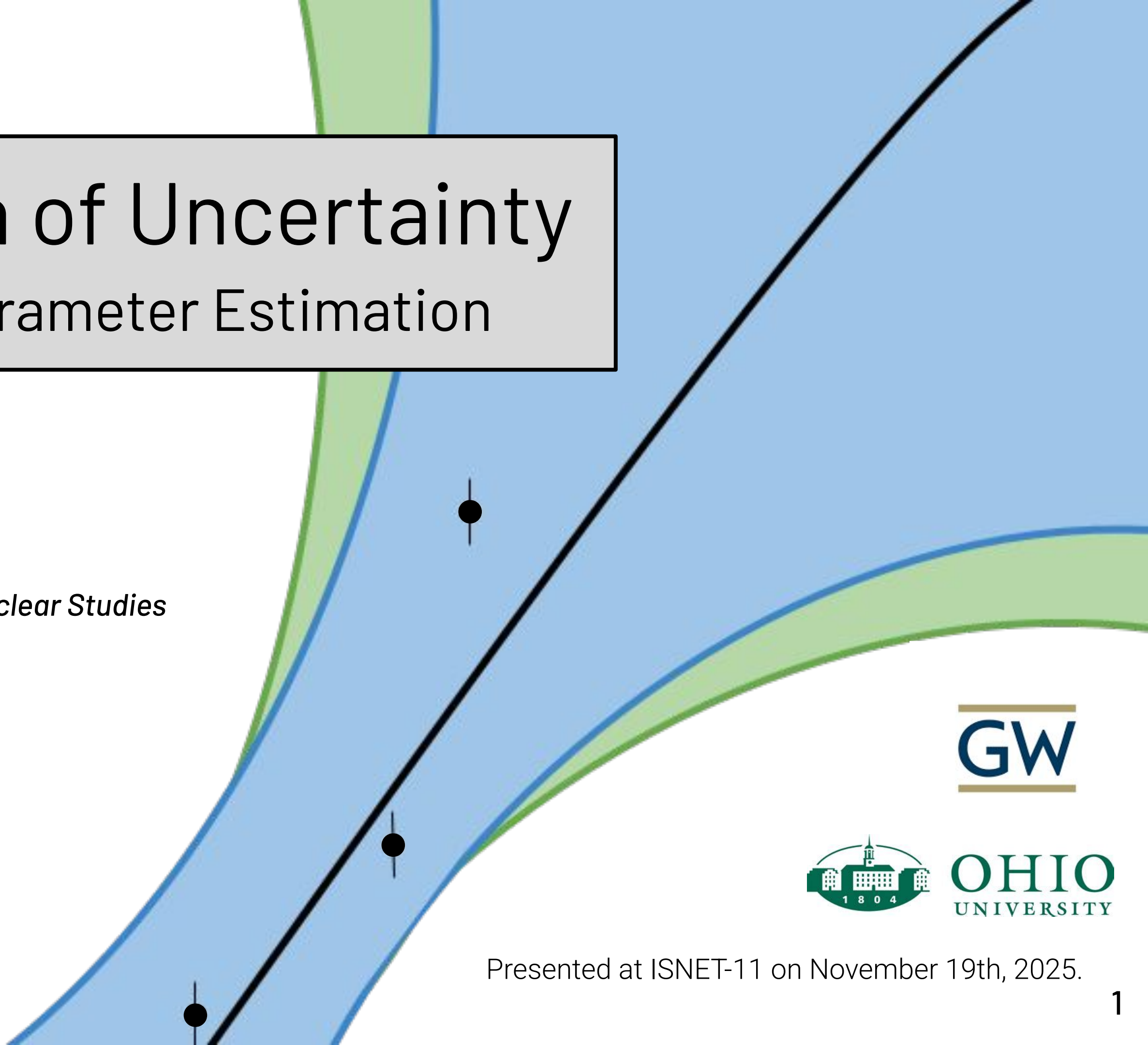
Overestimation of Uncertainty

In Bayesian EFT Parameter Estimation



Nathan Carter

Department of Physics and the Institute for Nuclear Studies

George Washington University

Washington, DC , USA

Collaborators:

Daniel Phillips

Richard Furnstahl

Jordan Melendez

Presented at ISNET-11 on November 19th, 2025.

Data from Above

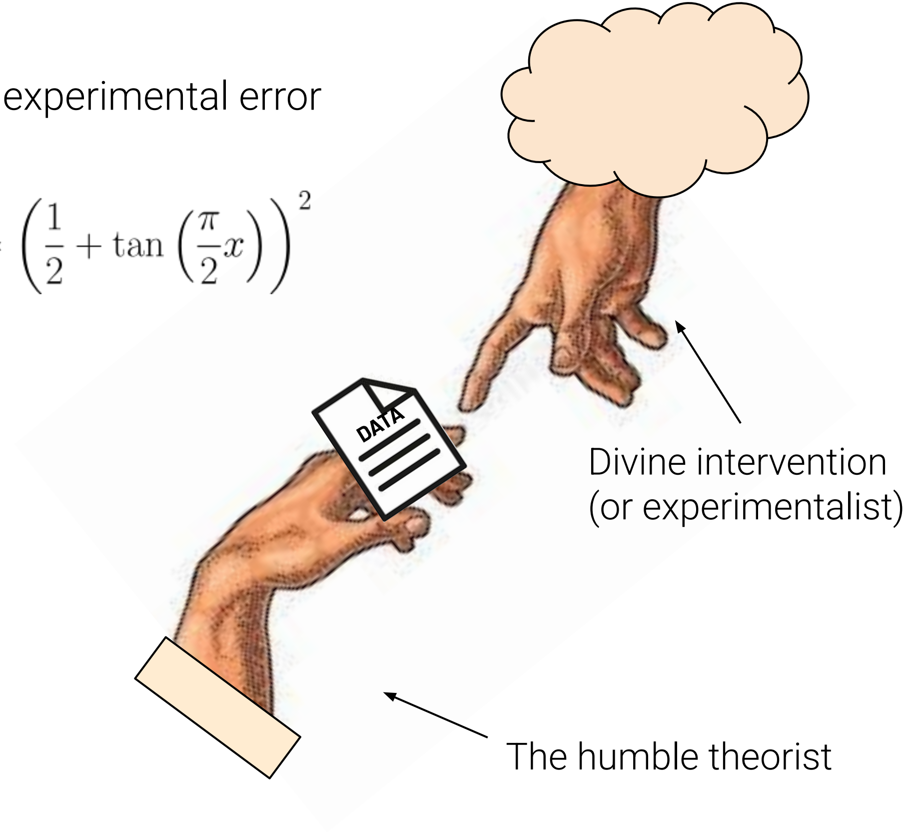
And on the 3rd day of ISNET-11, God said
“Let there be data.“ - Bayes 3:14159...

