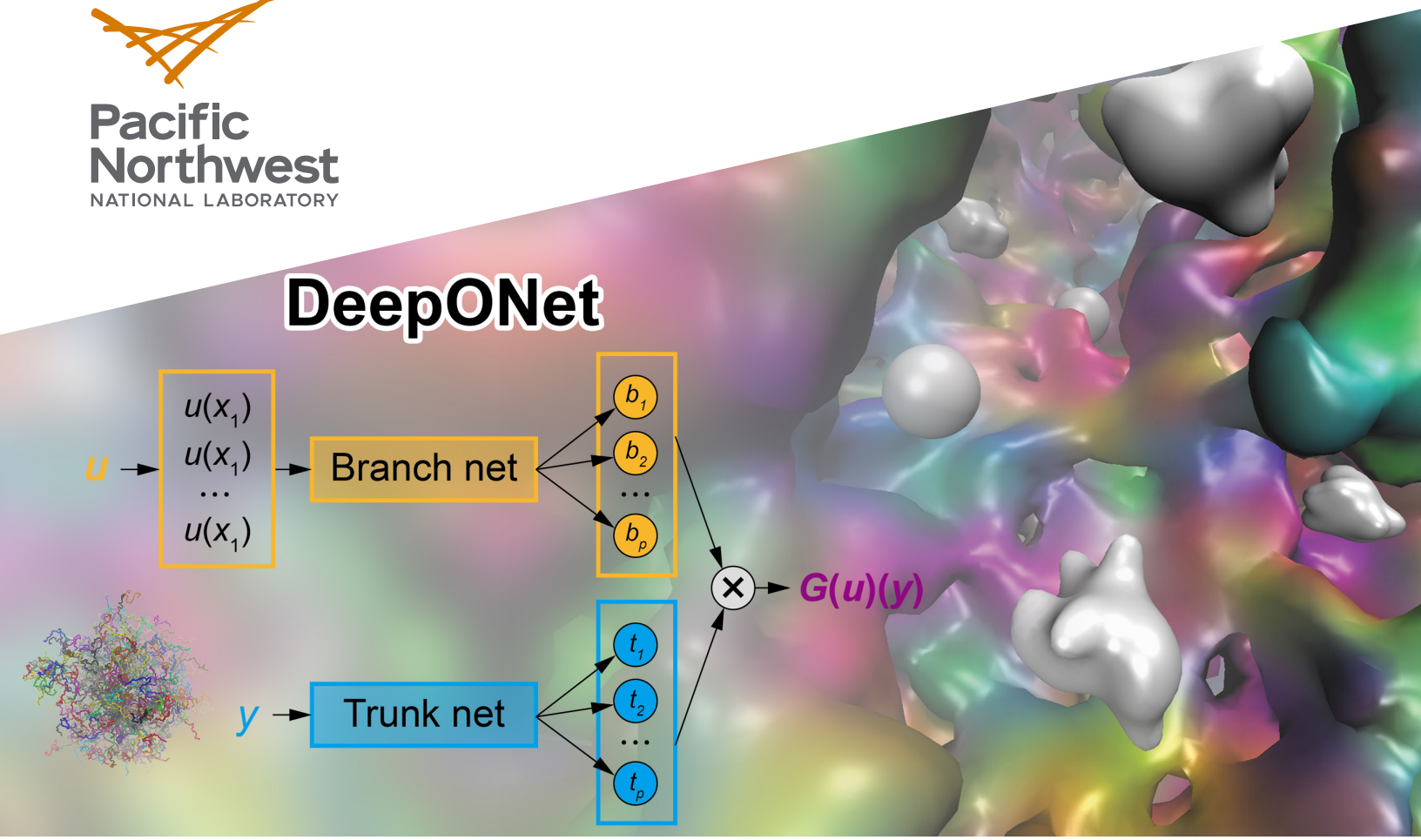
Pacific Northwest
NATIONAL LABORATORY




# PhILMS: Collaboratory on Mathematics and Physics-Informed Learning Machines for Multiscale and Multiphysics Problems

October 2020

U.S. DEPARTMENT OF ENERGY | Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830

## DISCLAIMER



PACIFIC NORTHWEST NATIONAL LABORATORY
*operated by*
BATTELLE
*for the*
UNITED STATES DEPARTMENT OF ENERGY
*under Contract DE-AC05-76RL01830*

**Printed in the United States of America**

**Available to DOE and DOE contractors from the**
**Office of Scientific and Technical Information,**
**P.O. Box 62, Oak Ridge, TN 37831-0062;**
**ph: (865) 576-8401**
**fax: (865) 576-5728**
**email: reports@adonis.osti.gov**

**Available to the public from the National Technical Information Service**
**5301 Shawnee Rd., Alexandria, VA 22312**
**ph: (800) 553-NTIS (6847)**
**email: orders@ntis.gov <https://www.ntis.gov/about>**
**Online ordering: http://www.ntis.gov**

# PhILMS: Collaboratory on Mathematics and Physics-Informed Learning Machines for Multiscale and Multiphysics Problems

ANNUAL REPORT

October 2020

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

## Executive Summary

### Motivation

Complex multiphysics systems are governed by the hidden physics of interfaces and inhomogeneous cascades of scales, (e.g., multifunctional materials, subsurface transport, reactive transport), hence requiring new data-driven multiscale modeling that is accurate, efficient, and easy to implement. In particular, inhomogeneous cascades of scales involve long-range spatiotemporal interactions and often lack proper closure relations to form complete and mathematically well-posed systems of governing equations. Existing multiscale and multiphysics approaches, along with classical integer-order partial differential equations (PDEs) have been ineffective in addressing nonlocal interactions, inhomogeneous cascades of scales, or propagation of uncertainty and stochasticity across scales. They are inadequate for solving *inverse* stochastic multiscale problems, especially in the context of noisy *multifidelity* data. Moreover, existing approaches do not scale well to big systems in real-world applications.

### Research Plan

In the Collaboratory on Mathematics and Physics-Informed Learning Machines for Multiscale and Multiphysics Problems (PhILMs) project, we are developing stochastic multiscale modeling frameworks in conjunction with emerging deep learning techniques to seamlessly fuse physical laws, including thermodynamics and multifidelity data for forward and inverse multiscale problems. Our approach is based on a synthesis of physics-based and data-driven tools and approaches, including nonlocal operators, multifidelity data and information fusion, deep neural networks (DNNs), meshless methods, uncertainty propagation, and stochasticity to simulate complex multiscale systems. Therefore, we are conducting research at the interface of mathematics, physics, data science, and deep learning. With an emphasis on predictability, reproducibility, *and* uncertainty quantification, the research plan of PhILMs includes:

- *Developing physics-informed learning machines by encoding conservation and thermodynamic laws as well as other prior physical knowledge into deep learning networks and analyzing their mathematical properties.*
- *Demonstrating the effectiveness of PhILMs in designing functional materials with tunable properties and extending PhILMs to other U.S. Department of Energy (DOE)-relevant multiscale problems (e.g., combustion, subsurface, and Earth systems)—all exhibiting inhomogeneous scaling cascades.*
- *Establishing scientific machine learning (ML) as a new meta-discipline at the interface of computational mathematics, data science, information fusion, and deep learning.*

Our integrated mathematical and computational activities can be classified into four research areas:

1. **RA-I**: PDE-based modeling of macroscales
2. **RA-II**: stochastic modeling of mesoscales
3. **RA-III**: bridging methods to connect the scales
4. **RA-IV**: deep learning approximations and algorithms to support RA-I to RA-III

### Coordination, Integration, and Evaluation Plan

We are developing PhILMs by coordinating and assimilating the four research areas. In **RA-I**, we are developing PDE-based, physics-informed neural networks (PINNs) for forward and inverse problems. We also develop appropriate representations and a unified theory for nonlocal calculus. A new exciting area is using DNNs to approximate nonlinear operators. In **RA-II**, we are learning PDEs from molecular simulations, applying operator regression to predict mesoscopic dynamics, and learning effective coarse-grain potentials for energy storage materials. In **RA-III**, we are considering concurrent coupling of heterogeneous domains using conservative PINNs, including nonlocal to local PDEs (e.g., fractional/local), and applying PINNs to learning non-Newtonian fluid dynamics and subsurface transport in porous media. In **RA-IV**, we develop foundational algorithms on optimization, uncertainty quantification, and generative adversarial networks (GANs). In addition, we develop new methods for data augmentation for faster training and a fundamental new theory for generalization. Ultimately, these advances will be combined with the knowledge and inherent structure of physical laws to develop PINNs for material design, enhancing efficiency in combustion and making predictions for subsurface systems and ice sheets.

