Draft version November 6, 2017
Typeset using LaTeX default style in AASTeX61

GAUSSIAN PROCESS MODELING: A BAYESIAN APPROACH TO DETECTING QUASI-PERIODIC OSCILLATIONS IN SOLAR FLARE DATA

Christopher A. Ick,[1] David W. Hogg,[2] and Daniela Huppenkothen[3]

[1] *New York University*
726 Broadway, Ofc 912
New York, NY 10003, USA
[2] *New York University*
726 Broadway, Ofc 1005
New York, NY 10003, USA
[3] *University of Washington*
3910 15th Ave NE
Office B356B
Seattle, WA 98195, USA

ABSTRACT

Solar flare analysis for quasi-periodic oscillations (QPOs) was formally done using Fourier power spectra analysis, limiting computational efficiency to $\mathcal{O}(n^2)$. We demonstrate a more effective and efficient method of modeling solar flare data using Gaussian Processes, a continuous-domain model, which scales at $\mathcal{O}(n)$ in one-dimensional models, and provide demonstrations of these methods by applying it to both simulated and real astronomical datasets. Using the package Celerite for Gaussian processes, we provide demonstrations of effectively modeling QPOs using a self-correllation function (a kernel function) to simulate solar flare data, and then to remodel and capture the original characteristics of the simulated flare data, using optimization and Markov-chain Monte Carlo sampling. After demonstrating the effectiveness of these fitting methods, we go on to apply these modeling/fitting methods to capture the characteristics of real solar flare data, and discuss the potential use of the algorithm for automated solar flare and QPO detection.

Keywords: keywords go here

Corresponding author: Is that me?
chris.ick@nyu.edu, david.hogg@nyu.edu, daniela.huppenkothen@nyu.edu

1. INTRODUCTION

Introduction goes here! Am I updating?

2. DATA

Data for this project was acquired from the Gamma-ray Burst Monitor (GBM) onboard the Fermi Gamma-ray Space Telescope, as well as the GOES-15 sattelite's x-ray imager. Solar flare data from GOES was provided by Andrew Inglis. The data is formatted in a .fits file, with thw first two columns corresponding to the flux of the 1-8Åand 0.5-4Åchannels respectively.

3. METHODS

Once our data is cleanly organized into a time series containing time, intensity, and error arrays, we can begin our modeling algorithm. We model the flare envelope using the flare model:

$$I(t) = A\lambda exp\left(\frac{-\tau_1}{t-t_s} - \frac{t-t_s}{\tau_2}\right) \tag{1}$$

where t is the time since trigger, t_s is the start tie, A is the pulse amplitude, τ_1 and τ_2 are characteristics of the pulse rise and decay respectively, and $\lambda = exp(2\frac{\tau_1}{\tau_2}]^{1/2}$ The free parameters for our fit are A, τ_1, and τ_2. We generate our initial guesses for these parameters using a simple function that takes the maximum of the flare as A, τ_1 to be $\frac{1}{3}$ of the length of our flare, and τ_2 to similarly be $\frac{2}{3}$ the length of our flare. To model a QPO, we use a covariance function, or kernel function $k(\tau)$, where τ_{ij} is the matrix $\tau ij = t_i - t_j$. To model a QPO, we utilize Celerite's built-in "terms" class to define a covariance function representing stochastically-driven, dampened harmonic oscillator, defined by the differential equation:

$$\left(\frac{d^2}{dt^2} + \frac{\omega_0}{Q}\frac{d}{dt} + \omega_0^2\right) y(t) = \epsilon(t) \tag{2}$$

Where ω_0 corresponds to the natural frequency of the undamped oscillator, Q is the quality factor, and $\epsilon(t)$ is a stochastic driving force. There is an additional scaling term S_0 in the PSD of this process that corresponds to the power at natural frequency, $\omega = \omega_0$. These are the free parameters of the kernel. These paremeters were set to constant values as an initial guess, based upon known values from a real solar flare.

Finally, we similarly model a red noise process using another model in the "terms" class in Celerite defined by:

$$k(\tau) = a * exp(-c\tau) \tag{3}$$

Where a and c are similar terms for correlational strength and timescale, that are similarly free, and guessed based on known values from a real solar flare.

From all of these models, we can write priors for our free parameters. For the same of convenience, we applied tophat priors for each of these parameters within reasonable known values of these. From here, we initialized the objects, and used them to construct a GP-class (Gaussian Process) object. This allows us to adjust the parameters of the model and parameters with a simple input function, as well as do measurements on the likelihood and prior of the entire fit.

From here, we ran a simple optimization scheme to compute the optimized values of each of the free parameters, by running scipy's optimization method, minimize, over the GP-class likelihood function, over the parameterspace defined by the GP objects parameter vector. By minimizing the negative log likelihood, we find the optimal values for a fit, which was primarily effective for the envelope paremters.

To better understand the parameter-space, and to maximize the probability of our fit, we resort to Markov chain Monte Carlo sampling (MCMC). (Explanation of MCMC) Using the emcee package, we initialize a set of walkers to explore the parameter-space of the model.

4. DISCUSSION

Let's talk about what happened.

REFERENCES

Something 2017