Generated from a specific function with a 5% experimental error on the points.

X	Y	σ_{exp}
0.03183	0.31694	0.01585
0.06366	0.33844	0.01692
0.09549	0.42142	0.02107
0.12732	0.57709	0.02885
0.15915	0.56218	0.02811
0.19099	0.68851	0.03443
0.22282	0.73625	0.03681
0.25465	0.8727	0.04364
0.28648	1.0015	0.050075
0.31831	1.0684	0.05342

Table 1: Data given.

$$g(x) = \left(\frac{1}{2} + \tan\left(\frac{\pi}{2}x\right)\right)^2$$



Statistical Modeling

With our God-given data, let g(x) represent the underlying theory, i.e. the true values.
Then,

$$g(x) = y_i - \delta y_{\text{expt}}(x)$$

$$\delta y_{\text{expt}}(x_i) \sim \mathcal{N}(0, [0.05 g(x_i)]^2)$$

$$g(x) = f_k(x_i; \mathbf{a}) + \delta f_k(x_i)$$

model discrepancy from truncation of the polynomial (might call it truncation uncertainty), can't calculate to infinite *k* unfortunately...

From EFT considerations, model discrepancy is dominated by the higher order terms,

$$\delta f_k(x) \sim \mathcal{N}(0, \bar{a}^2 x^{2k+2})$$

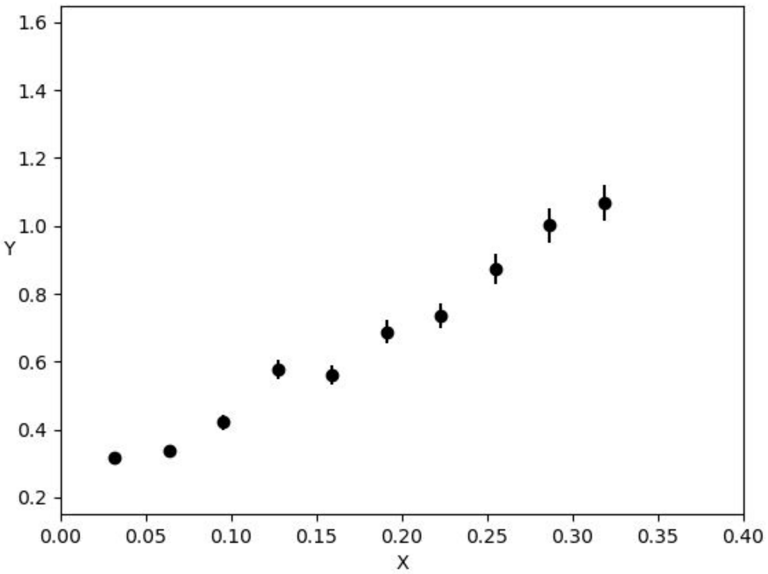


Figure 1: Data given.

Modeling like Brook Taylor Intended

How might we model this data?

Luckily, Taylor Series exist.

Loosely speaking, continuous functions can be represented as a polynomial, safe choice!

$$f_k(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k$$

Coordinates well with ideas of Low Energy Constants in Effective Field Theory for a small expansion parameter.

$$x \ll 1$$

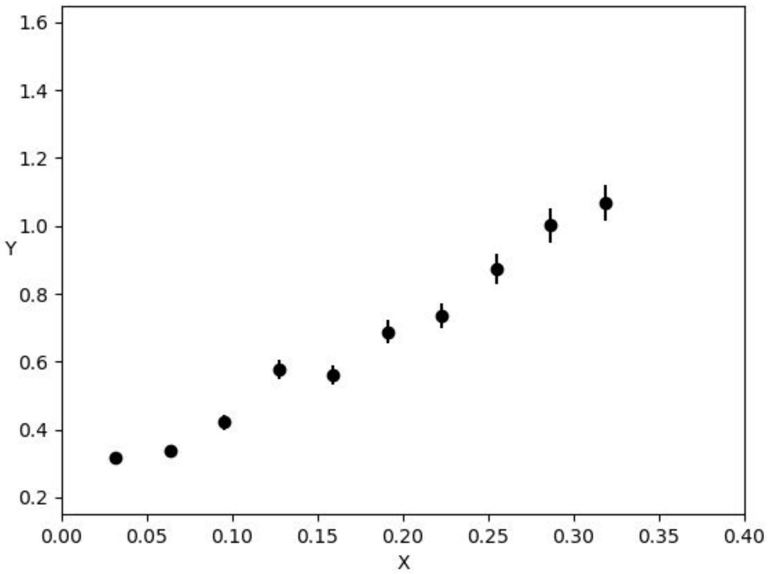


Figure 1: Data given.

Forming the Posterior

Bayes Theorem:

$$\mathrm{prob}(\mathbf{a}|D,k,I) \propto \underbrace{\mathrm{prob}(D|\mathbf{a},k,I)}_{\text{Likelihood}}\underbrace{\mathrm{pr}(\mathbf{a}|k,I)}_{\text{Prior}}$$

Assume *N* independent data points, then the likelihood can be represented by the product of *N* Gaussian distributions.

$$\mathrm{prob}(D|\mathbf{a},k,I) \propto \exp\left(\frac{-\chi^2}{2}\right)$$

where

$$\chi^2 = \sum_{i=1}^{N}\left(\frac{y_i - f_k(x_i;\mathbf{a}) - \delta f_k(x_i)}{\sigma_i}\right)^2$$

Expect "natural" parameters.

$$\mathrm{pr}(\mathbf{a}|I) = \exp\left(-\frac{1}{2}\frac{\mathbf{a}^2}{\bar{a}^2}\right)$$

The prior's parameter, $\bar{a}$, determines the extent which the parameters can deviate from their mean value of 0.

