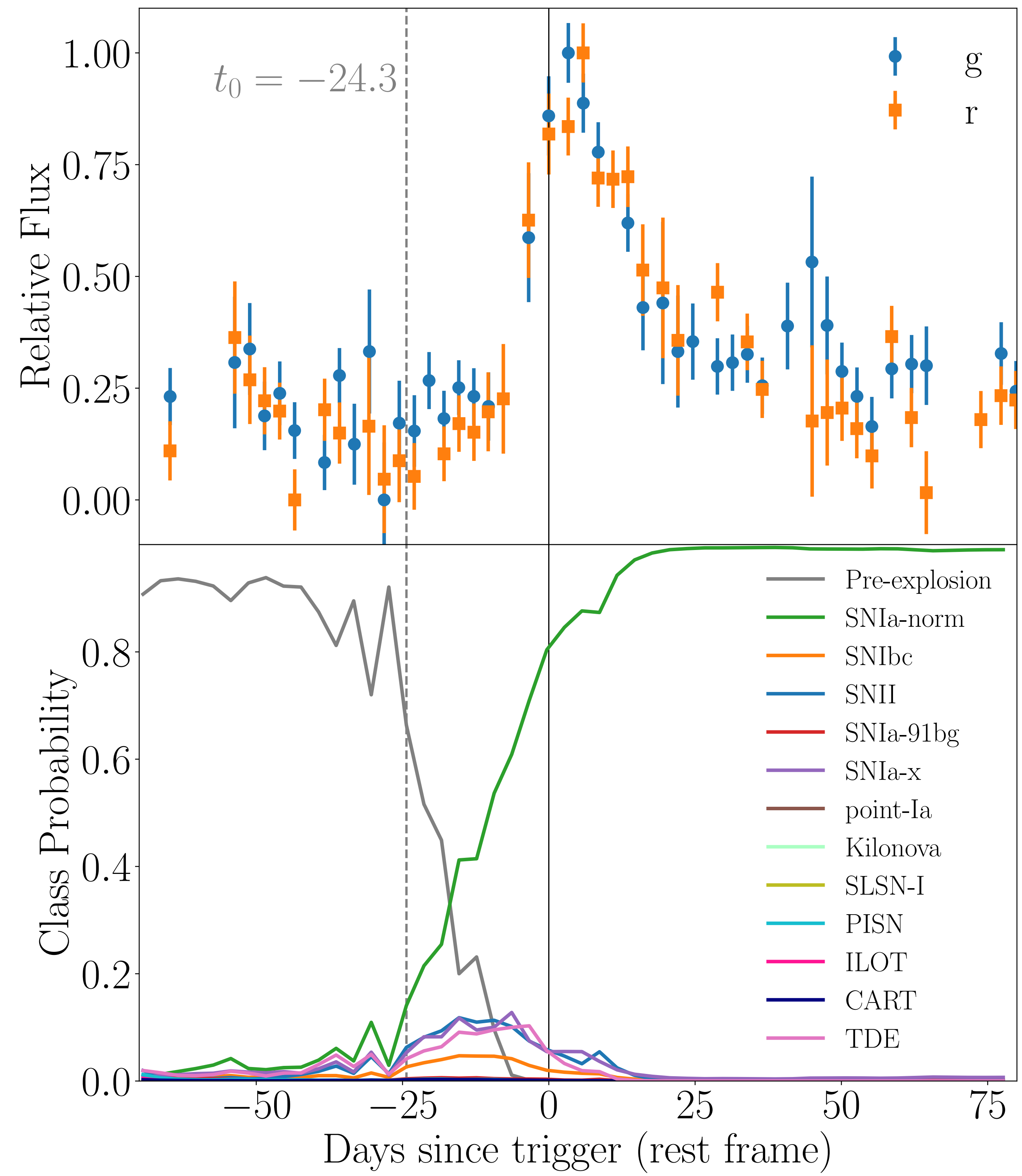


Transient Time Series Classification

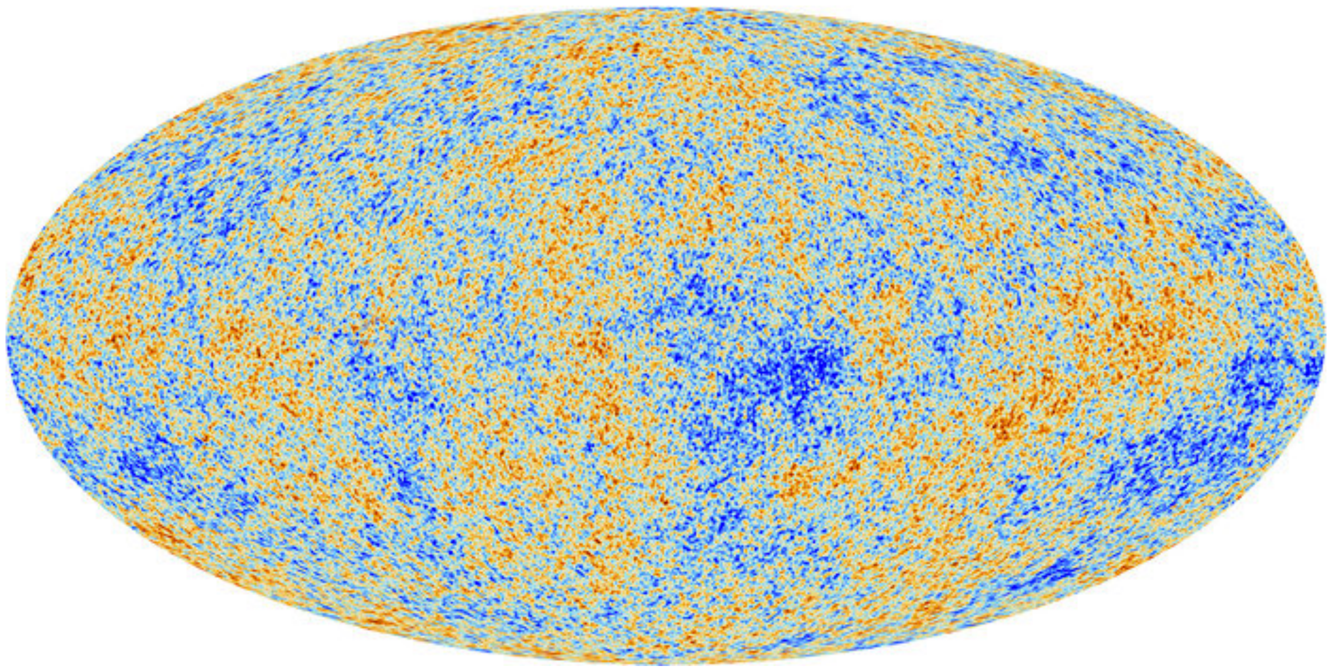
Deep Recurrent Neural Network



Muthukrishna et al. 2019 (arXiv:1904.00014) Deep Learning for Transient Classification

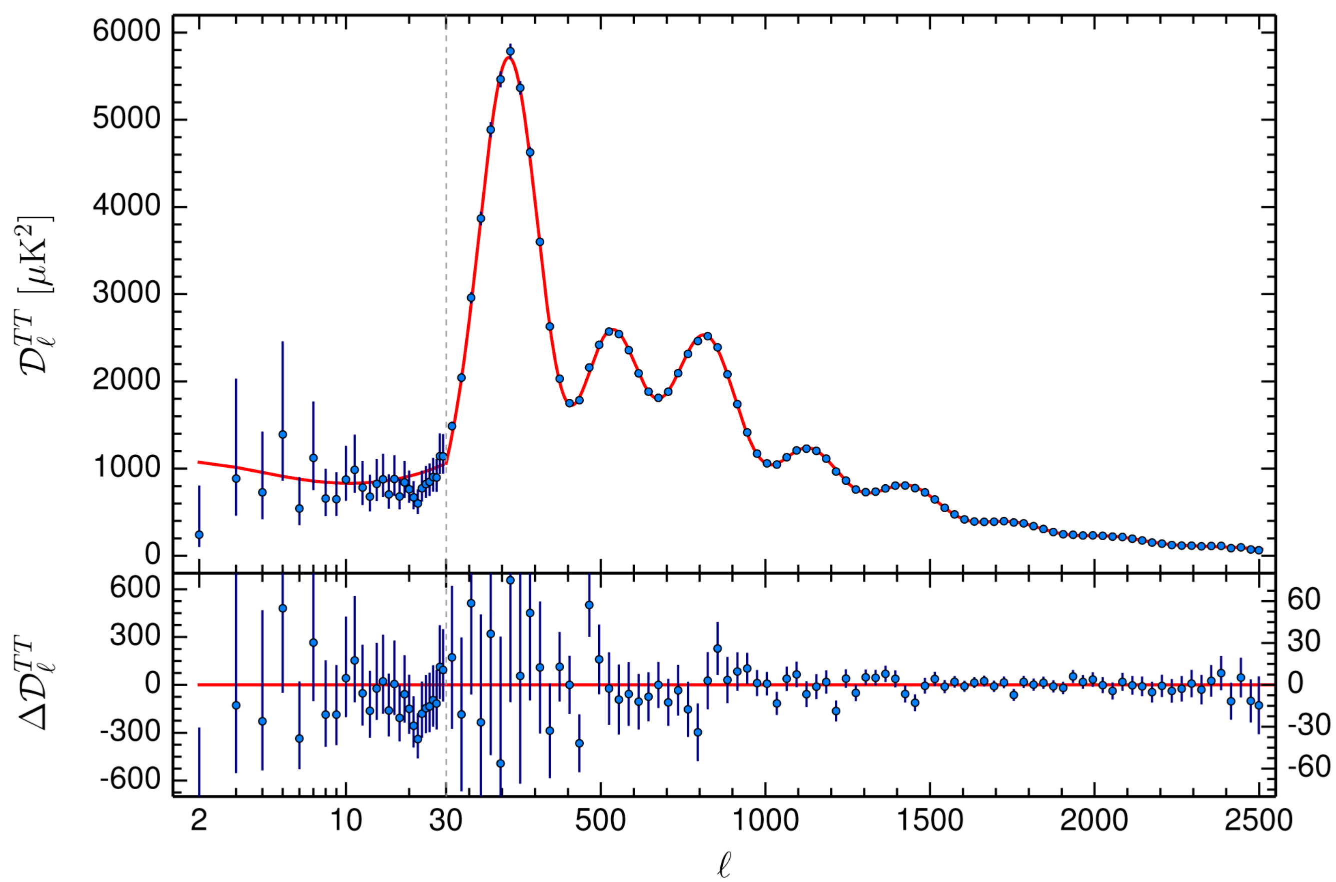


Spatial Variation



Cosmic Microwave Background (Planck) ~ Gaussian Random Field (mean = 2.7 K, std dev ~ 10^{-5})

