WHEN IS A NOISY SIGNAL OSCILLATING AND WHEN IS IT NOISY ?

SUPERVISED BY PETER SWAIN

WHEN IS A NOISY SIGNAL OSCILLATING AND WHEN IS IT NOISY ?

Topics to be covered

INTRODUCTION TO PROBLEM

BAYESIAN INFERENCE AND TOY EXAMPLE

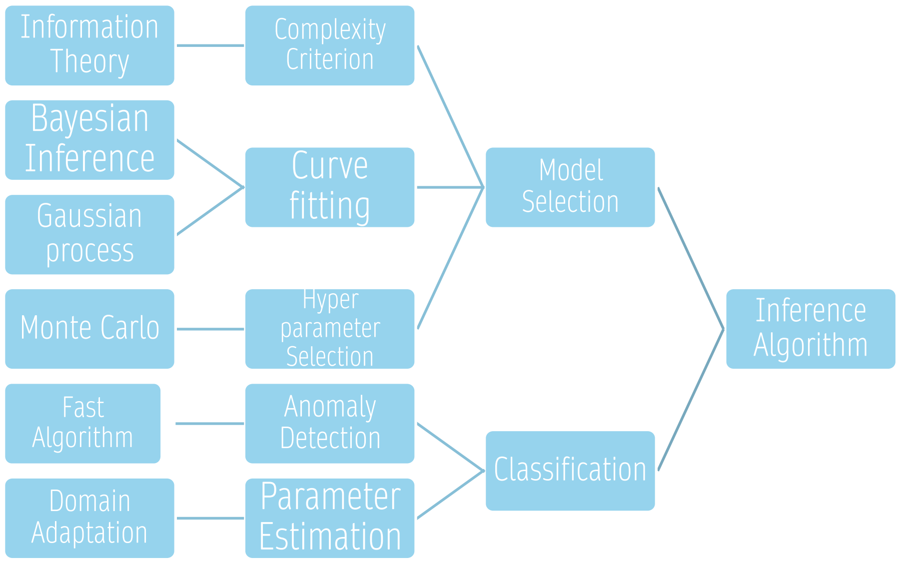
BASICS OF MODEL SELECTION

INTRODUCTION TO GAUSSIAN PROCESS

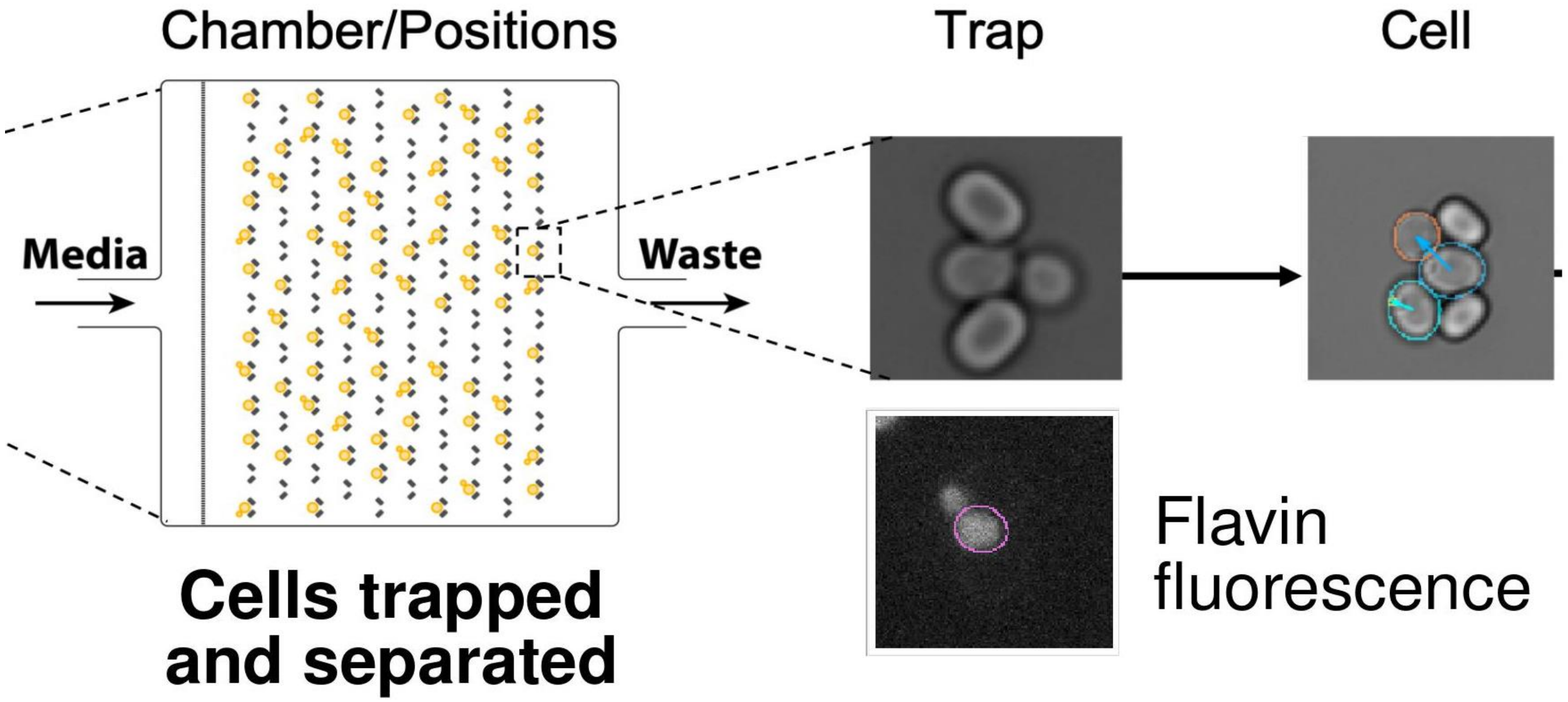
TOY EXAMPLE FOR GAUSSIAN PROCESS

WHAT PROJECT AIMS?

Prior of project

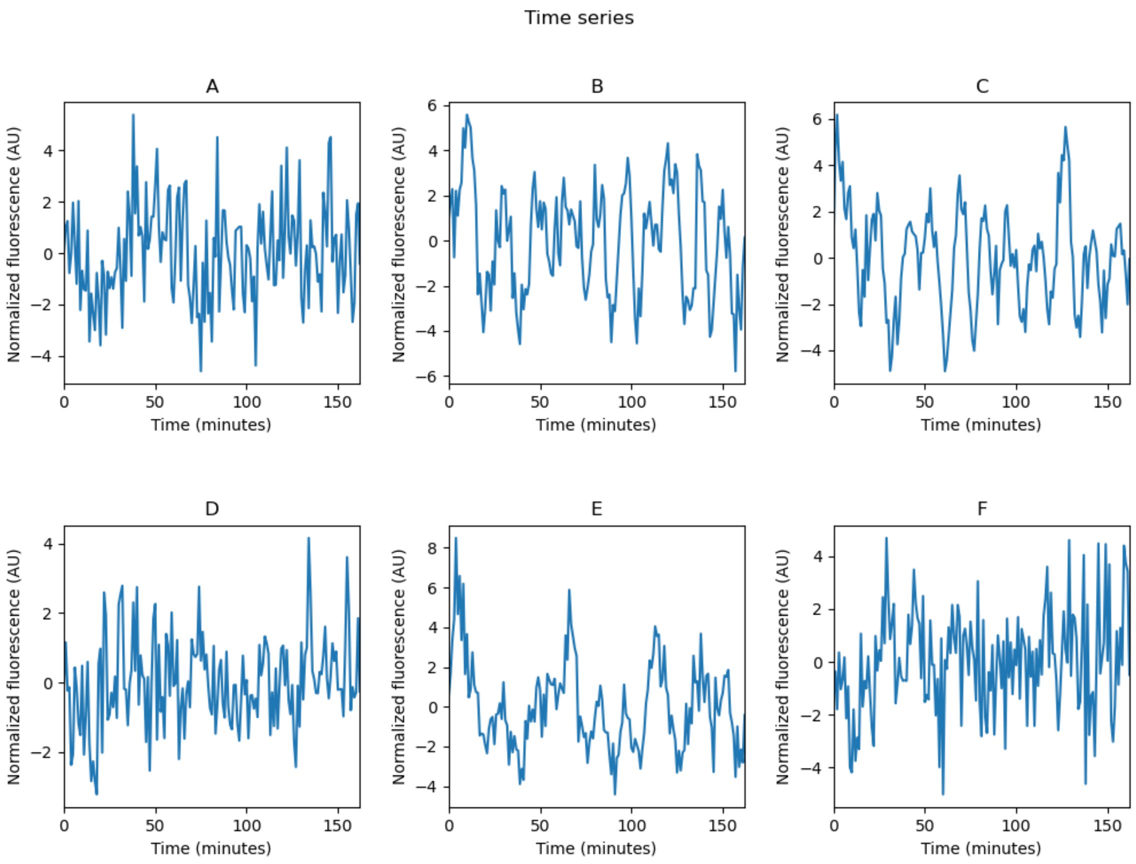


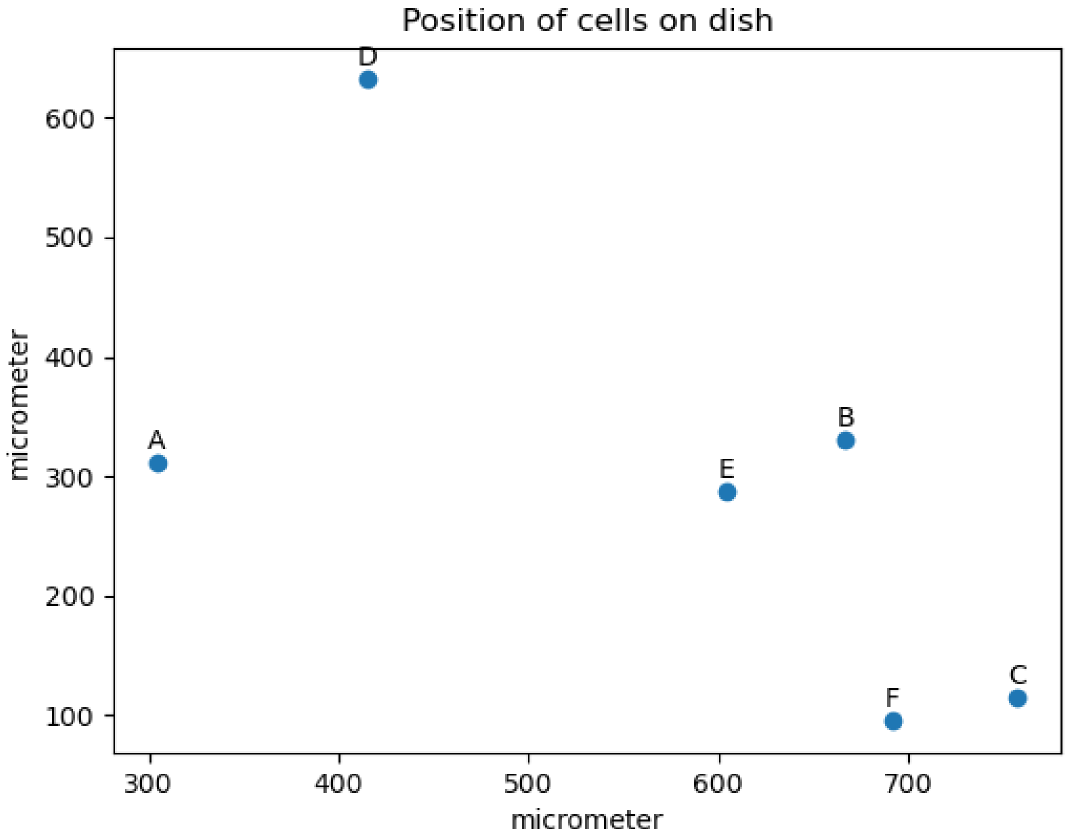
PETRI DISH BEING OBSERVED IN FLUORESCENCE ENVIRONMENT



Images acquired at regular intervals.

TIME SERIES DATA OF CELLS FROM PREVIOUS SLIDE
NO MODEL IS KNOWN OF UNDERLYING PROCESS

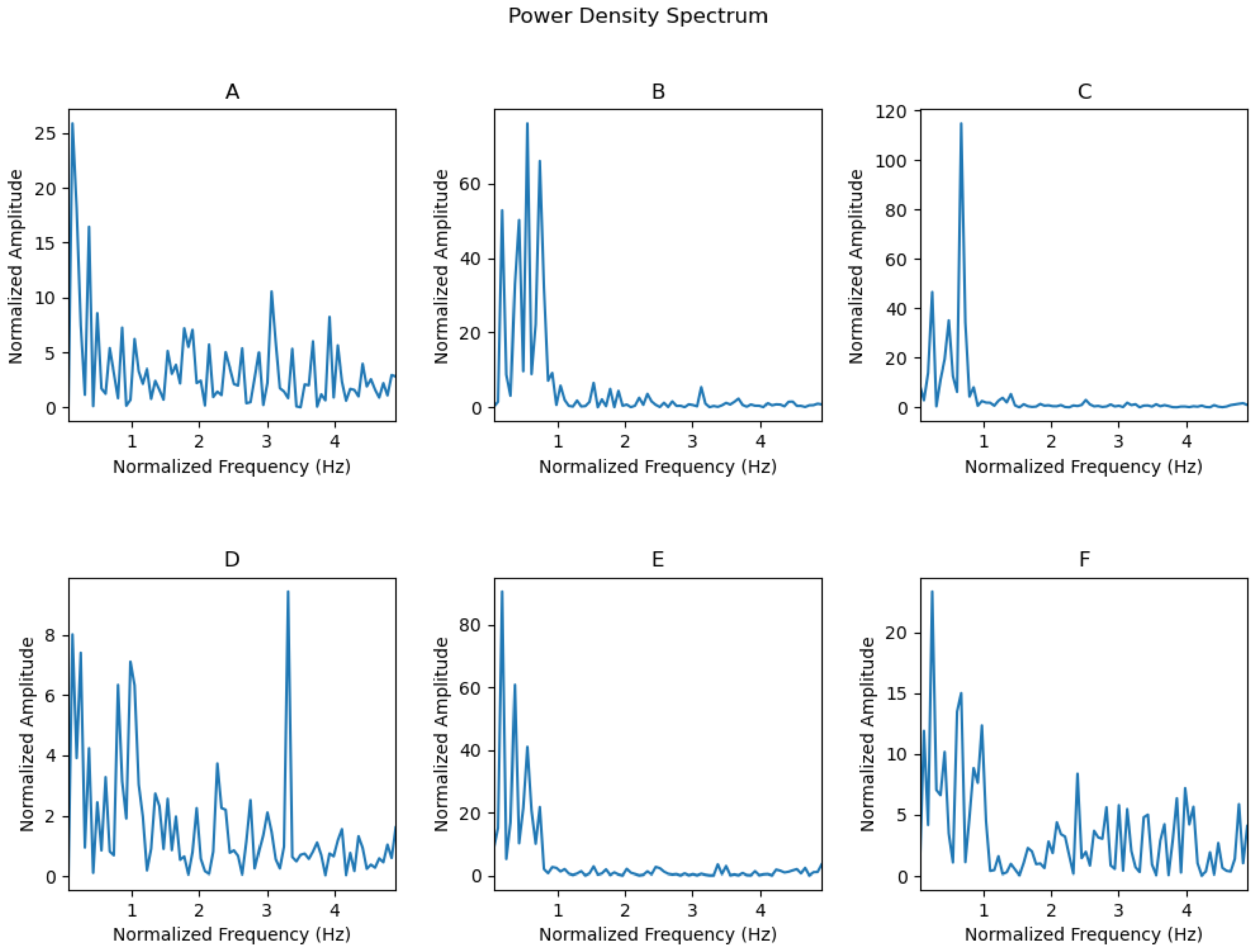




* Data Collected by Arin Wongprommon (PhD Student Swain Lab)

