Generalizing Optimization Surrogates based on Neural Networks

Jonas Langhabel[1], Jannik Wolff[2]
[1]langhabel@campus.tu-berlin.de, [2]jannik.wolff@outlook.com

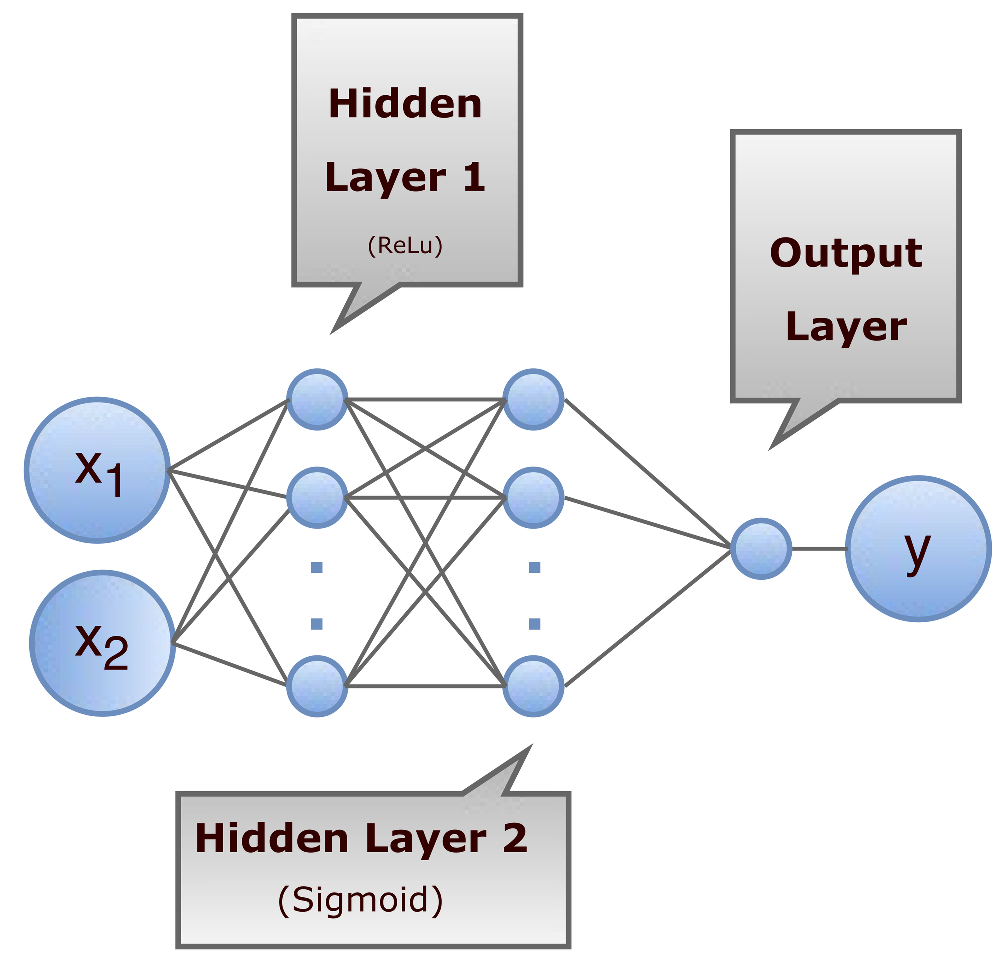
Introduction

We use neural networks to build a meta-model over multiple function approximation problems. This surrogate model relies on a minimum number of samples whose selection process has been orchestrated by an uncertainty-based approach considering the Exploitation-Exploration Dilemma.

Neural Network

- Weight Initialization: Truncated normal
- Adam Optimizer: Initial learning rate of 0.008
- Input range scaled to [0,1]
- 170-230 units/layer