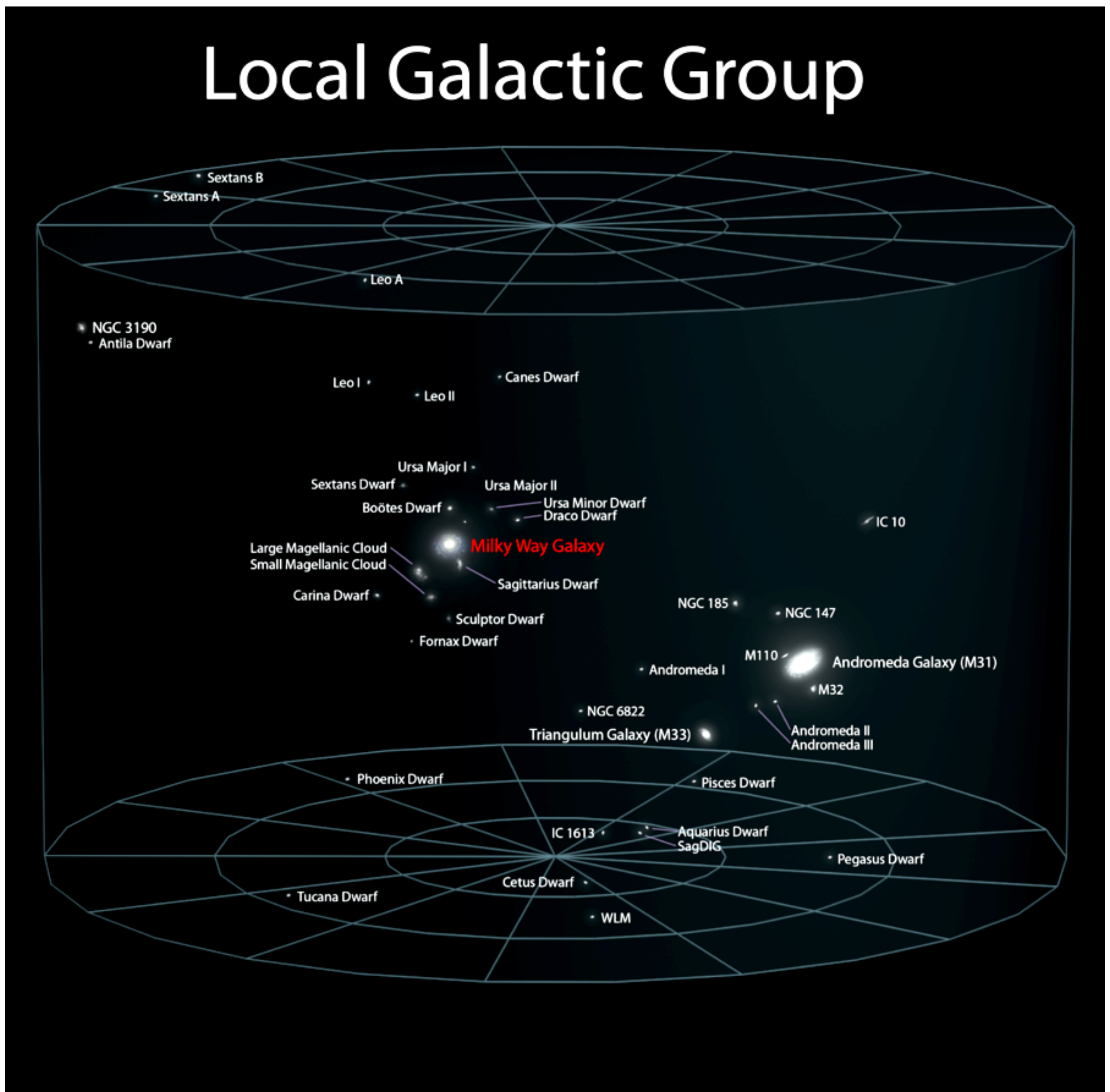
Power Spectrum (~Fourier Transform of Correlation Function) sensitive to cosmological parameters

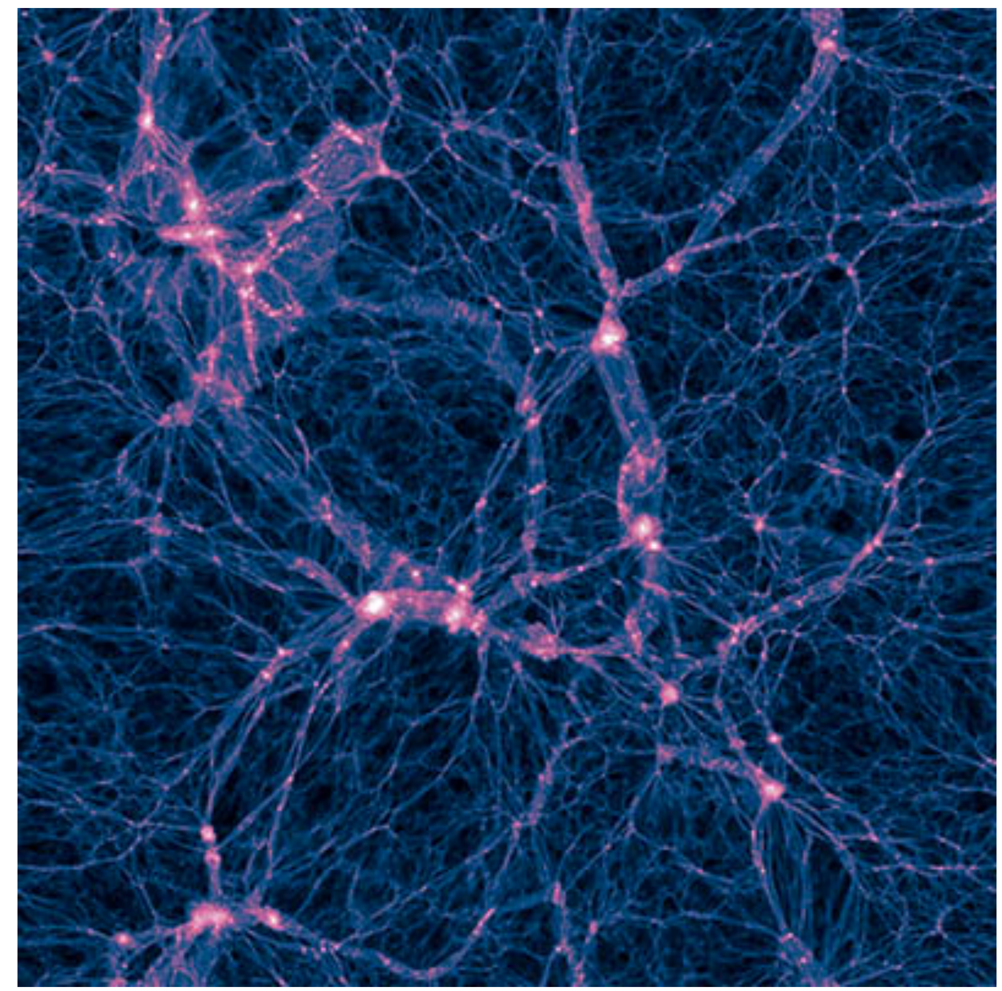


Astrostatistics Case Study:

Bayesian estimates of the Milky Way and Andromeda masses using high-precision astrometry and cosmological simulations

(Patel et al. 2017, arXiv:1703.05767)