FOURIER TRANSFORM OF TIME SERIES DATA

FFT



INTRODUCTION TO BAYESIAN FRAMEWORK

FFT

BAYES

Data/Evidence

Hypothesis/Model

Probability of $\mathcal{E}$ given $\mathcal{H}_i$

$$\mathcal{P}(\mathcal{E})\,\mathcal{P}(\mathcal{H}_i|\mathcal{E}) = \mathcal{P}(\mathcal{E}|\mathcal{H}_i)\mathcal{P}(\mathcal{H}_i)$$

$$\mathcal{P}(\mathcal{H}_i|\mathcal{E}) = \frac{\mathcal{P}(\mathcal{E}|\mathcal{H}_i)\mathcal{P}(\mathcal{H}_i)}{\mathcal{P}(\mathcal{E})}$$

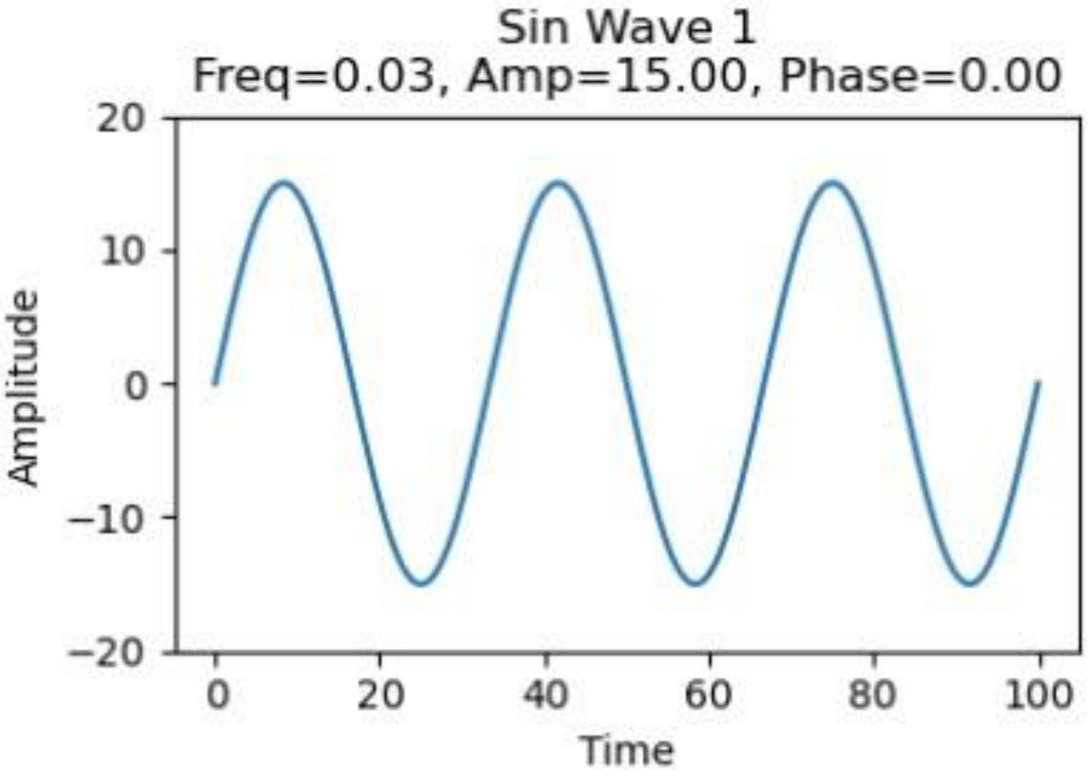
$$Posterior = \frac{likelihood \times Prior}{Marginal\ likelihood}$$

$Where \quad \mathcal{P}(\mathcal{E}) = \mathcal{P}(\mathcal{E}|\mathcal{H}_1)\mathcal{P}(\mathcal{H}_1) + \mathcal{P}(\mathcal{E}|\mathcal{H}_2)\mathcal{P}(\mathcal{H}_2) + \mathcal{P}(\mathcal{E}|\mathcal{H}_3)\mathcal{P}(\mathcal{H}_3) + \ldots$

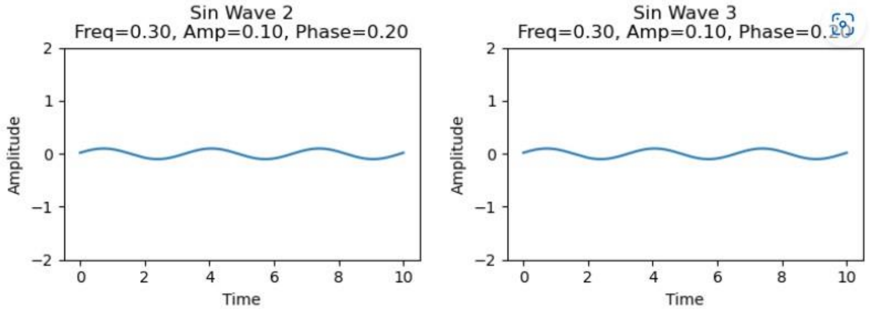
WE WILL UNDERSTAND BAYESIAN IMPLEMENTATION WITH TOY EXAMPLE OF SUPERIMPOSED SINUSOIDAL WAVES

