

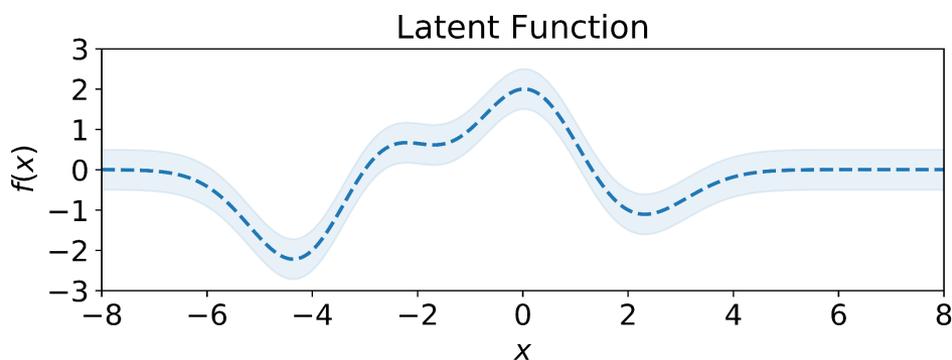
THE UNIVERSITY OF
SYDNEY



Bayesian Deconditional Kernel Mean Embeddings

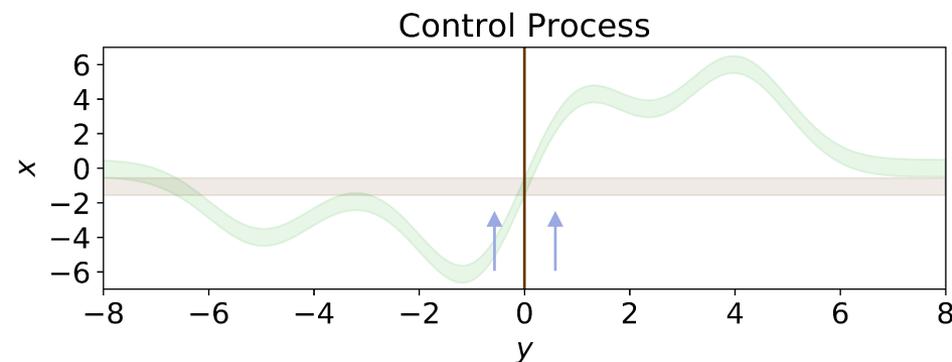
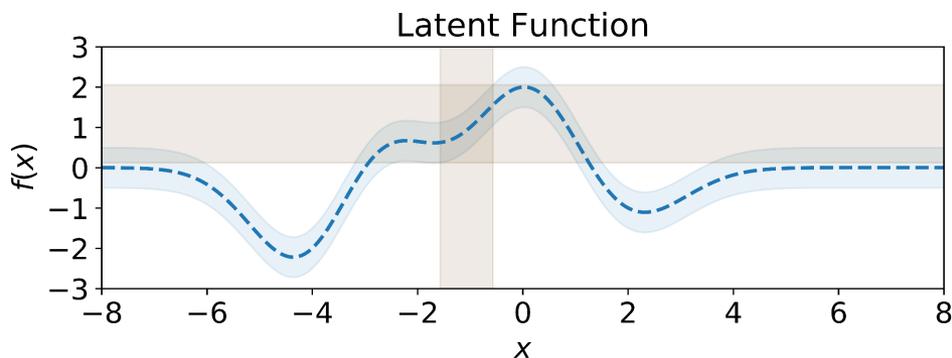
Kelvin Hsu and Fabio Ramos