In a similar fashion to the PhILMs project's objective of elucidating and integrating activity across different scales of complex multiphysics systems, the successful coordination of the four research areas requires a significant level of integration of activities from the participating institutions. We have established and maintained pair and triple collaborative interactions among the various PhILMs institutions so that we cultivate an exchange of strengths, ideas, and approaches. In addition, we have encouraged mutual visits to each principal investigator's (PI's) laboratories and student and postdoc exchanges.

Our evaluation plan defines benchmarks of increasing complexity to measure progress in each research area and calls for annual workshops and mini-symposia. We believe that the documented early adoption of ideas originating in the PhILMs project points towards the viability of our approach. In addition, we have formed a dissemination panel so that our work is presented across all the DOE laboratories to make sure that DOE scientists will be the first who can benefit from our findings. We have also successfully reached out to industrial partners. During the second year, we held regular webinars on Monday afternoons given by scientists, both external and internal to PhILMs, and attended by all researchers associated with PhILMs at Pacific Northwest National Laboratory (PNNL), Sandia National Laboratories (SNL), Stanford University, Massachusetts Institute of Technology (MIT), University of California, Santa Barbara (UCSB), and Brown University.

**Highlights of Accomplishments and Outcomes**

**Research Area I**

*PDE-based Modeling of Macroscales*: The Brown team has developed Galerkin-based PINNs with error control and Petro-Galerkin PINNs with domain decomposition while the SNL team has developed control volume PINNs and physics-informed graph neural networks (NNs). Brown, SNL, and UCSB have developed a unified nonlocal theory and new algorithms for learning nonlocal kernels. UCSB and SNL have developed data-driven numerical solvers for surface PDEs on manifolds. Brown and SNL have developed new types of DNNs that approximate nonlinear local and nonlocal operators, and SNL has developed data-driven exponential integrators in small data regimes for the drift-diffusion model governing semiconductor dynamics. Stanford has inferred the constitutive relationship for a generalized Maxwell viscoelasticity model by expressing both DNNs and numerical PDE solvers in terms of computational graphs. Brown and PNNL have applied PINNs to hidden fluid mechanics (HFM) (Science paper), to inferring the mechanical properties of 3D printed materials through multifidelity data, and to retrieving the effective permittivity of metamaterials. Brown and SNL have applied nonlocal PINN (nPINNs) to turbulence modeling of fully-developed channel flow, where a universal profile of the fractional order of the operator for the Reynolds stresses is discovered.

**Research Area II**

*Stochastic Modeling of Mesoscales:* SNL and UCSB have developed Generalized Moving Least Squares (GMLS-Nets) to infer PDE descriptions from molecular-level simulation data. A NN architecture generalizing convolutional NNs to scattered datasets was developed, allowing for an inference approach applied to an ensemble of material particles. Brown and MIT have developed a new method for inferring the governing stochastic ordinary differential equations (SODEs) by observing particle ensembles at discrete and sparse time instants (i.e., multiple "snapshots"). PNNL and Lehigh University have developed a new framework by integrating the polynomial-regression-based ML model with the course-grained (CG) molecular dynamics (MD) simulation to obtain the effective parameters of a multi-fluid system. Brown has applied the deep operator networks (DeepOnet), trained using dissipative particle dynamics (DPD) data, to model the growth of nanobubbles at extremely fast speeds. PNNL has obtained a new DNN potential of lithium electrode from Density Functional Theory (DFT) data to be used for atomistic simulations by enforcing the physical nature of interatomic bonding in order to improve model transferability.

**Research Area III**

*Bridging of Methods to Connect the Scales:* Brown has applied active- and transfer-learning to connect the macroscales (non-Newtonian Navier-Stokes) with mesoscales (DPD) and obtained a viscoelastic closure at the large scales that reflects the microstructure of the small scales. Brown has developed an extension of PINNs with domain decomposition, where state variables and fluxes are matched at the interfaces of the subdomains. This can be used for coupling mesoscale-macroscale dynamics. SNL applied operator regression developed in RA-I to simulate diverse multiscale phenomena. In numerical homogenization of multiphase continuum models from molecular data they extracted a multiphase continuum model from MD simulations of a bidisperse Leonard-

Jones fluid in a microchannel. In similar work they extracted thermodynamically consistent equations of state from molecular models in learning shock hydrodynamics of solids in the high energy density regime. PNNL has extended the PINN method to learn viscosity models of two non-Newtonian systems (polymer melts and suspensions of particles) using only velocity measurements. This method allows prediction of the stresses in non-Newtonian fluids with only the average streamwise velocity, a quantity easily measured in experiments. PNNL also developed a multiphysics-informed DNN method for simultaneously estimating space-dependent hydraulic conductivity, hydraulic head, and tracer concentration fields from sparse measurements in subsurface transport applications, which is of great interest to Hanford site monitoring plan.

**Research Area IV**

*Statistical Learning and Deep Learning Approximations and Algorithms:* We have developed new algorithms for DNNs, training, and for generalization. On algorithms, PNNL has developed a physics-informed DNN coupled with a conditional Gaussian process method for estimating the space-dependent coefficient and the state variables of a PDE from sparse measurements. Brown has proposed a potential flow generator with $L_2$ optimal transport regularity, which can be easily integrated into a wide range of generative models including different versions of GANs and flow-based models. Brown has also proposed a Bayesian PINN (B-PINN) to solve both forward and inverse nonlinear problems described by PDEs and noisy data; the Bayesian NN combined with a PINN for PDEs serves as the prior while the Hamiltonian Monte Carlo serves as an estimator of the posterior. PNNL, UCSB, and MIT have been collaborating to investigate deep learning generative modeling approaches for learning models for stochastic dynamical systems when there is access to observables or partial information on system state evolution. On training, SNL has presented a scheme exploiting a recent trend in nonlinear approximation theory, interpreting DNNs as an adaptive basis to identify an effective splitting. For regression and PINN loss functions, the model reduces to solving a least squares problem at each iteration, and has been shown to provide up to 100X speed-up. MIT has shed light on why min-max optimization is computationally intractable and the team has shown that the problem is PPAD-complete, which means that no algorithm can compute approximate local min-max equilibria in polynomial time. Brown further analyzed the gradient descent (GD) method to identify and quantify the root causes of plateau phenomenon (i.e., periods during which activation patterns remain constant). They proposed a new iterative training method, the Active Neuron Least Squares, characterized by the explicit adjustment of the activation pattern at each step, which is designed to enable a quick exit from a plateau. Stanford investigated several questions related to data augmentation, for which no theory exists to explain how and why certain augmentation strategies are better than others and to understand the possibilities and limitations of augmentation. Brown proposed two approaches of locally adaptive activation functions namely, layer-wise and neuron-wise locally adaptive activation functions, which improve the performance of DNNs and PINNs. Brown also quantified the generalization error in deep learning in terms of data distribution and NN smoothness and presented the first-ever theory on the convergence and generalization of PINNs for elliptic and parabolic PDEs. Stanford has studied an implicit regularization for DNNs driven by an Ornstein-Uhlenbeck like process in order to explain in a rigorous way why deep learning performs as well as it does. Their analysis focuses on deep networks, trained via stochastic gradient descent (SGD), but where the gradient updates are computed with respect to noisy training labels.

**Challenges and Implemented Course Corrections**

Alexandre Tartakovsky of PNNL is no longer with PhILMs due to a new job at University of Illinois at Urbana-Champaign but we have hired other experienced personnel and re-assigned tasks to cover this change. Panos Sitnis is the new center coordinator at PNNL.