Using both Bayes' Theorem and Marginalization, the posterior is

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \mathrm{pr}(g(x_{n+1})|a_{k+1}, f_k(x_{n+1}), I)\mathrm{pr}(a_{k+1}, f_k(x_{n+1})|D,k,I)da_{k+1}df_k(x_{n+1})$$

Now apply the polynomial model, then

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \mathrm{pr}\Big(g(x_{n+1})|\sum_{i=0}^{k+1} a_i x_{n+1}^i, I\Big)\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k|D,I)da_{k+1}da_0 \ldots da_k$$

Because of the relationship between g(x), the model, and the model discrepancy,

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \delta\Big(g(x_{n+1}) - \sum_{i=0}^{k+1} a_i x_{n+1}^i\Big)\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k|D,I)da_{k+1}da_0 \ldots da_k$$

Using both Bayes' Theorem and Marginalization, the posterior is

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \mathrm{pr}(g(x_{n+1})|a_{k+1}, f_k(x_{n+1}), I)\mathrm{pr}(a_{k+1}, f_k(x_{n+1})|D,k,I)da_{k+1}df_k(x_{n+1})$$

Now apply the polynomial model, then

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \mathrm{pr}\Big(g(x_{n+1})|\sum_{i=0}^{k+1} a_i x_{n+1}^i, I\Big)\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k|D,I)da_{k+1}da_0\ldots da_k$$

Because of the relationship between g(x), the model, and the model discrepancy,

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \delta\Big(g(x_{n+1}) - \sum_{i=0}^{k+1} a_i x_{n+1}^i\Big)\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k|D,I)da_{k+1}da_0\ldots da_k$$

Let us now assume the truncation term's parameter is independent of all the other parameters…

$$\mathrm{pr}(g(x_{n+1})|D,k,I) = \int \delta\Big(g(x_{n+1}) - \sum_{i=0}^{k+1} a_i x_{n+1}^i\Big) \underline{\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k|D,I)} da_{k+1} da_0 \ldots da_k$$

The red term then factorizes:

$$\boxed{\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k|D,I) = \mathrm{pr}(a_{k+1}|D,I)\mathrm{pr}(a_0, a_1, \ldots, a_k|D,I)}$$

Perform the integral over the parameters, **a**, and replace them by a single integral over a Gaussian distribution for the polynomial model. Insert the prior for the truncation term, such that

$$\mathrm{pr}(g(x_{n+1})|D,k,I) \propto \int \delta\Big(g(x_{n+1}) - f_k - a_{k+1} x_{n+1}^{k+1}\Big) \exp\left(-\frac{1}{2}\frac{(f_k - \langle f_k\rangle)^2}{2\sigma(x_{n+1})^2}\right) \exp\left(-\frac{a_{k+1}^2}{2\bar{a}^2}\right) df_k\, da_{k+1}$$

describes the posterior probability distribution using our polynomial model.

$$\mathrm{pr}(g(x_{n+1})|D,k,I) \propto \int \delta\Big(g(x_{n+1}) - f_k - a_{k+1}x_{n+1}^{k+1}\Big) \exp\left(-\frac{1}{2}\frac{(f_k - \langle f_k\rangle)^2}{2\sigma(x_{n+1})^2}\right) \exp\left(-\frac{a_{k+1}^2}{2\bar{a}^2}\right) df_k\, da_{k+1}$$

Now, since we defined $\sigma_{\mathrm{H.O.}} \equiv \bar{a}x^{k+1}$

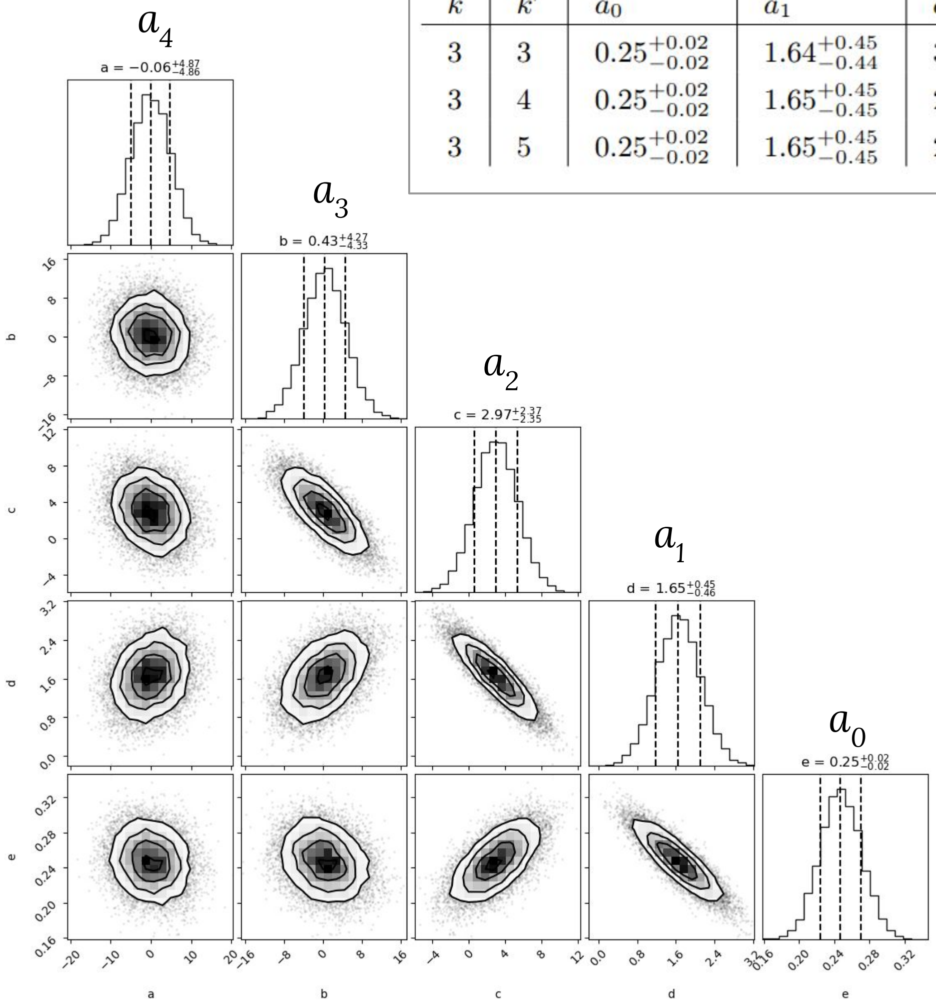
then upon performing the Gaussian integral, the full posterior will have the width

$$\boxed{\sigma_{\mathrm{full}} = \sqrt{\sigma_{H.O.}^2 + \sigma_{\mathrm{param}}^2}}$$

What is to say our original assumption was needed?

We can sample the original posterior before the assumption.