You want to model some latent phenomenon f

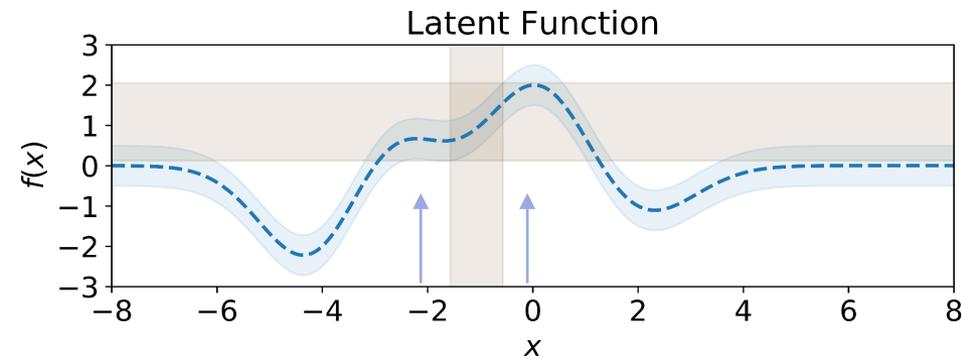


$$Z = f(x) + \text{noise}$$

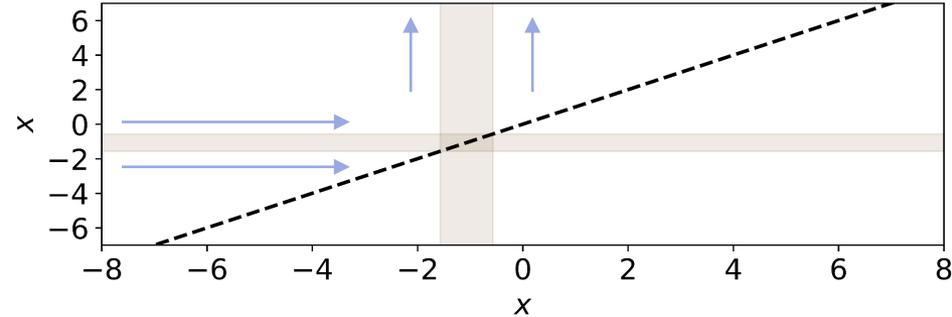
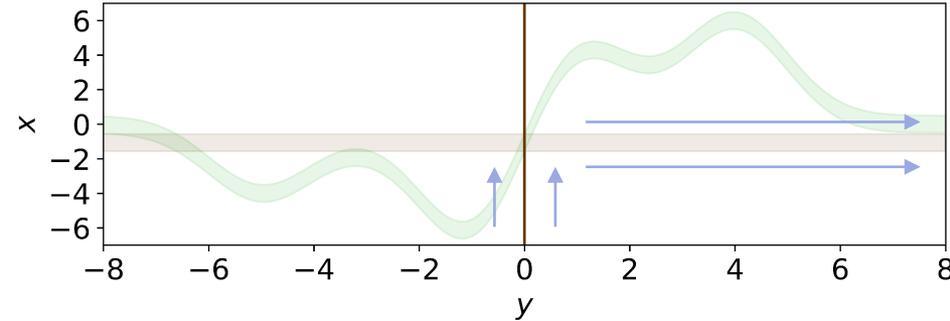
**But! When collecting observations,
you don't get to control x directly!
You only get to control y !**



...which gives some x that goes into f



Control Process

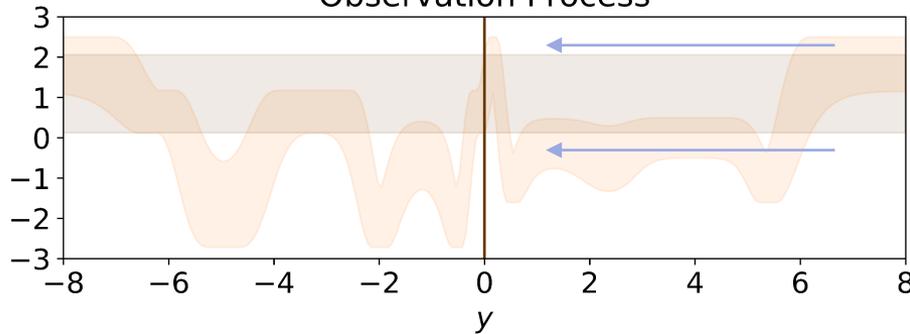


...which finally gives you observations z

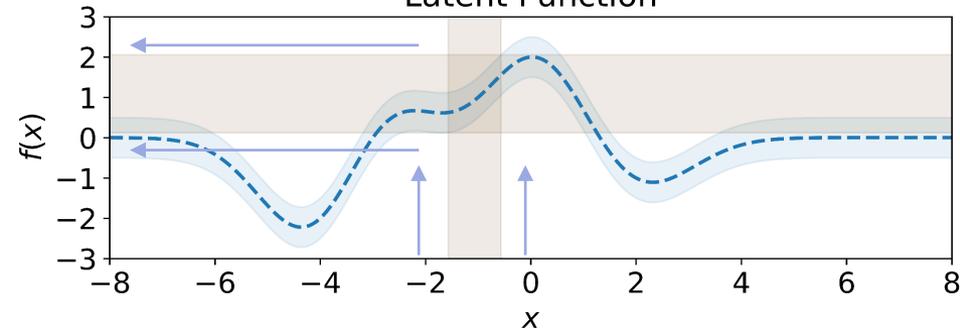


THE UNIVERSITY OF SYDNEY

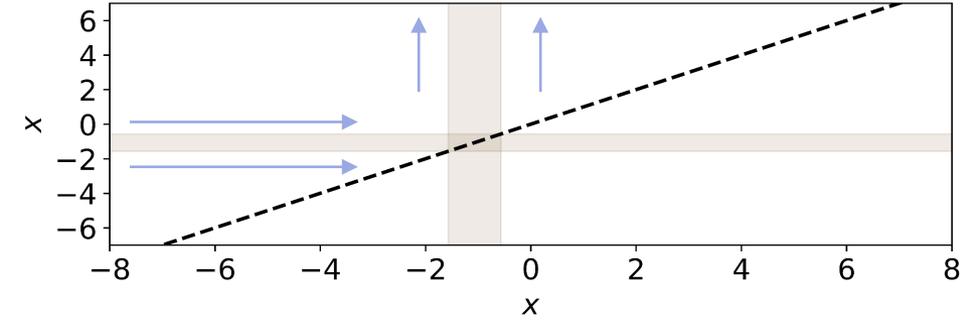
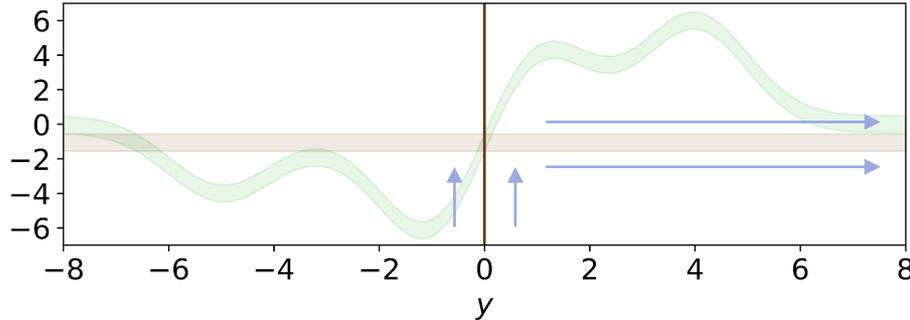
Observation Process