## Contents

# PhILMs Progress Report 2019-2020

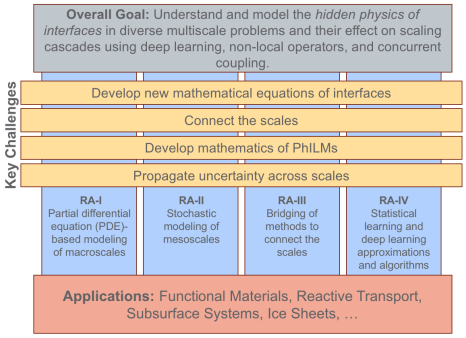



**Figure 1:** *Overview of the PhILMs proposal:* Relation between the overall goal of PhILMs, the thematic key challenges derived from this goal, and its four research areas motivated by exemplar DOE applications.

Most multiscale problems can be broadly classified into three categories: (i) data poor but complete knowledge of the governing physics; (ii) data rich but little knowledge of governing physics; and (iii) only moderate data and moderate knowledge of the governing physics. Existing multiscale and multiphysics approaches have been ineffective, especially for the latter two categories, in addressing key scientific challenges, such as inhomogeneous cascades of scales, propagation of uncertainty/stochasticity across scales, and nonlocal interactions. They are inefficient for solving inverse stochastic multiscale problems, especially in the context of noisy multifidelity data. In the PhILMs project, we address these shortcomings by seamlessly fusing physical laws and multifidelity data, for both forward *and* inverse multiscale and multiphysics problems, with the goal of discovering, modeling, and understanding the hidden physics of interfaces and beyond. Our integrated mathematical and computational activities can be classified into the following research areas (see Figure 1): (RA-I) PDE-based modeling of macroscales; (RA-II) stochastic modeling of mesoscales; (RA-III) bridging methods to connect the scales; and (RA-IV) statistical learning and deep learning approximations and algorithms. Physics-informed neural networks (PINNs) are a workhorse tool for the project, and new work in this year has emphasized operator regression. In the following, we summarize research accomplishments, highlights, and current directions from the *second year* of the project.

## 1 RA-I: PDE-Based Modeling of Macroscales

We have utilized the expressivity of DNNs to construct representations of the solution of PDE models with local and nonlocal interactions. In addition, we have started addressing the operator regression problem, both in local and nonlocal settings. In RA-I we present an overview of our main developments in algorithms, theory, and applications. Brown, PNNL, SNL, Stanford, and UCSB have contributed to RA-I.

### 1.1 Local Calculus – Algorithms

Following the PINN paradigm, several new algorithms were proposed this year collectively by the PhILMs PIs, including solving PDEs with error control, domain decomposition methods, graph-based methods, and data-driven solvers for surface PDEs on manifolds.

**Galerkin DNNs for Variational Equations with Error Control:** In [1] we use DNNs to approximate variational equations based on the adaptive construction of a sequence of finite-dimensional subspaces whose basis functions are realizations of a sequence of NNs. The finite-dimensional subspaces can be used to define a standard Galerkin approximation of the variational equation. This approach enjoys a number of advantages, including: the need to train a single, large network is replaced by the solution of a sequence of smaller problems combined with a Galerkin approach to select the best linear combination of the resulting corrections; the sequential nature of the algorithm offers a systematic approach to enhancing the accuracy of a given approximation (without having to resort to training a new, larger network); the sequential enhancements can be shown to provide a useful indicator for the error in the current NN approximation that can be used as a criterion for terminating the sequential updates; the basic approach is to some extent oblivious to the nature of the PDE under consideration beyond assuming that a stable variational formulation is available, meaning that the same approach is equally well-applicable to differing orders, including the case of singular data or singular solutions; and finally, some basic theoretical results are presented regarding the convergence (or otherwise) of the method which are used to formulate some basic guidelines for applying the method.

**Variational PINNs with Domain Decomposition:** In [2] we formulated a general Petrov-Galerkin framework for *hp*-variational PINNs based on the nonlinear approximation of DNNs and *hp*-refinement via domain decomposition and projection onto space of high-order polynomials. The trial space is the space of DNNs, which is defined globally over the whole computational domain, while the test space contains piecewise high-order polynomials. This approach is particularly effective for multiscale and multiphysics problems and it is amenable to parallelization.

**Compatible PDE Discretization Informed Learning:** Connections between PINNs and least squares finite element methods employing DNNs as a nonlinear basis [3, 4] prompted the synthesis of more advanced PDE discretizations with DNNs to engineer physics directly into architectures, rather than via penalty as done with PINNs. Work in this direction has led to the development of two new capabilities using classical numerical tools to remove numerical challenges while allowing novel imposition of properties related to conservation, entropy principles, thermodynamic consistency, and exact sequence structure.

**Control Volume PINNs (cvPINNs):** We incorporate the subdomain collocation method [5] into PINNs by minimizing a residual discretizing conservation laws over control volumes and applying Gauss' theorem providing reduced regularity requirements, a $\sim 10\times$ accuracy increase, removing the need for penalties to enforce conservation or boundary conditions to eliminate hyperparameters. It allows leveraging of classical finite volume tools for shock physics to enforce total variation bounds and entropy stability. This forms the foundation of learning equations of state from multiscale data in RA-III and leverages foundational development of fast optimizers discussed under RA-IV. Cross-center collaborations employing cvPINNs include: A. Tartakovsky (PNNL) for subsurface applications, J. Chen (SNL) for combustion problems learning for shock/ignition interactions, and E. Darve to compare performance against his ADCME framework. Internally at SNL, two projects have been funded separately from PhILMs adopting this approach to study pulsed power fusion (relevant to FES) and a national security application.

**Physics-informed Graph Neural Networks (pigNNs):** A parallel thrust unifies the discrete exterior calculus underpinning compatible PDE discretization and the combinatorial Hodge theory [6, 7]. Common mathematical structure has allowed the development of pigNNs, which exploit graph topology to develop architectures mimicking mimetic PDE discretization, providing rigorous guarantees of convergence, stability, and structure preservation. Initial proof of concept work from the past year formed the foundation for a successfully funded DOE early career project for N. Trask. In the coming year, we are planning a collaboration with P. Atzberger using the data-driven exterior calculus to study non-equilibrium dynamics associated with multiscale systems, targeting the combustion exemplar. The approach has been adopted by the electrical sciences department at SNL [8] demonstrating a potential capability to impact Basic Energy Sciences objectives in microelectronics [9].

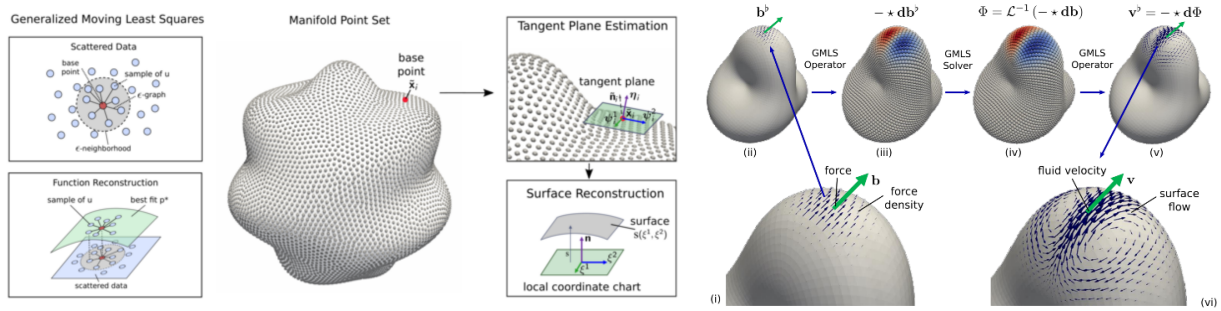



**Figure 2:** Data-driven Solvers for PDEs on Manifolds. Left: setup for approximating the geometry and operators on the surface using GMLS. Right: step for a solver developed for investigating the role of geometry in incompressible hydrodynamic flow responses on curved fluid interfaces.