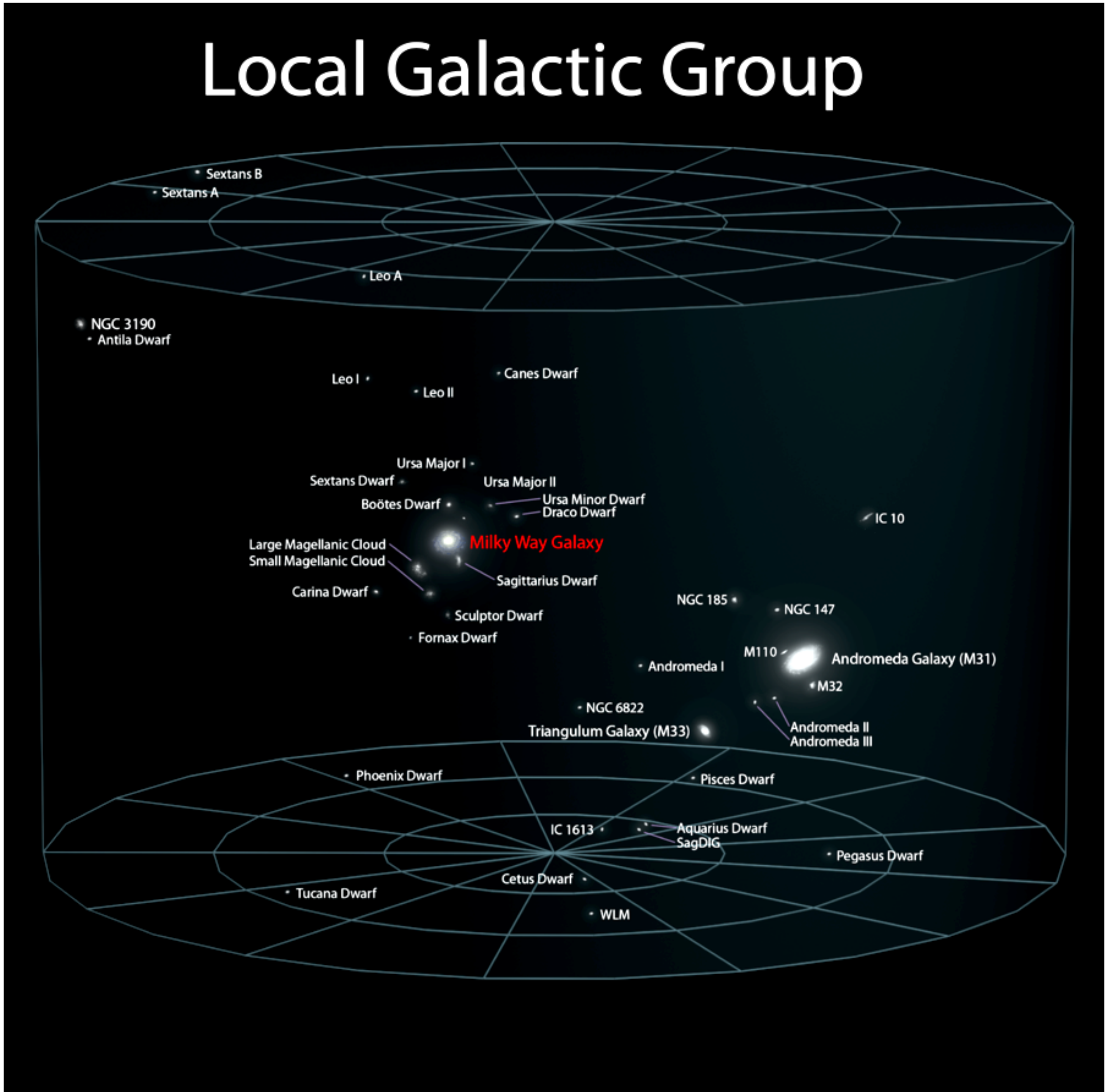


Illustris
Cosmological Simulation of Galaxy Formation

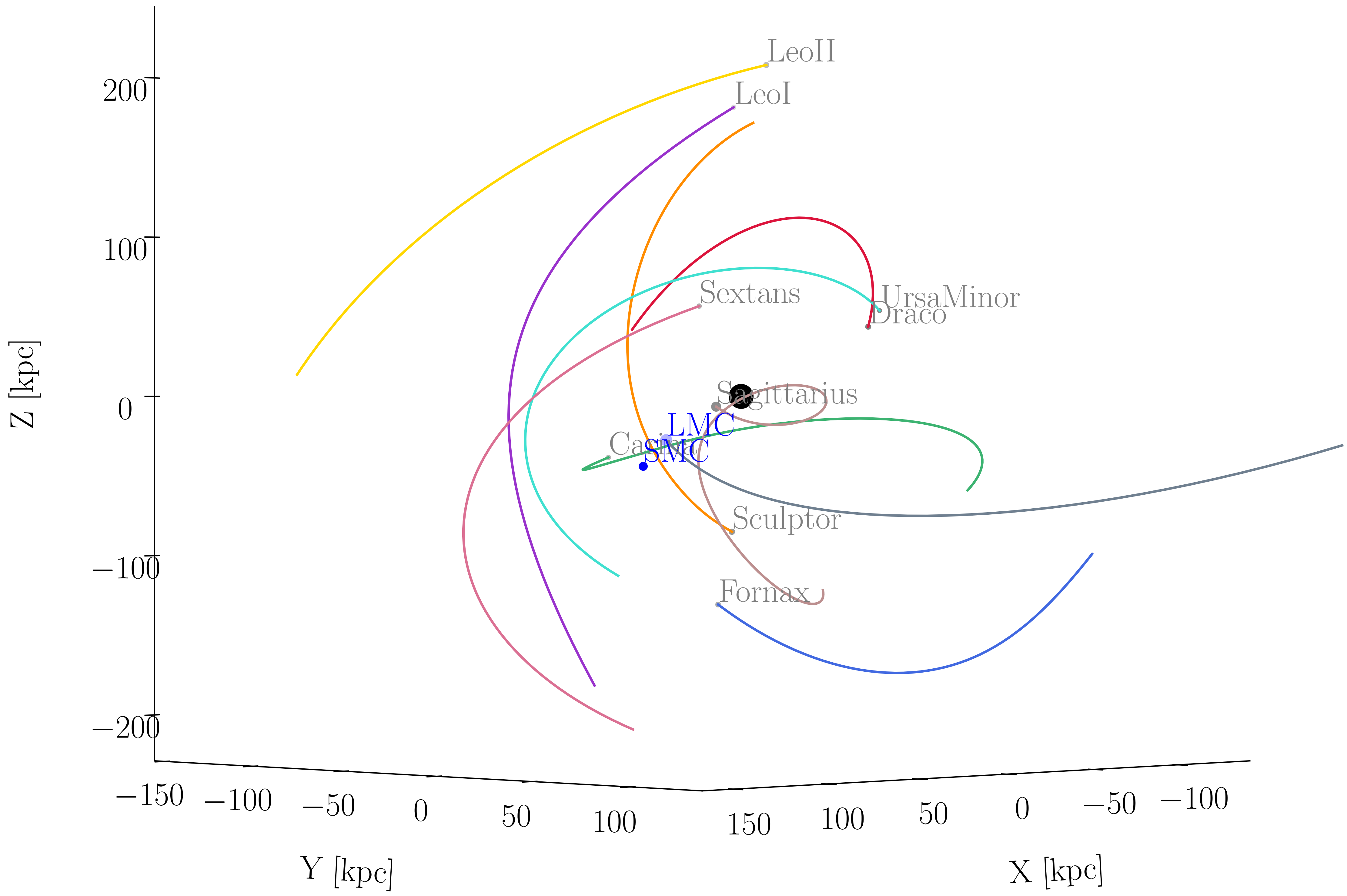
Illustris Cosmological Simulation
Movie

http://www.illustris-project.org/movies/illustris_movie_cube_sub_frame.mp4

Milky Way has satellite galaxies



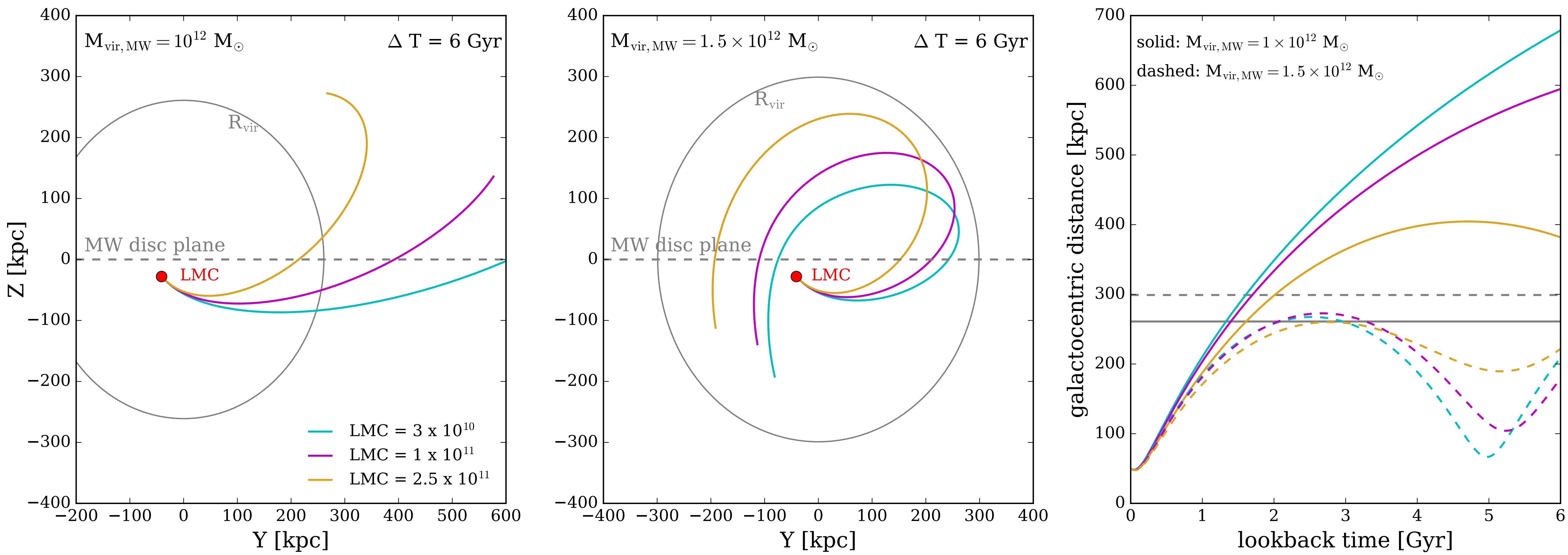
They are moving around



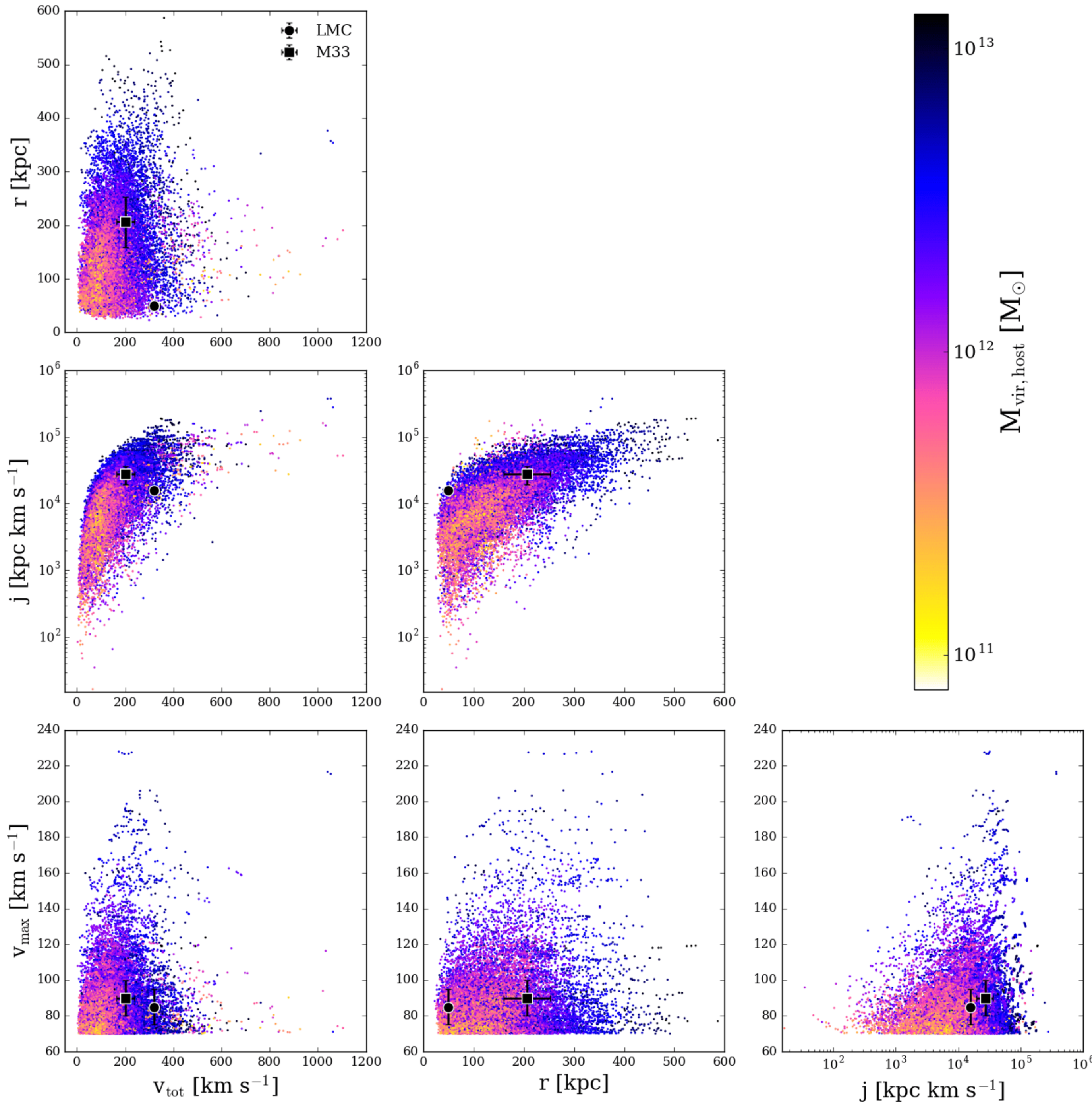
Their trajectories dependent on the Milky Way Mass

Low MW Mass

High MW Mass



Velocities (v), positions (r), momenta (j), of satellites are correlated with mass via galaxy formation physics in simulations (Prior)



x = latent (true) values of v, r, j

M_{vir} = Mass of Galaxy

Parameters are:

$\theta = (x, M_{vir})$

We can measure the (v, r, j) of MW's biggest satellite, Large Magellanic Cloud (LMC)

Table 1. Observational data ($\boldsymbol{d}$) for the LMC and M33 used to build likelihoods in the Bayesian inference scheme include the maximum circular velocity, current separation from the host galaxy and total velocity relative to the host galaxy.

	LMC μ	LMC σ	M33 μ	M33 σ
$v_{\rm max}^{\rm obs}$ (km s^{-1})	85[a]	10	90[b]	10
$r^{\rm obs}$ (kpc)	50	5	203	47
$v_{\rm tot}^{\rm obs}$ (km s^{-1})	321	24	202	38
$j^{\rm obs}$ (kpc km s^{-1})	15 688	1788	27 656	8219

Notes. [a]The maximal circular velocity of the LMC's halo rotation curve is adopted from Besla et al. (2012).
[b]M33's halo rotation curve maximum is duplicated from van der Marel et al. (2012b).
M33's position, velocity and their errors are adopted from Paper I (table 1), and references within.