Latent Function



Control Process

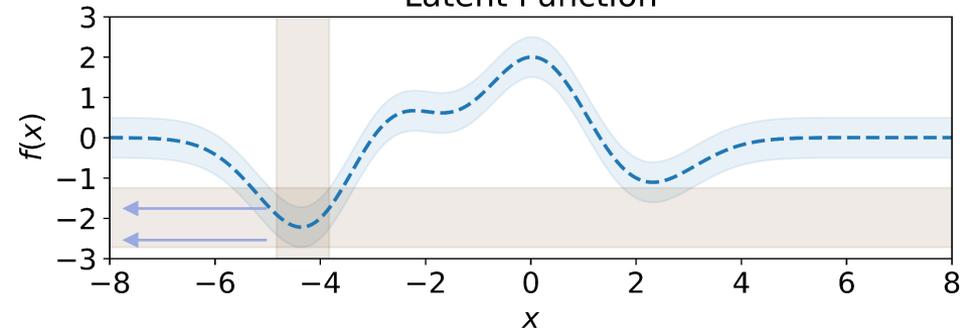
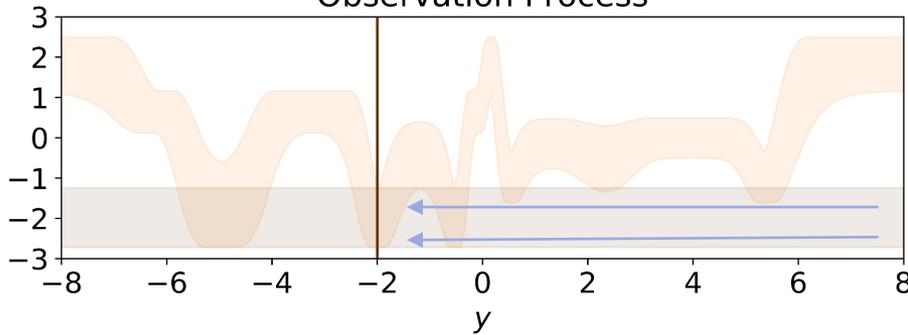


Overall: Query a y and observe a z

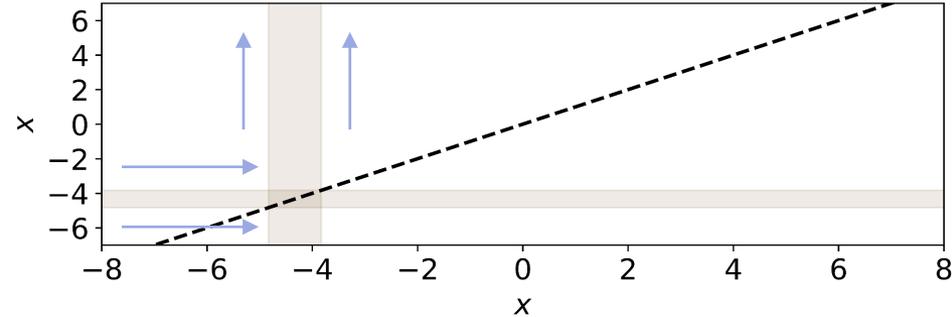
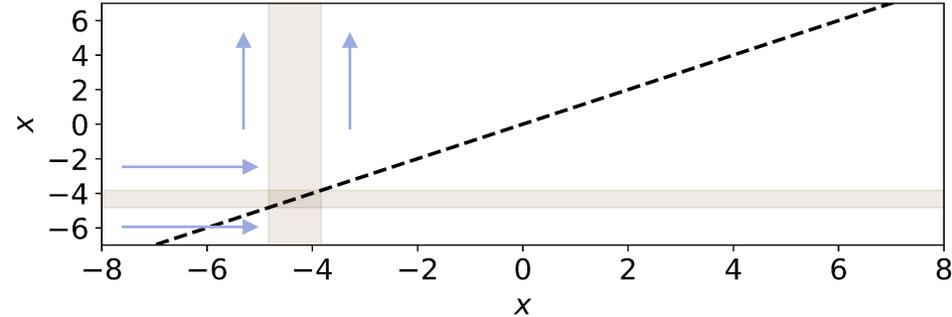


Observation Process

Latent Function



Control Process

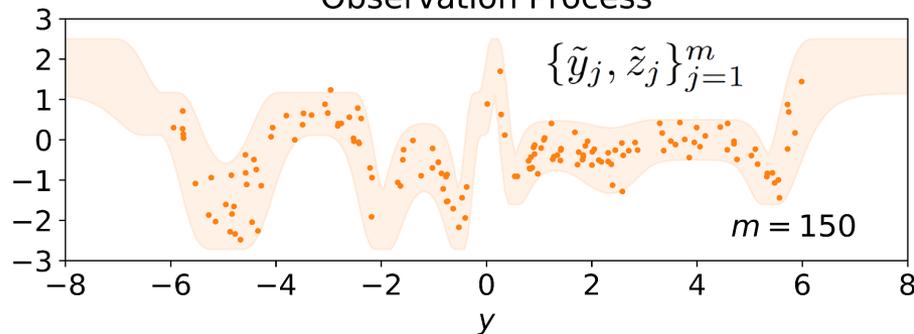


So, you end up with only pairs of (y, z)

