Multi-fidelity Learning with Heterogeneous Domains

Soumalya Sarkar
Autonomous and Intelligent Systems Dept.
United Technologies Research Center (UTRC)
East Hartford, CT
sarkars@utrc.utc.com

Michael Joly
Thermal and Fluid Science Dept.
UTRC
East Hartford, CT
jolymm@utrc.utc.com

Paris Perdikaris
MEAM Dept.
University of Pennsylvania
Philadelphia, PA
pgp@seas.upenn.edu

Abstract

Multi-fidelity learning is a significant tool in the intersection of machine learning and engineering sciences that enables learning a surrogate model from multiple fidelity of data sources with different accuracy and varying cost of data generation. One of the common assumptions in this tool is the presence of homogeneous input domain across all fidelities, which creates a limitation for surrogate modeling required in engineering design problem. This paper proposes an approach to extend this capability towards incorporating input domains which are heterogeneous in their dimensionality across the fidelities. The proposed approach is tested on an airfoil problem where a surrogate is generated from multiple fidelities of computational fluid dynamics (CFD) solvers with heterogeneous input domains and single target quantity of interest (QoI) and it is demonstrated to exhibit better predictive accuracy compared to the surrogate learned only from a sparse data generated by high fidelity solver.

1 Introduction

Designing complex physical systems generally go through an iterative process where physical solvers of different predictive accuracy and compute cost, and experiments/testing are invoked to generate supporting data on the design quantity of interest (QoI) or design objectives. This type of repetitive design loop involving aforementioned variable-fidelity data sources can become prohibitively expensive and it results in a slow development process in the fields of aerospace, material science and drug discovery. One approach to accelerate a design process is to learn surrogate models that can predict QoI from design variables much faster than physical solvers within an acceptable tolerance and run surrogate assisted optimization (SAO) [1]. Multi-fidelity learning is a tool that enables learning surrogate models from multiple fidelity of physical solvers where higher fidelity solvers generate more accurate QoI but they are computationally more expensive compared to lower fidelity solvers, sometimes by many orders. To give an example with respect to a compute of a hundred CPU cluster, a 3-dimensional CFD evaluation of the flow field around a standard airfoil can take about few minutes to complete at Reynolds Averaged Navier Stokes (RANS) fidelity and up to few days to weeks at Large Eddy Simulation (LES) fidelity.

A common machine learning approach for learning probabilistic multi-fidelity surrogates is based on a multi-variate Gaussian process and a simple linear auto-regressive correlation model to propagate information across fidelities [2–6]. Recent works in this area [7–9] have demonstrated the remarkable

Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019), Vancouver, Canada.

effectiveness of this approach in accelerating the simulation-based design of complex dynamic systems (e.g., shape optimization of super-cavitating hydrofoils, high dimensional design optimization of NASA rotor-37 airfoil) primarily using low cost and low fidelity solvers, guided by a small number of judiciously selected evaluations of expensive high-fidelity models.

In many physical systems design optimization, although the target QoIs remain the same across all information sources employed within each discipline, every source admits its own parameterization (e.g., design variable and operating conditions). Unfortunately, this parametric and geometric heterogeneity cannot be directly addressed using state-of-the-art multi-fidelity learning formulations and requires the development of a new mathematical framework for synthesizing variable fidelity data that originate from disjoint input domains, i.e., multi-source data in general.

This paper proposes an approach of learning a recursive mapping to extend the capability of multi-fidelity learning towards incorporating input domains which are heterogeneous in their dimensionality across the fidelities. The proposed approach is tested on an airfoil problem where a multi-fidelity surrogate is generated from multiple fidelities of computational fluid dynamics (CFD) solvers with variable input domains and single target QoI and it is demonstrated to exhibit better predictive accuracy compared to the surrogate learned only from sparse data generated by higher fidelity solver.