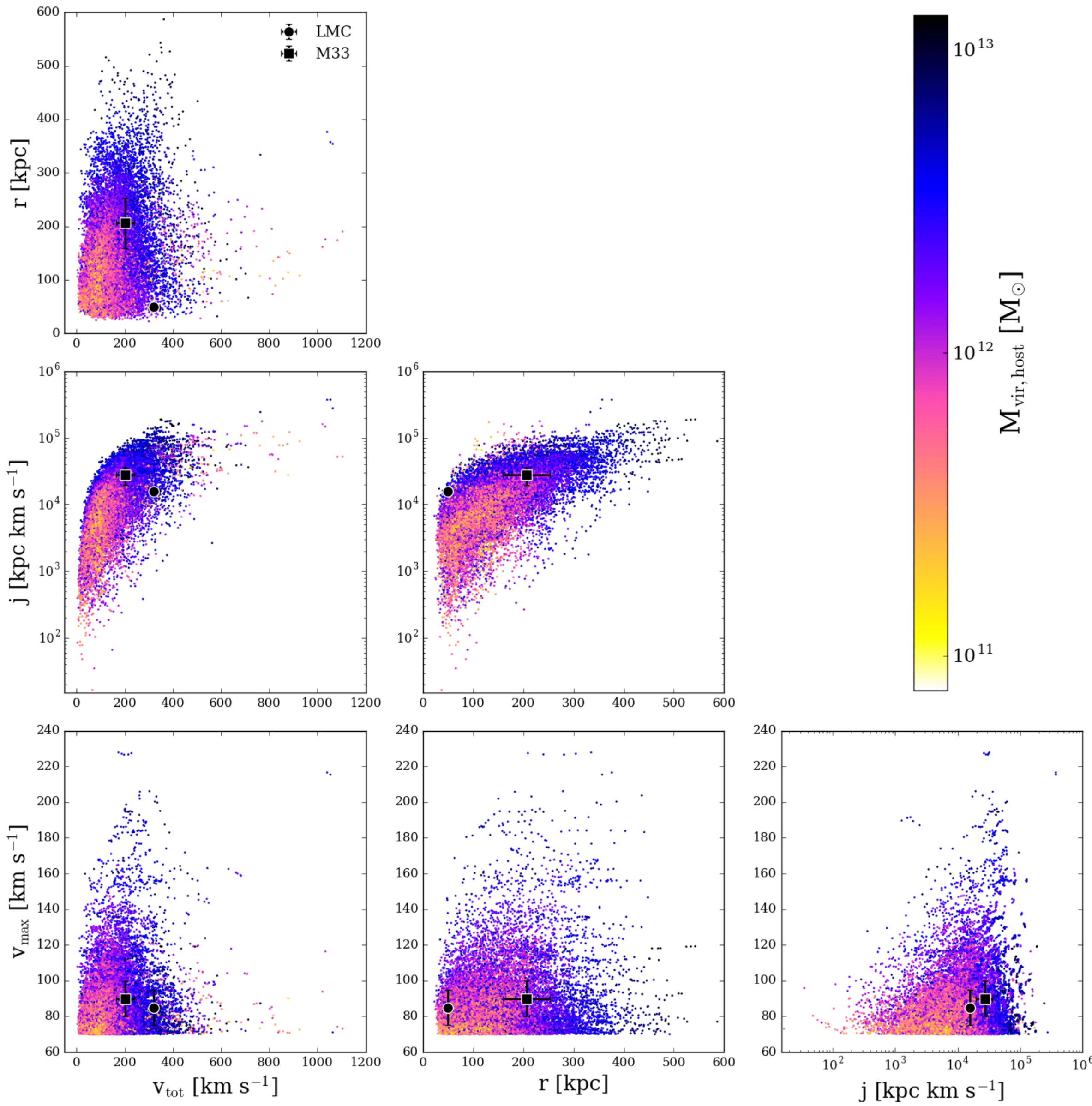
$$\mathcal{L}(\boldsymbol{x}\,|\,\boldsymbol{d}) = N(v_{\rm max}^{\rm obs}\,|\,v_{\rm max}, \sigma_v^2) \times N(r^{\rm obs}|r, \sigma_r^2) \times N(v^{\rm obs}|v_{\rm tot}, \sigma_{\rm v}^2), \quad (8)$$

where

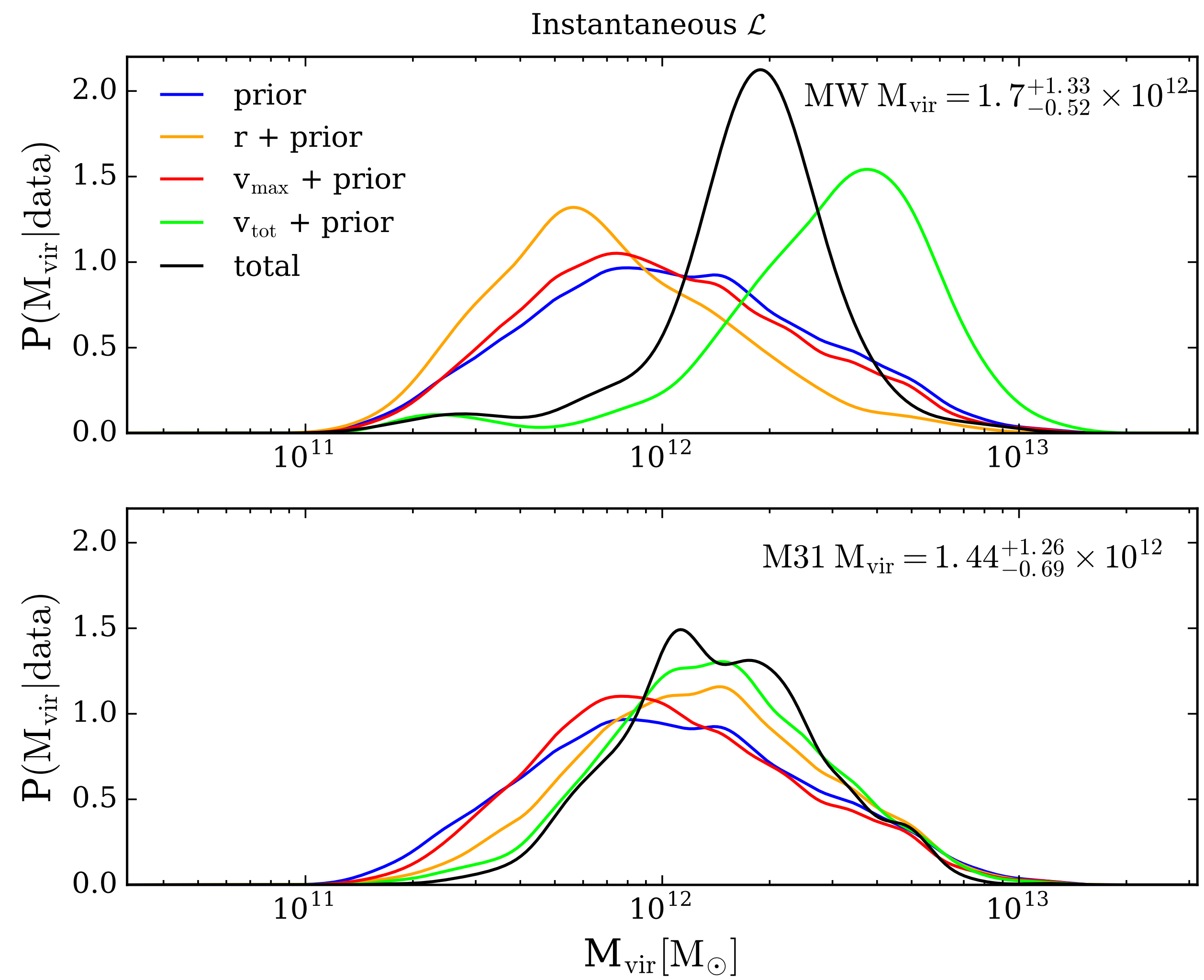
$$N(y|\,\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y-\mu)^2}{2\sigma^2}\right] \quad (9)$$

Astrostatistics Case Study:

Bayesian estimates of the Milky Way and Andromeda masses using high-precision astrometry and cosmological simulations (Patel et al. 2017, arXiv:1703.05767)