Sampling…



k	k'	a_0	a_1	a_2	a_3	a_4	a_5
3	3	$0.25^{+0.02}_{-0.02}$	$1.64^{+0.45}_{-0.44}$	$3.0^{+2.3}_{-2.3}$	$0.35^{+4.3}_{-4.3}$	–	–
3	4	$0.25^{+0.02}_{-0.02}$	$1.65^{+0.45}_{-0.45}$	$2.99^{+2.34}_{-2.38}$	$0.32^{+4.37}_{-4.28}$	$0.06^{+4.9}_{-4.9}$	–
3	5	$0.25^{+0.02}_{-0.02}$	$1.65^{+0.45}_{-0.45}$	$2.95^{+2.34}_{-2.38}$	$0.32^{+4.37}_{-4.28}$	$0.07^{+4.80}_{-4.90}$	$0.02^{+4.93}_{-4.98}$

We use the python module "emcee" (an affine-invariant MCMC Ensemble sampler).

Use the prior's parameter, $\bar{a}$, of 5.

(For analysis, stop at **k = 4**)

Results of Emcee

Central value (**dark blue line**) and the total uncertainty (**blue 1σ band**).

Let's deconstruct this.

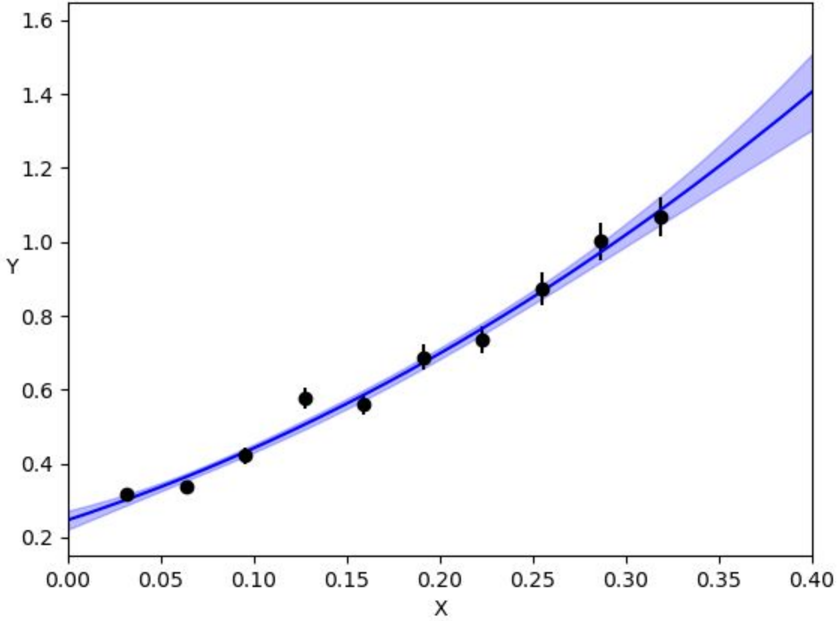


Figure 2: Posterior Probability Distribution for the full model.

Calculating Truncation Uncertainty (Model Discrepancy) and Parametric Uncertainty

After obtaining the samples, we construct the parametric uncertainty by setting the truncation term's parameter to 0 and with those samples compute the sum

$$f_k(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k$$

(for each sample!). Then, calculate the mean and standard deviation of $f_k(x; \mathbf{a})$ over the samples. The effect of higher-order coefficients will be indirectly included since those higher-order coefficients were considered during sampling process.

The truncation uncertainty is then, in a more general case to as we defined before,

$$\sigma_{\mathrm{H.O.}} = \sqrt{\langle a_{k+1}^2 x^{2k+2} \rangle}$$

where the average is taken over the samples.

Parametric Uncertainty vs. Truncation Uncertainty

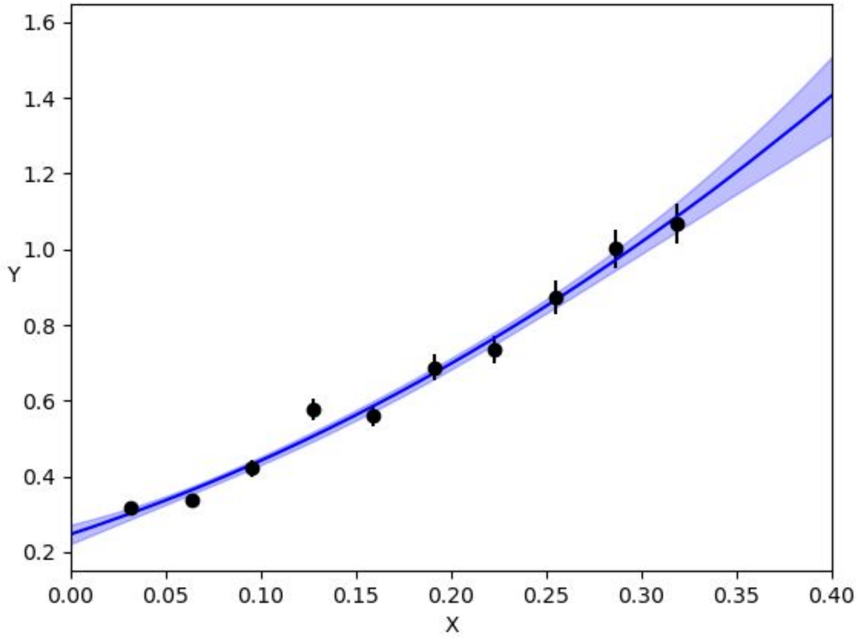


Figure 2: Posterior Probability Distribution for the full model.

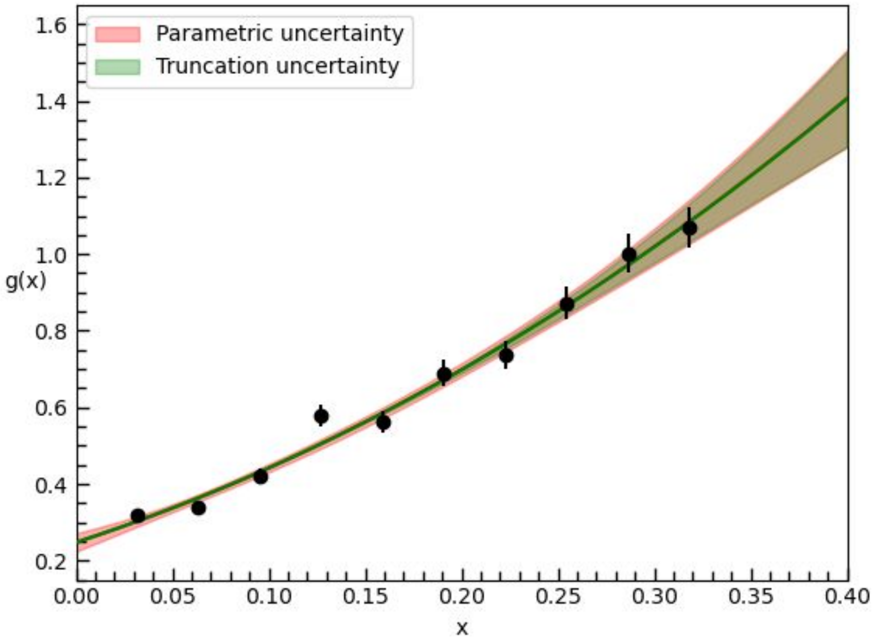


Figure 3: Posterior Probability Distribution for the Truncation model (green) and the Parametric model (red).