FFT

BAYES



Primary Signal

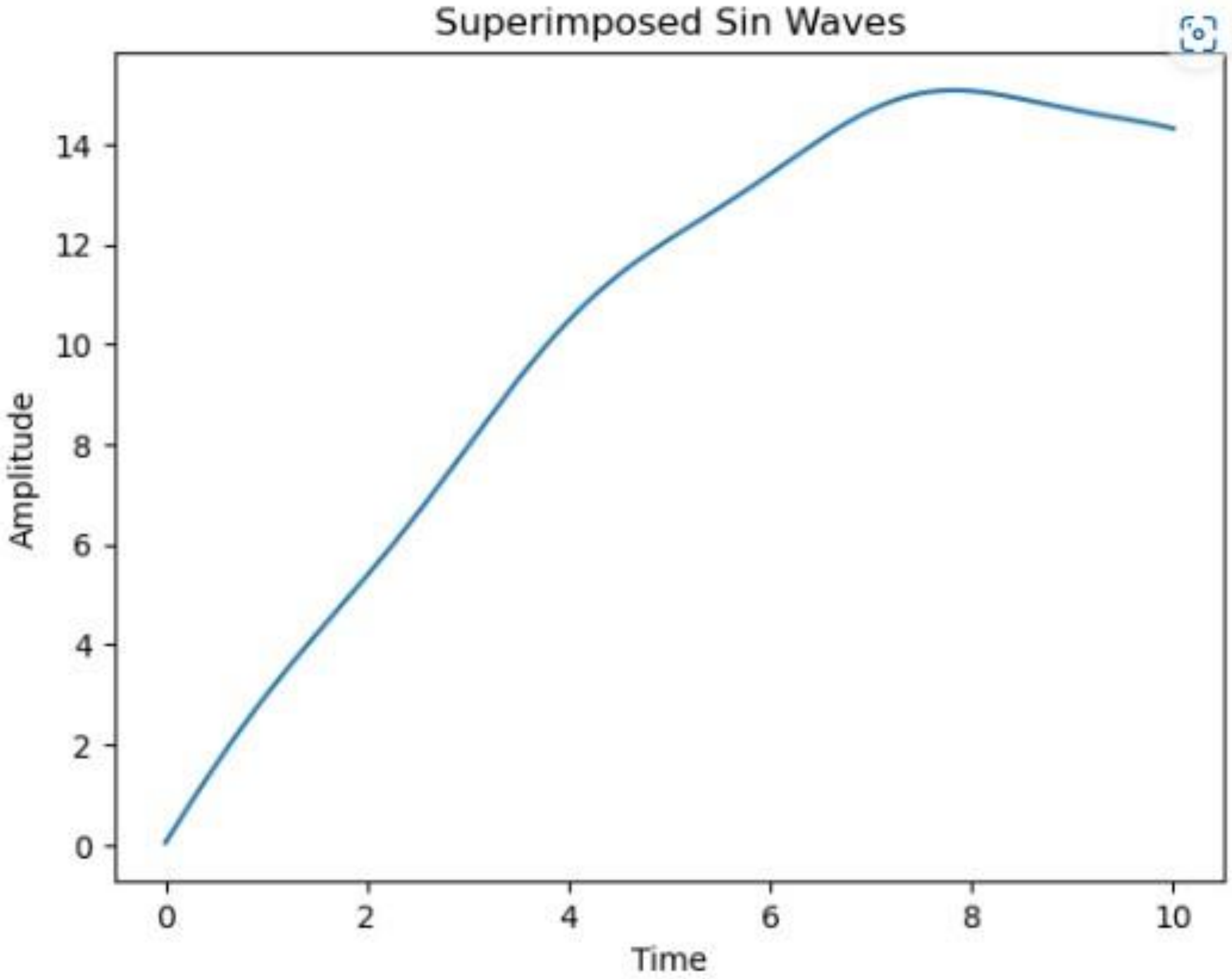


Secondary Signal

OBSERVATION TIME WINDOW ON SUPERIMPOSED SIGNAL

FFT

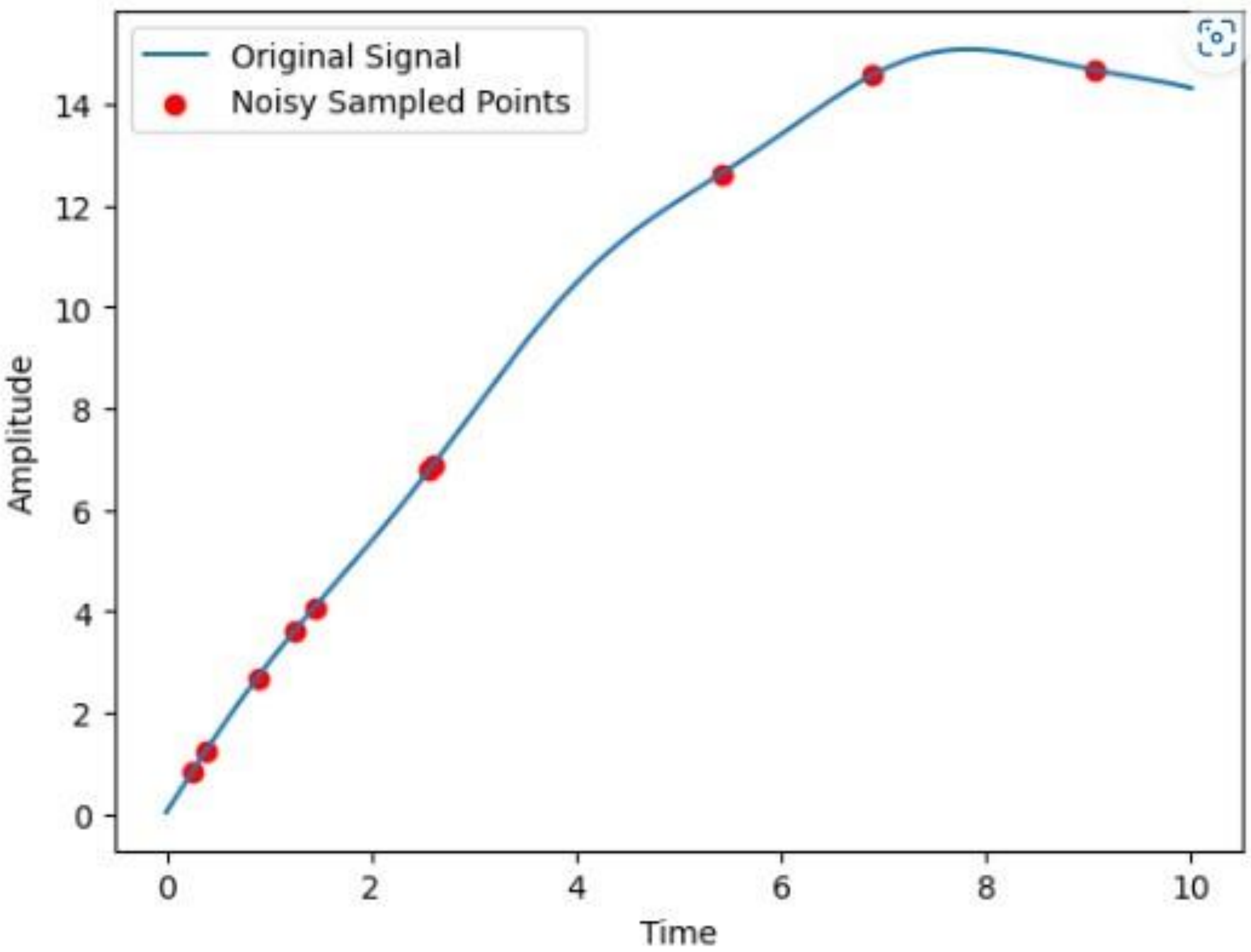
BAYES



RANDOM POINTS ON CURVE WILL ACT LIKE DATA OBSERVED IN EXPERIMENT

FFT

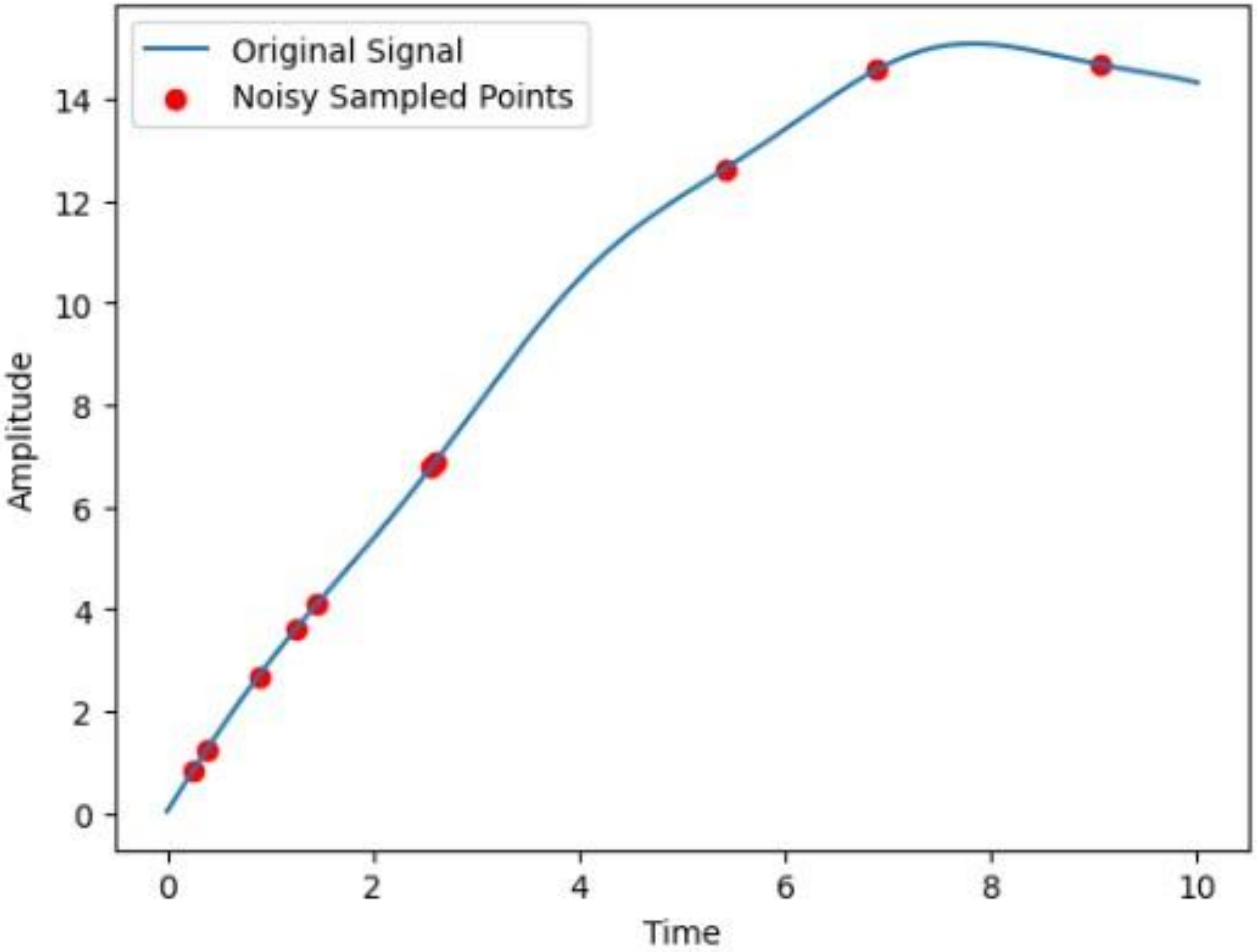
BAYES



ADDING NOISE TO DATA

FFT

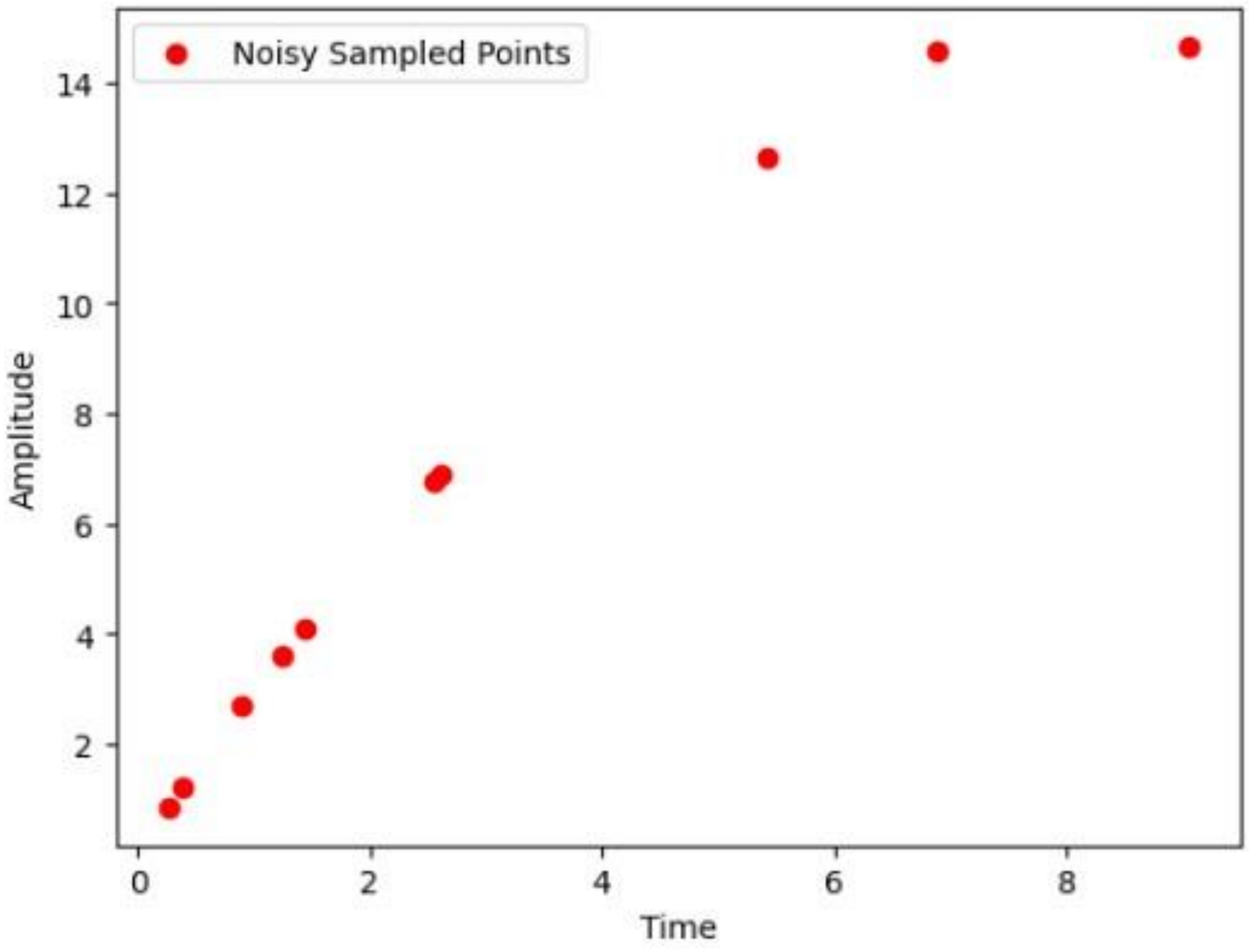
BAYES



OBSERVED NOISY DATA POINTS

FFT

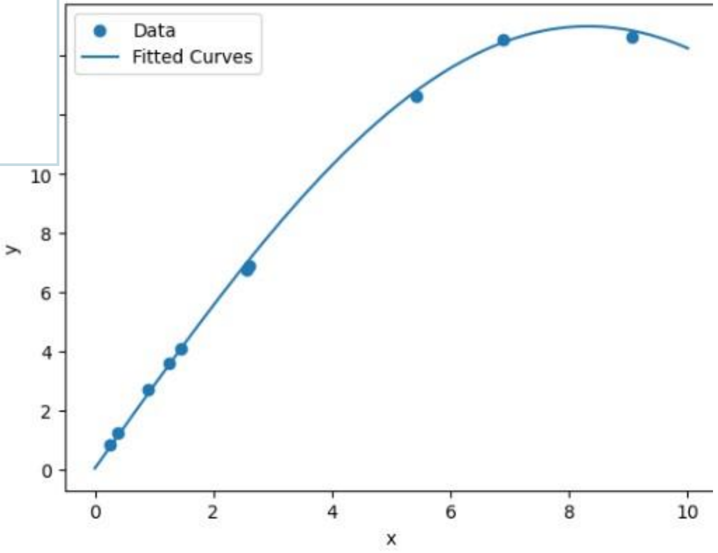
BAYES



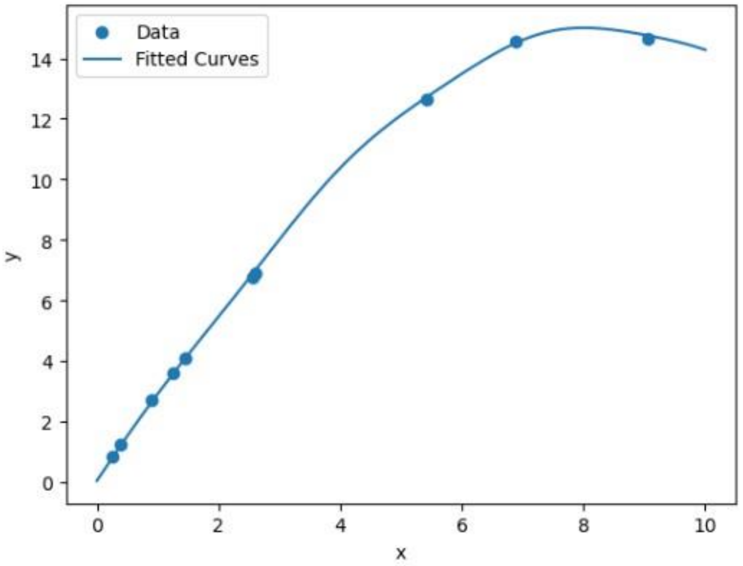
FITTING WITH DIFFERENT NUMBER OF SINUSOIDAL

FFT

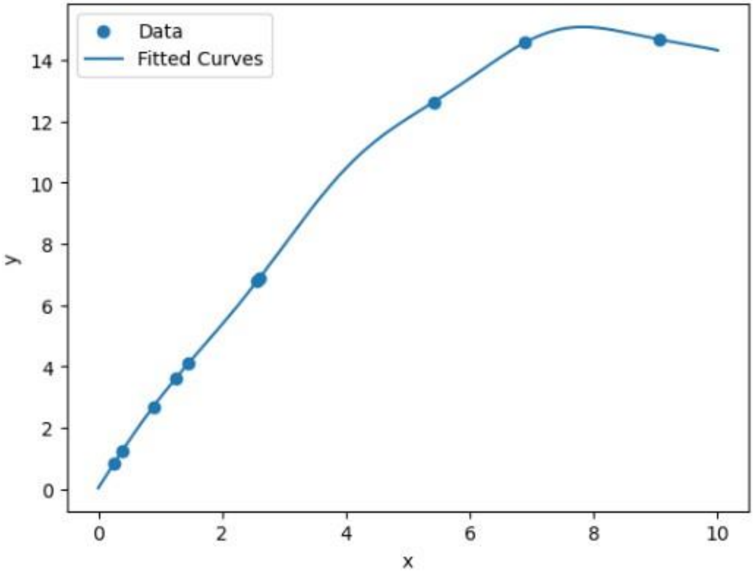
BAYES



Fit with 1 oscillator



Fit with 2 oscillator

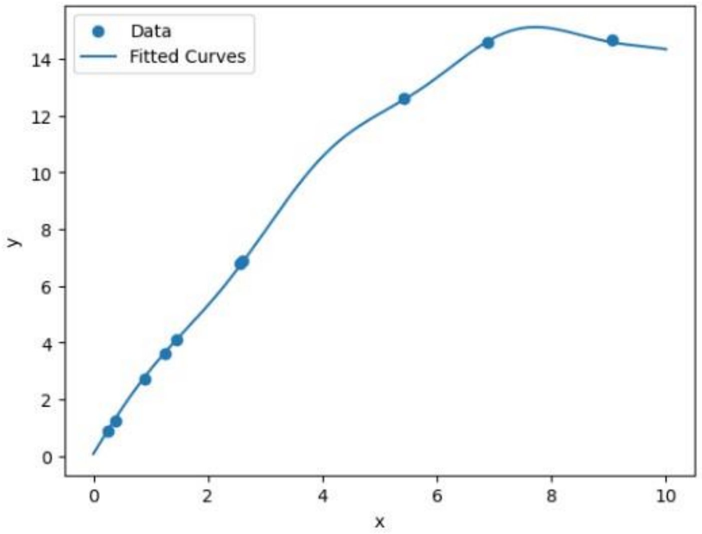


Fit with 3 oscillator

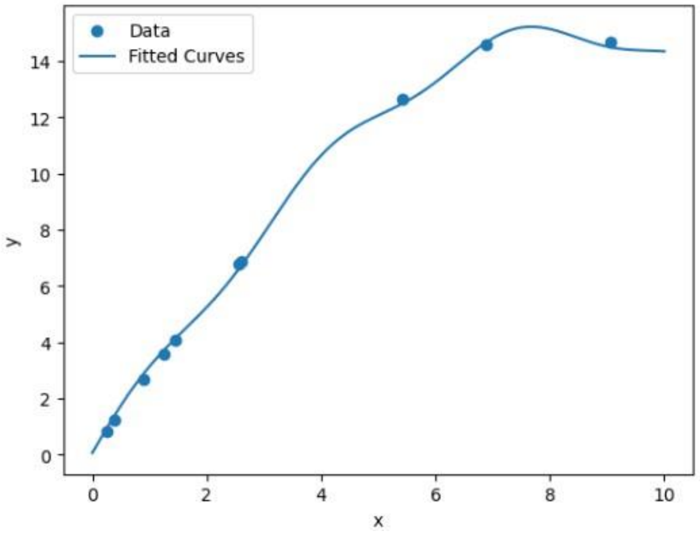
FITTING WITH DIFFERENT NUMBER OF SINUSOIDAL

FFT

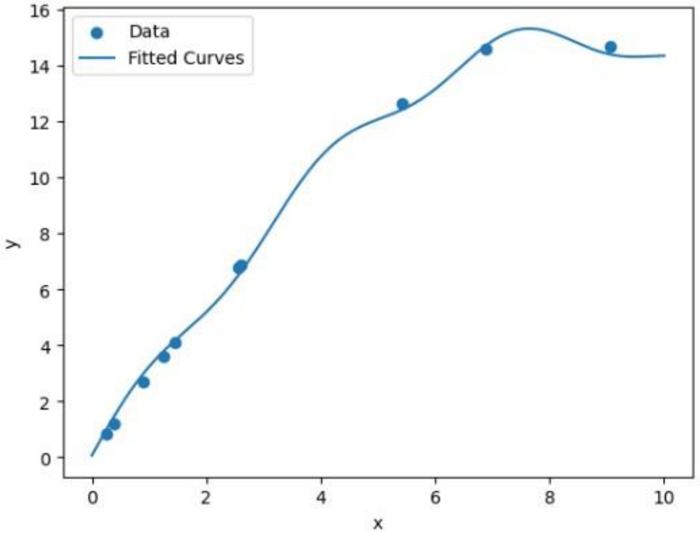
BAYES



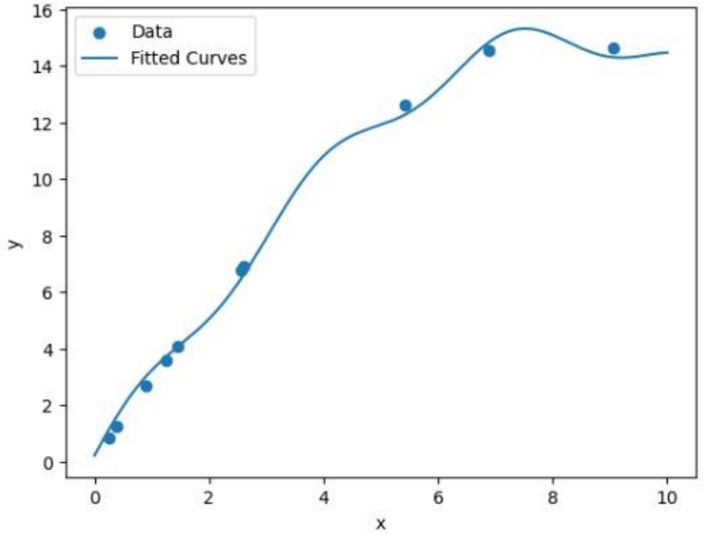
Fit with 4 oscillator



Fit with 5 oscillator



Fit with 6 oscillator



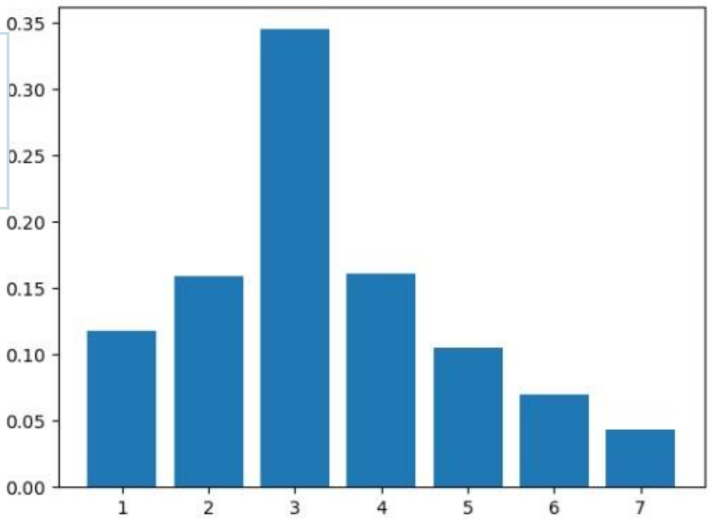
Fit with 7 oscillator

CHANGING BELIEF USING BAYES THEOREM

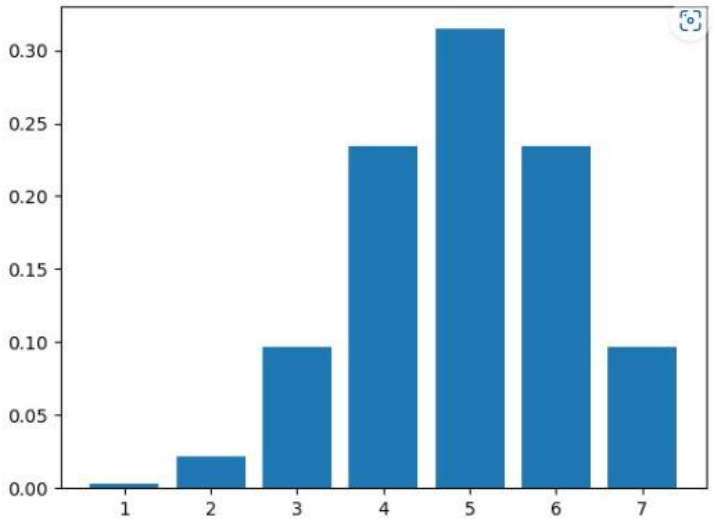
FFT

BAYES

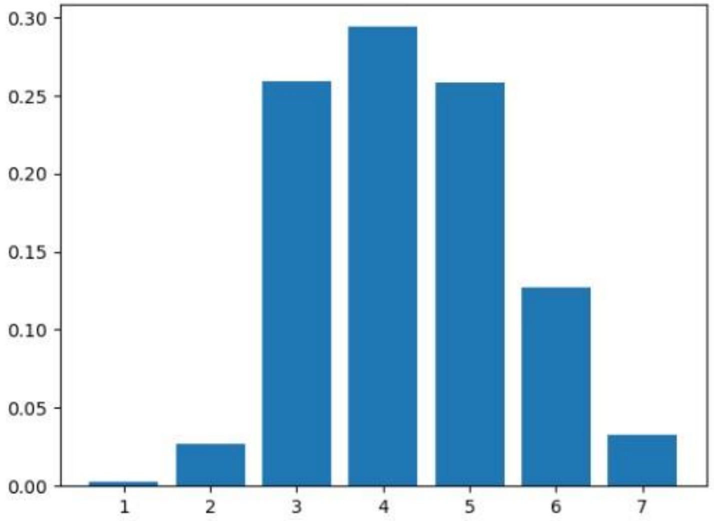
$$Posterior = \frac{likelihood \times Prior}{Marginal\ likelihood}$$



Likelihood