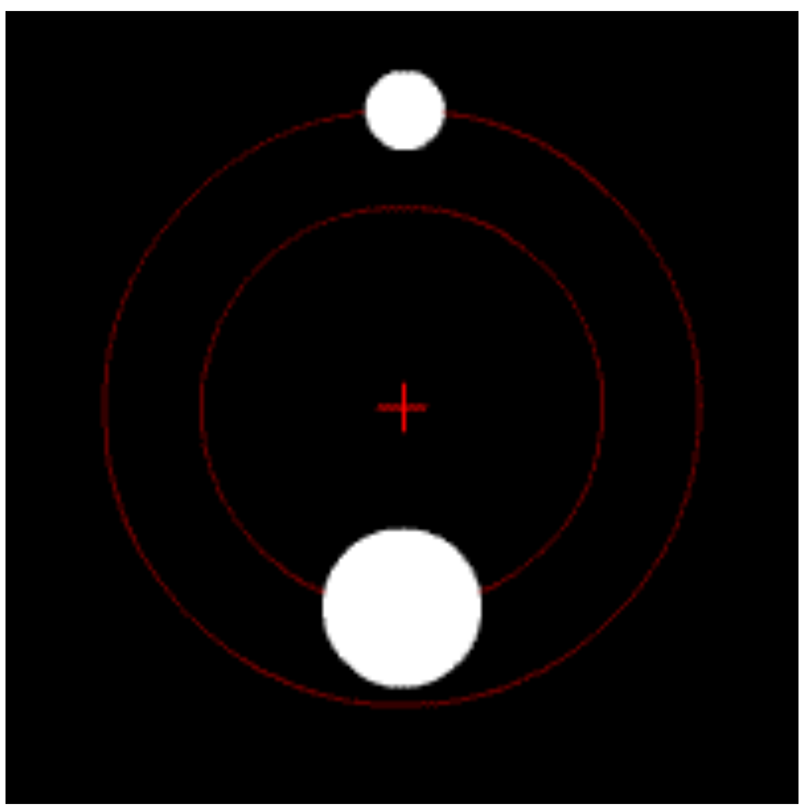
Simulation —> Prior

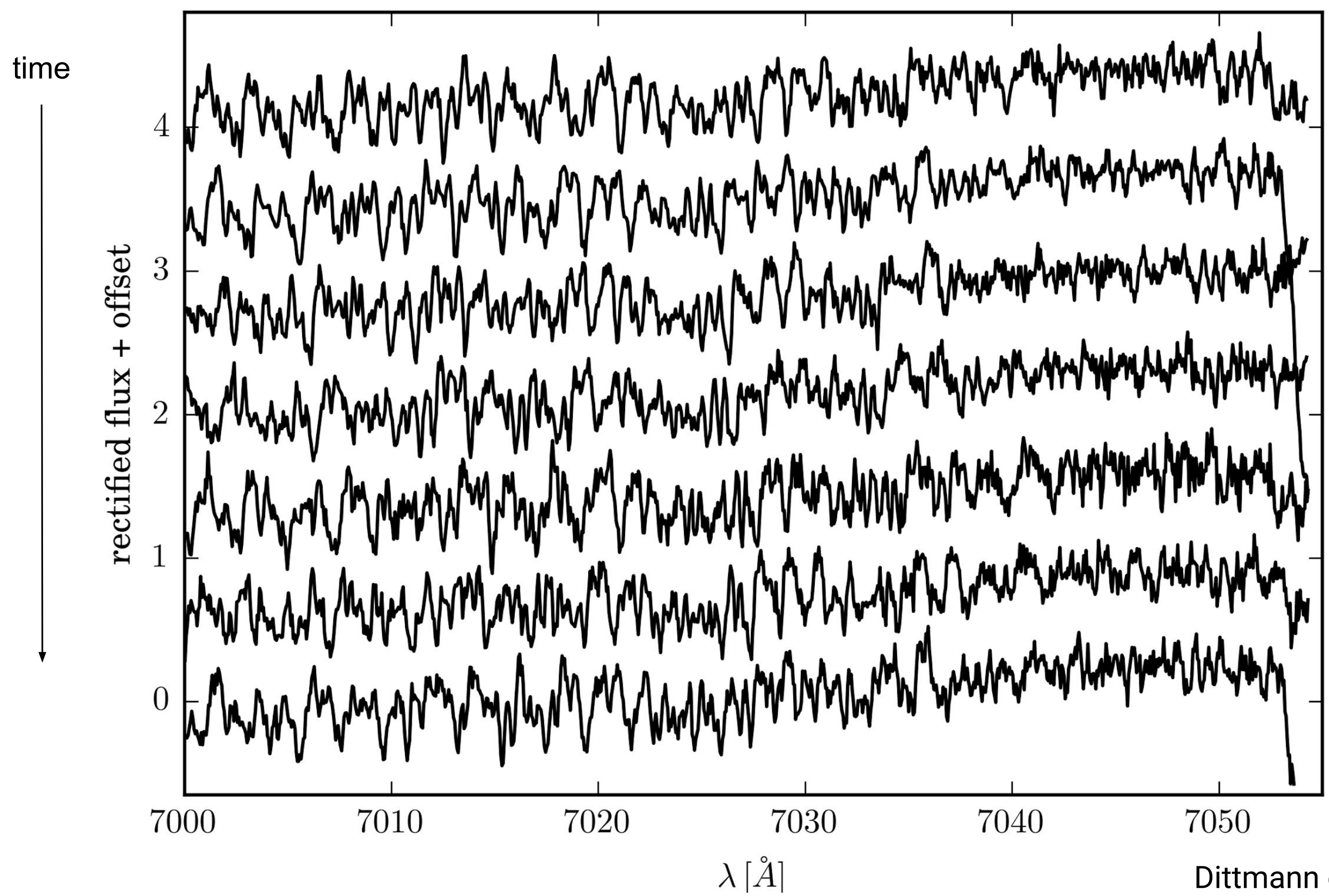


- Bayesian Inference
- Importance Sampling
- Kernel Density Estimation

Astrostatistics Case Studies: Disentangling Time Series Spectra with Gaussian Processes: Applications to Radial Velocity Analysis (Czekala et al. 2017, ApJ, 840, 49. arXiv:1702.05652)

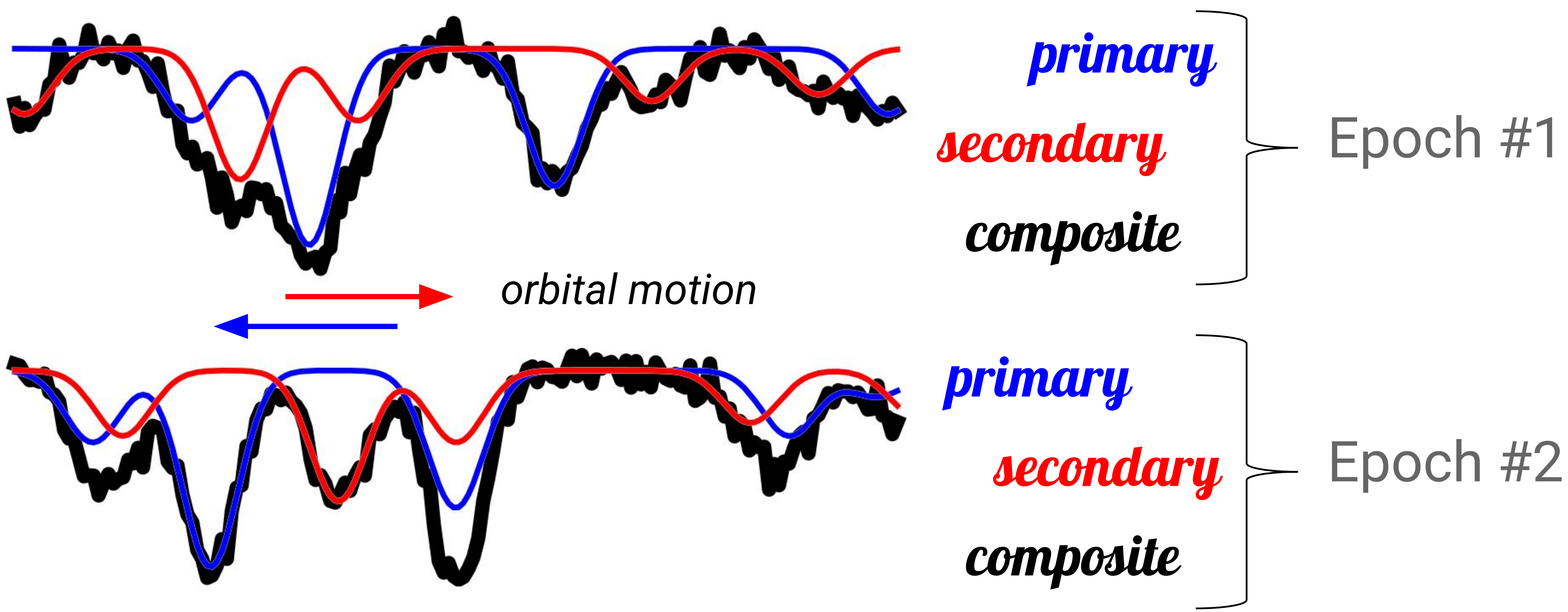
Raw Observations of the LP661-13 M4 Binary





Dittmann et al. 17
Czekala et al. 17a

Spectroscopic Binary Stars

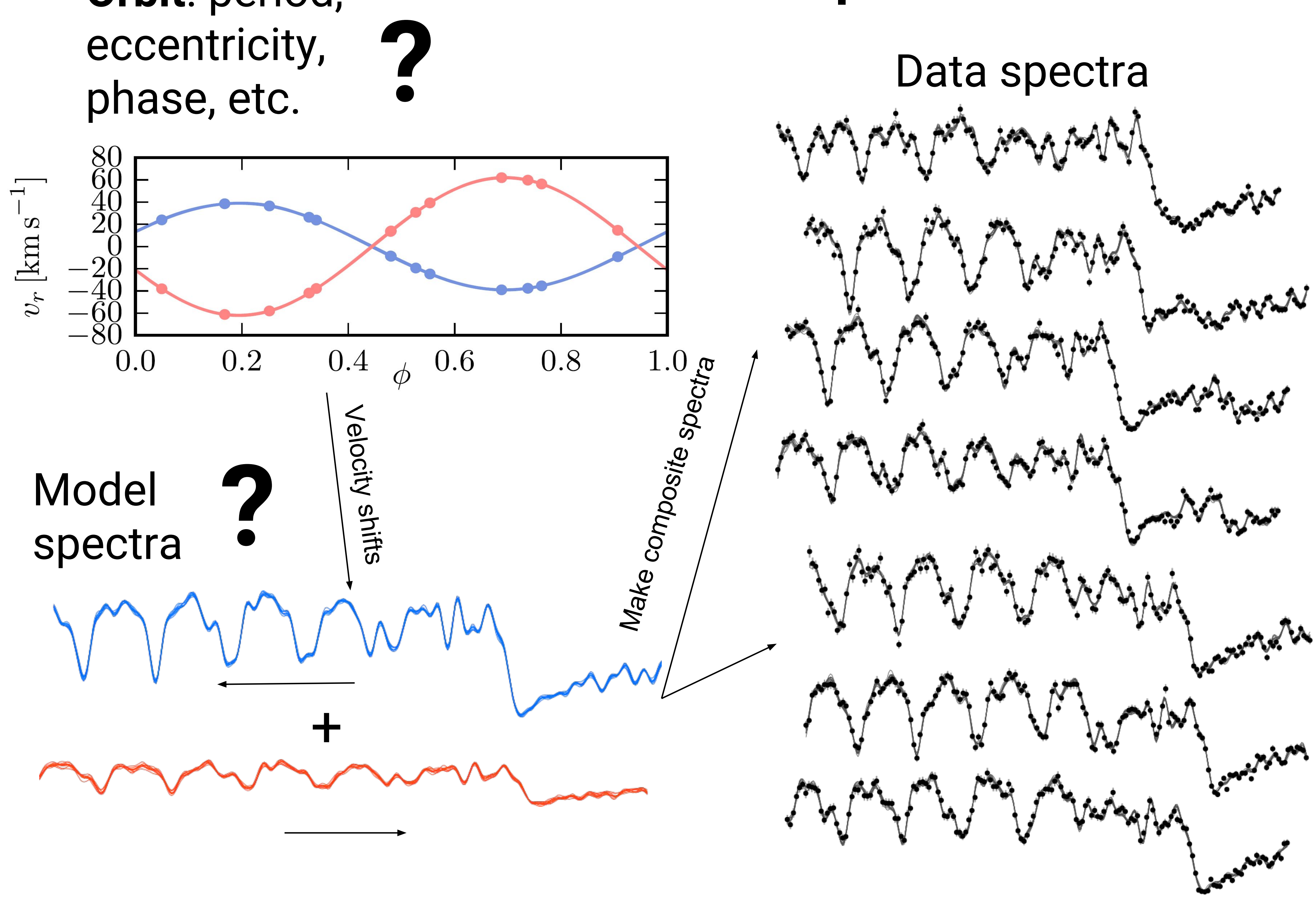


(Epoch = Time)

We only observe the "noisy" sum of two (latent) spectra.
Latent (underlying) spectra are unknown functions
Observed spectrum = Measured Data