**Data-Driven Numerical Solvers for Surface PDEs on Manifolds:** Many problems arising in scientific ML need effective methods for handling processes that occur within manifolds. P. Atzberger (UCSB) and N. Trask (SNL), along with P. Kuberry (SNL) and B. Gross (UCSB), developed a new class of solvers using exterior calculus approaches from differential geometry and discretizations based on a local data-driven GMLS approximations (see fig. 2). Here, data from a point cloud sampling of the local manifold geometry and vector fields were used to construct quantities arising in differential geometry and exterior calculus. Among other advantages, the use of local least-squares problems helped to circumvent technical challenges associated with more analytic approaches which must grapple with complicated analytic derivations and coordinate-based expressions. For example, to handle the incompressibility of flows for surface geometries, a generalized Hodge Decomposition was developed in this exterior calculus framework providing an approach to handle the constraints without the need for additional solves (pressure term), see fig. 2. This was based on a generalized curl operator with a local representation learned using GMLS approaches. The numerical methods provide new capabilities for solving general surface PDEs and other processes occurring on manifolds on complicated shapes. The methods were demonstrated in the context

of solving surface incompressible hydrodynamics to investigate the roles played by the geometry for surface flow responses. More details can be found in the paper [10]. Related ideas were further developed to obtain new architectures for NNs, referred to as GMLS-Nets in papers [11, 12].

### 1.2 Local Calculus – Applications

**PINNs for Fluids, Solids, and Optics:** For centuries, flow visualization has been the art of making fluid motion visible in physical and biological systems. Although such flow patterns can be, in principle, described by the Navier-Stokes equations, extracting the velocity and pressure fields directly from the images is challenging. We addressed this problem by developing hidden fluid mechanics (HFM), a physics-informed deep-learning framework capable of encoding the Navier-Stokes equations into the DNNs while being agnostic to the geometry or the initial and boundary conditions [13]. We demonstrated HFM for several physical and biomedical problems by extracting quantitative information for which direct measurements may not be possible. HFM is robust to low resolution and substantial noise in the observation data, which is important for potential applications. In fig. 3 we apply HFM to infer the velocity and pressure fields from dye visualizations only in an arbitrary domain in the wake of flow past a cylinder.

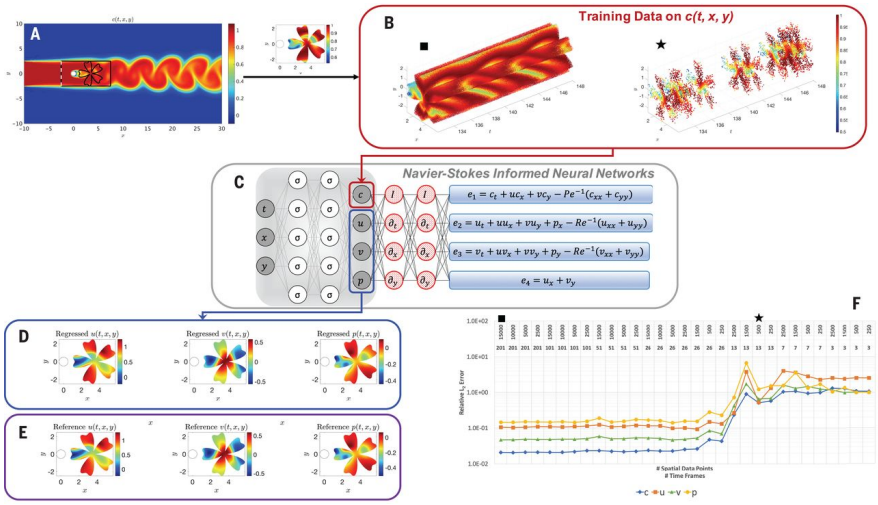



**Figure 3:** Arbitrary training domain in the wake of a cylinder. (A) Domain where the training data for concentration and reference data for velocity and pressure are generated by using direct numerical simulation. (B) Training data on concentration $c(t, x, y)$ in an arbitrary domain in the shape of a flower located in the wake of the cylinder. The solid black square corresponds to a very refined point cloud of data, whereas the solid black star corresponds to a low-resolution point cloud. (C) A PINN (left) takes the input variables $t$, $x$, and $y$ and outputs $c$, $u$, $v$, and $p$. By applying AD on the output variables, we encode the transport and Navier-Stokes equations in the PINNs $e_i$, $i = 1, ..., 4$ (right). (D) Velocity and pressure fields regressed by means of HFM. (E) Reference velocity and pressure fields obtained by cutting out the arbitrary domain in (A) are used for testing the performance of HFM. (F) Relative $L_2$ errors estimated for various spatiotemporal resolutions of observations for $c$. On the top line, we list the spatial resolution for each case, and on the line below, we list the corresponding temporal resolution over 2.5 vortex shedding cycles.

An example of using multifidelity and physics as a low-fidelity model was presented in [14]. Specifically, we were able to extract the mechanical properties of materials through deep learning from instrumented indentation. We have formulated algorithms for solving inverse problems by recourse to single, dual, and multiple indentation and demonstrate that these algorithms significantly outperform traditional brute force computations and function-fitting methods. The predictive capabilities and advantages of these multifidelity methods have been assessed by direct comparisons with experimental results for indentation for different commercial alloys, including two wrought aluminum alloys and several three-dimensional printed titanium alloys.

In another application of PINNs to inverse problems we considered nano-optics and metamaterials in [15]. We successfully apply mesh-free PINNs to the difficult task of retrieving the effective permittivity parameters of a number of infinite-size scattering systems that involve many interacting nanostructures as well as multi-component nanoparticles. The development of physics-informed deep learning techniques for inverse scattering can enable the design of novel functional nanostructures and significantly broaden the design space of metamaterials by naturally accounting for radiation and infinite-size effects beyond the limitations of traditional effective medium theories.

**NN-Based Constitutive Modeling:** We consider a parametric-free, data-driven approach for constitutive modeling. We substitute the constitutive relation by a DNN, which is then coupled with physical laws described by

PDEs. The advantage of this approach is two-fold. Firstly, we can leverage the expressivity of DNN for approximating constitutive relations without making assumptions about their forms. Secondly, we can preserve the known physics to the largest extent by incorporating constraints in the form of differential or algebraic equations (e.g., conservation laws). The constitutive relation is optimized only indirectly through the solution of the PDE. Therefore, we cannot simply train DNNs using common deep learning software. In our approach, we formulate the inverse problem as a PDE-constrained optimization problem and iteratively update the NN weights and biases until the discrepancy between the hypothetical and true observations becomes sufficiently small.

We have developed an effective and novel approach—Physics Constrained Learning (PCL)—that enables us to calculate gradients with respect to the weights and biases of DNNs that are embedded in a system of PDEs. In PCL, both DNNs and numerical PDE solvers are expressed in terms of computational graphs. Thus, we can use reverse-mode AD to calculate the gradients algorithmically. In our implementation of PCL, we use the implicit function theorem to efficiently back-propagate gradients through implicit and iterative numerical solvers. As an example of a successful outcome, we have trained a DNN to learn a generalized Maxwell viscoelasticity model (see fig. 4). We have also demonstrated our method for coupled systems of multi-phase flow and geomechanics equations. We found that NN-based constitutive relations are superior to other competing methods.

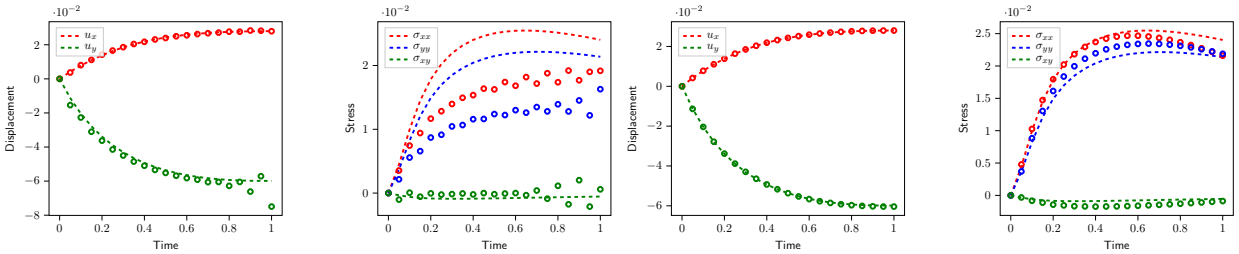



**Figure 4:** First and second plots from left: Spatially Varying linear elasticity; Third and fourth plots from left: NN-based Viscoelasticity. Displacement and stress tensors of the left top point. The dashed lines are true values, and the dots are reproduced values using the calibrated models.Compare the stress in the second and fourth plots. The DNN is more accurate than the conventional approach.