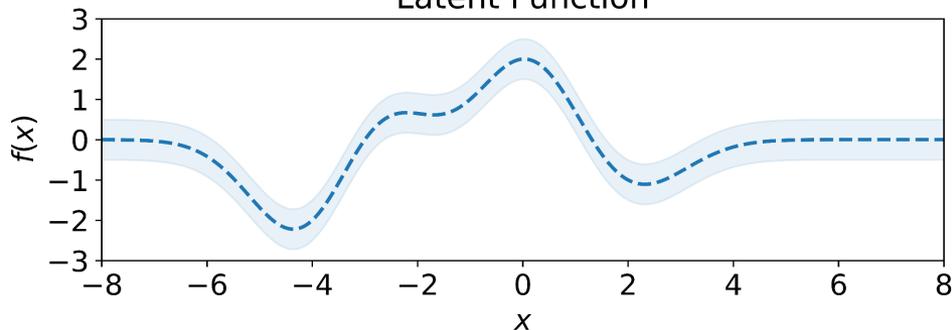


THE UNIVERSITY OF SYDNEY

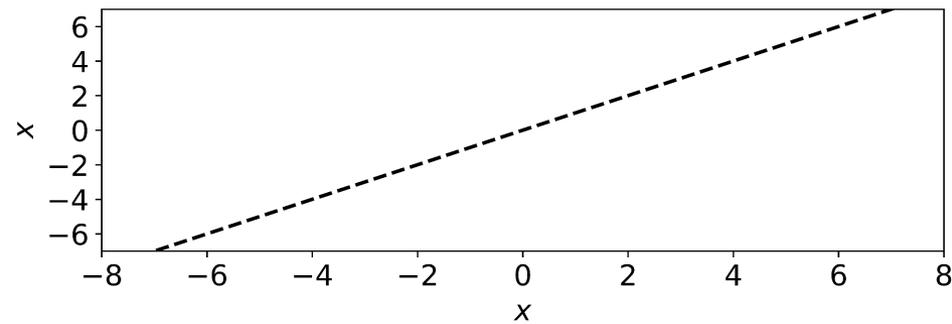
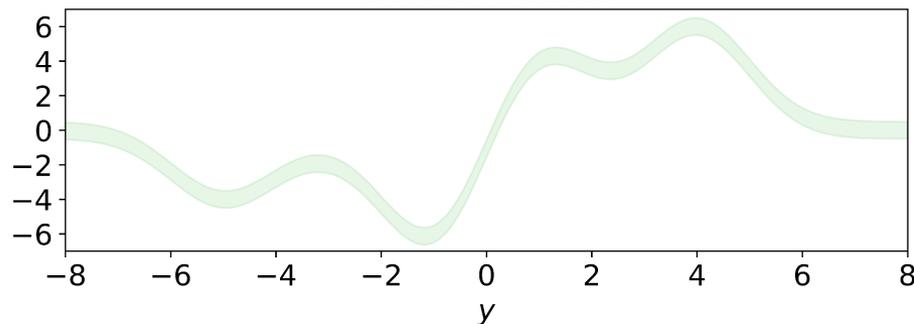
Observation Process



Latent Function



Control Process



• Observed Data

Conditional mean of targets Z given y

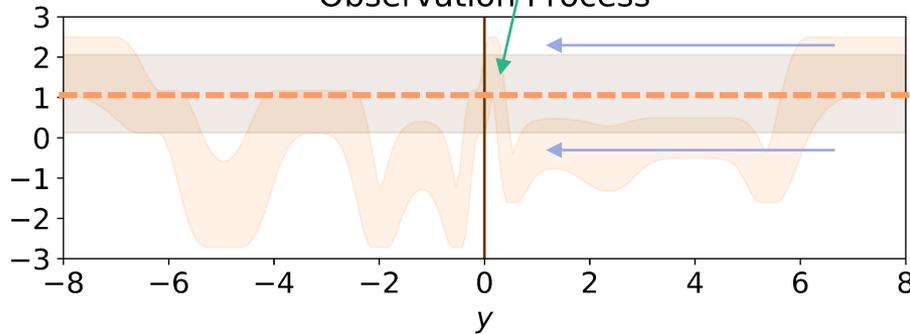


Conditional Mean

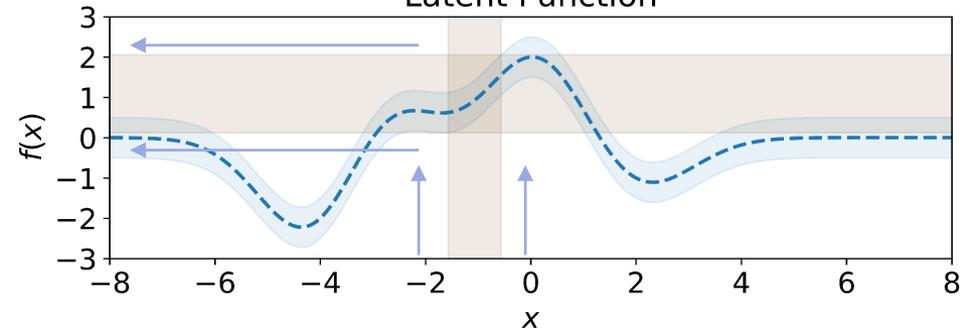
$$\mathbb{E}[Z|Y = y] = \mathbb{E}[f(X)|Y = y] =: g(y)$$

