Gaussian Process Regression for Gravitational Waves

Seán White

Supervisor: Dr. Sarp Akcay

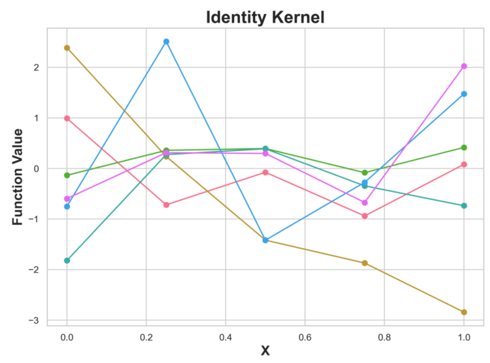
Introduction

- State of the art Gravitational Waveforms from binary black hole mergers are generated using computationally expensive numerical relativity **(NR)** simulations, which solve Einstein's field equations.
- Recently less computationally demanding waveform models have been developed.
- My Goal: Build a Gaussian Process Regression **(GPR)** model predicting the unfaithfulness (mismatch) of these waveform models to NR simulations.
- My GPR based predictor allows us to choose which model is the most faithful to NR in a given region of the parameter space.

GPR Fundamentals: The Prior Distribution

The prior distribution of functions in GPR represents our initial assumptions of our data.

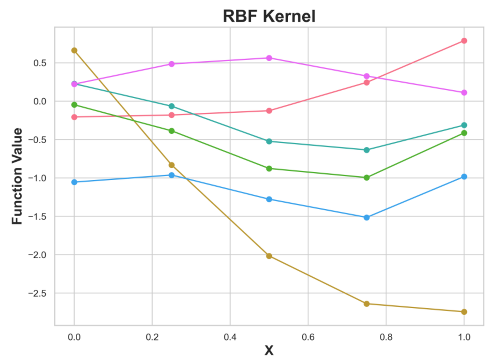
Assuming a prior distribution of : $f(x) \sim \mathcal{N}(0, k(x,x))$, the samples below show how the distribution varies with different kernel functions



Identity Kernel

$k(x,x') = I$

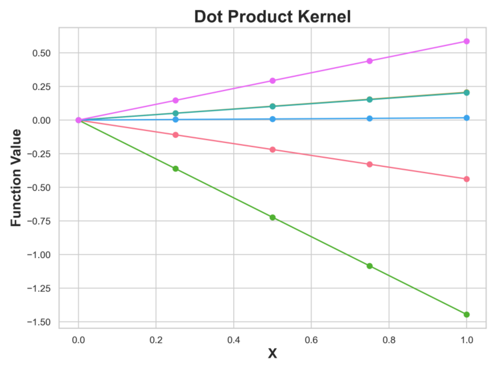
No correlation between points.



Radial Basis Function Kernel

$k(x,x') = \sigma^2 \exp\left(-\frac{|x-x'|^2}{2l^2}\right)$

High correlation for nearby points results in smoother functions



Dot Product Kernel

$k(x,x') = x \cdot x'$

Linear relationships

The Radial Basis Function (RBF) is the Kernel I use

Updating the Prior to Obtain the Posterior

We condition the prior distribution with "true" data. This results in the posterior distribution.

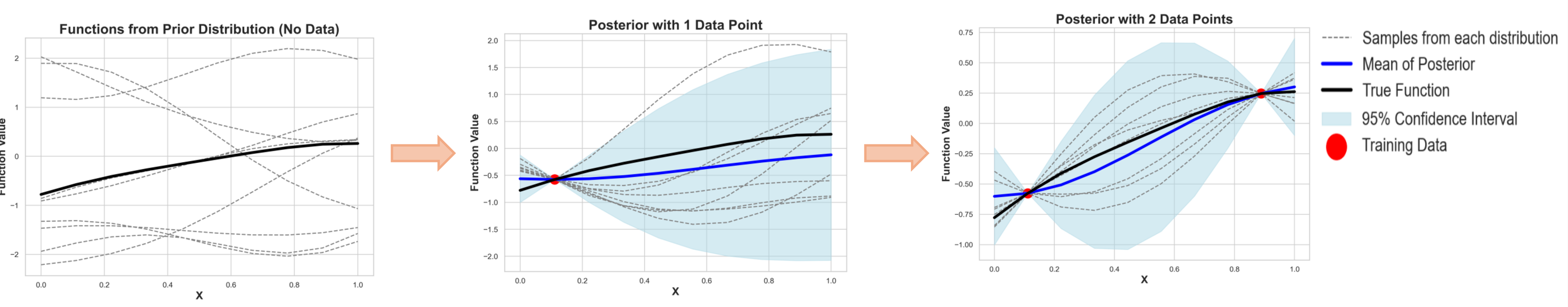
X:training points $\quad$ X_*: test points