### 1.3 Nonlocal Calculus – Theory

**A Unified Nonlocal Calculus to Support ML of Nonlocal Models:** Nonlocal models (NLMs) are integro-differential equations with potential advantages over PDEs, including reduced regularity requirements for simpler descriptions of discontinuous solutions (e.g., in fracture mechanics), natural descriptions of long-range molecular interactions (such as electrokinetic and surface tension effects), and an explicit lengthscale $\delta$ which may be used to model subgrid microstructures. To support ML algorithms for NLMs and document their effectiveness, in [16] we have developed a unified theory that rigorously bridges the well-established fractional and nonlocal calculi, and provides a universal form of nonlocal operators that induce well-posed problems and can be used in learning algorithms to guarantee that the resulting model is well-defined and physically consistent. Our main result states that there exists a universal operator that describes a broad class of diffusion processes including the well-known fractional and classical diffusion: $\mathcal{L}_k[u] = \int_{B_\delta(x)}(u(y) - u(x))k(x,y)dy$, where $B_\delta(x)$ is a ball centered at $x$ of radius $\delta$, i.e., the lengthscale. Other key contributions in [16] include: the proof of equivalence of several widely used definitions of fractional vector calculus operators; the proof that the same operators are special instances of nonlocal operators for which there exists a well-established theory and that their truncated versions (i.e., with finite nonlocal interactions) converge to the non-truncated ones as the interactions become infinite; and the extension of results of the classical variational theory to a new class of nonlocal operators. Our more recent efforts include the extension of these results to vector-valued functions [17] for operators currently used for robust turbulence modeling (see section 1.4); the study of equivalence, in the sense of energy, of tempered and truncated fractional operators, allowing reproduceable actions of expensive infinite-range operators with cheaper finite-range operators [18].

**Nonlocal PINNs for Kernel Learning:** Building on the work developed in [19, 20] we proposed a nPINNs framework. First, the kernel $k$ is approximated by a DNN multiplied by a fractional kernel, to accommodate for singularities: $k(x,y) = k_{NN}(x,y;\theta,\alpha) = \rho(x,y;\theta)\,|x-y|^\alpha\,\mathcal{X}(y \in B_\delta(x))$, where $\theta$ represents weights and biases of the network and $\alpha$ is a tunable diffusion rate. Then, the optimal $\theta$ and $\alpha$ are obtained by minimizing $\sum_{i=1}^{N}\sum_{j=1}^{J_i}(\mathcal{L}_{NN}u_i(x_{j,i}) - f_i(x_{j,i}))^2$, where $\{x_{j,i}\}_{i=1}^{J_i}$ are the collocation points corresponding to the $i$th pair, $\mathcal{L}_{NN}$ is the nonlocal operator corresponding to the kernel $k_{NN}$ and $f$ the RHS of the equation $\mathcal{L}u = f$. Numerical

results for manufactured kernels with truncated interactions show that nPINNs are robust and that the DNN is able to implicitly identify the truncation parameter $\delta$, thus improving computational cost.

**Foundational Meshfree and Nonlocal Theory:** A thrust has been to further develop meshfree and nonlocal discretizations. Notable works this fiscal year include a collaboration with P. Atzberger (UCSB) resulting in publications on GMLS-Nets at NeurIPS and AAAI [21, 22], meshfree treatment of exterior calculus operators to handle flows on manifolds [23], work using graphs to develop the first conservative meshfree discretization [24] which formed the foundation for pigNNs discussed previously, and a variety of nonlocal discretizations [25, 26].

**Fractional PINNs:** In [27], we extended PINNs to fractional PINNs (fPINNs) to solve space-time fractional advection-diffusion equations, and we systematically studied their convergence, hence explaining both fPINNs and PINNs for the first time. Specifically, we demonstrated their accuracy and effectiveness in solving multidimensional forward and inverse problems with forcing terms whose values are only known at randomly scattered spatio-temporal coordinates (black-box forcing terms). A novel element of the fPINNs is the hybrid approach that we introduced for constructing the residual in the loss function using both AD for the integer-order operators and numerical discretization for the fractional operators. We solved several in three dimensions to identify the fractional orders, diffusion coefficients, and transport velocities and obtained accurate results given proper initializations even in the presence of significant noise.

### 1.4 Nonlocal Calculus – Applications

**nPINNs for Turbulence Modeling:** We have applied the nPINNs algorithm developed for nonlocal diffusion [19, 20] to parameter identification for a new nonlocal wall-bounded turbulence model. Based on [16] we introduced the universal operator $\mathcal{L}^{\delta,\alpha(y^+)}$ that recovers finite-range nonlocal operators, classical derivatives, and fractional Caputo derivatives. This parameterized nonlocal operator has been used to generalize the local shear stress equation, and nPINNs has been extended to learn its space-dependent parameter $\alpha(y^+)$, the diffusion rate as a function of the wall distance, and the nonlocal lengthscale $\delta$ (fig. 5.) This framework has thus discovered a universal behavior and contributes to the understanding of nonlocal interactions in turbulence. With the purpose of learning nonlocal closure models for large eddy simulations, we are currently using several tools developed in the first two years of PhILMs. The work above and in section 1.3 provides the groundwork for NLM identification. The extension of the nonlocal calculus in [17] indicates the existence of a universal nonlocal vector Laplacian. The nonlocal operator regression in section 1.5 showed that fractional operators can be mimicked by cheaper short-range operators and the incorporation of cvPINNs and fast optimizers improve accuracy and performance. Thus, we may learn nonlocal generalizations of the Smagorinsky model. Our strategy consists in approximating the kernel by a DNN and training it based on DNS data from the Johns Hopkins turbulence database [29] and DNS data for multiphase flows provided by J. Chen (SNL) using the Exascale Computing Project code S3D [30] to impact the combustion exemplar.

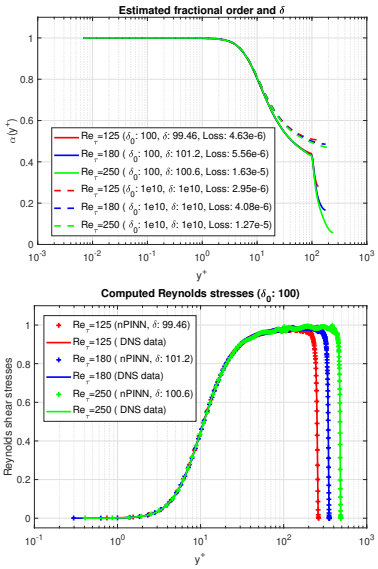



**Figure 5:** Top: Universal profile of fractional order in wall units $y^+$. We observe a *universal* profile for the rate $\alpha(y^+)$ with respect to the Reynolds number. Bottom: Computed shear stresses exhibiting a nearly perfect match with the DNS data from [28].

### 1.5 Operator Regression

We have posed as a major thrust for learning data-driven models the abstract operator regression problem, whereby one seeks to learn an operator $\mathcal{L}$ from observations of the form $\mathcal{D} = \{u_i, \mathcal{L}[u_i]\}_{i=1}^{N_{\text{data}}}$, where $u_i$ are functions sampled from a relevant Banach space $V$. Many problems related to system identification, learning closures, parameter estimation, multiscale modeling, etc. may be posed in terms of this abstract setting which encapsulates three synergistic efforts across the center. The work at Brown on DeepONets [31] is the first ever rigorous approach to operator regression and has addressed over 20 different applications of learning explicit and implicit operators. Researchers at SNL we are considering two sub-thrusts targeting learning of both local and nonlocal models; each approach learns physics without the need for an explicit dictionary of admissible model forms while working with varying choice of operator parameterization with different target applications. This provides data-driven models that impose minimal inductive biases and can capture physics such as anomalous diffusion which cannot be captured with PDE models. In RA-III we present examples of learning multiphase nanoflows atomistic simulations.

