astro-ph seminar 2025.06.18 (Ryou Ohsawa)

# gallifrey: JAX-based Gaussian process structure learning for astronomical time series

C. Boettner, 2025, accepted for publication in A&A (arXiv:2505.20394)


*Context.* Gaussian processes (GPs) have become a common tool in astronomy for analysing time series data, particularly in exoplanet science and stellar astrophysics. However, choosing the appropriate covariance structure for a GP model remains a challenge in many situations, limiting model flexibility and performance.
*Aims.* This work provides an introduction to recent advances in GP structure learning methods, which enable the automated discovery of optimal GP kernels directly from the data, with the aim of making these methods more accessible to the astronomical community.
*Methods.* We present `gallifrey`, a JAX-based Python package that implements a sequential Monte Carlo algorithm for Bayesian kernel structure learning. This approach defines a prior distribution over kernel structures and hyperparameters, and efficiently samples the GP posterior distribution using a novel involutive Markov chain Monte Carlo procedure.
*Results.* We applied `gallifrey` to common astronomical time series tasks, including stellar variability modelling, exoplanet transit modelling, and transmission spectroscopy. We show that this methodology can accurately interpolate and extrapolate stellar variability, recover transit parameters with robust uncertainties, and derive transmission spectra by effectively separating the background from the transit signal. When compared with traditional fixed-kernel approaches, we show that structure learning has advantages in terms of accuracy and uncertainty estimation.
*Conclusions.* Structure learning can enhance the performance of GP regression for astronomical time series modelling. We discuss a road map for algorithmic improvements in terms of scalability to larger datasets, so that the methods presented here can be applied to future stellar and exoplanet missions such as PLATO.



**Key words.** Asteroseismology – Methods: data analysis – Methods: statistical – Planets and satellites: detection – Techniques: photometric – Techniques: spectroscopic


## Background

Gaussian Process has been widely used to approximate a sequence of measurements.

The "smoothness" of the sequence is controlled via a kernel (correlation).

All the expectations can be provided from a Normal distribution when the kernel is fixed.

$$f(x) \sim \mathcal{GP}(0, k(x, x')) \longrightarrow \boldsymbol{f}(\boldsymbol{x}) \sim \underline{\mathcal{N}\left(\mathbf{0}, K(\boldsymbol{x}, \boldsymbol{x})\right)}$$

Gaussian (normal) distribution

Kernel function

$$P\left(\underline{\boldsymbol{y}^*} | \underline{\boldsymbol{x}^*}, \underline{\boldsymbol{x}}, \underline{\boldsymbol{y}}\right) = \mathcal{N}(\underline{\boldsymbol{\mu}}, \underline{\boldsymbol{\Sigma}} + \underline{\sigma_n^2} \boldsymbol{I}).$$

observed data

new coordinates

expected values

Kernel function

observed uncertainties

GP provides a useful Bayesian solution, but choosing an appropriate kernel function remains a difficult task.

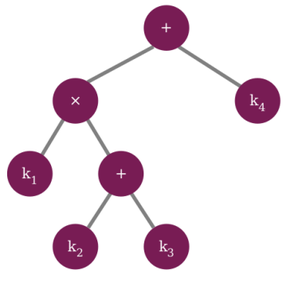



Fig. 1: Example of a kernel representation

## Key point

Saad et al. (2023) proposed a novel method to optimize a GP kernel, where they introduce a binary tree (Fig. 1) to represent the structure of a kernel. A kernel is procedurally sampled from a prior distribution.

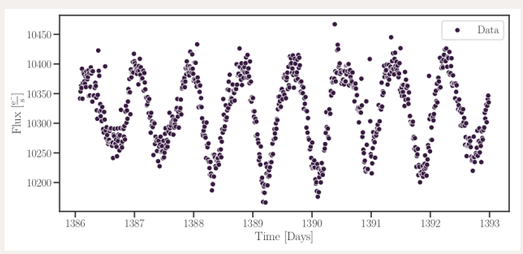
They use a hybrid approach in optimization:

Kernel structure: Sqeuential Monte Carlo (population-based Monte Carlo method)
Kernel hyperparameters: Hamiltonian Monte Carlo (high performance MCMC)

The kernel structure and GP parameters are simultaneously optimized in a Bayesian framework.