Mean of $f(X_*)$

$$\mu(X_*) = K(X_*,X)\left(K(X,X)+\sigma^2 I\right)^{-1} f$$

Variance of $f(X_*)$

$$\Sigma(X_*) = K(X_*,X_*) - K(X_*,X)\left(K(X,X)+\sigma^2 I\right)^{-1} K(X,X_*)$$



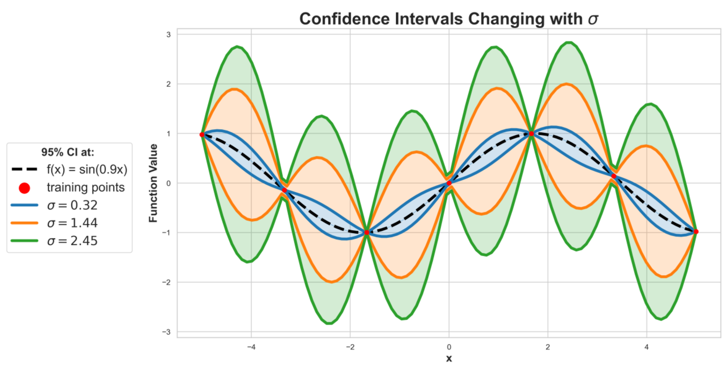
Prior samples: Initial assumptions

One data point reduces uncertainty near the point

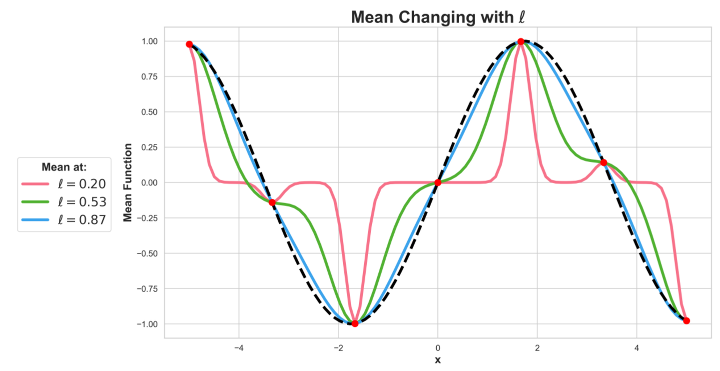
More data further refines the function and reduces uncertainty.

Hyper - Parameter Optimisation

The RBF depends on l and σ which effect the mean and variance of the posterior.



Small variance shrinks the confidence interval leading to less variability in predictions



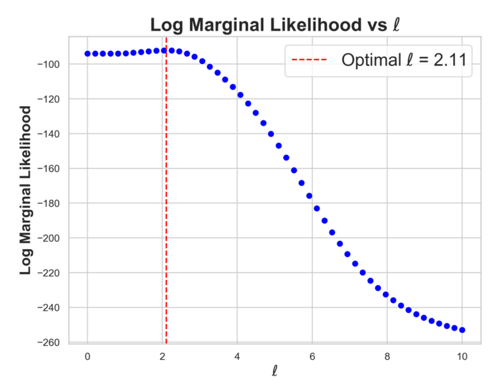
A smaller l results in sharper localised changes, larger the l the smoother the changes

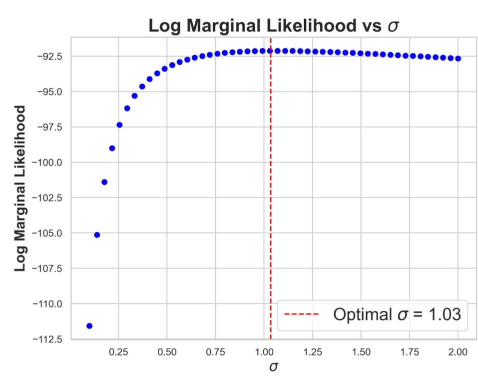
We can optimise l and σ to determine values of l and σ that best fit the data. Optimised parameters ensure that the predictions are neither too rigid nor too noisy.