2 Method

Let's assume that $x_L \in \mathbb{R}^n$ is the input to lower fidelity solver g_L and the corresponding QoI is $y_L \in \mathbb{R}$. Similarly, for higher fidelity solver g_H, the input is $x_H \in \mathbb{R}^m$ and it generates a output $y_H \in \mathbb{R}$. Generally in practice, $m > n$ and the volume of samples available from g_L is higher than g_H as g_H is computationally more expensive. By taking inspiration from the area of heterogeneous transfer learning (HTL) [10], we propose to learn an asymmetric mapping $\varphi : \mathbb{R}^n \mapsto \mathbb{R}^m$ from x_L domain to x_H domain. See figure 1 for a graphical representation of our proposed approach. This framework can be recursively applied when the number of fidelities is more than 2. Once the dimensionality of lower fidelities are brought up to that of highest fidelity, an auto-regressive co-kriging framework can be applied as follows. The mapping φ is learned as a multi-input multi-output regressor on the data with corresponding heterogeneous input pairs across fidelities. Preferred models for learning φ should be able to handle low sample-to-variable ratio.

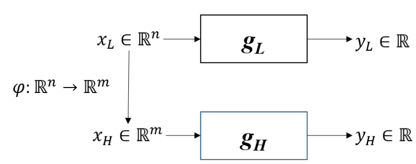

Figure 1: Heterogeneous multi-fidelity modeling

The auto-regressive multi-fidelity learning here is based on Gaussian Processes (GP) [2], which is a non-parametric regression approach commonly used in supervised statistical learning. GPs are a class of stochastic processes that assume a multivariate, jointly Gaussian probability distribution for any finite collection of random variables. For example, if the function $f(\mathbf{x})$ for an input $\mathbf{x}$ is considered a random variable, then, for a finite sub-collection $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$, the corresponding function outputs $\{f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_T)\}$ are assumed to have a multivariate jointly Gaussian distribution:

$$\begin{bmatrix} f(\mathbf{x}_1) \\ \vdots \\ f(\mathbf{x}_T) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(\mathbf{x}_1) \\ \vdots \\ m(\mathbf{x}_T) \end{bmatrix}, \begin{bmatrix} k(\mathbf{x}_1,\mathbf{x}_1) \cdots k(\mathbf{x}_1,\mathbf{x}_T) \\ \vdots \\ k(\mathbf{x}_T,\mathbf{x}_1) \cdots k(\mathbf{x}_T,\mathbf{x}_T) \end{bmatrix} \right) \tag{1}$$

where the underlying Gaussian process is completely characterized by a mean function : $m(\mathbf{x}) \triangleq E[f(\mathbf{x})]$, and a covariance function : $k(\mathbf{x},\mathbf{x}') \triangleq E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]$. The skeleton of our approach is based on the autoregressive co-kriging approach put forth by Kennedy and O'Hagan [3]. The multi-fidelity GP approach is presented in a recursive form as:

$$f_t(\mathbf{x}) = \rho_{t-1} f_{t-1}(\mathbf{x}) + \delta_t(\mathbf{x}), \quad 2 \le t \le Q \tag{2}$$

where ρ_{t-1} quantifies the correlation between $f_t(\mathbf{x})$ and $f_{t-1}(\mathbf{x})$, and the Gaussian process $\delta_t(x)$ represents the discrepancy between $f_t(\mathbf{x})$ and $f_{t-1}(\mathbf{x})$. In this formulation, t refers to the fidelity level

Table 1: descriptions of inputs, fidelity solvers for the output QoI of pressure drop (P_T)

Fidelity	input (x)	fidelity solver (f), run time
Low (L)	$x_L \in \mathbb{R}^2$, rms of camber and thickness	Euler, 82 seconds
High (H)	$x_H \in \mathbb{R}^6$, 3 values for each of camber and thickness	RANS, 385 seconds

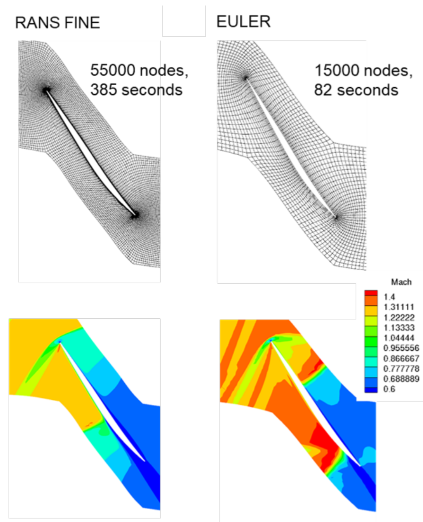

Figure 2: Left: Fine mesh RANS with 55000 nodes (top) and its velocity field (bottom), Right: Euler with 15000 nodes (top) and its velocity field (bottom)