Boettener implemented Saad's algorithm in Python and presented some astrophysical applications.

a transit in a PLATO-like light curve / a transit on a varying baseline / transmission spectroscopy of HATS-46 b



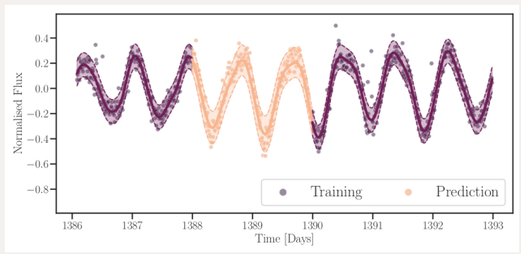




Fig 2&3: Example of Gaussian Process. (left) A light curve of TIC 10863087 observed by TESS. (right) An example of interpolation with GP. The orange data are masked and interpolated using the remaining data. The uncertainties are indicated by the shaded area.

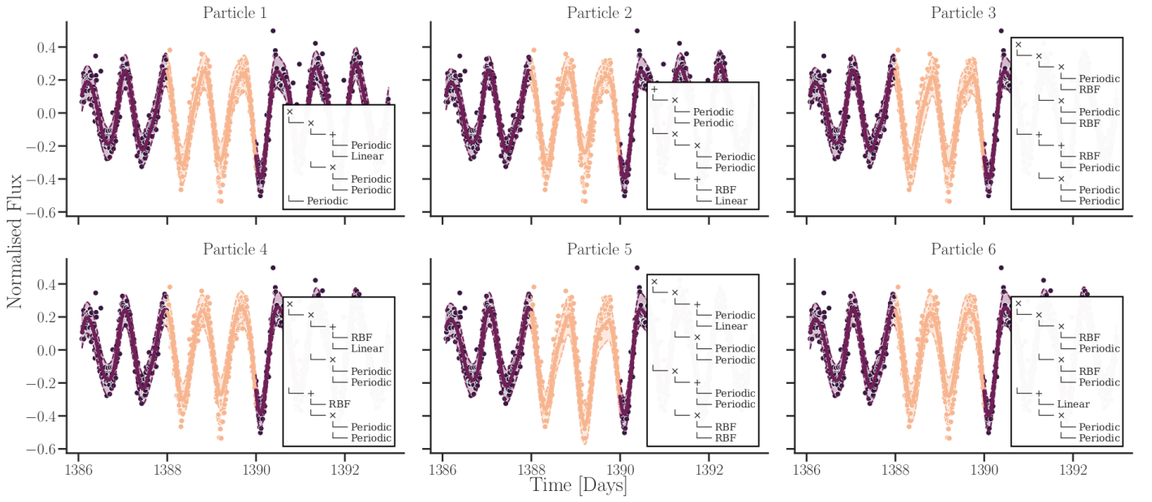



Fig. 5: Sample of predictions for individual particles from the final SMC ensemble used for Fig. 3, including their respective kernel structures. Particles exhibit a variety of kernel structures and produce varying predictions, which leads to a more robust overall prediction when combined. Note that while the depth of every kernel displayed here corresponds to the maximum tree depth, $D_{\max} = 3$, the learned kernels can end up shallower if simpler kernel structures are preferred.