Forward Model = Generates Data

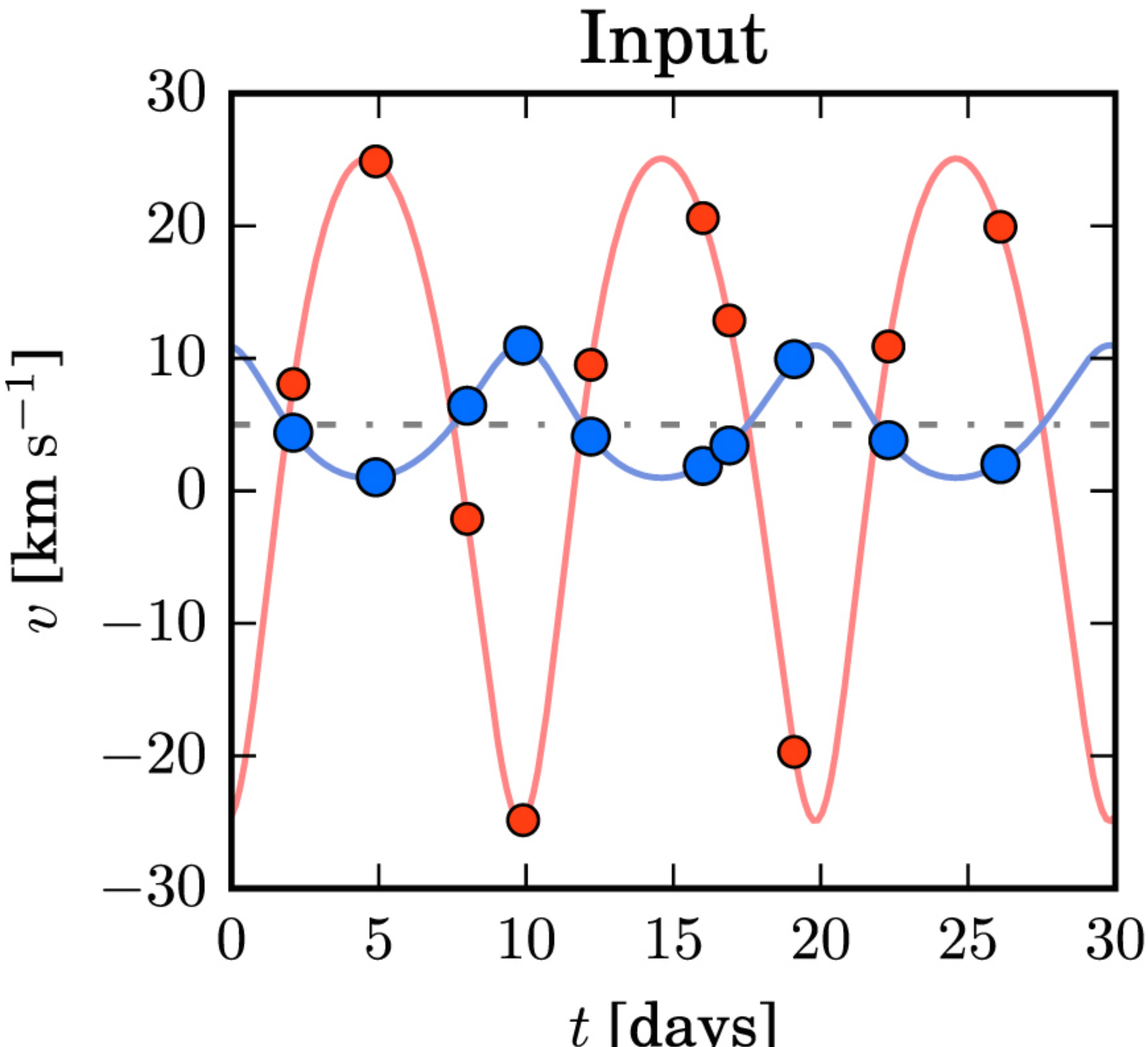
Problem setup



https://www.youtube.com/watch?v=kHjN42ft6aU

Goal: Go Backwards and Infer the Component Spectra & Orbital Parameters from noisy, observed (composite) spectra time series

Orbital Parametric Model



Seven Parameters:

- Mass Ratio
- Velocity Amplitude
- eccentricity
- Arg of Periastron
- Epoch of Periastron
- Orbital Period
- Systemic Velocity

Nonparametric Bayes

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***

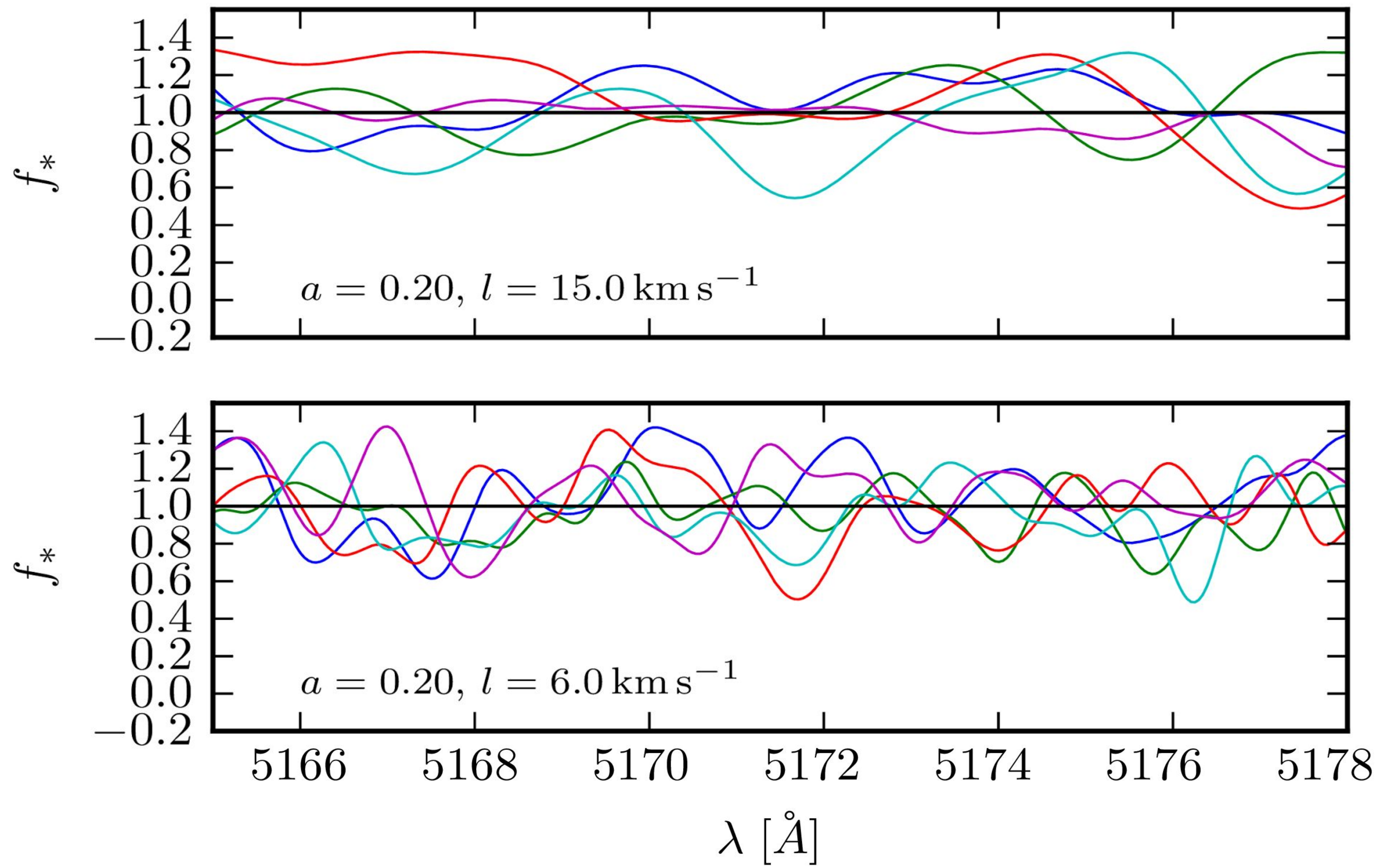
For a covariance kernel, we will use the commonly used squared exponential kernel, which relates pixels in the spectrum based upon their distance in log-wavelength (∝ velocity)

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

Gaussian Process = a prior on functions (latent spectra)

Gaussian Process model for a single, stationary star

(Zoomed) draws from the prior



l

l

Inference = Which function is most consistent with the data?

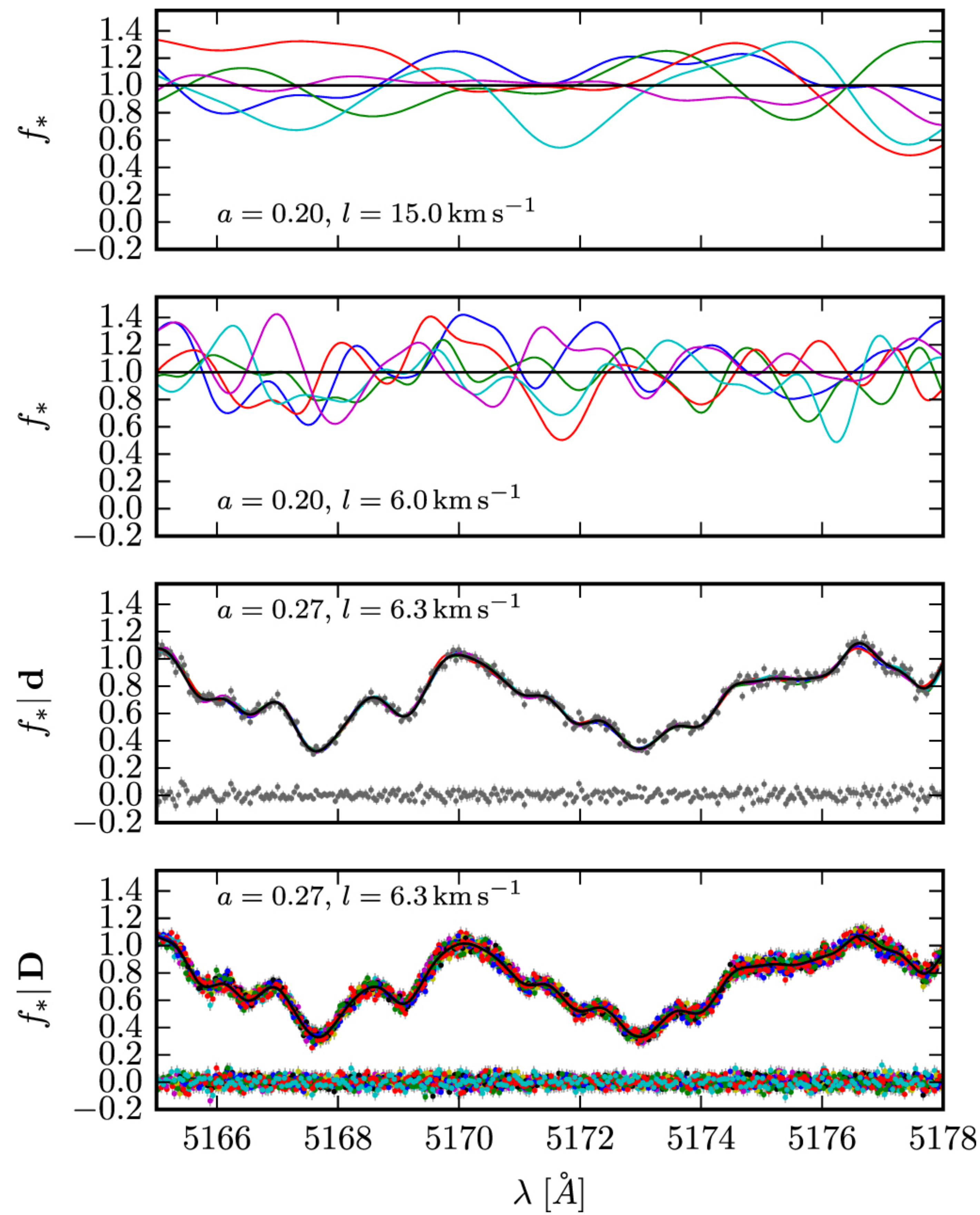
Gaussian Process: Priors & Posteriors

GP prior
(long length scale)

(short length scale)

GP Posterior
(conditioned on data spectrum d)