Deconditional Mean

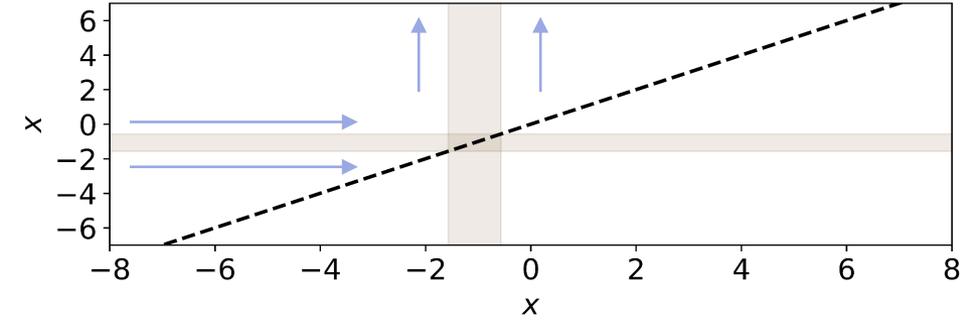
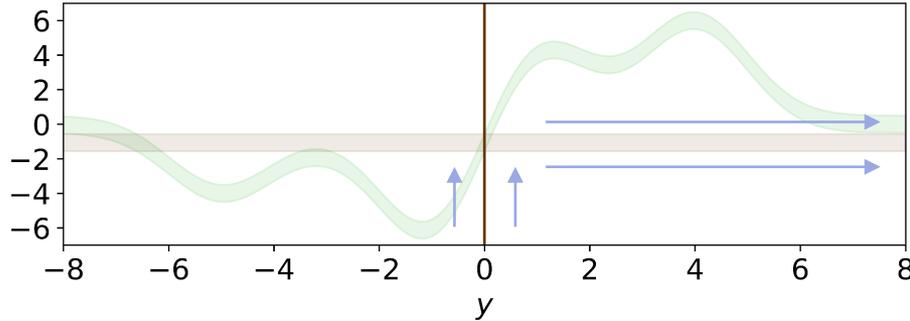
Observation Process



Latent Function

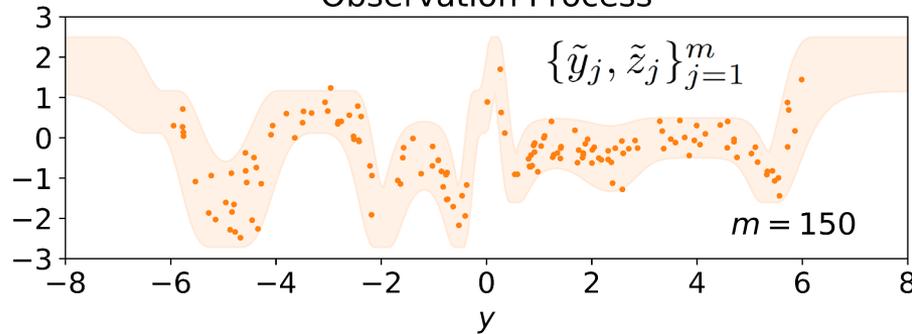


Control Process

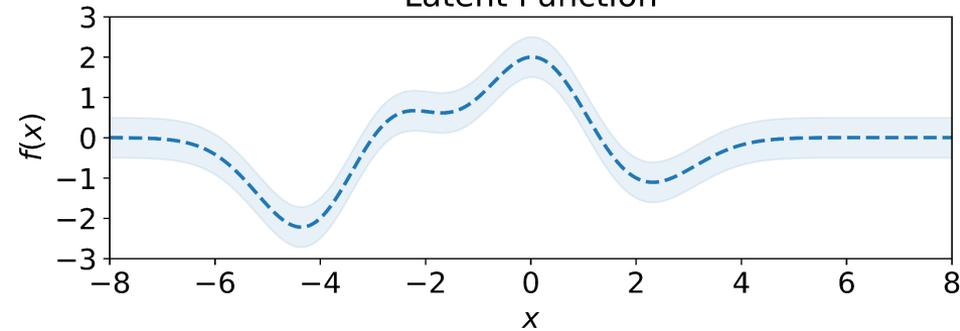


Afterwards, you obtain pairs of (x, y) from a separate process or simulation

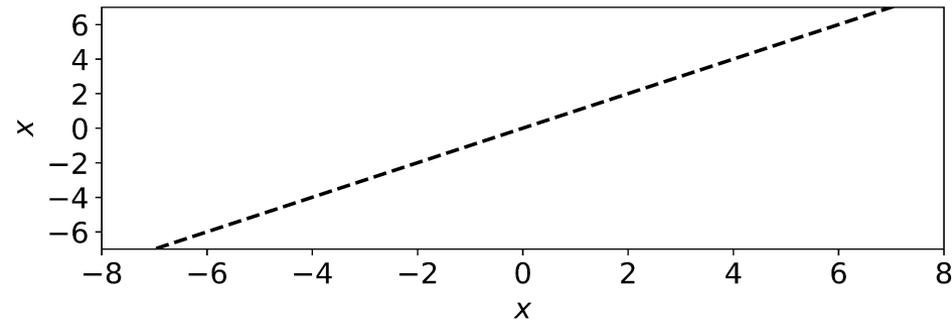
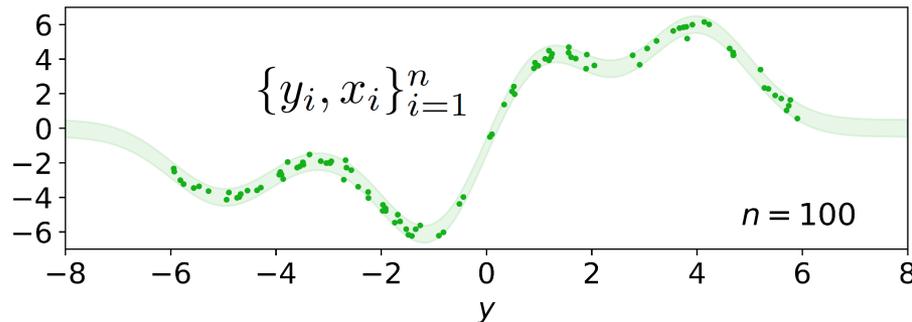
Observation Process