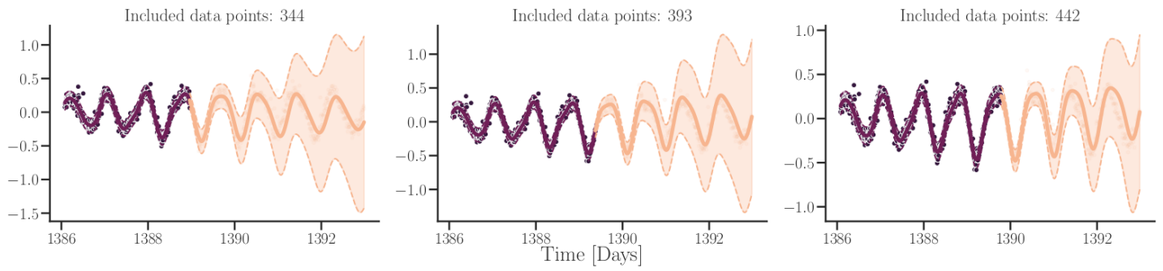
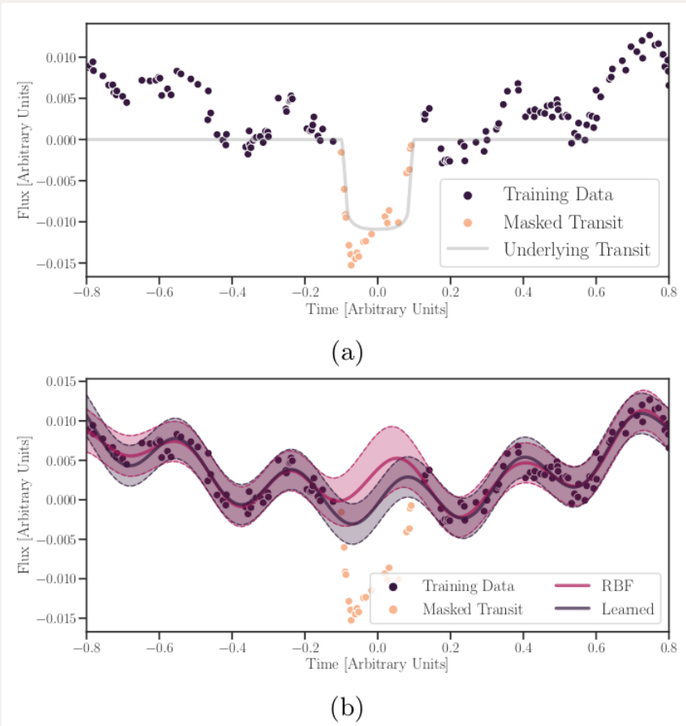



Fig. 6: Snapshots of SMC rounds during the structure learning process for producing the forecast shown in Fig. 4. Purple data points have already been seen during the annealing schedule, while yellow regions are predictions. The algorithm learns more intricate patterns and reduces prediction uncertainty as more data points are added.



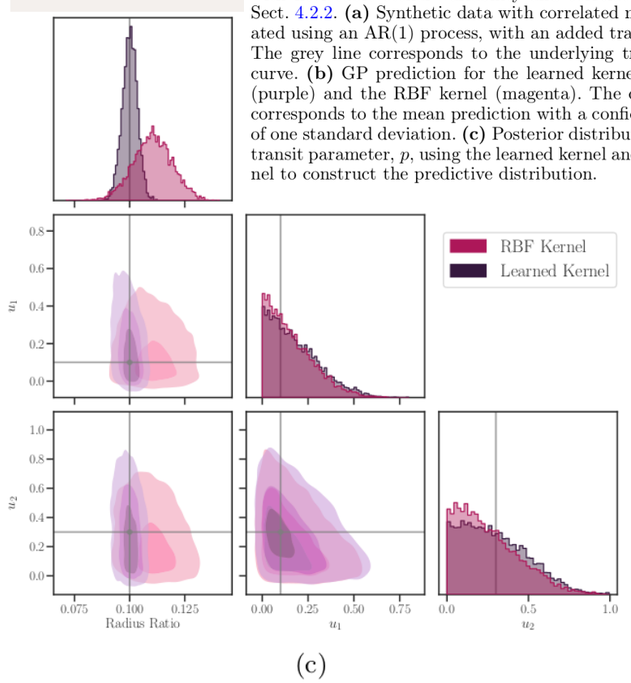




Fig. 8: Transit parameter fit for the synthetic light curve with noticeable stellar variability and correlated noise in Sect. 4.2.2. **(a)** Synthetic data with correlated noise generated using an AR(1) process, with an added transit signal. The grey line corresponds to the underlying transit light curve. **(b)** GP prediction for the learned kernel structure (purple) and the RBF kernel (magenta). The central line corresponds to the mean prediction with a confidence band of one standard deviation. **(c)** Posterior distribution for the transit parameter, $p$, using the learned kernel and RBF kernel to construct the predictive distribution.