Prior



Posterior

$$\mathcal{P}(\mathcal{H}_i|\mathcal{E}) = \frac{\mathcal{P}(\mathcal{E}|\mathcal{H}_i)\mathcal{P}(\mathcal{H}_i)}{\mathcal{P}(\mathcal{E})}$$

$Where \quad \mathcal{P}(\mathcal{E}) = \mathcal{P}(\mathcal{E}|\mathcal{H}_1)\mathcal{P}(\mathcal{H}_1) + \mathcal{P}(\mathcal{E}|\mathcal{H}_2)\mathcal{P}(\mathcal{H}_2) + \mathcal{P}(\mathcal{E}|\mathcal{H}_3)\mathcal{P}(\mathcal{H}_3) + \ldots$

WHAT IS HIERARCHICAL **BAYESIAN MODEL** SELECTION?

FFT

BAYES

MODEL SELECTION

Hyper-Parameters

Parameter Value

$$\mathcal{P}(\boldsymbol{\theta}|\mathcal{E}) = \frac{\mathcal{P}(\mathcal{E}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})}{\mathcal{P}(\mathcal{E})}$$

Where $\mathcal{P}(\mathcal{E}) = \int \mathcal{P}(\mathcal{E}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})d\boldsymbol{\theta}$

$$\mathcal{P}(\boldsymbol{w}|\mathcal{E},\boldsymbol{\theta}) = \frac{\mathcal{P}(\mathcal{E}|\boldsymbol{w},\boldsymbol{\theta})\mathcal{P}(\boldsymbol{w}|\boldsymbol{\theta})}{\mathcal{P}(\mathcal{E}|\boldsymbol{\theta})}$$

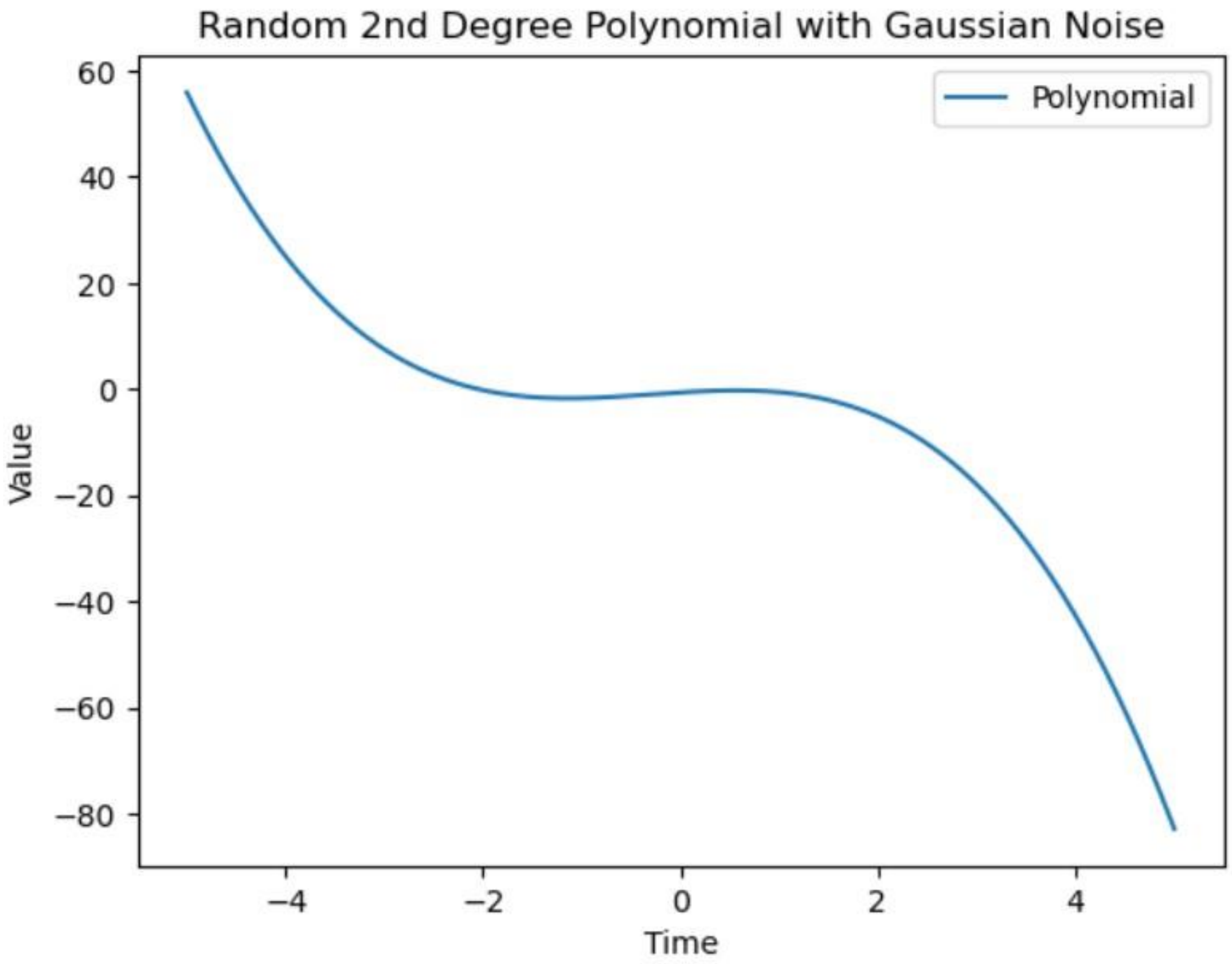
Where $\mathcal{P}(\mathcal{E}|\boldsymbol{\theta}) = \int \mathcal{P}(\mathcal{E}|\boldsymbol{w})\mathcal{P}(\boldsymbol{w}|\boldsymbol{\theta})d\boldsymbol{w}$