ranging from 2 to Q (number of fidelities), where $f_t(\mathbf{x})$ is the higher fidelity output and $f_{t-1}(\mathbf{x})$ is the lower fidelity output. We use a noisy GP model with zero mean for each fidelity, where the outputs at each fidelity level are noise corrupted, i.e. $\mathbf{y}_t \sim \mathcal{N}(0, \sigma^2_{\epsilon_t})$. The kernel function is parameterized by a set of hyper-parameters $\boldsymbol{\theta}$ and the GP priors belongs to the class of automatic relevance determination (ARD) Matérn covariance functions [2].

3 Results

The proposed approach is tested on the problem of building multi-fidelity surrogate to obtain a predictive model between geometric design variables of a transonic NASA rotor 37 airfoil and relative total pressure drop (P_T) across it. P_T is obtained from post-processing of the CFD solution and it is an important performance parameter to optimize its shape for maximizing efficiency and work.

Table 1 describes the inputs, solvers and respective computational complexity for the QoI of P_T for each fidelity. In this case, solvers are run on 2 dimensions over the cross section at 70% span of the airfoil and the airfoil shape is parameterized by spline distribution of camber and thickness. Fidelity-specific mesh resolutions and velocity fields at $2D$ are presented in figure 2. To visualize the fluid dynamics loss mechanism within the airfoil passage, plotting the Mach number is instrumental to identify the shock and separation behaviors that influence P_T.

From the Euler solver, 200 samples are generated with 2-dimensional lower fidelity input x_L (rms of camber and thickness) and target QoI P_T. As the run time of RANS solver is 4.5 times larger, around 68 samples with 6-dimensional higher fidelity input (3 spline coefficients for each of camber and thickness) and target QoI P_T could be generated to represent a realistic scenario. RANS-generated target values are considered to be high fidelity (HF) ground truth and Euler-generated target values are denoted as low fidelity (LF) ground truth corresponding to their respective input samples. Here the fidelity levels are not coupled, but correlated.

The asymmetric mapping $\varphi : \mathbb{R}^2 \mapsto \mathbb{R}^6$ from x_L to x_H is learned by a multi-input multi-output Gaussian process model trained on the $x_L - x_H$ pairs at same corresponding input configuration across fidelities. To validate the fact that the multi-fidelity (MF) surrogate improves the overall QoI

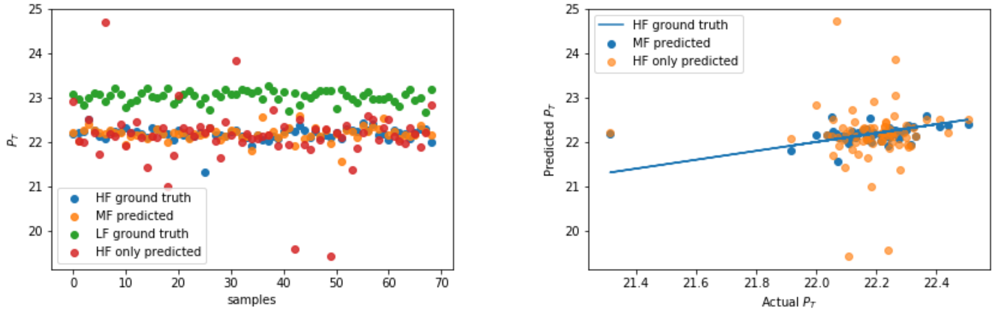

Figure 3: Left: sample wise ground truths of P_T and P_T prediction from single fidelity surrogate model learned from only HF data, and MF surrogate model learned by proposed approach , Right: Actual vs predicted scatter from single fidelity surrogate model learned from only HF data, and MF surrogate model learned by proposed approach