Parametric Uncertainty vs. Truncation Uncertainty

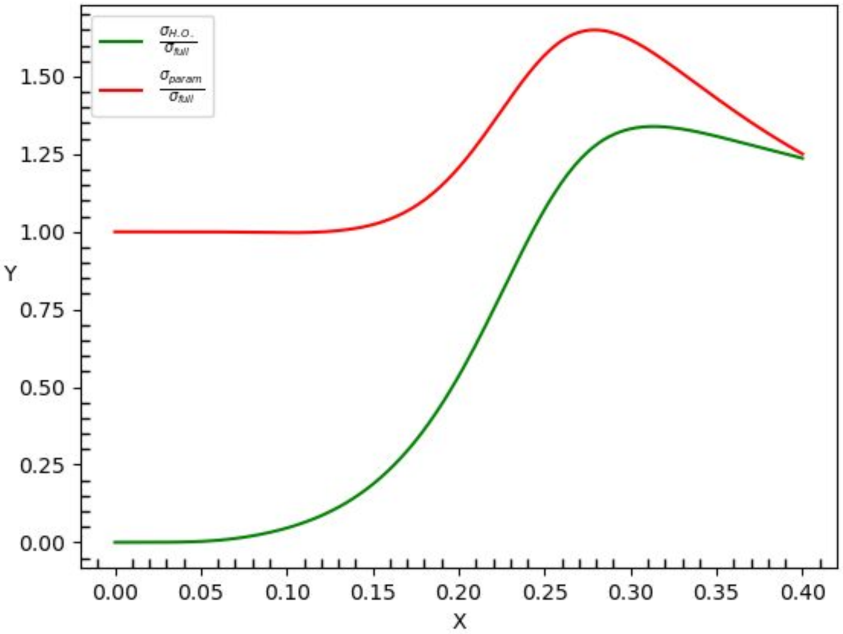


Figure 4: Truncation error divided by the full uncertainty (green) and the Parametric uncertainty divided by the full uncertainty (red).

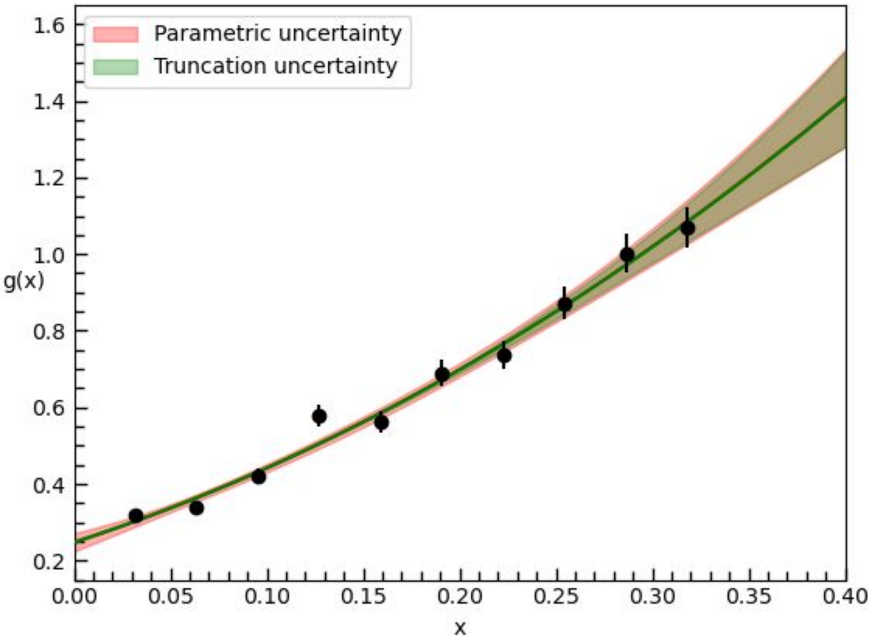


Figure 3: Posterior Probability Distribution for the Truncation model (green) and the Parametric model (red).

Correlation of Uncertainties

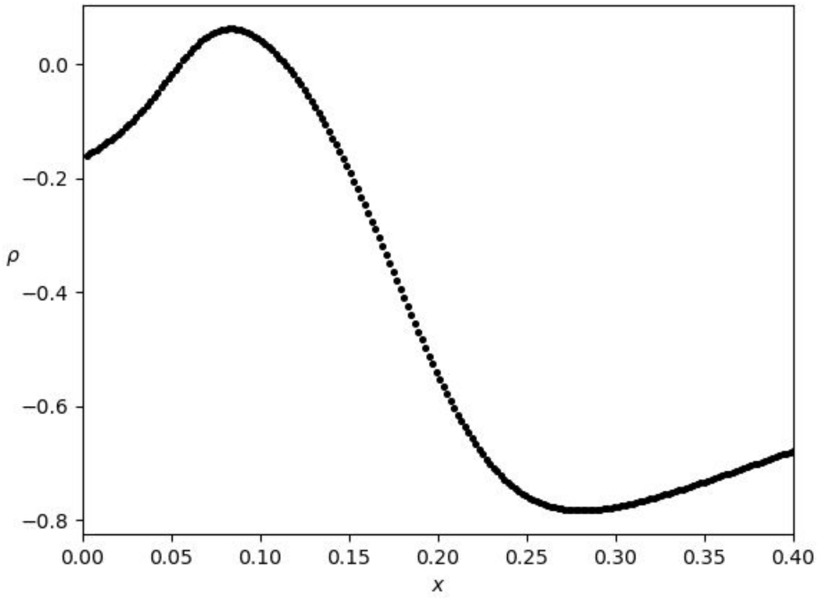


Figure 5: Correlation coefficient for the domain.