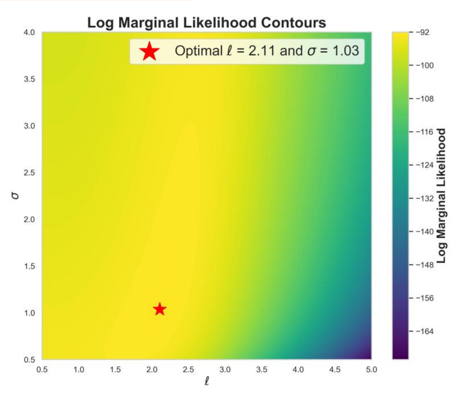
Maximising the Log Likelihood

By maximising the log-likelihood ($\log p(f \mid X, \theta)$) with respect to l and σ, we optimize the kernel hyperparameters to increase the probability that the true function is well-represented by the posterior distribution.

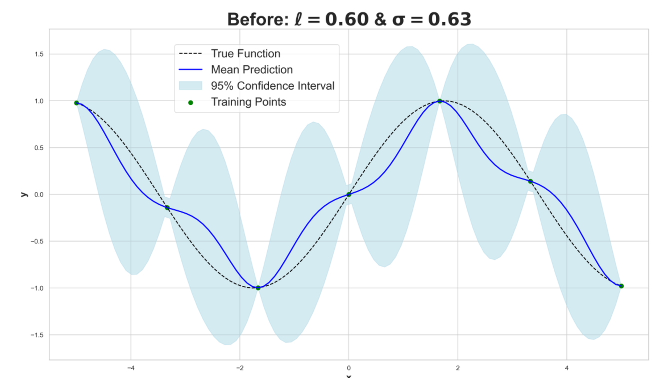
$$\log p(f \mid X, \theta) = -\frac{1}{2} y^T K^{-1} f - \frac{1}{2}\log\det(K) - \frac{n}{2}\log(2\pi)$$

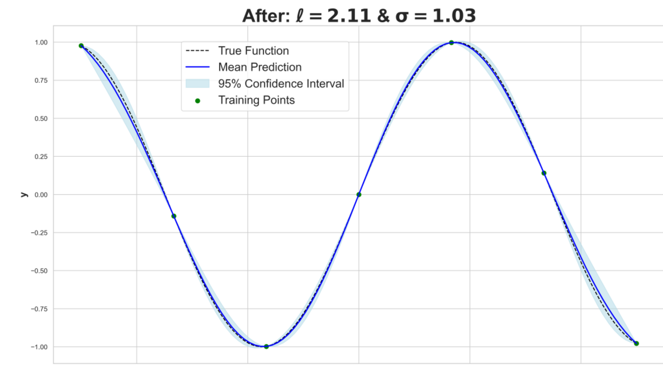






Effect of Parameter Optimisation on GPR

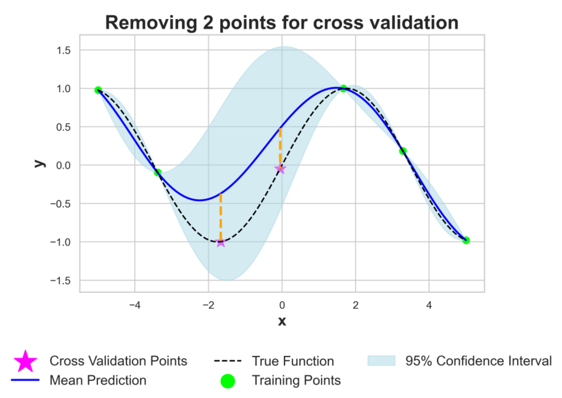




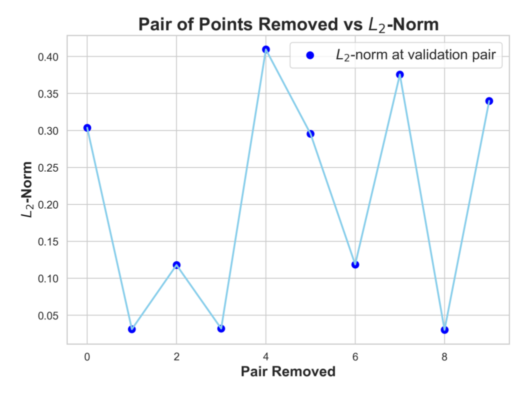
Cross Validation with the L_2-norm

For a set of N training points, selecting all pairs (e.g. N=5 $\Rightarrow \binom{5}{2} = 10$ pairs),the model's performance is evaluated using the L_2-norm aggregated over all the validation pairs