**DeepONet–Learning Nonlinear Operators Based on the Universal Approximation Theorem of Operators:** While it is widely known that DNNs are universal approximators of continuous functions, a less known and perhaps more powerful result is that a NN with a single hidden layer can approximate accurately any nonlinear continuous operator. To realize this theorem in practice and its extension to DNNs, in [31] we proposed deep operator networks (DeepONets) to learn operators accurately and efficiently from a relatively small dataset. A DeepONet consists of two sub-networks, one for encoding the input function at a fixed number of sensors $x_i, i = 1, \ldots, m$ (branch net), and another for encoding the locations for the output functions (trunk net). We performed systematic simulations for identifying two types of operators (i.e., dynamical systems and PDEs), and demonstrated that DeepONet significantly reduces the generalization error compared to the fully-connected networks. We also derived theoretically the dependence of the approximation error in terms of the number of sensors (where the input function is defined) as well as the input function type, and we verified the theorem with computational results. More importantly, we observed high-order error convergence in our computational tests, namely polynomial rates (from half order to fourth order) and even exponential convergence with respect to the training dataset size.

**Local Operator Regression:** In [32] we learned an update operator of a parabolic system of PDEs $\partial_t X = \mathcal{L}[X]$, where $\mathcal{L}$ is unknown. We adapted a modal perspective and parameterized via a DNN parameterizing a symbol in Fourier space and a second network parameterizing a nonlinear operator in physical space $\mathcal{L} = \mathcal{F}^{-1} \circ \mathcal{NN}_1 \circ \mathcal{F} \circ \mathcal{NN}_2$, where $\mathcal{F}$ denotes the Fourier transform. The Fourier transform acts as an encoding allowing parameterization of general pseudo-differential operators without the need to assume a model form or dictionary. In addition to providing performance/accuracy gains, the Fourier symbol exposes a structure to introduce physical-biases, allowing extraction of coarse-grained models extracted from MD data.

**Nonlocal Operator Regression:** The nonlocal calculus in section 1.3 inspired a framework to learn the nonlocal kernel $k$ that best describes a specific application. In [33] we proposed an operator-regression strategy that allows extraction of NLMs with partially negative kernels that are guaranteed to be solvable even in small-data regimes. Numerical tests showed that this learning algorithm can identify NLMs that accommodate the high-frequency response of a given material, a feature that is usually obtained by augmenting second-order elliptic PDEs with high-order derivatives [34]. Furthermore, this scheme is able to learn finite-range kernels that reproduce the action of (infinite-range) fractional operators. This has the potential of reducing the computational cost of simulations involving fractional operators, see section 1.4. Extensions of this method have been adopted at SNL by Dr. Silling for NLM learning from MD data for national security applications.

**Data-Driven Exponential Integrators in Small Data Regimes:** We targeted learning dynamics in data-sparse regimes, considering as model system the drift-diffusion model governing semiconductor dynamics in the presence of radiation where data are extremely sparse mandating exploitation of physics knowledge; a $pn$-junction semiconductor device with minority carriers serves as a canonical DOE application. We aimed to learn a linear dynamical system of the form $\boldsymbol{x}_{k+1} = \mathbf{A}\boldsymbol{x}_k + \mathbf{B}\boldsymbol{g}_k$. The ambipolar diffusion equation [35] provides a reasonably accurate first approximation of photocurrent effects: $\frac{\partial \delta p}{\partial t} - D_p \frac{\partial^2 \delta p}{\partial \delta x^2} + \mu_p E_n \frac{\partial \delta p}{\partial \delta x} + \frac{\delta p}{\tau_p} = g(x,t)$. Motivated by this form, we can refine the structure of the linear model above to exploit the parabolic structure of both the ambipolar diffusion equation and of exponential time integrators (ETI) [36] to inform the modification of the linear system above, which becomes $\boldsymbol{x}_{k+1} = (\mathbf{I} + \mathbf{\Phi}(\Delta_{\text{ETI}} t)\mathbf{K})\,\boldsymbol{x}_k + \mathbf{\Phi}(\Delta_{\text{ETI}} t)\boldsymbol{g}_k$, where $\mathbf{\Phi}(\Delta_{\text{ETI}} t)$ is the dynamic part of the model, advancing its state at time step $\Delta_{\text{ETI}} t$, and $\mathbf{K}$ is a spatial operator. The goal is to learn $\mathbf{\Phi}(\Delta_{\text{ETI}} t)$ and $\mathbf{K}$ from a small number of experimental data. The preliminary results we have obtained show that by including additional information about the underlying physics we can build data-driven models with acceptable accuracy from sparse data sets. The ETI model can be viewed as a hybrid between equation-based and equation-free techniques. In [37] we explored a purely equation-free approach based on dynamic mode decomposition ideas. In contrast to the ETI model, which is strongly connected to the underlying physics model, a dynamic mode decomposition approach is more flexible as it does not require significant prior knowledge of a physics model. However, in the absence of sufficient data this also limits its applicability to a smaller range of inputs.

## 2 RA-II: Stochastic Modeling of Mesoscales

We report here different approaches for modeling mesoscale phenomena: learning PDEs from MD simulation data, learning the stochastic dynamics from unpaired data, applying operator regression to predict nanobubble growth dynamics, and learning effective coarse-grained potentials for energy storage materials. PNNL, SNL, Brown, MIT, and UCSB have contributed to topics in RA-II.

## 2.1 Algorithms

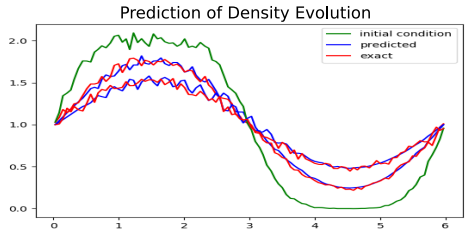



**Figure 6:** For an ensemble of particles with the initial distribution above (green), GMLS-Nets are capable of predicting the density with high level of accuracy (blue).

**Learning PDEs from Molecular-Level Simulation and Fine–Grained Continuum Models:** Over the past year, P. Atzberger (UCSB) and N. Trask (SNL) have been collaborating to investigate GMLS (GMLS-Nets) to infer PDE descriptions from molecular-level simulation data. Molecular datasets pose challenges given the scattered nature of sampled configuration data. A NN architecture generalizing convolutional NNs to scattered datasets was developed, allowing for an inference approach applied to an ensemble of material particles. This approach was used within a loss function that models a supervised learning problem where GMLS-Nets are required to learn a differential operator that makes predictions of the future density. An open source software was also developed and is available for both PyTorch at `https://github.com/atzberg/gmls-nets` and for TensorFlow at `https://github.com/rgp62/gmls-nets`. An example of inferring a PDE describing the Brownian diffusion of an ensemble of particles is shown in fig. 6 (see [11, 12] for more details). What is notable is our use of a free-form class of operators. We do not rely on a dictionary of operators as in other leading methods, yet we still retain inductive bias toward learning succinct operators similar to differential operators through the GMLS polynomial bases employed. Our approach can also be applied more generally for other molecular systems involving more complex interactions to infer materials fluxes and other transport phenomena.

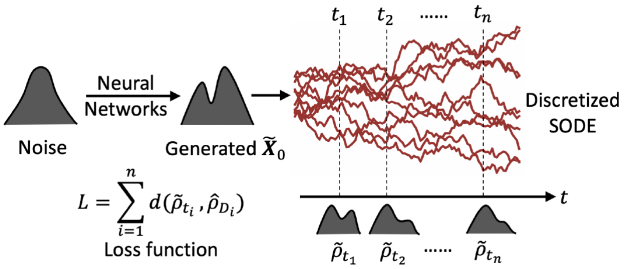



**Figure 7:** Schematic of the generative model for ensemble-regression. We first use a feed-forward NN to map the input Gaussian noise to the output, whose distribution $\tilde{\rho}_0$ is intended to approximate the initial distribution $\rho_0$. Subsequently, we apply the discretized SODE with trainable parameters to generate sample paths $\tilde{\boldsymbol{X}}_t$ for $t > 0$ with $\tilde{\boldsymbol{X}}_0$ as the initial condition (brown curves). We then measure the differences between the distributions of generated samples $\tilde{\boldsymbol{X}}_t$ and the snapshots from data as our loss function.