A WALKTHROUGH TO THE BASICS OF MODEL SELECTION

FFT

BAYES

MODEL SELECTION

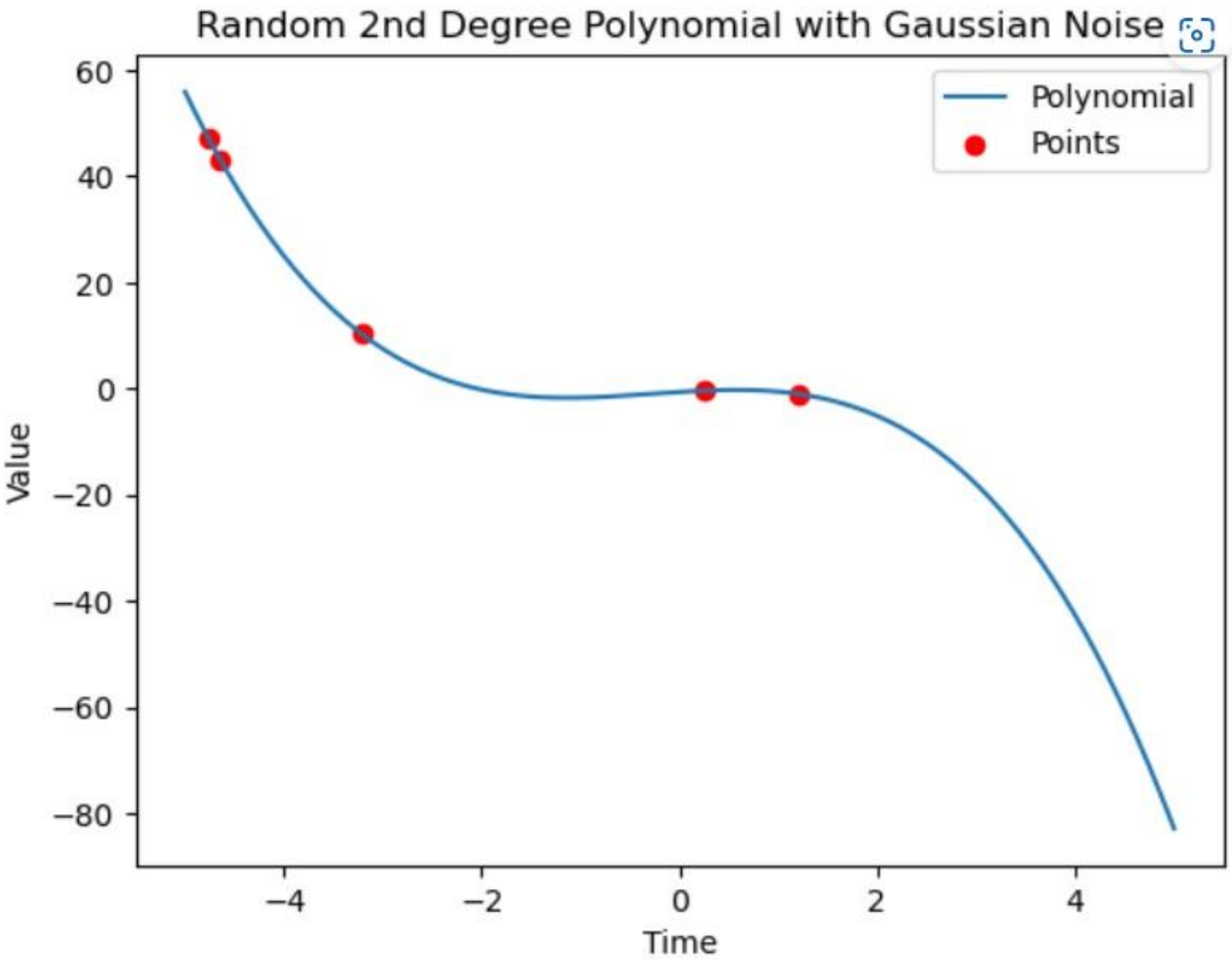


CHOSING RANDOM POINTS ON CURVE

FFT

BAYES

MODEL SELECTION

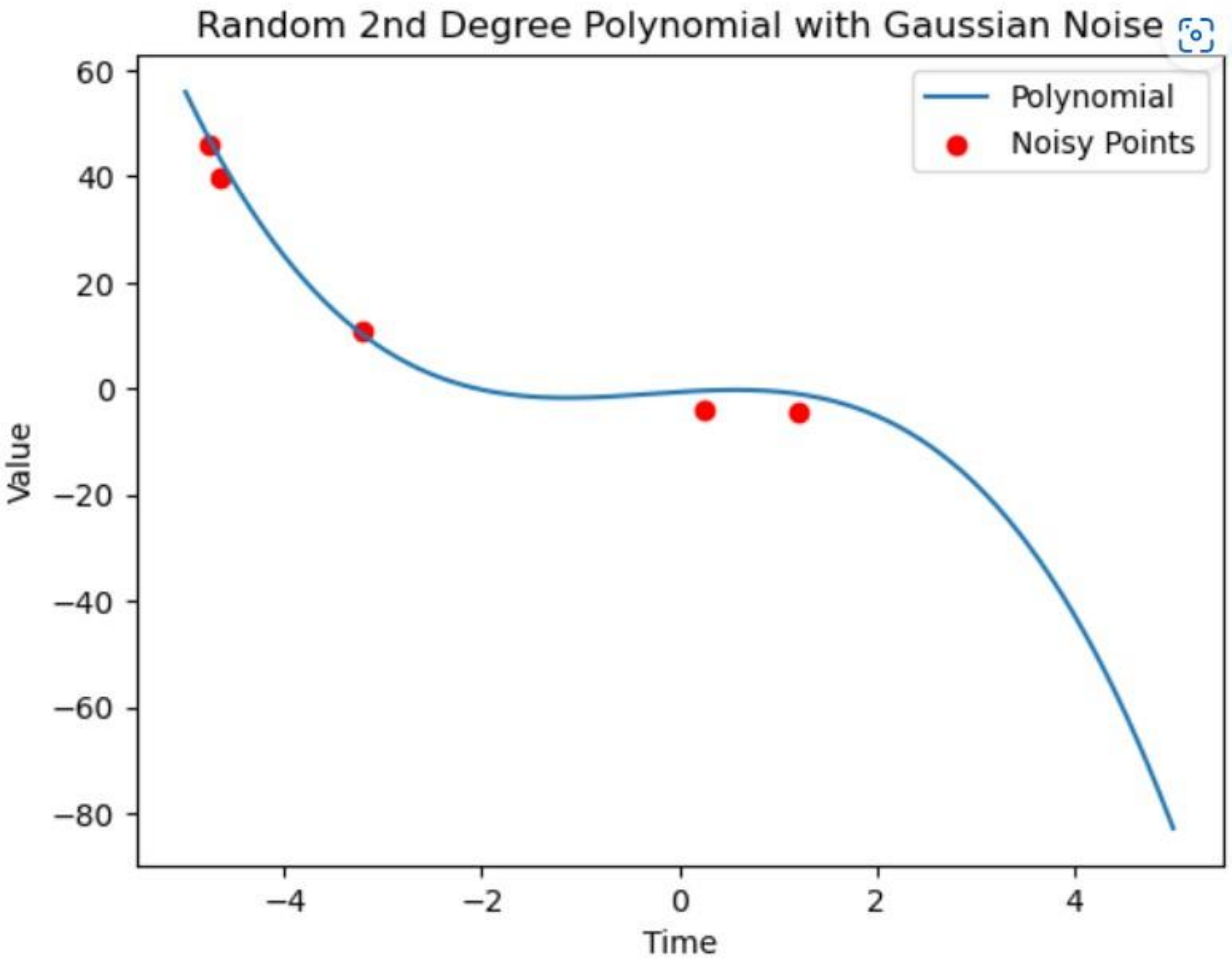


ADDING GAUSSIAN NOISE TO DATA

FFT

BAYES

MODEL SELECTION

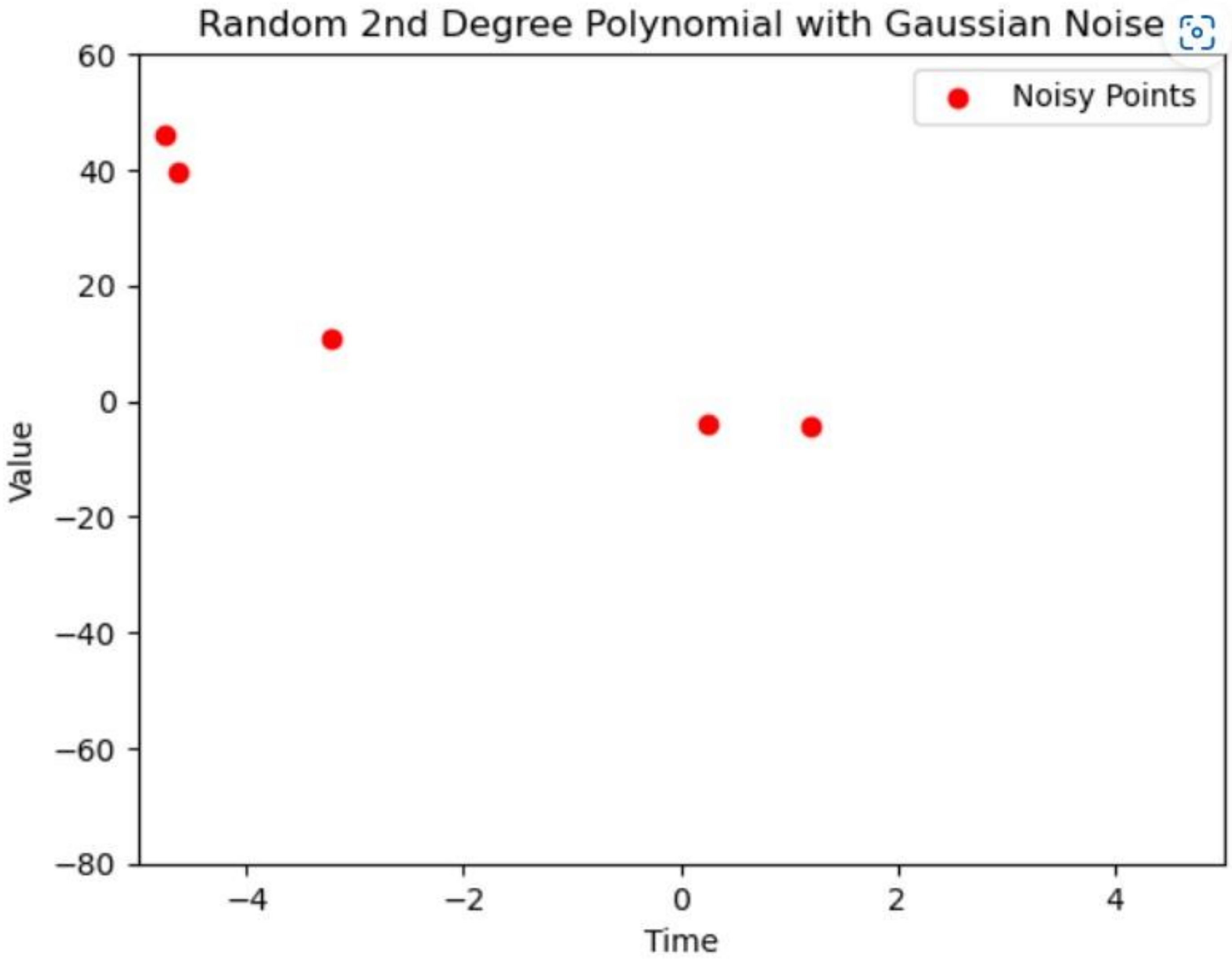


FINAL DATA TO BE USED IN ANALYSIS

FFT

BAYES

MODEL SELECTION

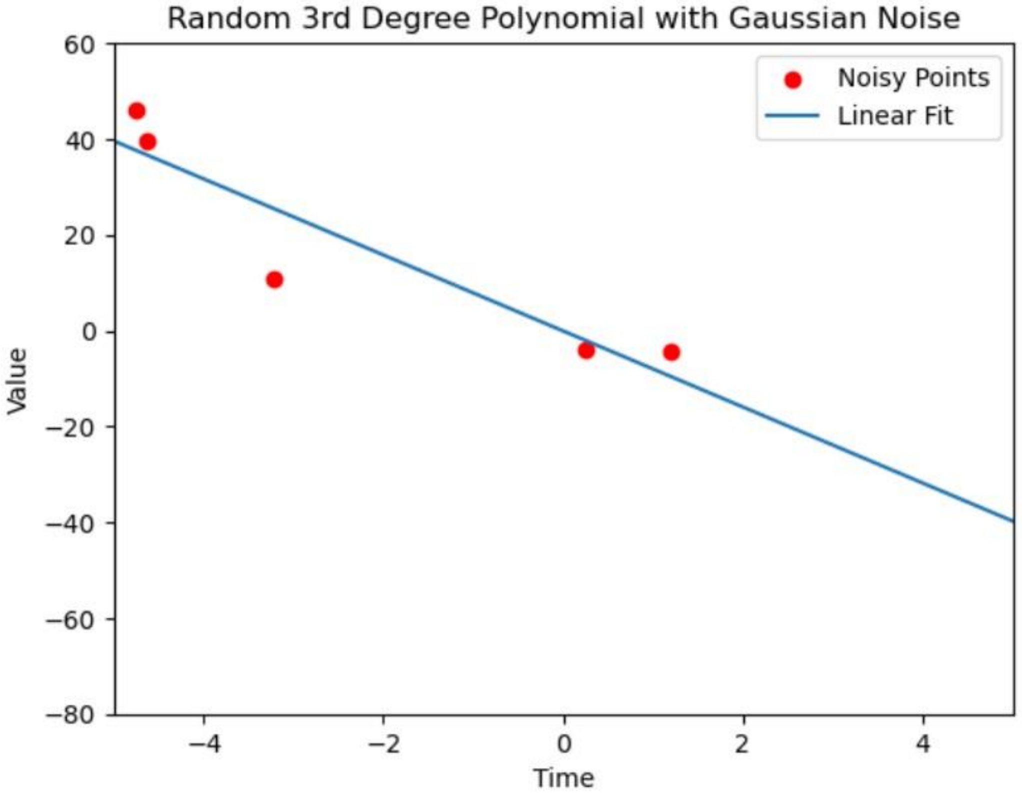


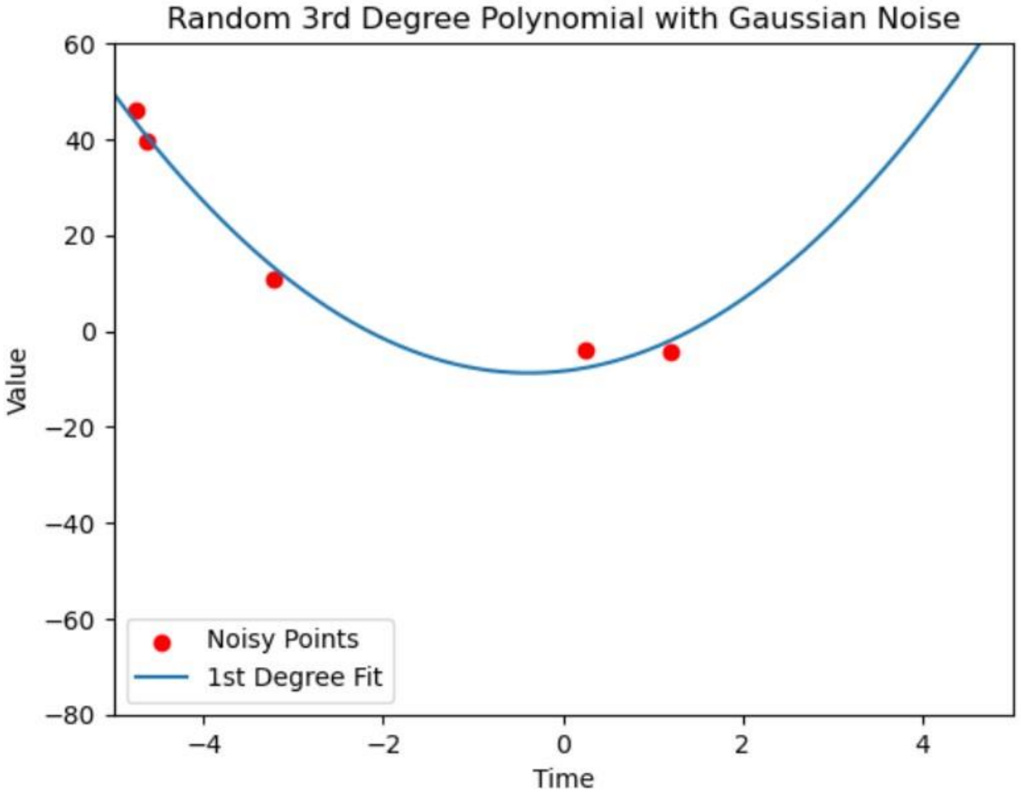
FITTING TO LINEAR AND 1ST DEGREE POLYNOMIAL

FFT

BAYES

MODEL SELECTION



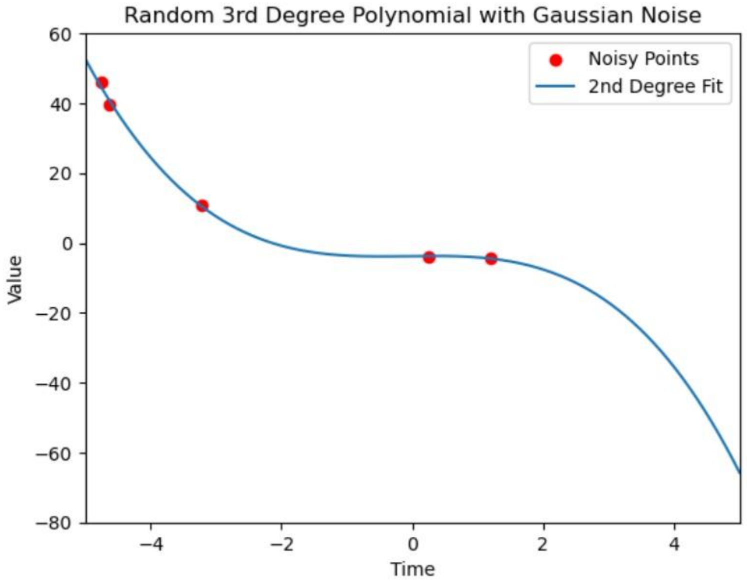


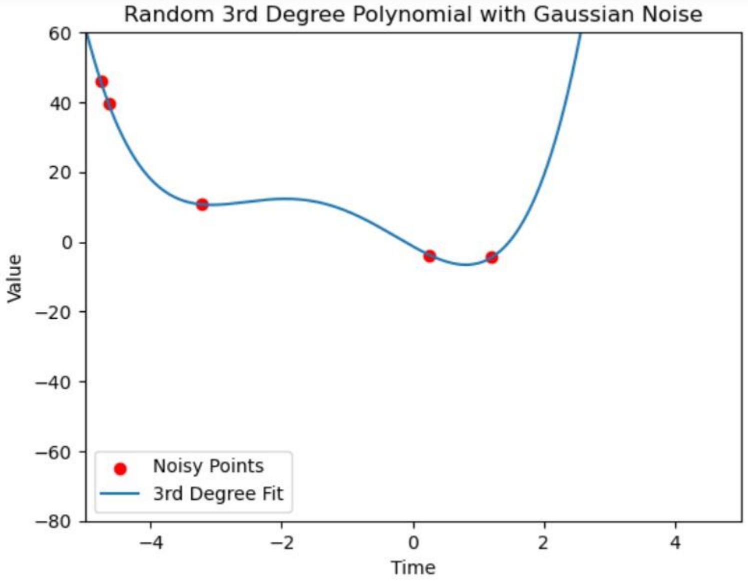
FITTING TO HIGHER DEGREE POLYNOMIAL

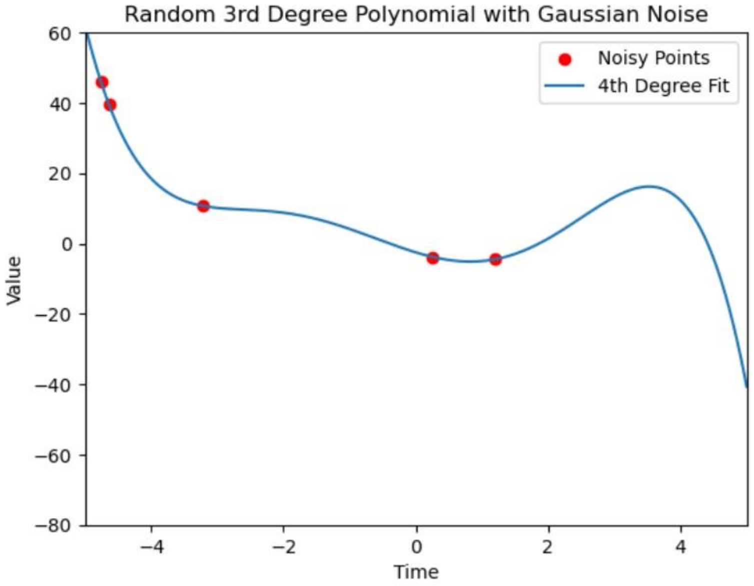
FFT

BAYES

MODEL SELECTION







USING BAYESIAN INFORMATION CRITERION (BIC) TO PENALIZE UNNECESSARY COMPLEXITY OF MODEL

FFT

BAYES

MODEL SELECTION

```
likelyhood - Linear Fit: -16.2692
likelyhood - 1st Degree Fit: -12.9734
likelyhood - 2nd Degree Fit: -11.0972
likelyhood - 3rd Degree Fit: -11.0543
likelyhood - 4th Degree Fit: -14.6380
```