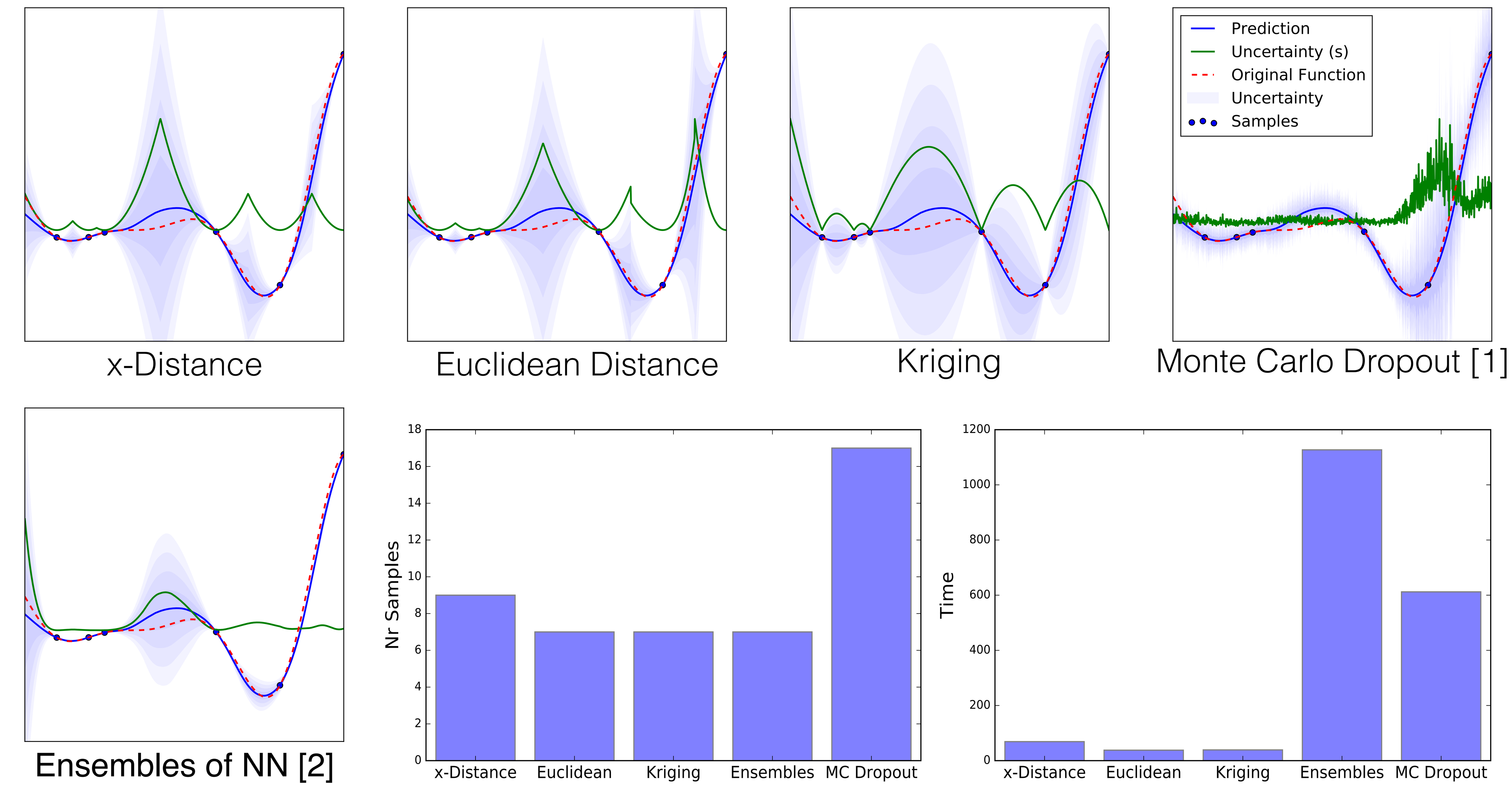
NN



Uncertainty

We use a variety of methods based on the samples and on the NN surrogate for computing the model's uncertainty regarding its output. Using ensembles of NNs yields the most promising results.