Inference of latent spectrum



Known Unknowns

$$\Theta$$

7-dim Orbital Parameters = Period, Phase, eccentricity, Velocity Amplitude

$$f(\lambda), g(\lambda)$$

(∞-dim) Latent Functions = the unobserved component spectra of the primary (f) and secondary (g) stars

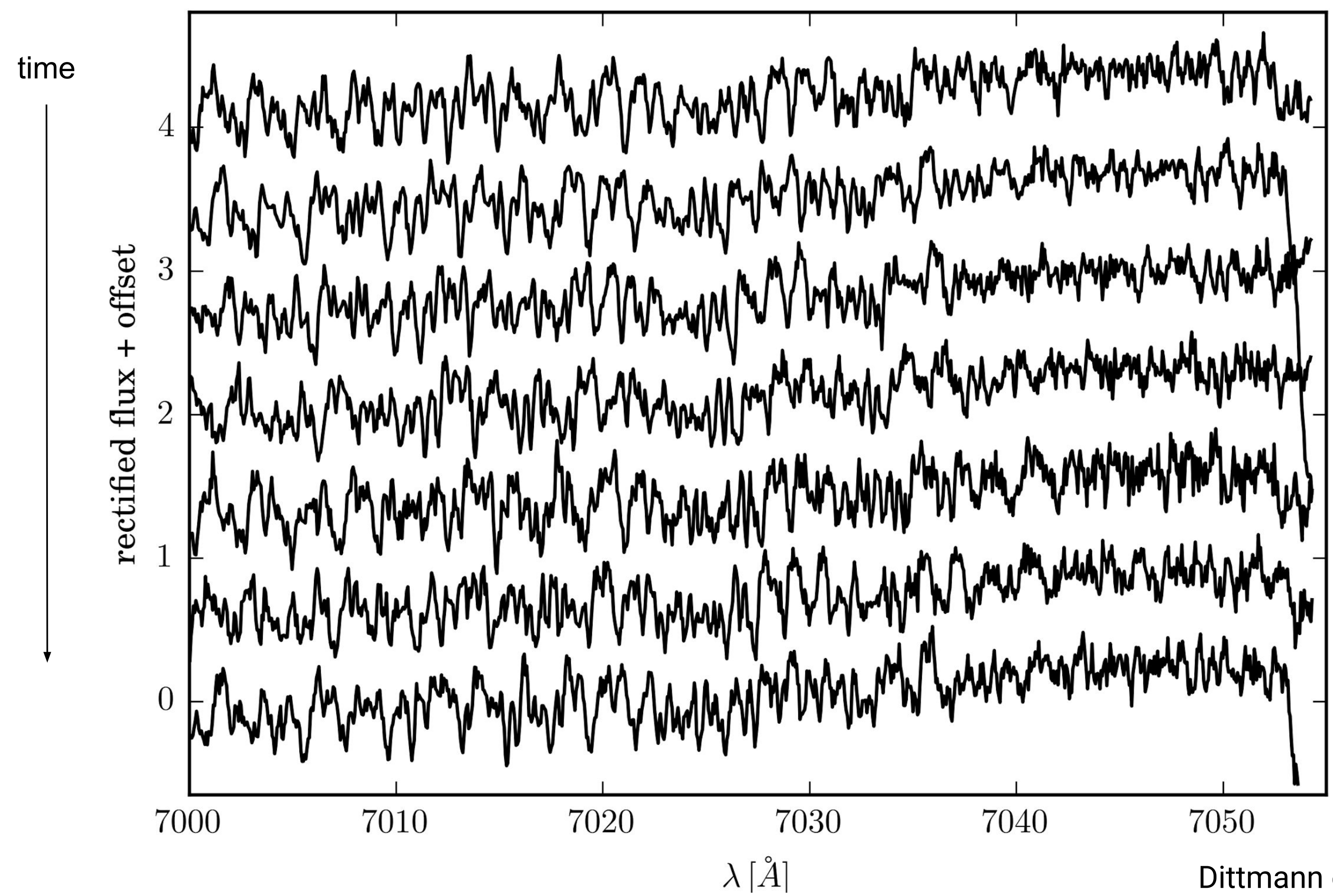
$$\boldsymbol{\alpha} = (a_f, l_f, a_g, l_g)$$

4-dim GP hyperparameters = controlling the amplitude and smoothness of Gaussian Process prior on latent spectra

Knowns (Data)

Raw Observations of the LP661-13 M4 Binary

$$\boldsymbol{D} =$$



Dittmann et al. 17
Czekala et al. 17a

Bayesian Inference

$$P(A|\,B) = \frac{P(B|A)P(A)}{P(B)}$$

In this case:

$$P(\boldsymbol{\Theta}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\alpha}|\boldsymbol{D}) \propto$$
$$P(\boldsymbol{D}|\,\boldsymbol{\Theta}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\alpha}) \times P(\boldsymbol{\Theta}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\alpha})$$

a probability density on (4+7+∞)-dim parameter space

Bayesian Computation

1. Run Markov Chain Monte Carlo (MCMC)
   (e.g. *emcee* affine-invariant ensemble sampler)
   on the 4+7 small dimensional marginal posterior

$$P(\boldsymbol{\Theta}, \boldsymbol{\alpha} | \boldsymbol{D}) = \int d\boldsymbol{f} \int d\boldsymbol{g} \, P(\boldsymbol{\Theta}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\alpha} | \boldsymbol{D})$$

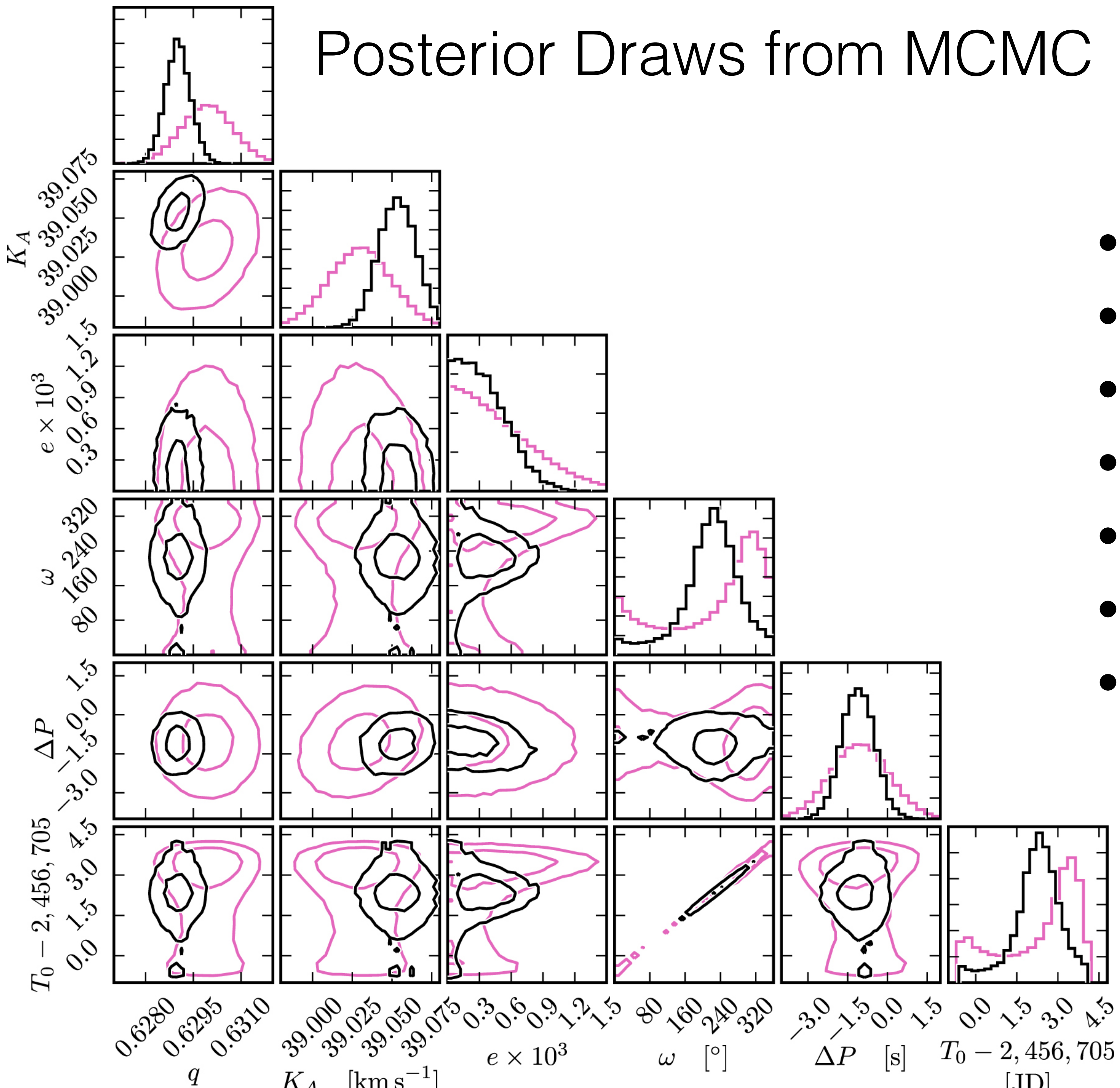
MCMC generates samples: $\boldsymbol{\Theta}_i, \boldsymbol{\alpha}_i \sim P(\boldsymbol{\Theta}, \boldsymbol{\alpha} | \boldsymbol{D})$

2. Draw high-dim (**f**, **g**) spectra from the posterior predictive distribution

$$\boldsymbol{f}_i, \boldsymbol{g}_i \sim P(\boldsymbol{f}, \boldsymbol{g} | \boldsymbol{\Theta}_i, \boldsymbol{\alpha}_i, \boldsymbol{D})$$

