CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.2020-0039

Vol. **1**, No. 4, pp. 715-739
December 2020

MFPC-Net: Multi-Fidelity Physics-Constrained Neural Process

Yating Wang[1] and Guang Lin[1,*]

[1] *Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA.*

Received 20 July 2020; Accepted 4 October 2020

Abstract. Recently, there are numerous work on developing surrogate models under the idea of deep learning. Many existing approaches use high fidelity input and solution labels for training. However, it is usually difficult to acquire sufficient high fidelity data in practice. In this work, we propose a network which can utilize computational cheap low-fidelity data together with limited high-fidelity data to train surrogate models, where the multi-fidelity data are generated from multiple underlying models. The network takes a context set as input (physical observation points, low fidelity solution at observed points) and output (high fidelity solution at observed points) pairs. It uses the neural process to learn a distribution over functions conditioned on context sets and provide the mean and standard deviation at target sets. Moreover, the proposed framework also takes into account the available physical laws that govern the data and imposes them as constraints in the loss function. The multi-fidelity physics-constrained network (MFPC-Net) (1) takes datasets obtained from multiple models at the same time in the training, (2) takes advantage of the available physical information, (3) learns a stochastic process which can encode prior beliefs about the correlation between two fidelity with a few observations, and (4) produces predictions with uncertainty. The ability of representing a class of functions is ensured by the property of neural process and is achieved by the global latent variables in the neural network. Physical constraints are added to the loss using Lagrange multipliers. An algorithm to optimize the loss function is proposed to effectively train the parameters in the network on an ad hoc basis. Once trained, one can obtain fast evaluations of the entire domain of interest given a few observation points from a new low-and high-fidelity model pair. Particularly, one can further identify the unknown parameters such as permeability fields in elliptic PDEs with a simple modification of the network. Several numerical examples for both forward and inverse problems are presented to demonstrate the performance of the proposed method.

AMS subject classifications: 65Z, 62M

Key words: Multi-fidelity, physics-constrained, deep learning, neural process.

*Corresponding author. *Email addresses:* `guanglin@purdue.edu` (G. Lin), `wang4190@purdue.edu` (Y. Wang)

1 Introduction

Many applications in science and engineering may encounter parameterized computational models, such as optimization of complex systems with uncertainty or inverse modeling. A common challenge lies in obtaining sufficient high-fidelity data, which requires expensive measurements or simulations for a large number of parameter instances. Many approaches have been proposed to develop surrogate models, such as global and local model reduction techniques [1,2,4,6–9,15]. The idea is to construct computational cheap reduced order models, however, the simulation data obtained may be low-fidelity and these approaches may face the challenge in generalization.

On the other hand, numerous works have been proposed to learn surrogate models given some data. In the case when only the data are available without knowing the underlying physics, one can treat it as a pure data-driven supervised learning task using methods like Gaussian process (GP) [5,20]. To obtain reliable results, one usually needs sufficient high-fidelity training data. However, if the random coefficients in the equation possess uncertainties and are high dimensional, or if the underlying map is nonlinear, GP may be inefficient and computationally expensive due to the curse of dimensionality. In the past few years, deep neural networks (DNNs) [22] have attracted increasing interest due to the universal function approximation property and their ability to model high-dimensional input-output relationships. Numerous studies are proposed to solve partial differential equations and have shown great performance in various applications. For example, learning the evolution operator of the PDE from data [27], approximating important physical quantities of the PDE [12], solving heterogeneous elliptic problems on varied domains [26], designing efficient algorithms to handle multiscale multiphase flow problems [25] and using the idea of multiscale model reductions for learning [3,23,24].

When the physical equations/laws are also available, it is important to incorporate these information in the design of surrogate deep neural networks to speed up learning. Physics informed neural networks (PINN) [18,19] were proposed to realize the idea and have been successfully applied to solve PDE problems subject to the law of physics that governs the data. On the other hand, for the case without data labels in the training, there have also been developments on physics-constrained surrogates for stochastic PDEs [28] just using the information on physical relationships. These approaches either require a large amount of high-fidelity data to train, or need to know the full high-fidelity information of coefficients in the model, such as random coefficients in the equation throughout the entire domain. This is usually hard to realize in practical problems, for example, in some optimization problems when large amounts of high-fidelity data are required, or when the coefficients are very difficult to measure.

A more practical case is that only a limited number of high-fidelity observations are available, but sufficient low-fidelity approximations can be obtained via fast simulation. Multi-fidelity methods combine high-and low-fidelity data, and can be employed to efficiently achieve desired accuracy [10,17]. In [16], the authors present a composite neural network (MPINNs) which can be trained utilizing multi-fidelity data. The architecture of