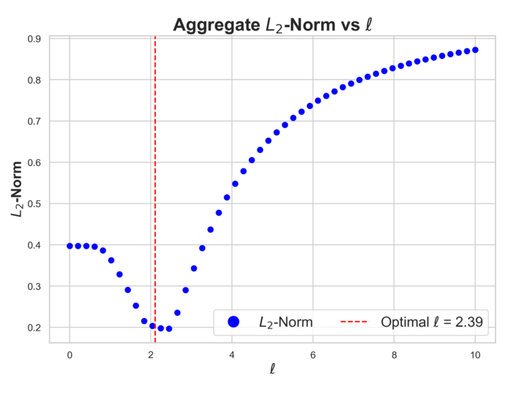
$$L_2\text{-norm of } f(X_*): \quad |f(X) - f(X_*)|_2 = \sqrt{\sum_{i=1}^{n}\left(f(X_i) - \mu(X_{*i})\right)^2}$$



Finding L_2-norm at cross validation points



L_2-norm at each pair of validation points



L_2-norm averaged over validation pairs

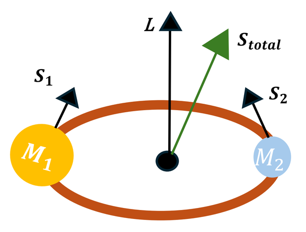
Applying to Gravitational Waves

We reduce the intrinsic 8 dimensional parameter space of a binary black hole to 4 dimensions:

Total Mass: $M_1 + M_2$

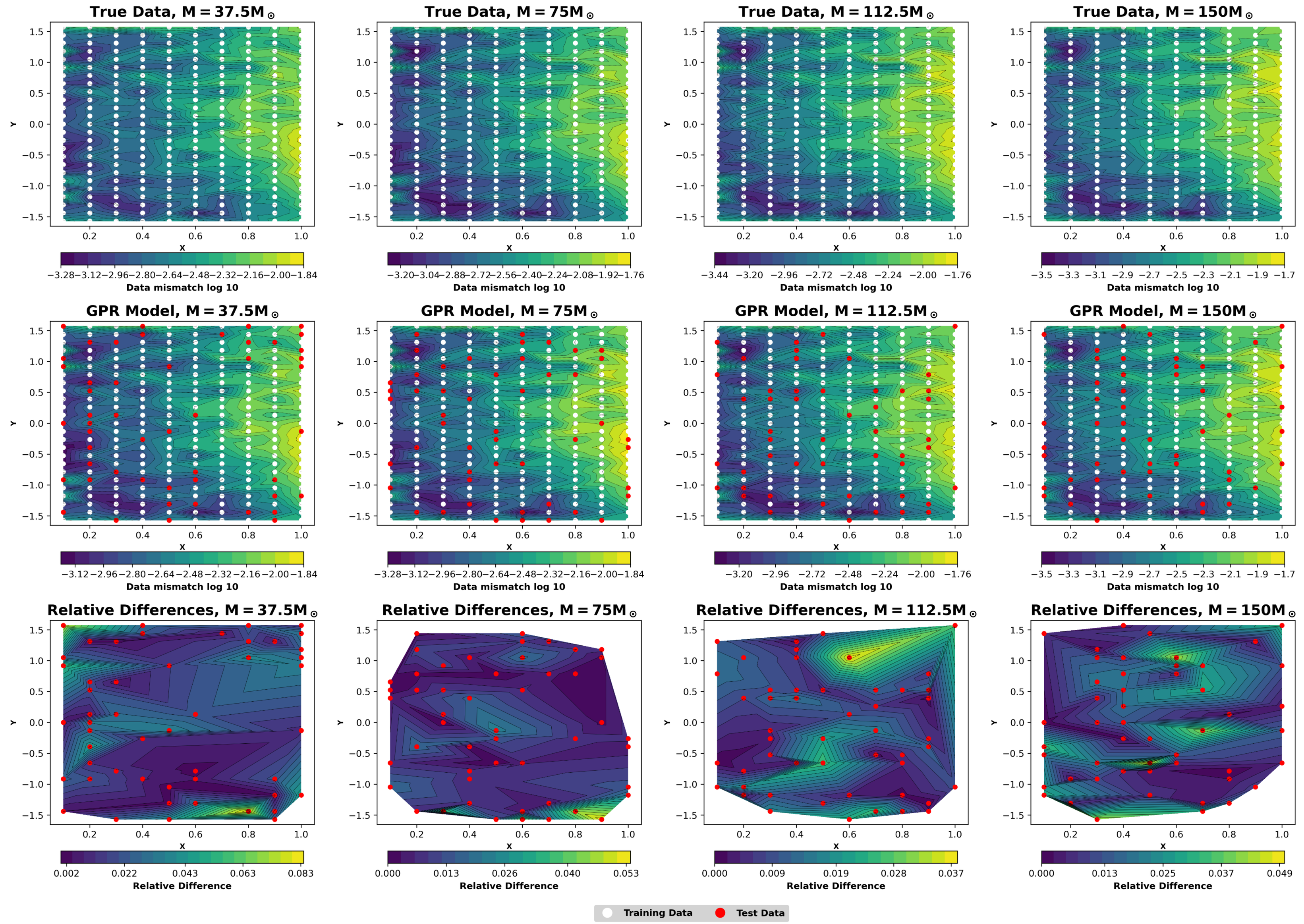
Symmetric Mass Ratio: $\frac{q}{(1+q)^2}$, $q = \frac{M_2}{M_1}$