Application to the Mid-M-Dwarf Binary LP661-13

Posterior Draws from MCMC



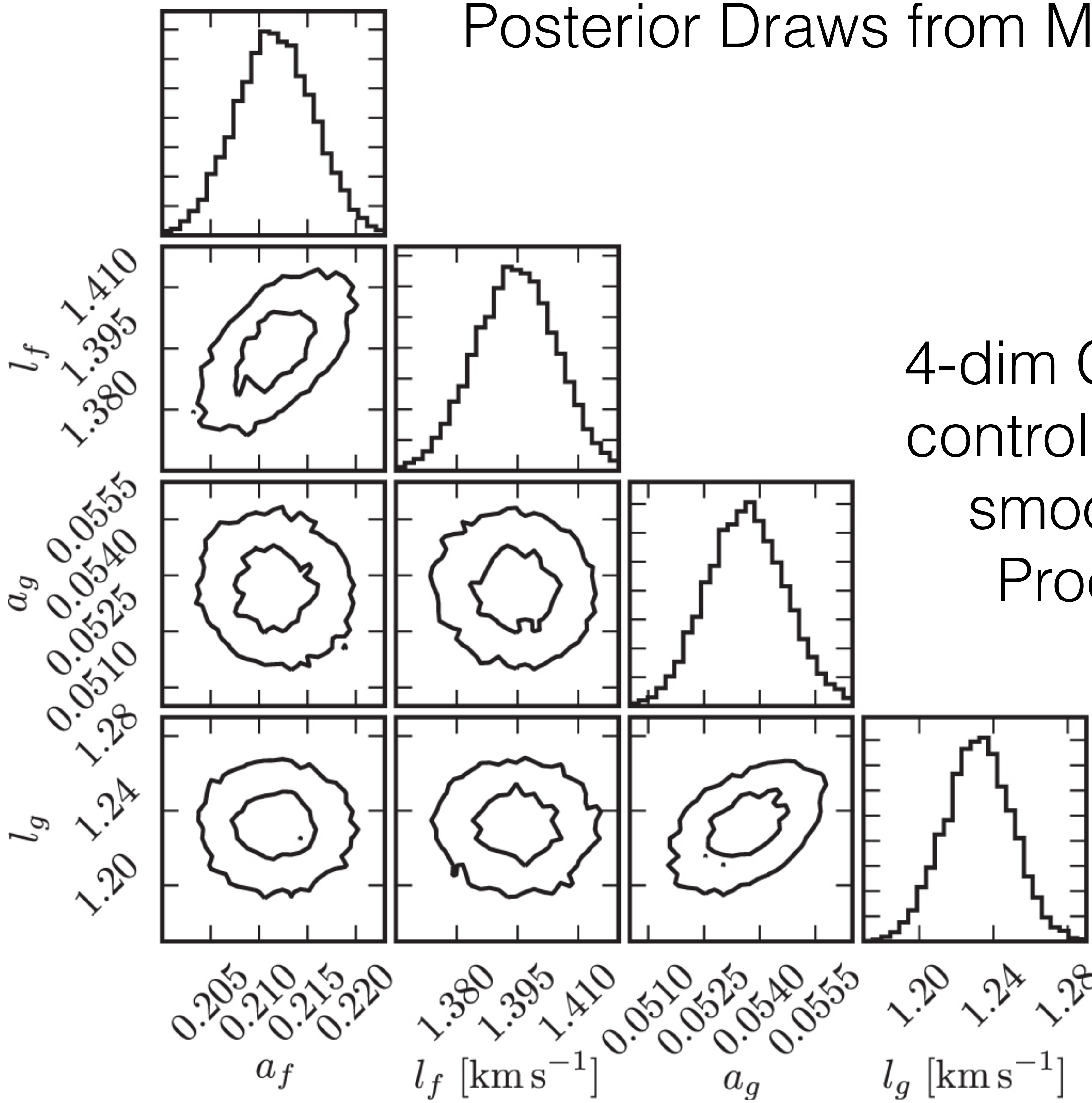
Seven Orbital Parameters:

- Mass Ratio
- Velocity Amplitude
- eccentricity
- Arg of Periastron
- Epoch of Periastron
- Orbital Period
- Systemic Velocity

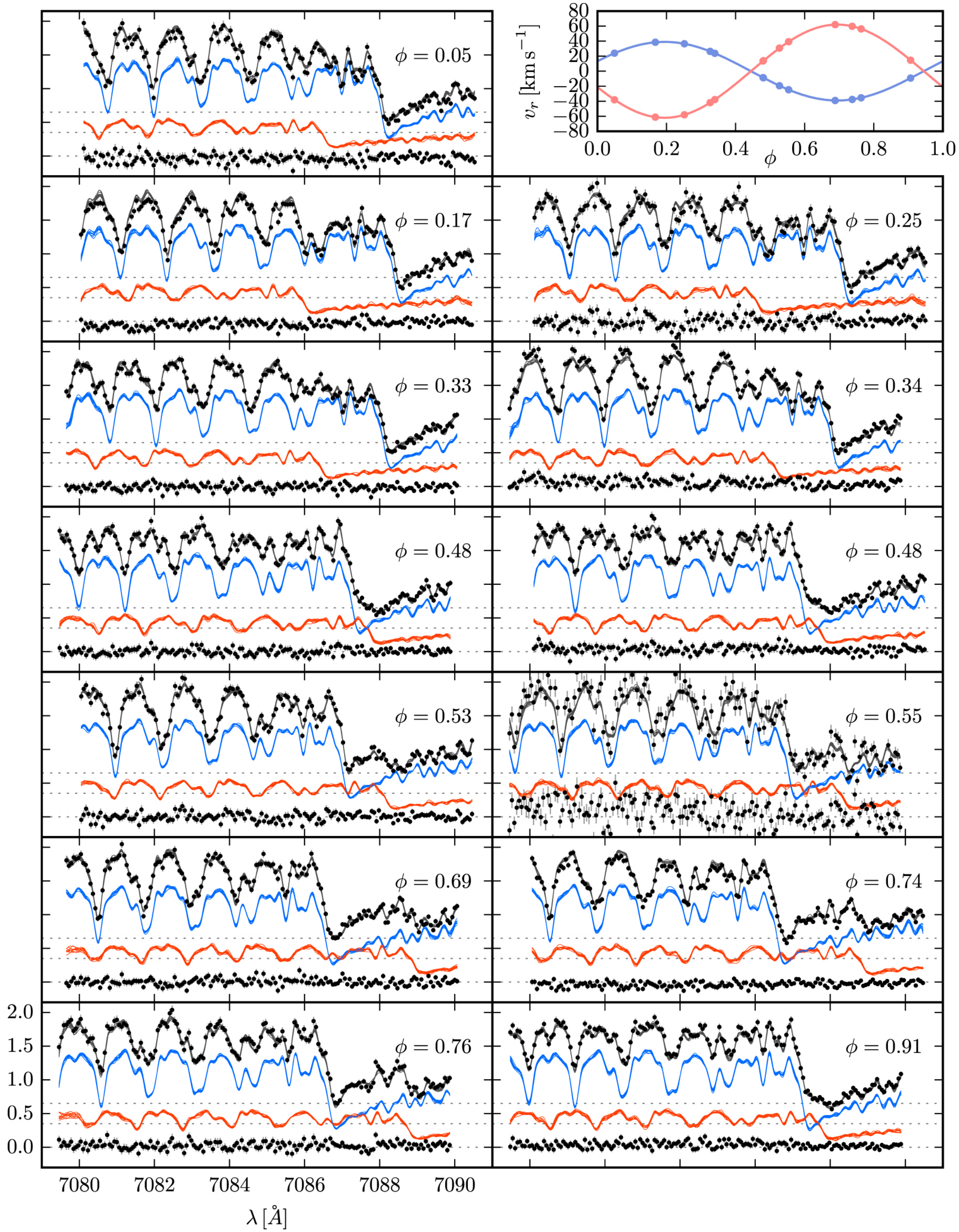
(Purple: Conventional Analysis)

Application to the Mid-M-Dwarf Binary LP661-13

Posterior Draws from MCMC $\boldsymbol{\alpha} = (a_f, l_f, a_g, l_g)$



4-dim GP hyperparameters = controlling the amplitude and smoothness of Gaussian Process prior on latent spectra



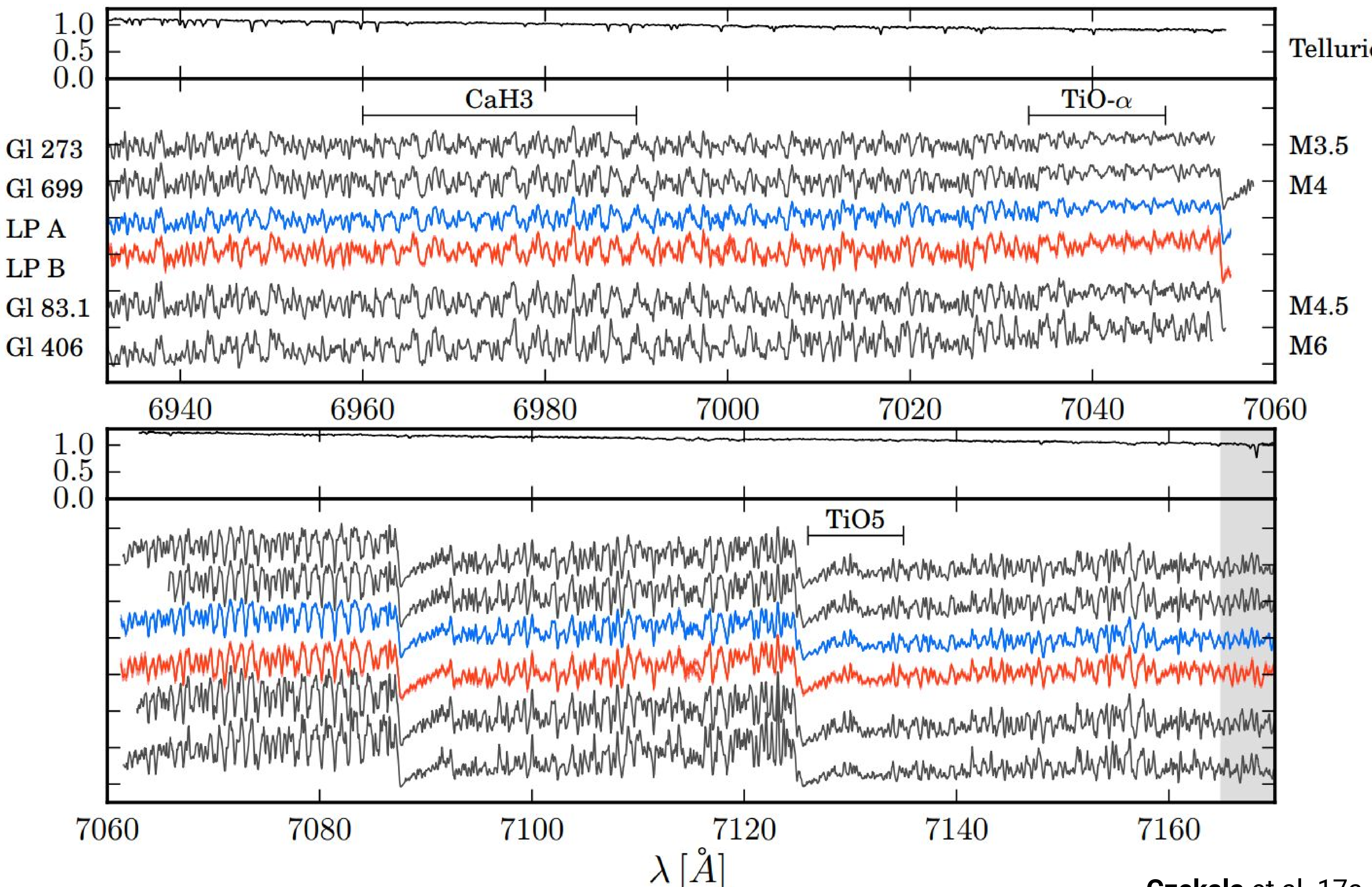
Posterior Inference of Component Spectra (**f**, **g**)

Compared to 10 epochs of observed spectra (**data**)

Model Checking! Checking Fit against Data

Model Checking!
Checking Fit against Domain Knowledge (astrophysics)!

Disentangled spectra match other single standard stars



Czekala et al. 17a

Astrostatistics Case Study: Disentangling Time Series Spectra with Gaussian Processes: Applications to Radial Velocity Analysis (Czekala et al. 2017, arXiv:1702.05652)

http://psoap.readthedocs.io/en/latest/

- Statistics:
  - Parametric Modelling (Stellar Orbit Parameters)
  - Nonparametric Modelling (Gaussian Process Spectrum)
  - Bayesian Inference (probability of unknowns given data)
  - Markov Chain Monte Carlo (computing posterior probability)
- Astronomy:
  - Applications to Radial Velocity Analysis of Stars/Exoplanets