**Generative Ensemble-Regression: Learning Stochastic Dynamics from Discrete Particle Ensemble Observations:** In [38] we propose a new method for inferring the governing SODEs by observing particle ensembles at discrete and sparse time instants, (i.e., multiple "snapshots"), as shown in fig. 7. Particle coordinates at a single time instant, possibly noisy or truncated, are recorded in each snapshot but are unpaired across the snapshots. By training a generative model that generates "fake" sample paths, we aim to fit the observed particle ensemble distributions with a curve in the probability measure space, which is induced from the inferred particle dynamics. We employ different metrics to quantify the differences between distributions, like the sliced Wasserstein distances and the adversarial losses in generative adversarial networks (GANs). We refer to this approach as generative "ensemble-regression", in analogy to the classic "point-regression", where we infer the dynamics by performing regression in the Euclidean space, (e.g., linear/logistic regression). We illustrate the ensemble-regression by learning the drift and diffusion terms of particle ensembles governed by stochastic ordinary differential equations with Brownian motions and Lévy processes up to 20 dimensions. We also discuss how to treat cases with noisy or truncated observations, as well as the scenario of paired observations, and we prove a theorem for the convergence in Wasserstein distance for continuous sample spaces. In [39], we scale up to continuum by considering the Fokker-Planck Equation and PINNs, and develop a general framework that introduces a new loss function using the Kullback-Leibler divergence to connect the stochastic samples with the Fokker-Planck equation, to simultaneously learn the equation and infer the multi-dimensional probability density function (PDF) at all times.

## 2.2 Applications

**Polynomial-Regression-Based Simulation of Liquid/Liquid Interfaces:** For a multiple-fluid system the ability of coarse-grained (CG) models to predict accurately the interfacial properties depends crucially on the accuracy of the CG force field (FF). ML approaches can be employed to design and optimize the CG FF parameters. X. Yang (Lehigh), P. Gao (PNNL) and A. Tartakovsky (PNNL) [40] have developed a new framework that integrates the polynomial-regression-based ML model with CG MD simulations to obtain the effective parameters. The input parameters for training the ML model are the output properties calculated from the atomistic simulations, and the output for the ML model are the FF parameters used for the CG simulations.

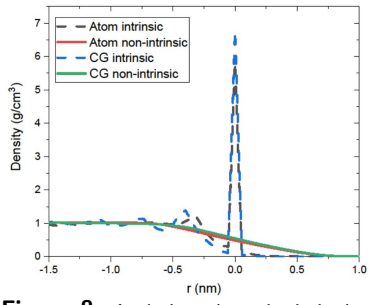



**Figure 8:** Intrinsic and non-intrinsic density profiles of a water droplet in hexane for atomistic and CG simulations.

As an application, we have developed a new CG FF for water-hexane mixture by a polynomial-regression-based ML approach in the Shinoda-DeVane-Klein CG FF framework. A standard response-surface approach does not provide a unique set of parameters as it results in a loss function with multiple shallow minima. We have developed a probabilistic ML approach where we compute the PDF of parameters that minimize the loss function. The PDF has a well-defined peak corresponding to a unique set of parameters in the CG FF that reproduces the desired properties of a liquid-liquid interface. The CG model was trained with interfacial tension data from a high-curvature planar water-hexane interface and tested for a system with different curvatures. The new CG FF is not only able to reproduce the interfacial tensions of planar and curved interfaces in atomistic MD simulations, but also improves the prediction of both the intrinsic and non-intrinsic density profiles on the water-hexane interface (see fig. 8).

**Operator Learning for Predicting Nano-Bubble Growth:** At the mesoscopic scale where the continuum assumption breaks down, we developed an extended Dissipative Particle Dynamics (DPD) system where many-body DPD is used to simulate the liquid and standard DPD is used to simulate the gas. This coupled method showed advantages in pressure and surface tension description. Considering the mesoscale thermal fluctuations, we run several simulations and used the POD method to extract the mean trajectory and the time-varying noise level. The two parts are then learned by DeepONet and LSTM, respectively. This trained model, which can predict mesoscale bubble evolution, contains the information of both the ensemble average and the thermal fluctuation level.

**ML Potentials for Energy Storage Materials:** Energy storage materials are an integral part of modern society. The properties and dynamics of electrode, electrolyte, and their interface play an important role in the performance of energy storage and conversion devices such as batteries and fuel cells. CG simulations are needed to study the relationship between material structure and properties since *ab initio* calculations are usually very expensive. The CG molecular models are constructed using data from large-scale atomistic computer simulations that depend heavily on the accuracy of the interatomic potentials. Most classical interatomic potentials are based on physical intuition and contain few adjustable parameters. However, they can lack accuracy, especially for systems exhibiting polarization effects like the electrolyte-electrode interface. ML potentials, which include many-body interactions implicitly, can achieve highly accurate interpolation results for training data coming from DFT. P. Gao (PNNL) and A. Tartakovsky (PNNL) have built a new NN ML potential of lithium electrode to be used for atomistic simulations. To improve its transferability the physical nature of interatomic bonding was enforced. This is achieved by combining physics-based bond-order potential with a NN regression. The physics-based bond-order potential parameters were added to the loss function to ensure that the variations of the ML potential parameters relative to their database-averaged bond-order potential parameter values are small. A MD simulation with such ML potential can accurately reproduce the macroscopic properties as computed by a DFT calculation. Figure 9 shows the stress-strain curve of a lithium nanowire. The MD simulation with ML potential can reproduce the yielding behavior as well as Young's modulus.

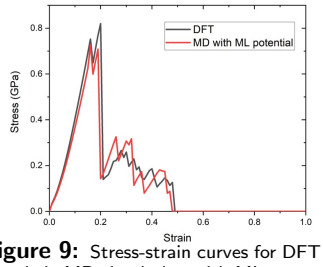



**Figure 9:** Stress-strain curves for DFT and atomistic MD simulation with ML potential.

## 3 RA-III: Bridging of Methods to Connect the Scales

We have continued to develop methods for the concurrent coupling of heterogeneous domains, including coupling across scales, to discover the hidden physics models at the interfaces and provide the "glue" functions/functionals to seamlessly connect the cascade of scales. Brown, PNNL, and SNL have contributed to RA-III.

### 3.1 Algorithms

**Active- and Transfer-Learning Applied to Microscale-Macroscale Coupling to Simulate Viscoelastic Flows:** In [41] we applied active- and transfer-learning to polymer flows for the multiscale discovery of effective constitutive approximations required in viscoelastic flow simulation. The result is macroscopic rheology directly connected to a microstructural model. Micro and macroscale simulations are adaptively coupled using Gaussian process regression to run the expensive microscale computations only as necessary. This active-learning guided multiscale method can automatically detect the inaccuracy of the learned constitutive closure and initiate

simulations at new sampling points informed by proper acquisition functions, leading to an autonomic microscale-macroscale coupled system. The effective closure learned in a channel simulation is then transferred directly to the flow past a circular cylinder, where the results show that only two additional microscopic simulations are required to achieve a satisfactory constitutive model to once again close the continuum equations. This new paradigm of active- and transfer-learning for multiscale modeling is readily applicable to other microscale-macroscale coupled simulations of complex fluids and other materials.

**Conservative PINNs on Discrete Domains for Conservation Laws:** In [42] we consider multiscale problems and propose a conservative PINN (cPINN) formulation on discrete domains for nonlinear conservation laws. Here, the term discrete domain represents the discrete sub-domains obtained after division of the computational domain, where PINN is applied and the conservation property of cPINN is obtained by enforcing the flux continuity in the strong form along the sub-domain interfaces. In case of hyperbolic conservation laws, the convective flux contributes at the interfaces, whereas in case of viscous conservation laws, both convective and diffusive fluxes contribute. One can also employ a DNN in the domain, where the solution may have complex structure, whereas a shallow NN can be used in the sub-domains with relatively simple and smooth solutions. Another advantage of the proposed method is the additional freedom it gives in terms of the choice of optimization algorithm and the various training parameters like residual points, activation function, width and depth of the network, etc. Various test cases ranging from scalar nonlinear conservation laws, like Burgers or Korteweg-de Vries equations, to systems of conservation laws, like compressible Euler equations, are solved.