Latent Function



Simulation Process



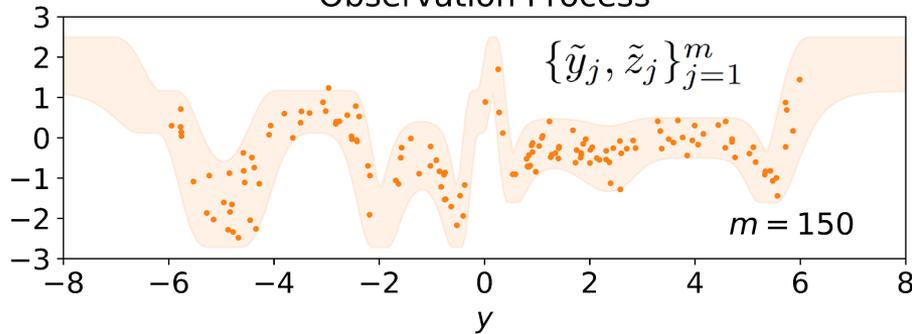
• Observed Data • Simulated Data

Your Goal: Recover f without observing direct pairs of (x, z)

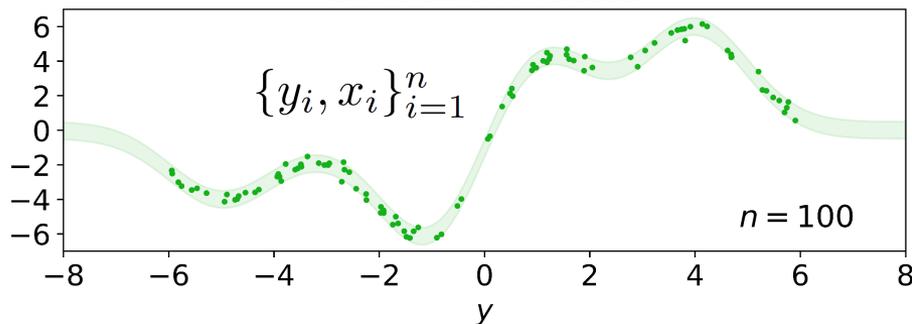


THE UNIVERSITY OF SYDNEY

Observation Process

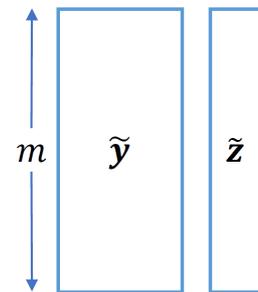


Simulation Process

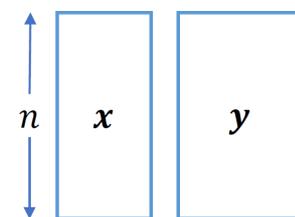


• Observed Data • Simulated Data

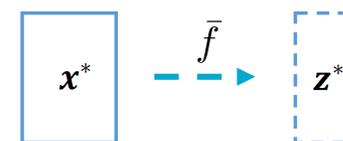
Task or Original Dataset



Transformation Dataset



Recover Latent Function



Deconditional Mean Embeddings

First View: Reversing Conditional Means of Functions



$$(C_{X|Y})^T f = g$$

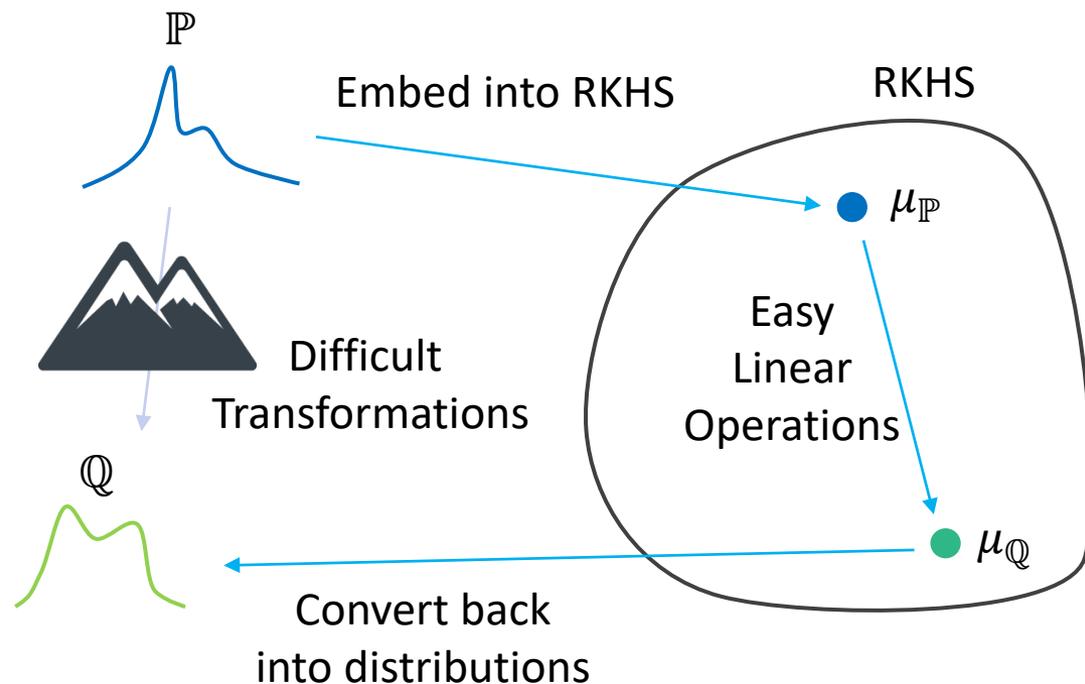
Conditional Mean Operator

$$\mathbb{E}[f(X)|Y = y] =: g(y)$$

Deconditional Mean Operator