$$\chi_\perp = \frac{|S_{1,\perp} + S_{2,\perp}|}{M^2} \qquad \chi_\parallel := \frac{|S_{1,\parallel} + S_{2,\parallel}|}{M^2}$$

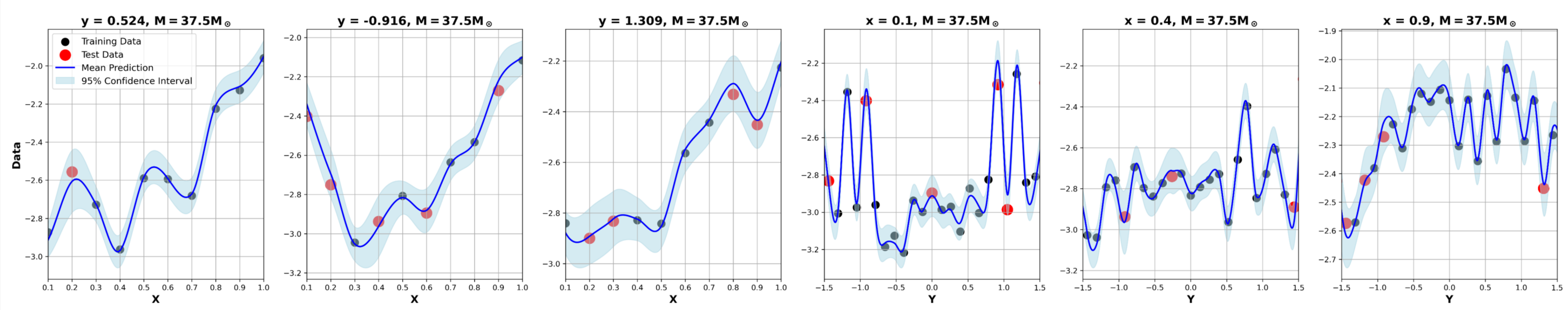


1. $S_{total} = S_1 + S_2$
2. $S = (S_x, S_y, S_z)$:Spin Vectors
3. M: Masses
4. $L = (L_x, L_y, L_z)$:Orbital angular momentum

4D GPR on Gravitational Wave Mismatch



1D Cross-cuts



Next Steps

1. Improve my GPR model by using different Kernels and optimisation techniques.
2. Compare my GPR model with functional fits (interpolants)[4].
3. Use GPR directly on the 8 dimensional intrinsic parameter space.

References

1. Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.
2. Wang, J. (2023). *An intuitive tutorial to Gaussian process regression*. University of Waterloo.
3. Neptune.ai. (2024). *Hyperparameter Tuning in Python: A Complete Guide*. Retrieved from https://neptune.ai/blog/hyperparameter-tuning-in-python-complete-guide.
4. Hoy, C., Akçay, S., Mac Uilliam, J., & Thompson, J. E. (2024). Incorporating model accuracy into gravitational-wave Bayesian inference.

sean.white2@ucdconnect.ie