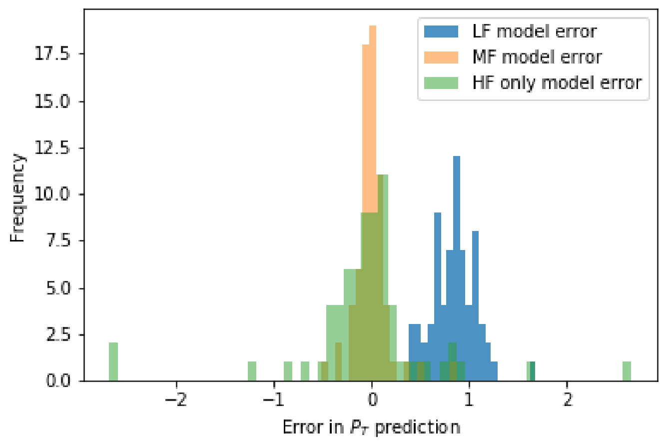

Figure 4: Histograms of error in P_T prediction w.r.t the HF ground truth by surrogate models learned only from LF data, single fidelity surrogate model learned only from HF data, and MF surrogate model learned by proposed approach

prediction, it is compared with a similar complexity Gaussian process model trained only on HF ground truths. This validation exercise is performed by a 10-fold cross validation set up.

Figure 3 presents sample-wise prediction of P_T aggregated over 10-fold cross validation from different models and the scatter around HF ground truth. Although P_T prediction from Euler solver has large discrepancy compared to HF ground truth as RANS solver captures more flow physics, it enriches the multi-fidelity (MF) surrogate by reducing its prediction scatter significantly. Small prediction scatter is required for incorporating the surrogates reliably into a SAO. Figure 4 presents that the standard deviation of prediction error is significantly improved by 3 times from 0.682 (HF only surrogate), which is more than 3% of mean P_T, to 0.226 ($\leq 1\%$ of mean P_T) by our MF surrogate model even in the presence of limited HF ground truth and heterogeneous input domains between low and high fidelity.

4 Conclusions and Future Work

This paper proposes an asymmetric mapping approach for bringing heterogeneous input domains across multiple fidelities into a homogeneous description to fundamentally extend the capability of multi-fidelity learning. Our proposed MF learning framework is tested on an airfoil problem where a multi-fidelity surrogate is generated from multiple fidelities of CFD solvers (e.g., RANS and Euler) with variable input domains and single target quantity of interest (QoI) and it is demonstrated to exhibit 3 times lower predictive uncertainty compared to the predictive surrogate learned only from a sparse set of data generated by higher fidelity solver. Considering this work an initial attempt towards tackling heterogeneous multi-fidelity learning, several areas are pursued as ongoing research:

1. Learning mapping between heterogeneous domains when input correspondence is ill-defined.
2. Heterogeneous multi-fidelity learning in the presence of output heterogeneity.

References

[1] Alan Díaz-Manríquez, Gregorio Toscano, Jose Hugo Barron-Zambrano, and Edgar Tello-Leal. A review of surrogate assisted multiobjective evolutionary algorithms. *Computational Intelligence and Neuroscience*, 2016.

[2] Carl E. Rasmussen and Christopher K.I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, Cambridge, MA, USA, 2005.

[3] Marc C. Kennedy and Anthony O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 2000.

[4] Loic Le Gratiet and Josselin Garnier. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.

[5] Paris Perdikaris, Daniele Venturi, and George Karniadakis. Multifidelity information fusion algorithms for high-dimensional systems and massive data sets. *SIAM Journal on Scientific Computing*, 38(4):B521–B538, 2016.

[6] Kirthevasan Kandasamy, Gautam Dasarathy, Jeff Schneider, and Barnabás Póczos. Multi-fidelity Bayesian optimisation with continuous approximations. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1799–1808, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.

[7] Paris Perdikaris and George E. Karniadakis. Model inversion via multi-fidelity bayesian optimization: a new paradigm for parameter estimation in haemodynamics, and beyond. *Journal of The Royal Society Interface*, 13(118):20151107, 2016.

[8] Paris Perdikaris, Daniele Venturi, Johannes O. Royset, and George E. Karniadakis. Multi-fidelity modelling via recursive co-kriging and gaussian-markov random fields. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2179):20150018, 2015.

[9] Sudeepta Mondal, Michael Joly, and Soumalya Sarkar. Multi-fidelity global-local optimization of a transonic compressor rotor. In *ASME Turbo Expo 2019: Turbomachinery Technical Conference and Exposition*. American Society of Mechanical Engineers, 2019.

[10] Oscar Day and Taghi M. Khoshgoftaar. A survey on heterogeneous transfer learning. *Journal of Big Data*, 4(1):29, Sep 2017.