$$(C'_{X|Y})^T g = f$$

Kernel Mean Embeddings



$$C'_{X|Y} = (C_{X|Y} C_{YY})^T (C_{X|Y} C_{YY} (C_{X|Y})^T)^{-1}$$

Nonparametric Bayes' Rule

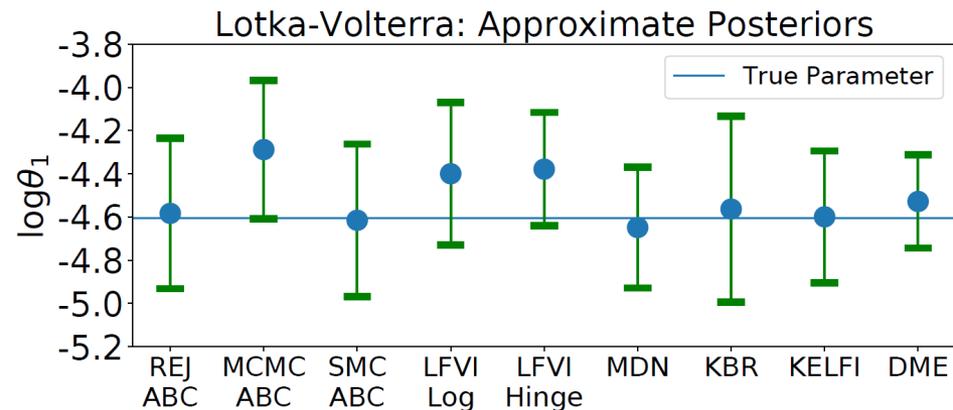
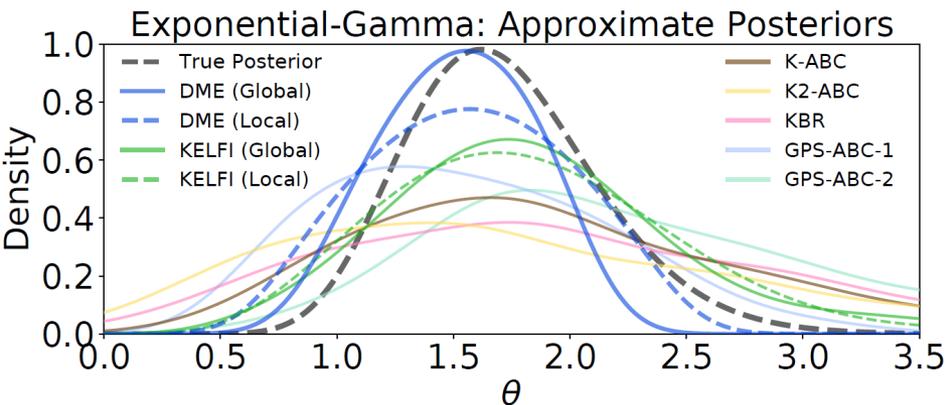


$\mathbb{P}_{Y|X}?$ $\mathbb{P}_{X|Y}$ \mathbb{P}_Y

$$C'_{X|Y} = (C_{X|Y}C_{YY})^T (C_{X|Y}C_{YY}(C_{X|Y})^T)^{-1}$$

If $\mathbb{P}_{X|Y}$ and \mathbb{P}_Y play the roles of likelihood and prior respectively, then when is $C'_{X|Y}$ the same as $C_{Y|X}$, which encodes the corresponding posterior $\mathbb{P}_{Y|X}$?

Likelihood-Free Inference

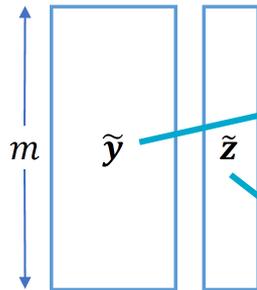


Task Transformed Gaussian Processes

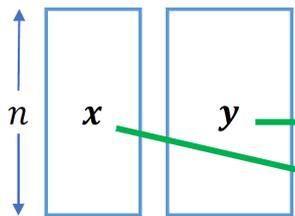
Second View: Transforming Task of Regressing Z on Y to Z on X