The correlation coefficient can be calculated as,

$$\rho(x) = \frac{\left\langle (f_k(x) - \langle f_k(x) \rangle)(a_{k+1}x^{k+1}) \right\rangle}{\sigma_{\mathrm{HO}}\sigma_{\mathrm{param}}}$$

If we assume the data has no influence on the truncated portion, i.e. they are independent, then

$$\boxed{\sigma_{\text{full}} = \sqrt{\sigma_{H.O.}^2 + \sigma_{\text{param}}^2}}$$

However, considering correlations,

$$\boxed{\sigma_{\text{full}} = \sqrt{\sigma_{H.O.}^2 + 2\rho(x)\sigma_{H.O.}\sigma_{\text{param}} + \sigma_{\text{param}}^2}}$$

Quadrature vs. Considering Correlation

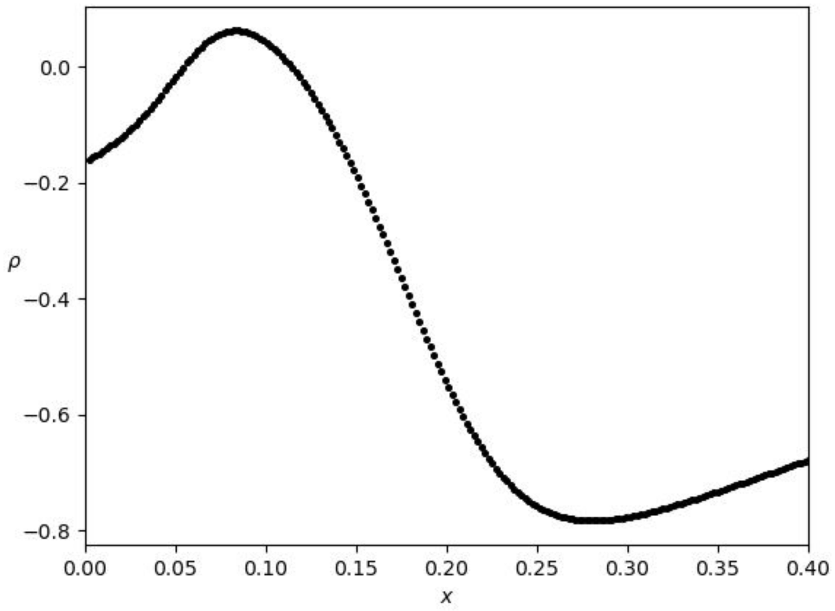


Figure 5: Correlation coefficient for the domain.

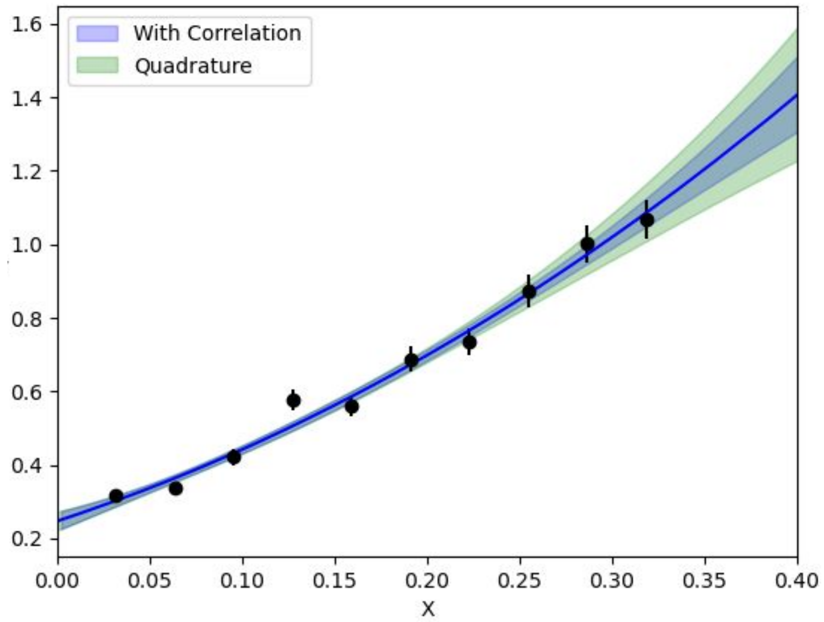


Figure 6: Posterior probability distribution considering correlation (blue) and considering no correlation (green)

Conclusions

Model discrepancy or parameter uncertainty alone cannot give an accurate picture of the full uncertainty of a model prediction, as considering only one can lead to overestimation of the full model uncertainty, assuming appropriate correlations.

In the case of the toy model shown here, the deviation from data that a particular set of these parameters produces is anti-correlated with the truncation error.

Compute the joint posterior of the model parameters and the discrepancy term, i.e. $\mathrm{pr}(a_{k+1}, a_0, a_1, \ldots, a_k | D, I)$, and the truncation uncertainty will be properly accounted for.

Appendix A: 1σ to 3σ Correlation Bands

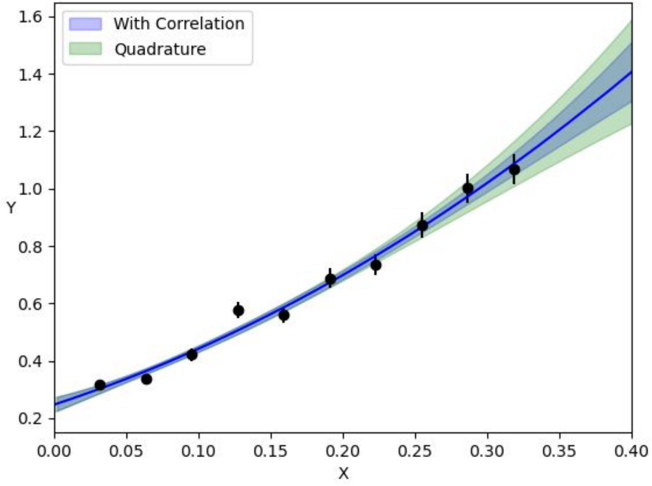


Figure A1: Posterior probability distribution considering correlation in truncation error (blue) and considering no correlation (green) quoting the ***1σ band.***

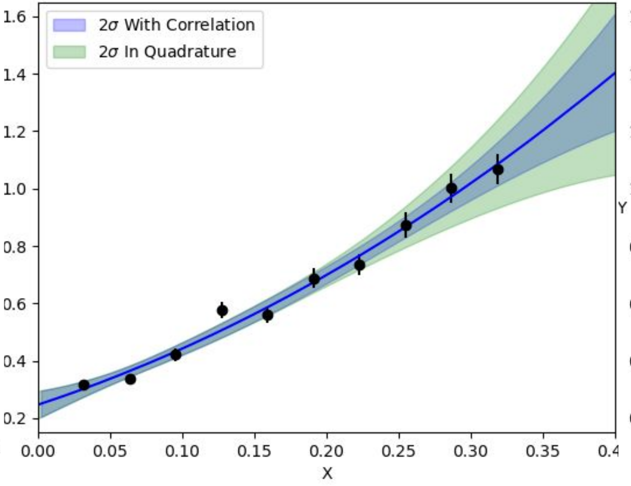


Figure A2: Posterior probability distribution considering correlation in truncation error (blue) and considering no correlation (green) quoting the ***2σ band****.*