Uncertainty

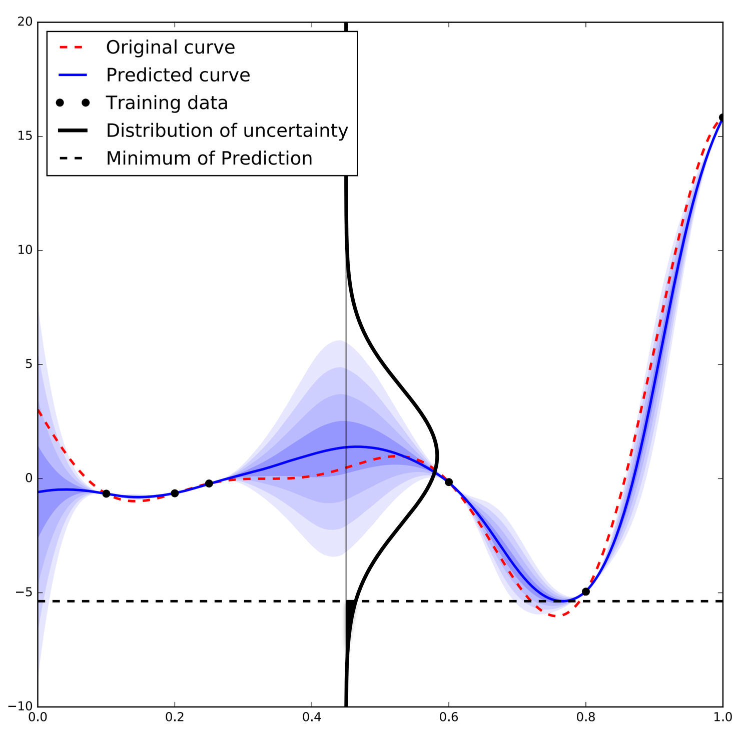


The Exploitation-Exploration Dilemma

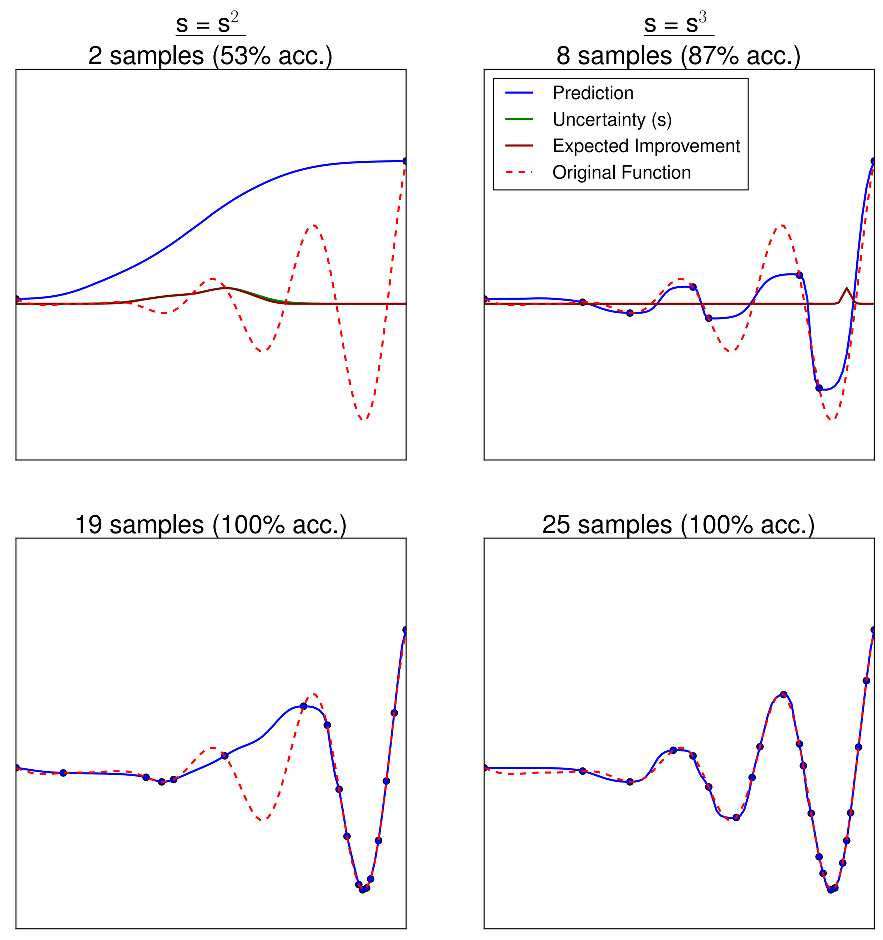
Collecting samples by evaluating the underlying function lies at the heart of optimizing a surrogate. This is assumed to be very expensive, so we use the two-stage approach as used by Forrester [3] to efficiently balance exploitation and exploration based on uncertainty estimates.

E & E

$$E[I(x)] = \begin{cases} (y_{min} - \hat{y}(\mathbf{x}))\Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right) + s\phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right) & \text{if } s > 0 \\ 0 & \text{if } s = 0 \end{cases}$$