*Higher is better

Penalty for extra parameter

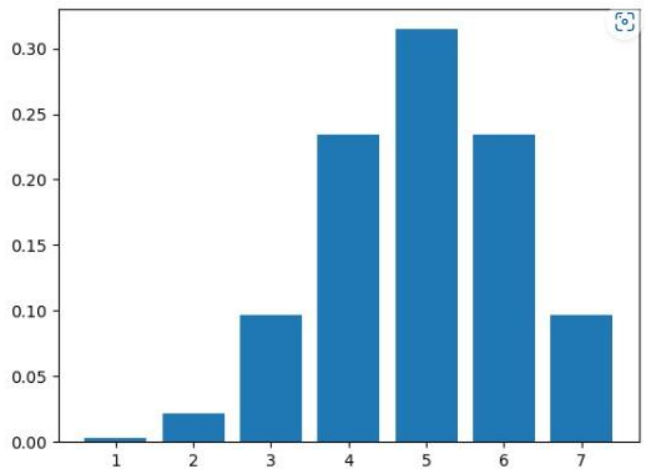
$$\mathrm{BIC} = k\ln(n) - 2\ln(\widehat{L}).$$

```
BIC Score - Linear Fit: 35.7574
BIC Score - 1st Degree Fit: 30.7750
BIC Score - 2nd Degree Fit: 28.6321
BIC Score - 3rd Degree Fit: 30.1559
BIC Score - 4th Degree Fit: 38.9327
```

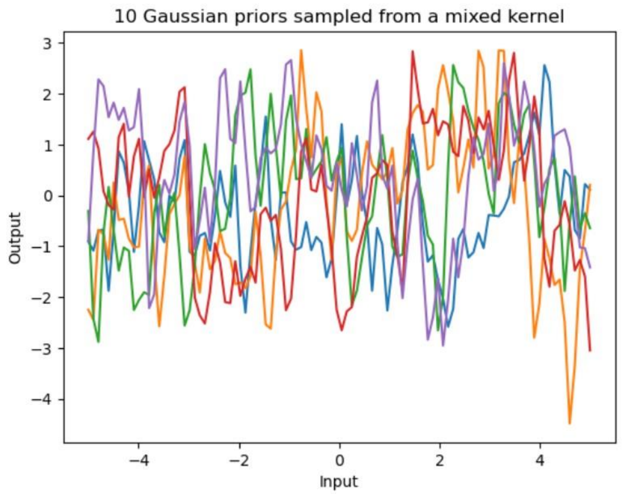
*Lower is better

GAUSSIAN PROCESS REGRESSION

SWITCHING FROM WEIGHT TO FUNCTION SPACE VIEW



Prior of Normal Bayesian inference



Prior of Gaussian Process

- E.g. of the kernel of gaussian process

$$K(x, x_*) = \theta_1^2 \exp\left(-\frac{(x - x_1)^2}{2\theta_2^2}\right)$$

- x_i are the points on x axis, they are domain of kernel functions, which in turn gives element of covariance matrix

$$N(\mu, \Sigma) = N\left(\begin{bmatrix}\mu_X \\ \mu_{x_i}\end{bmatrix}, \begin{bmatrix}K(x,x) & K(x,x_i) \\ K(x_i,x) & K(x_i,x_i)\end{bmatrix}\right)$$

TWO STAGE OPTIMIZATION TO FIND POSTERIOR DISTRIBUTION AND PREDICTING RESULTS

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS

OPTIMIZATION :

$\mathrm{GP}\left(\mu\ , \acute{\Sigma}_{x,x}\right)$

- Bayesian margenal likelyhood $\mathcal{P}(\boldsymbol{\mathcal{E}}|\boldsymbol{\theta}) = \int \mathcal{P}(\boldsymbol{\mathcal{E}}|\boldsymbol{w})\mathcal{P}(\boldsymbol{w}|\boldsymbol{\theta})d\boldsymbol{w}$

- Minimisation with respect to hyper-parameter (Parameter of kernels)

$$\frac{\partial}{\partial\theta_i} log(\boldsymbol{\mathcal{E}}|\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\mathcal{E}}^T K^{-1}\frac{\partial K}{\partial\theta_i}K^{-1}\boldsymbol{\mathcal{E}} - \frac{1}{2}\mathrm{tr}\left(K^{-1}\frac{\partial K}{\partial\theta_i}\right)$$

PREDICTION :

$$\begin{bmatrix}\boldsymbol{\mathcal{E}}\\ \boldsymbol{\mathcal{E}}_*\end{bmatrix} \sim N\left(0, \begin{bmatrix} K(x,x) & K(x,x_*)\\ K(x_*,x) & K(x_*,x_*)\end{bmatrix}\right)$$

HIERARCHICAL STRUCTURE OF GAUSSIAN PROCESS

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS

Hyper-Parameters

Parameter Value

Minimisation w.r.t Q

- Gradient decent
- Monte Carlo

$$\mathcal{P}(\boldsymbol{\theta}|\mathcal{E}) = \frac{\mathcal{P}(\mathcal{E}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})}{\mathcal{P}(\mathcal{E})}$$

Where $\mathcal{P}(\mathcal{E}) = \int \mathcal{P}(\mathcal{E}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})d\boldsymbol{\theta}$

Gaussian Process Prior

$$\mathcal{P}(\boldsymbol{w}|\mathcal{E},\boldsymbol{\theta}) = \frac{\mathcal{P}(\mathcal{E}|\boldsymbol{w},\boldsymbol{\theta})\mathcal{P}(\boldsymbol{w}|\boldsymbol{\theta})}{\mathcal{P}(\mathcal{E}|\boldsymbol{\theta})}$$

Analytical Integral (function of Q)

Where $\mathcal{P}(\mathcal{E}|\boldsymbol{\theta}) = \int \mathcal{P}(\mathcal{E}|\boldsymbol{w})\mathcal{P}(\boldsymbol{w}|\boldsymbol{\theta})d\boldsymbol{w}$

SAMPLING PRIOR FUNCTIONS FROM RBF KERNEL

FFT

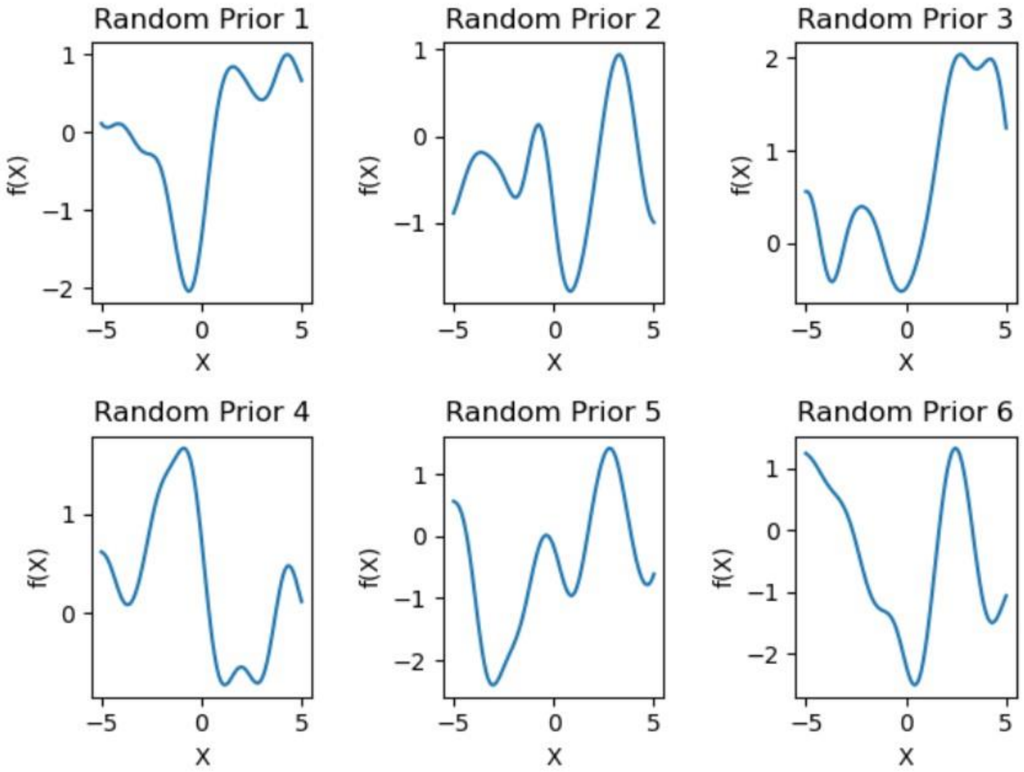
BAYES

MODEL SELECTION

GAUSSIAN PROCESS

RBF Kernel

$$K(x, x_*) \;=\; \theta_1^2 \exp\left(-\frac{(x - x_1)^2}{2\theta_2^2}\right)$$



SAMPLING PRIOR FUNCTIONS FROM MIXED KERNELS

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS

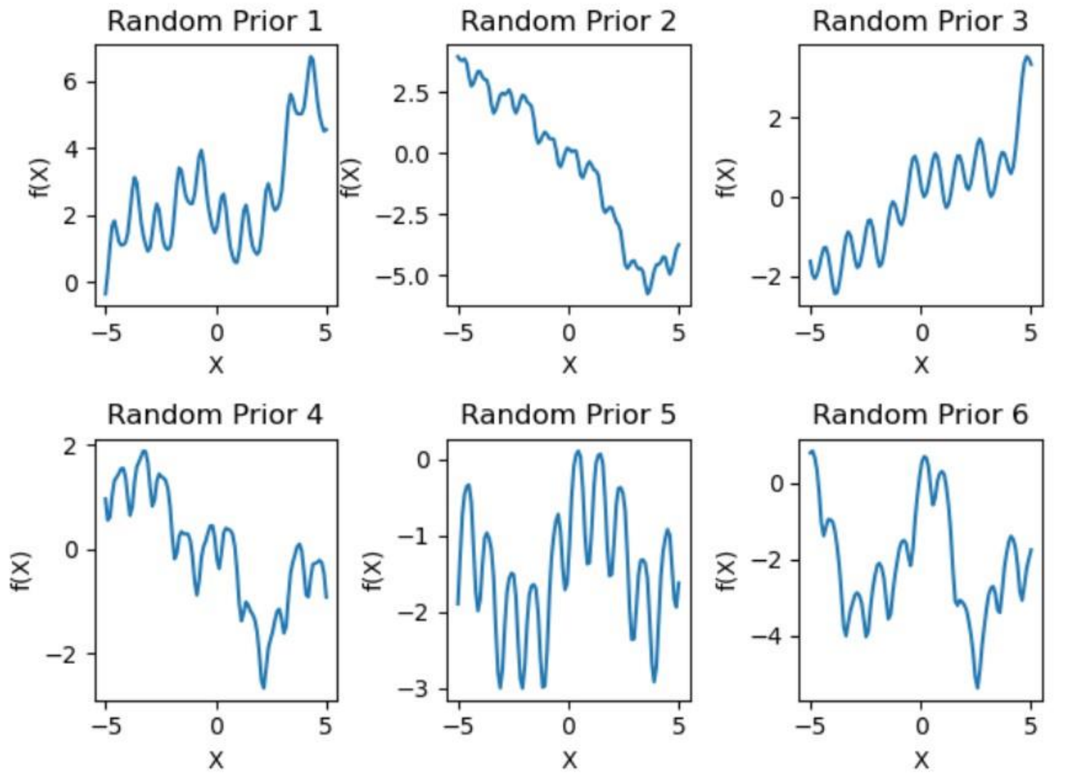
Mixed Kernel

$$K_1(x, x_*) \;=\; \theta_1^2 \exp\left(-\frac{(x-x_1)^2}{2\theta_2^2}\right)$$

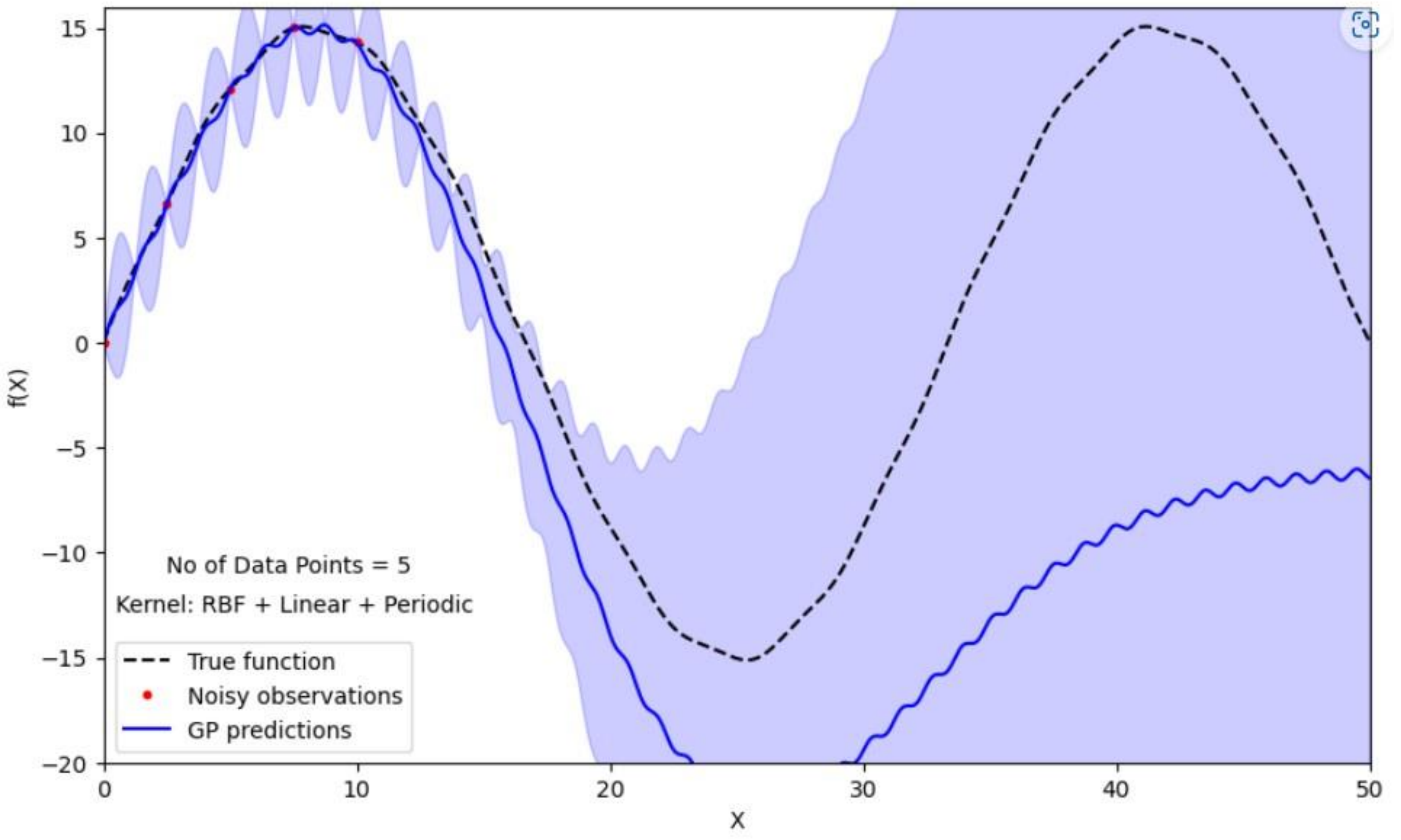
$$K_2(x, x_*) \;= \theta_3^2 \exp\left(-\frac{(x-\acute{x})^2}{2\theta_4^2} - \frac{2sin^2(\pi(x-\acute{x}))}{\theta_5^2}\right)$$

$$K_3(x, x_*) \;= \theta_6^2 \left(1 + \frac{(x-\acute{x})^2}{2\theta_8\theta_7^2}\right)^{-\theta_8}$$