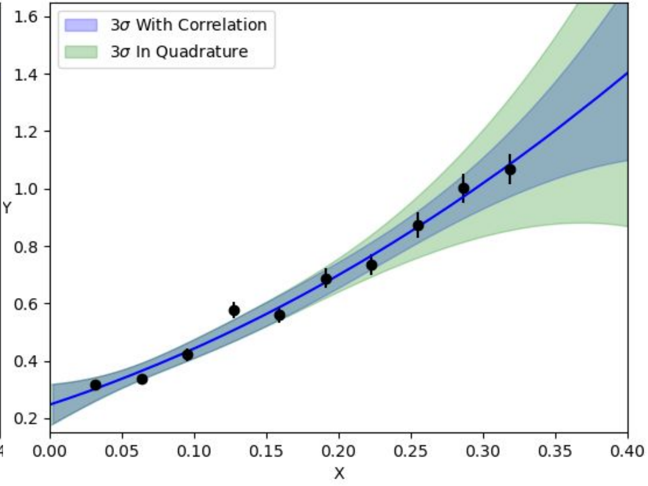


Figure A3: Posterior probability distribution considering correlation in truncation error (blue) and considering no correlation (green) quoting the ***3σ band****.*

Appendix B: Ratio Comparison with Correlation Plot

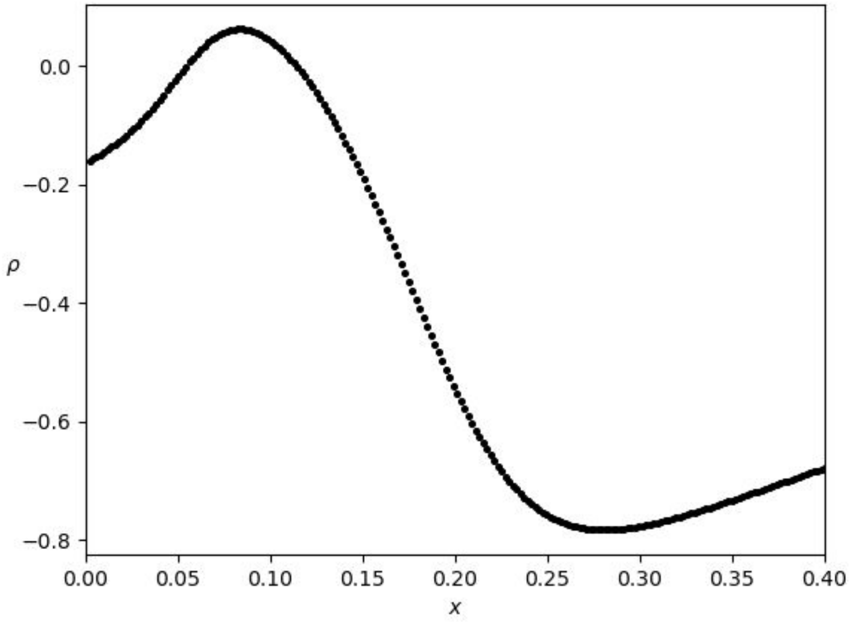


Figure B1: Correlation coefficient for the domain.

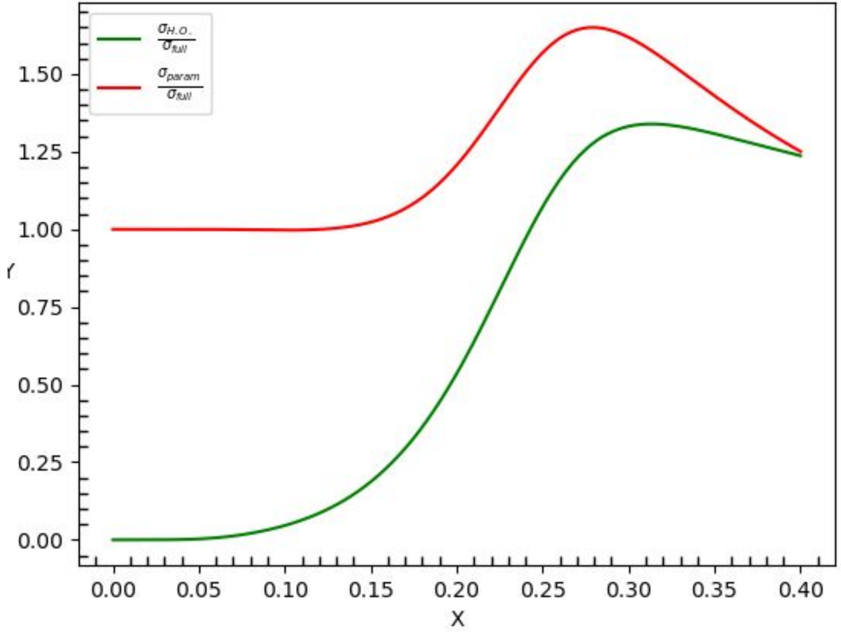


Figure B2: Posterior probability distribution considering correlation in truncation error (blue) and considering independence in correlation (green), with the true function (red)

Appendix C: Ratio Comparison with Correlation/Quadrature

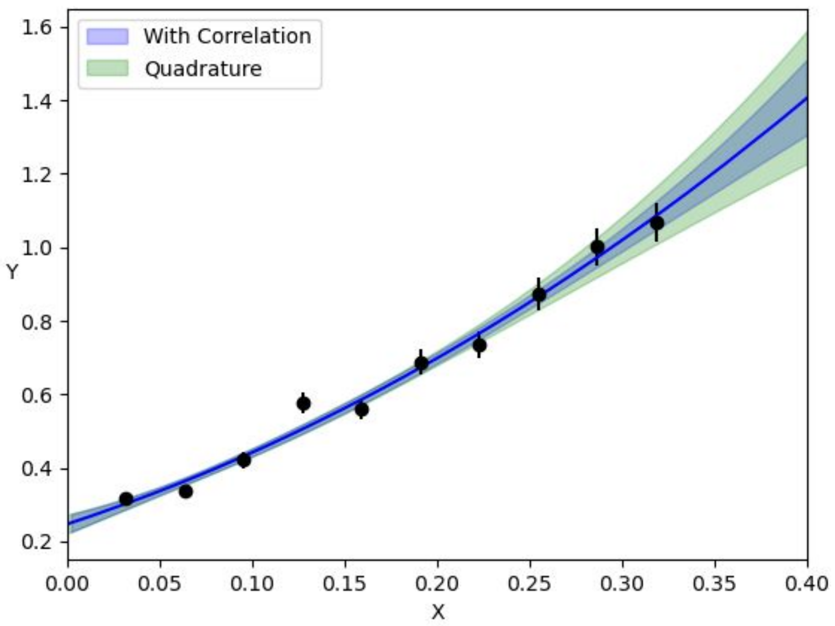


Figure C1: Posterior probability distribution considering correlation in truncation error (blue) and considering no correlation (green)