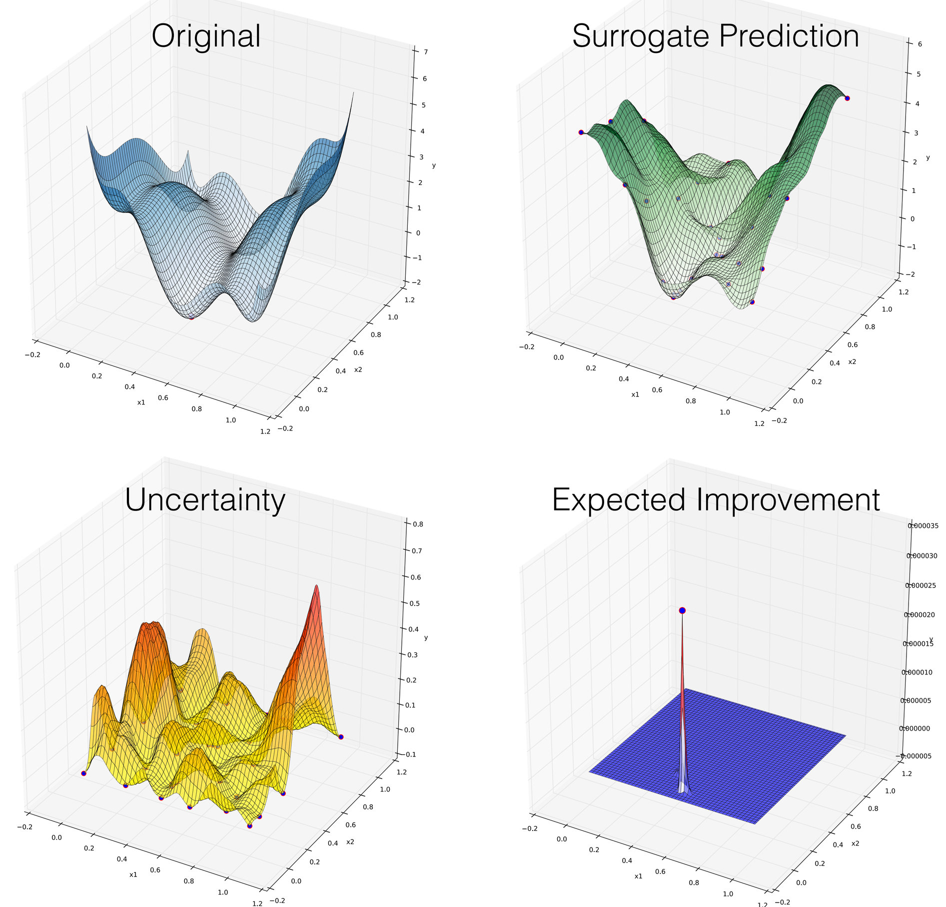


1D



(6x - 2)² ·sin(12x - 4)

2D

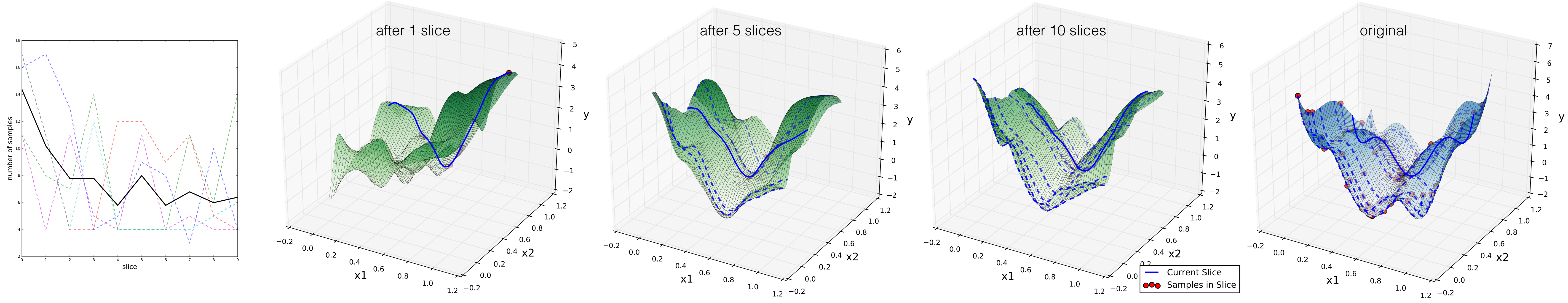


Six Hump Camel: 23 samples taken

Generalization

We want to solve novel optimization problems by generalizing from similar previously solved problems. The 2D-function below displays *slices* of 1D optimization problems. We generalize over x_2 to solve novel 1D-problems faster using a surrogate as a meta-model.

Generalization



References:
1 Gal Y, Ghahramani Z (2015). Dropout as a Bayesian Approximation. *arXiv*:1506.02142
2 Jiang Y (2007). Uncertainty in the Output of Artificial Neural Networks. *IEEE Transactions on Medical Imaging*, 22:7
3 Forrester AIJ, Keane AJ (2009). Recent advances in surrogate-based optimization. *Prog. Aerosp. Sci.*, 45:50–79