$$K \;=\; K_1 + K_2 + K_3$$



GAUSSIAN PROCESS REGRESSION RESULT WITH 5 DATA POINTS ON PREVIOUS TOY EXAMPLE



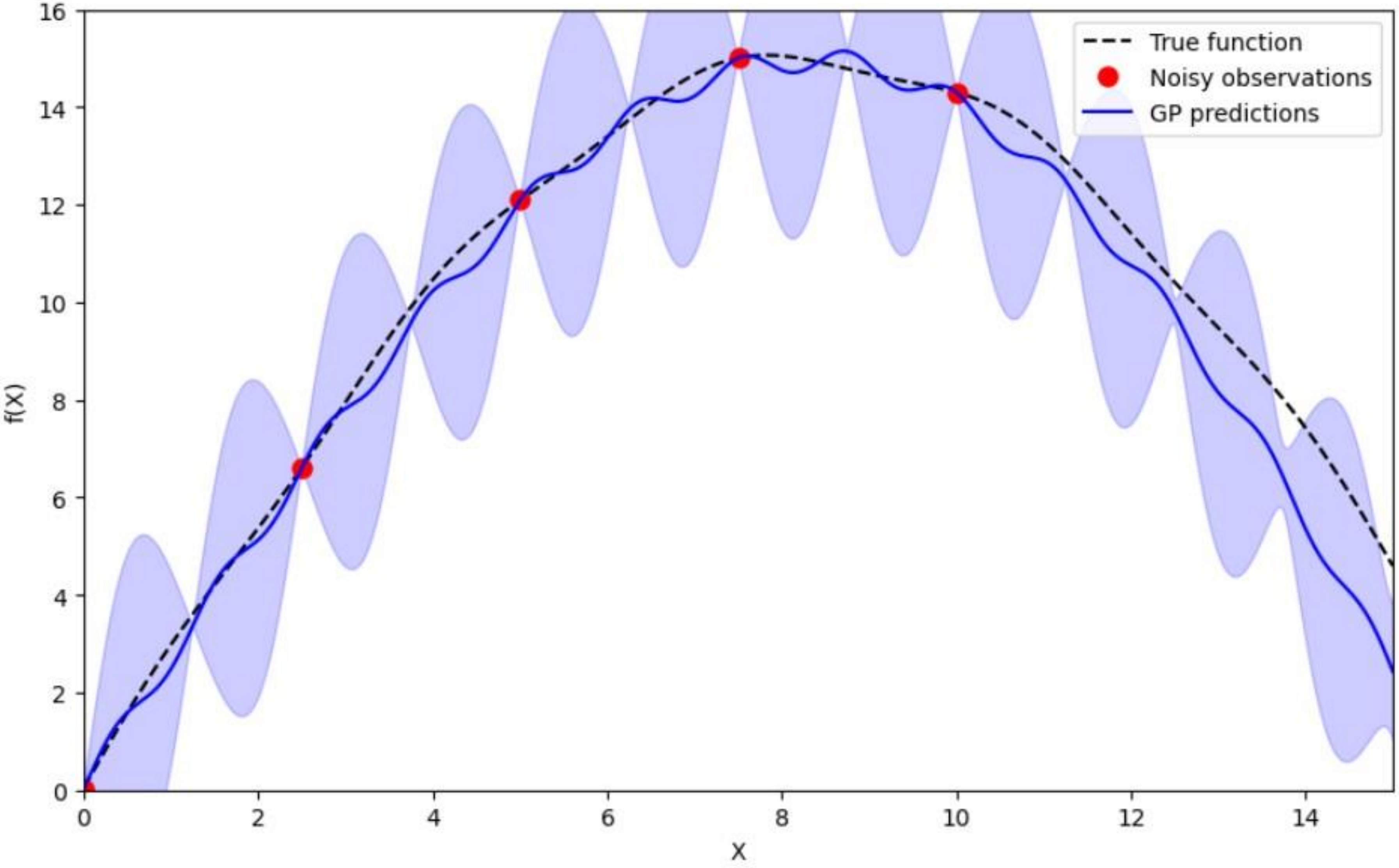
LOOKING CLOSER AT CONFIDENCE INTERVALS OF GPR RESULTS

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



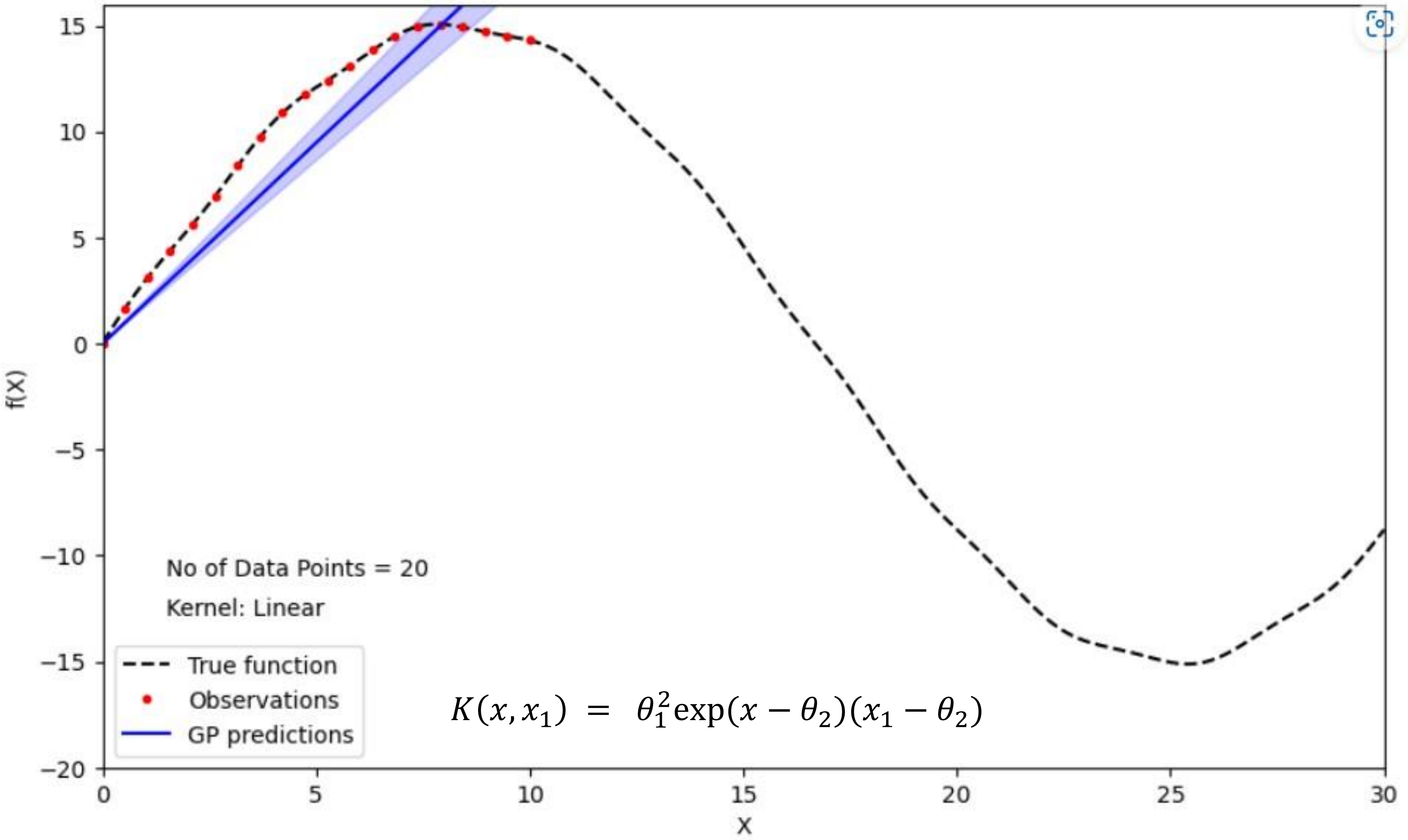
GPR RESULT WITH LINEAR KERNEL

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



$$K(x, x_1) = \theta_1^2 \exp(x - \theta_2)(x_1 - \theta_2)$$

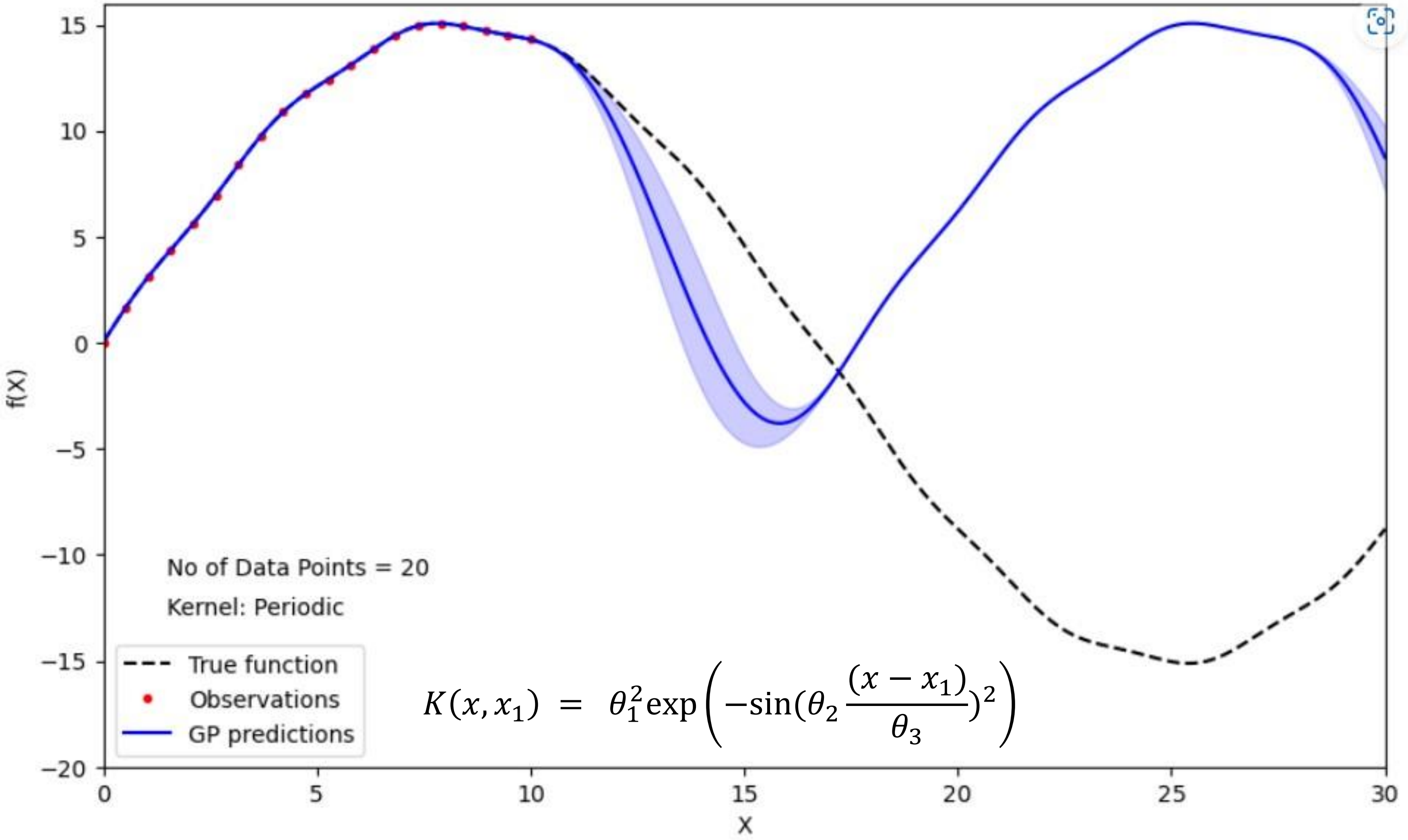
GPR RESULT WITH PERIODIC KERNEL

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



$$K(x,x_1) = \theta_1^2 \exp\left(-\sin(\theta_2 \frac{(x-x_1)}{\theta_3})^2\right)$$