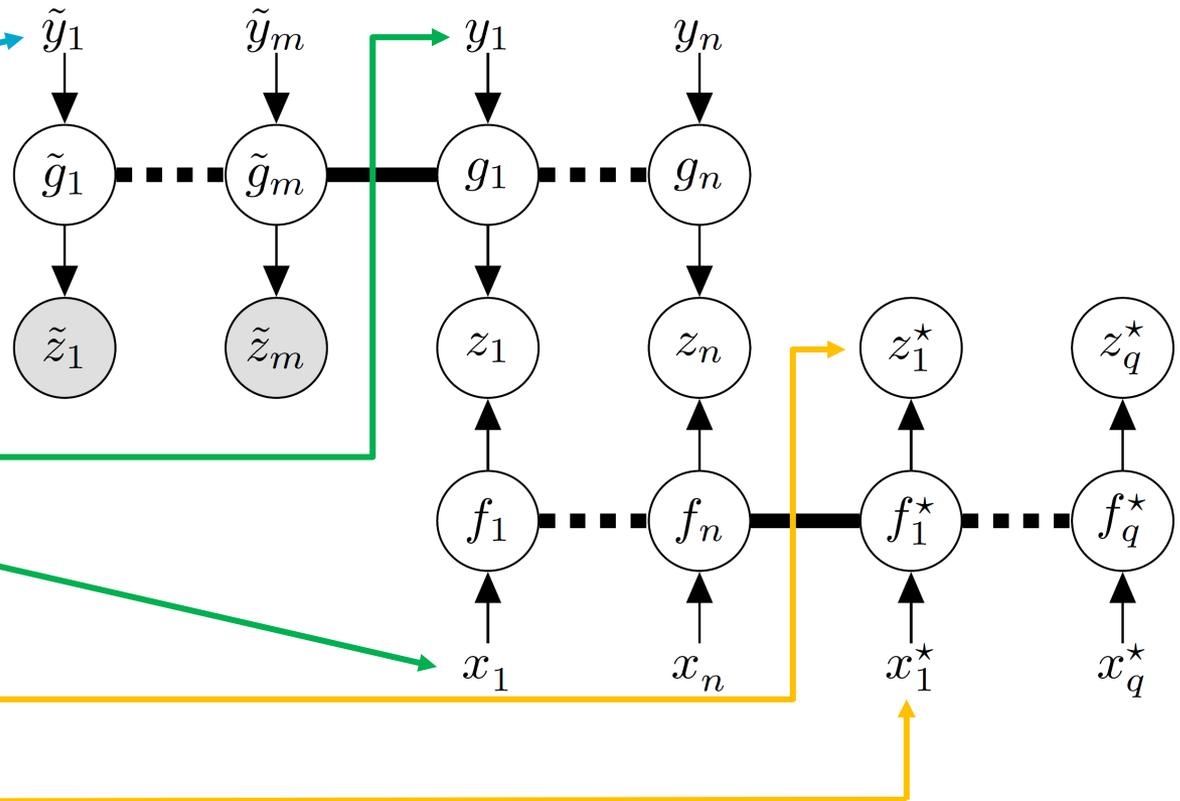
Task or Original Dataset



Transformation Dataset



Recover Latent Function

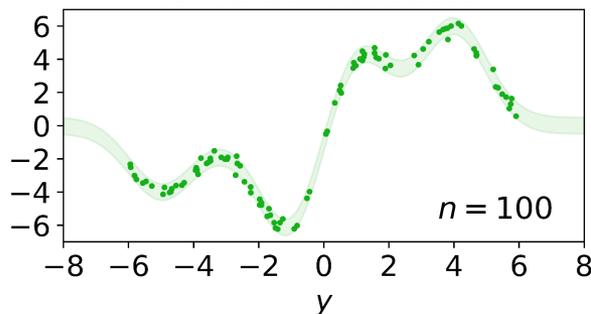


Uncertainty Propagation and Hyperparameter Learning

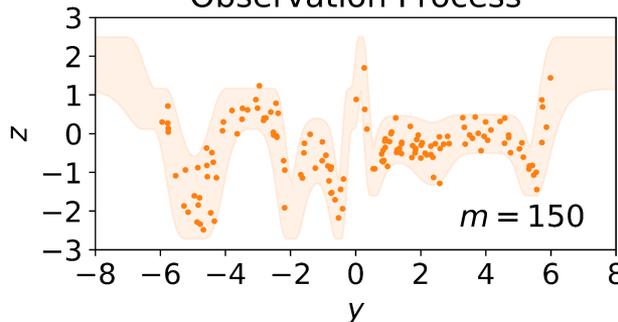


Task Transformed Regression

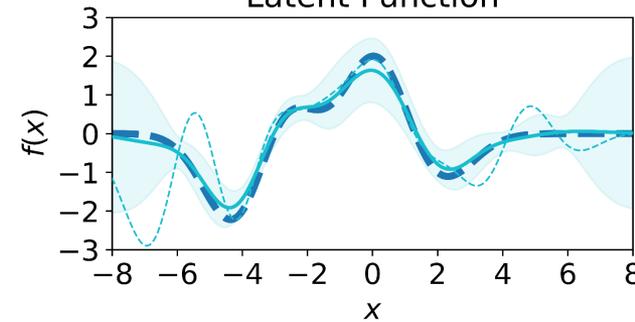
Simulation Process



Observation Process



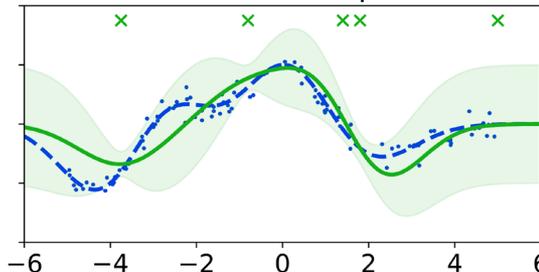
Latent Function



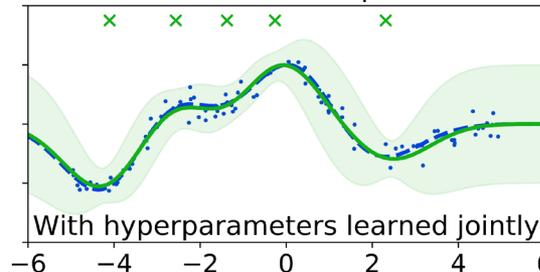
• Simulated Data • Observed Data - - True $f(x)$ - - - DME-TTGP (Init) - - - DME-TTGP (Learn)

Sparse Representation Learning

DME-TTGP: Random Representation

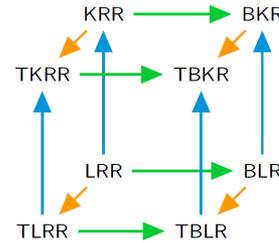


DME-TTGP: Learned Representation



- - - True Function • Generated Data x Sparse Representation

Thank You!



Please come to our poster to find out more interesting connections between the two views and more

Poster 222