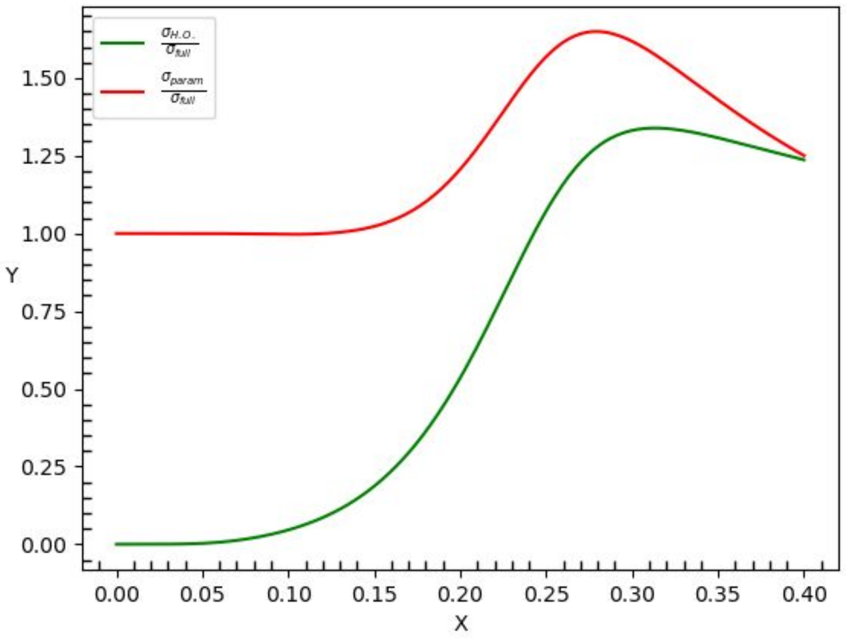


Figure C2: Posterior probability distribution considering correlation in truncation error (blue) and considering independence in correlation (green), with the true function (red)

Appendix D: True Function With Figure 5

$$g(x) = \left(\frac{1}{2} + \tan\left(\frac{\pi}{2}x\right)\right)^2$$

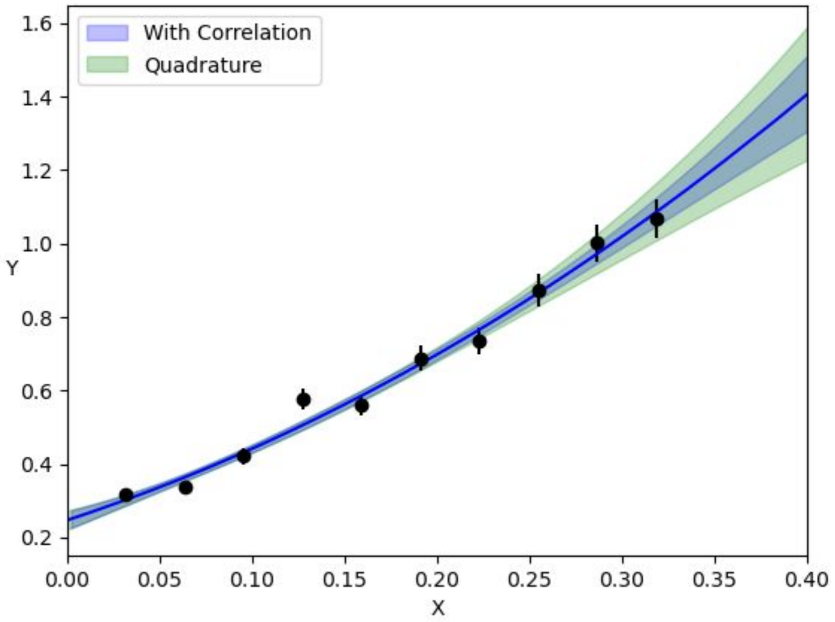


Figure D1: Posterior probability distribution considering correlation in truncation error (blue) and considering no correlation (green)

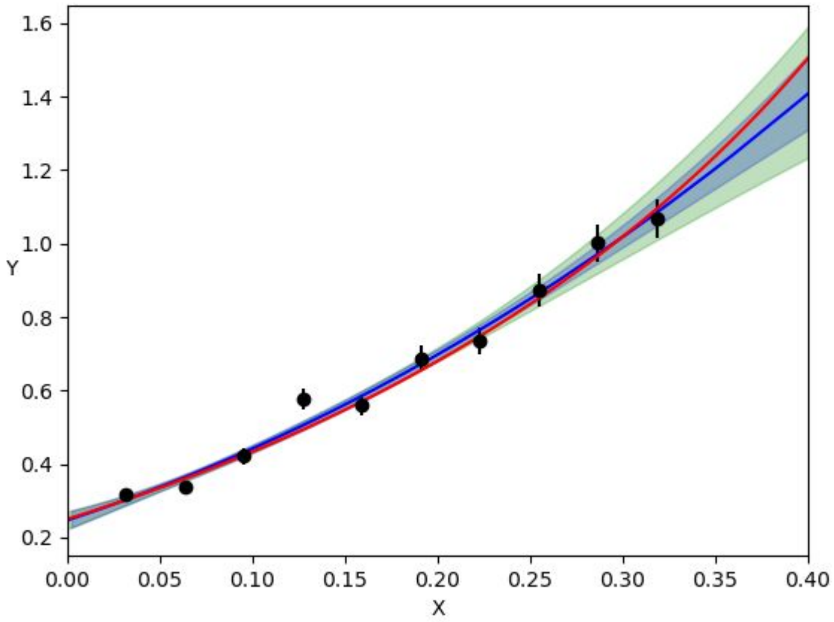


Figure D2: Posterior probability distribution considering correlation in truncation error (blue) and considering independence in correlation (green), with the true function (red)