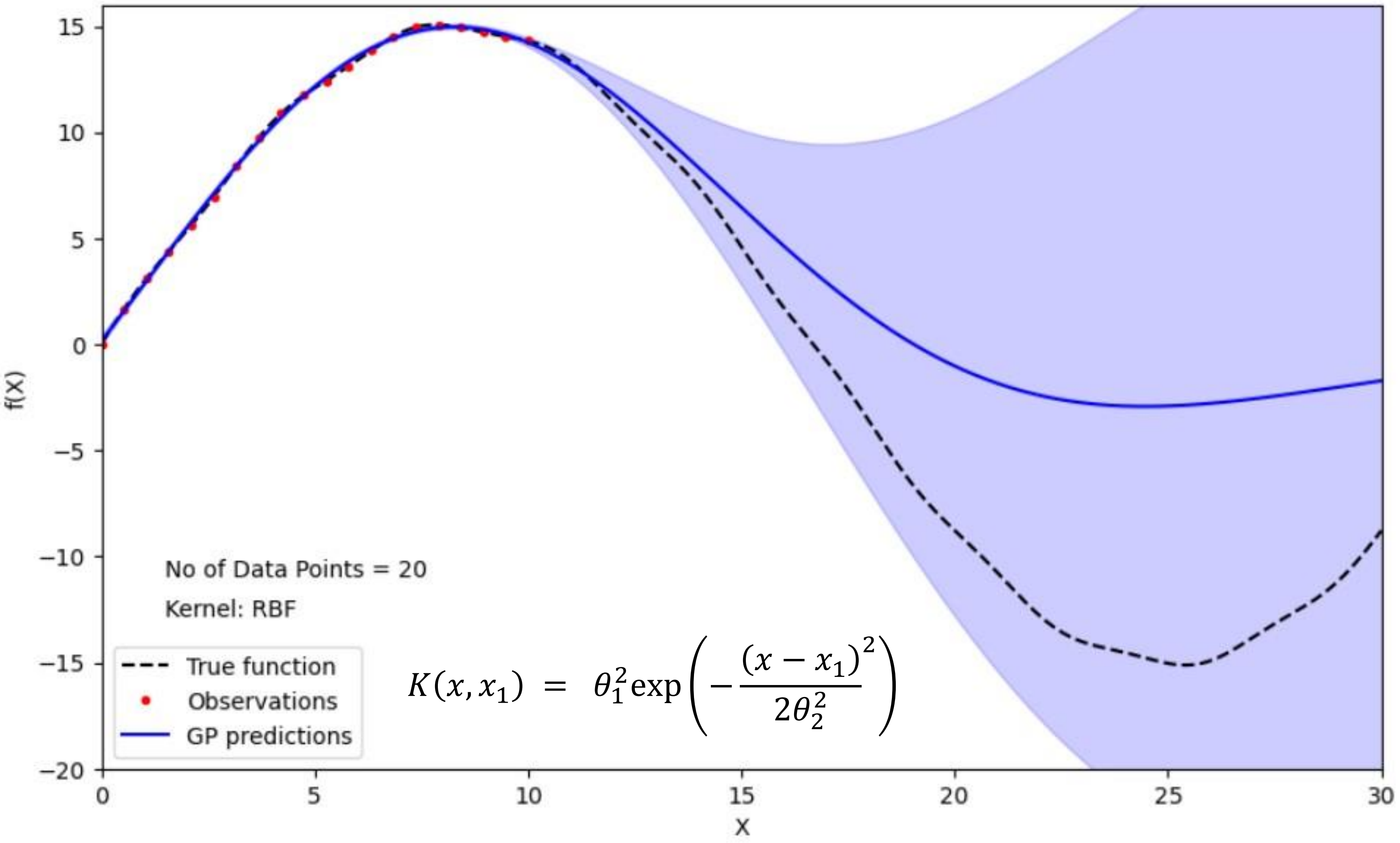
GPR RESULT WITH RBF KERNEL

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



$$K(x, x_1) = \theta_1^2 \exp\left(-\frac{(x - x_1)^2}{2\theta_2^2}\right)$$

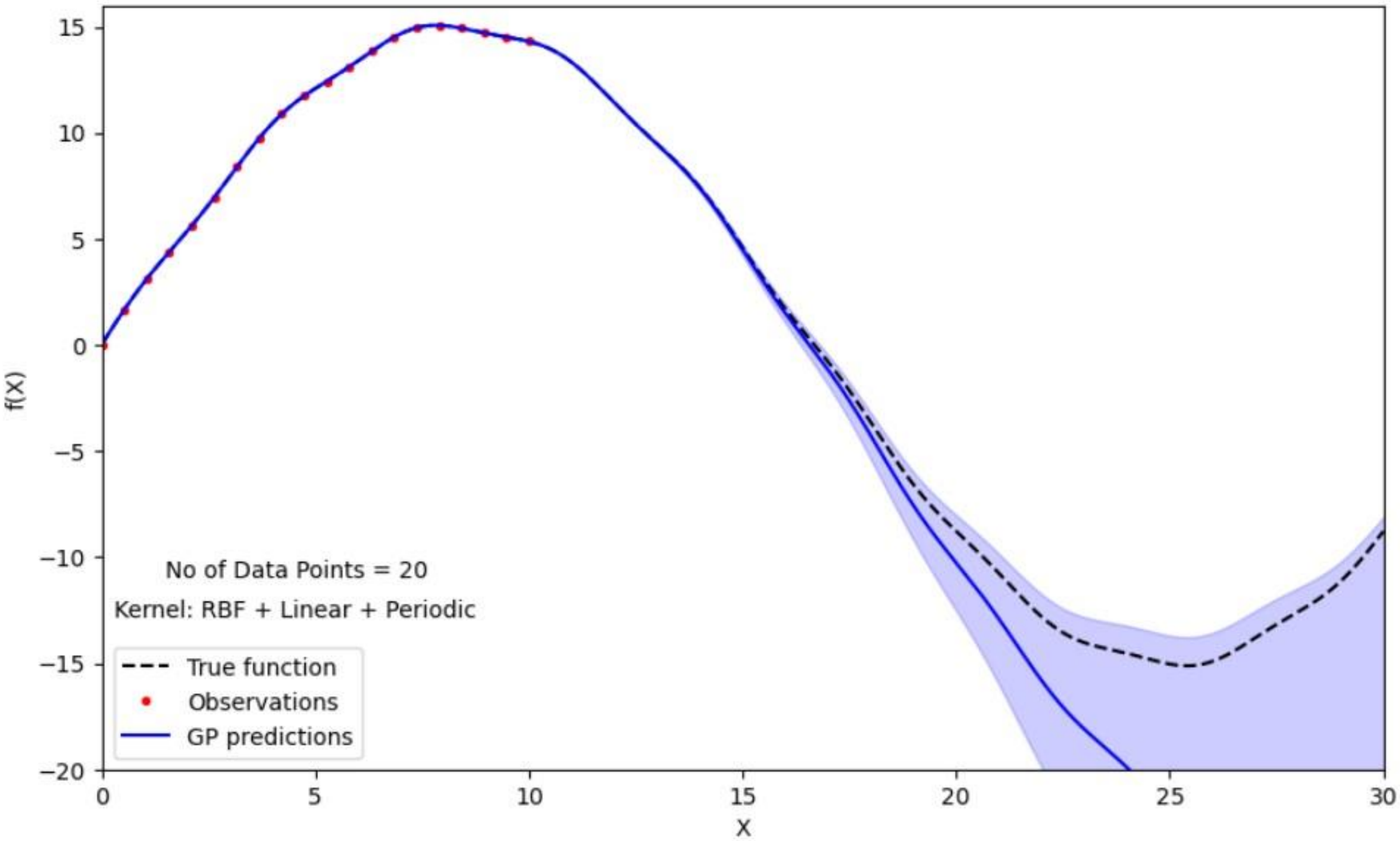
GPR RESULT WITH MIXED KERNEL

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



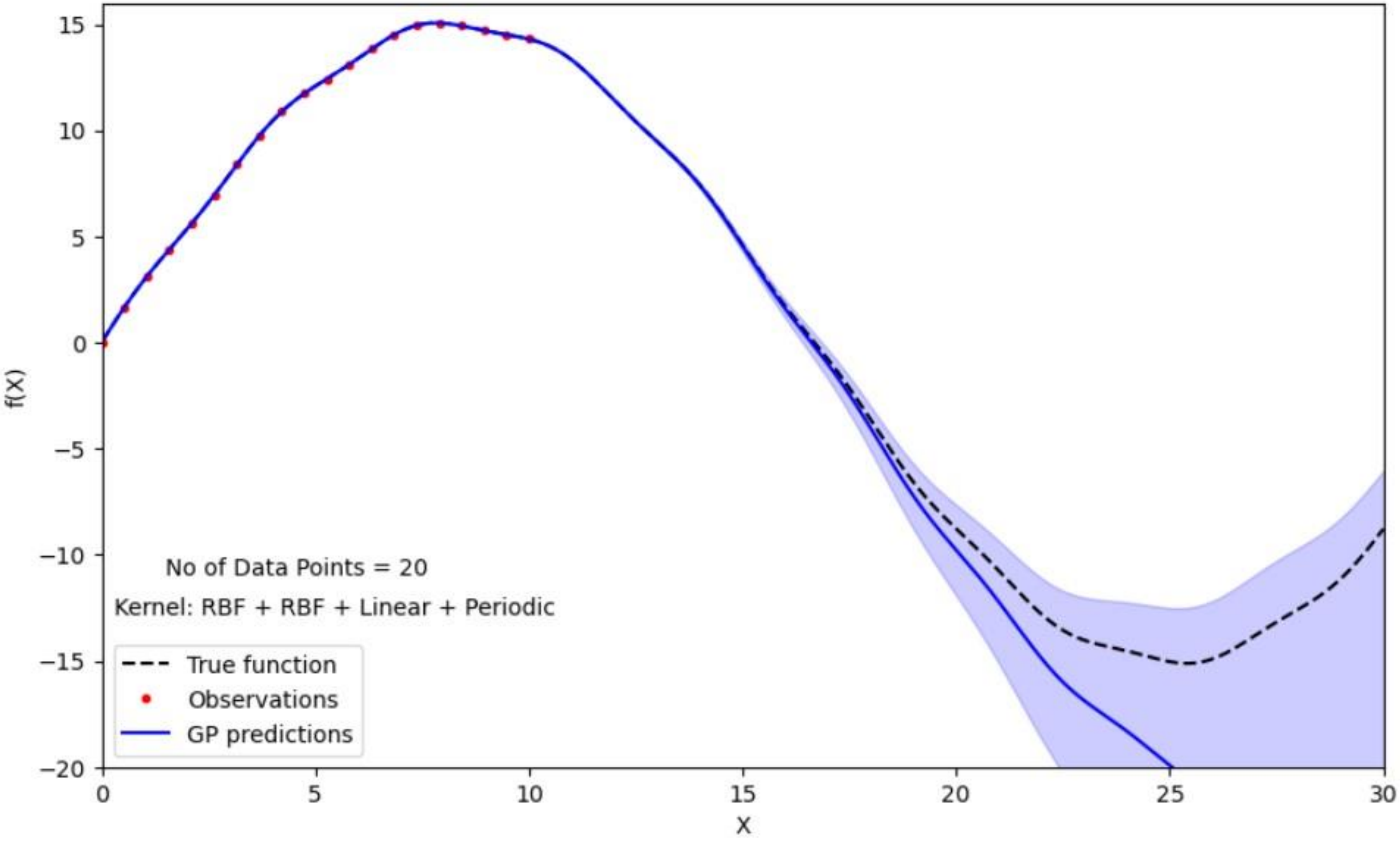
RESULTS WITH EVEN COMPLEX KERNELS

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



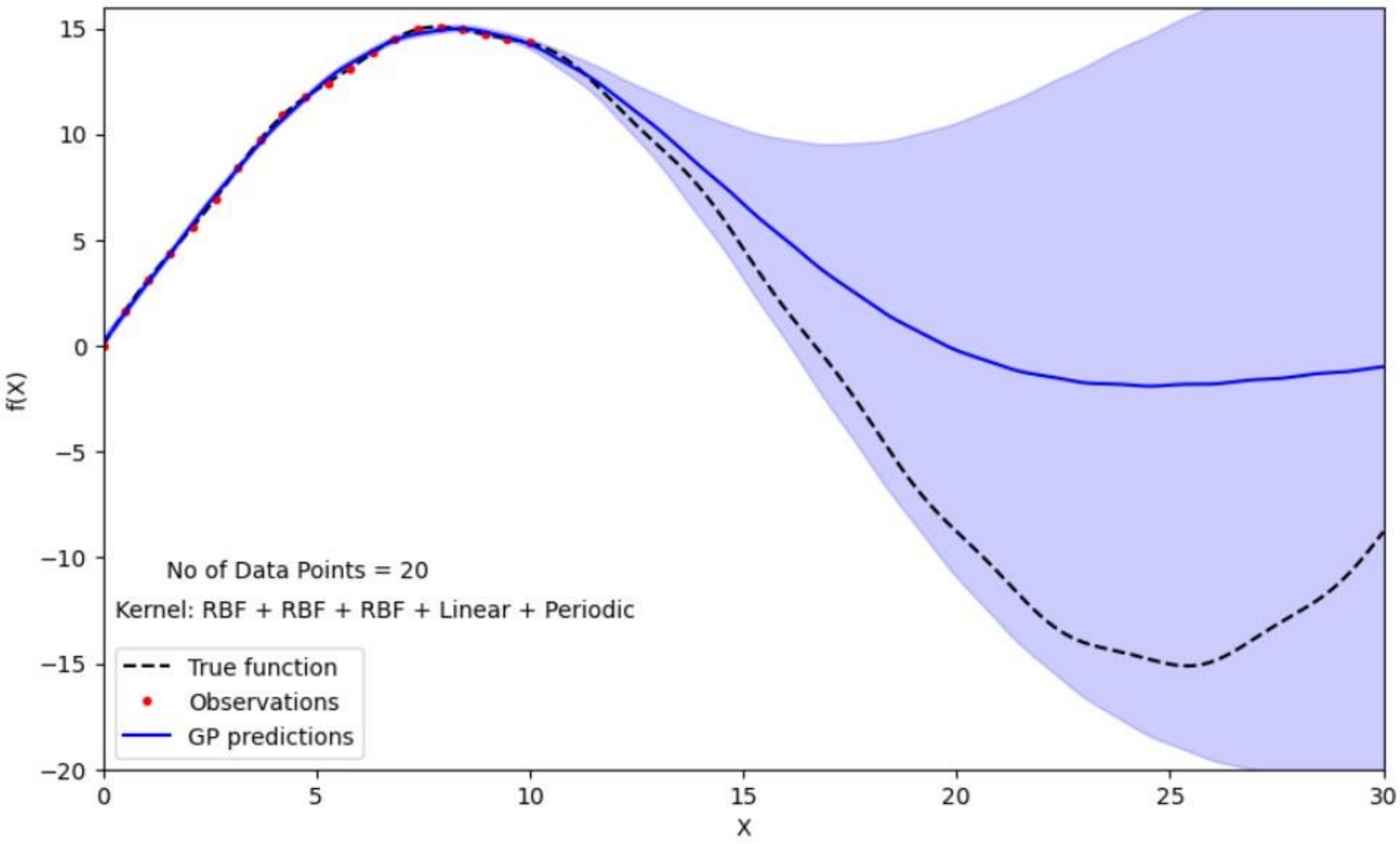
*RESULTS WITH EVEN COMPLEX KERNELS

FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS



MODEL SELECTION WITH GAUSSIAN PROCESS REGRESSION

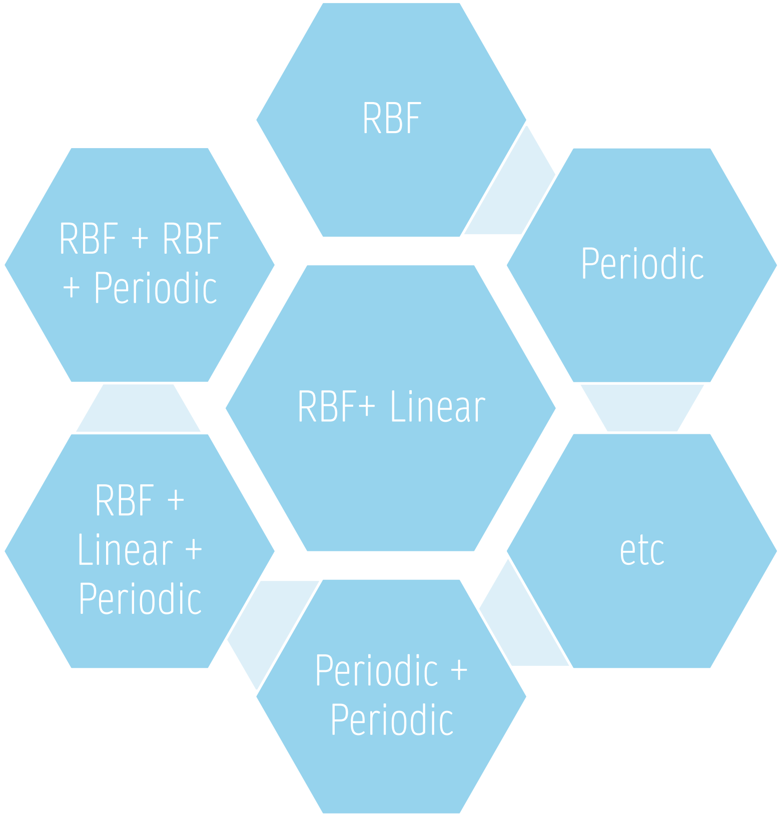
FFT

BAYES

MODEL SELECTION

GAUSSIAN PROCESS

MODEL SELECTION :



Factors to consider

- Likelihood
- Information Criterion
- Prior Knowledge
- Intuition

AIM OF THE PROJECT

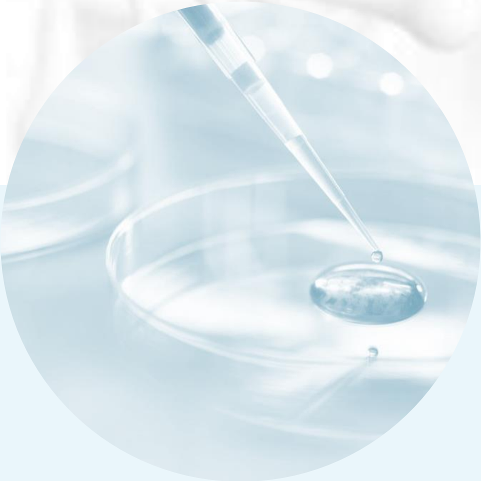
- FFT
- BAYES
- MODEL SELECTION
- GAUSSIAN PROCESS
- CONCLUSION

Train model → Domain Adaptation → Classification

- White Noise
- Model selection
- Hyperparameter interpretation

- Model adaptation
- Hyperparameter Bound

When Is A Noisy Signal Oscillating And When is it Noisy?



THANK YOU

RITESH BALAYAN
S2456148@ED.AC.UK