### 3.2 Applications

**Employing Tools from RA-I to Consider Multiscale Phenomena**

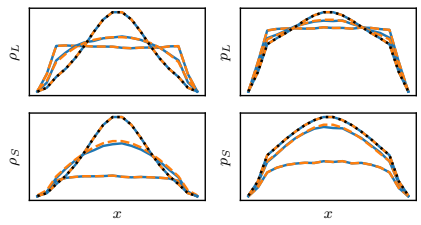



**Figure 10:** The extracted data-driven model for concentration and velocity is able to achieve predictive solutions for long-term integration and generalizes well beyond data used in training. No dictionary is involved in training; instead a combination of physical biases and modal encoding provide a predictive "black-box" model.

*Numerical Homogenization of Multiphase Continuum Models from Molecular Data:* We summarize here, an application of the operator regression framework from RA-I [32] to extract a multiphase continuum model from MD simulations of a bidisperse Leonard-Jonesian fluid in a microchannel (see fig. 10). With a short burst of MD simulations, we can learn the long-term dynamics of the overall system, allowing accurate predictions of velocity and concentration profiles at times an order of magnitude longer than those used in training. The resulting model generalizes well, as training over varying physical parameters (e.g., particle density, volume fractions) leads to accurate predictions beyond those included in the MD training data.

*Data-Driven Extraction of Thermodynamically Consistent Equations of State from Molecular Models:* In this work, we assume that the EOS model is unknown a priori, and we seek to learn a black-box relationship $p = \mathcal{NN}(\rho, E)$. This is particularly relevant to learning shock hydrodynamics of solids in the high energy density regime. To achieve this, it is necessary to impose constraints enforcing compatibility with the first and second laws of thermodynamics. In fig. 11, we demonstrate fitting shock hydro models to MD simulations of copper and DSMC simulations of Argon. Enforcing thermodynamic constraints allows generalization and extrapolation of PINNs to unseen regimes while preserving hyperbolicity of the governing system. We are leveraging a constrained spline approach suggested in a comprehensive survey by M. Gulian [43] on constrained Gaussian process regression to enforce physical constraints on the EOS predicted by the PINN.

**Identification of Nonlocal Lernels for Multiscale Simulations of Heterogeneous Materials:** The applications of NLMs to problems involving strong heterogeneities and interfaces is hampered by the lack of a rigorous nonlocal interface theory. In [44, 45, 46] we investigated two approaches to handle material interfaces and heterogeneities; the key is the identification of an appropriate nonlocal kernel that yields a well-posed, physically consistent, NLM.

*Physically Consistent Kernels for Multiscale Nonlocal Simulations:* In [44] we developed an optimization-based method to couple NLMs with different length scales and model parameters. Through minimization of the nonlocal energy of the system, our method identifies the optimal kernel across the interface by enforcing consistency with physical laws at the limit of vanishing nonlocality. This work contributed to the design of a new general framework for nonlocal domain decomposition [47].

*Improved Fractional Kernels for Subsurface Flow Simulations:* In [46] we introduced new fractional models of

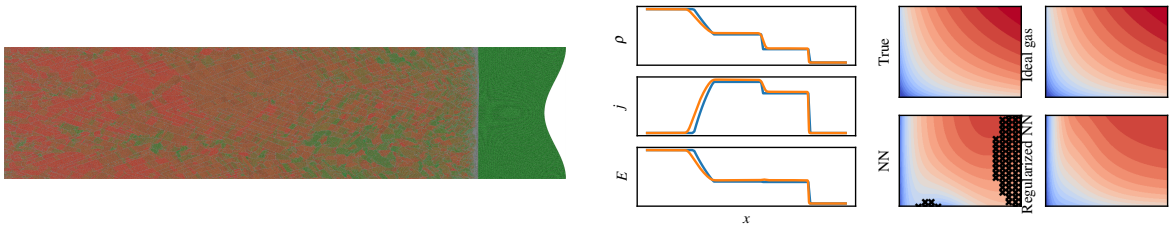



**Figure 11:** MD simulation of a normal shock propagating from left to right through a copper bar. Bottom: Extracted shock hydrodynamics model surrogate used to learn a corresponding equation of state. Bottom-left: cvPINNs allow use of standard finite volume tools to obtain non-oscillatory shock treatment and find physically relevant entropy solution. Bottom-right: Extracted EOS from 4 DSMC simulations of Argon gas. Assuming an appropriate a priori model form (i.e., estimating $\gamma$ in ideal gas law) provides a good EOS prediction. Using a black box DNN generalizes poorly, providing nonphysical EOS, which cause the system to lose hyperbolicity (marked with x's). Imposing thermodynamic constraints recovers generalizability without a priori known model form.

improved descriptive power, with kernels $k(x,y) = \rho(x,y)/|x-y|^{n+s(x,y)}\mathcal{X}(|x-y| < \delta(x))$, which allow descriptions of small-scale subsurface behavior at a much larger scale. We are employing these NLMs in the cross-center collaboration on model learning via cvPINNs with A. Tartakovsky, mentioned in RA-I.For the same model, in collaboration with the MMICC center Aeolus, we designed an optimization-based learning algorithm [45] for model parameters.

**PINNs for Modeling Non-Newtonian Fluid Dynamics:** In complex non-Newtonian fluids of suspended particles or polymers, the apparent viscosity of the fluid increases with increasing volume fraction of the suspended particles and decreasing shear rate. Simulations that resolve the system at the particle scale are very expensive, so there is a need for continuum-scale models of these systems. We have extended the PINN method to learn viscosity models of two non-Newtonian systems (polymer melts and suspensions of particles) using only velocity measurements [48]. The PINN-inferred viscosity models agree with the empirical models for shear rates with large absolute values but deviate for shear rates near zero where the analytical models have an unphysical singularity. This work allows prediction of the stresses in non-Newtonian fluids with only the average streamwise velocity, a quantity easily measured in experiments.

**PINNs for Multiphysics Data Assimilation - Subsurface Transport in Porous Media:** Data assimilation for parameter and state estimation in subsurface transport problems remains a significant challenge due to the sparsity of measurements, the heterogeneity of porous media, and the high computational cost of forward numerical models [49, 50]. We have developed a multiphysics-informed DNN method [51] for simultaneously estimating space-dependent hydraulic conductivity $K(\vec{x})$, hydraulic head $h(\vec{x})$, and tracer concentration $C(\vec{x})$ fields from sparse measurements. We model $K(\vec{x})$, $h(\vec{x})$, and $c(\vec{x})$ with three individual feed-forward DNNs, and we train them by using the residual of the governing PDEs in addition to the given measurements. For comparison, we tested the pure "data-based" DNN approach that trains the networks with only measurements with $L_2$ regularization to suppress over-fitting. PINN provides a much better estimation of $K$ and $C$ than the data-based DNNs. We have also applied PINNs to the unsteady advection-dispersion equation, and the PINN method agrees well with the analytical solutions even for problems with high Péclet number.

## 4 RA-IV: Statistical Learning and Deep Learning Approximations and Algorithms

We have continued to develop foundational algorithms on optimization, uncertainty quantification, GANs, as well as new methods for data augmentation for faster training, and fundamental new theory for generalization. Brown, Stanford, MIT, and PNNL have contributed to this research area.

### 4.1 Algorithms

**PINNs Coupled with Conditional Gaussian Process:** We have developed a physics-informed DNN coupled with a conditional Gaussian process ML method for estimating space-dependent coefficient $\kappa(\mathbf{x})$, and state of a PDE from sparse measurements. Different from standard PINNs, we approximate the space-dependent coefficient $\kappa(\mathbf{x})$ by a conditional Gaussian process $\kappa(\mathbf{x},\boldsymbol{\xi})$ such that any realization of $\kappa(\mathbf{x},\boldsymbol{\xi})$ matches the measurements $\{\kappa_i\}_{i=1}^N$. Instead of solving the three-term optimization problem $\arg\min_{\boldsymbol{\xi},\boldsymbol{\theta}} \sum_{i=1}^M (|u_i - \tilde{u}(\mathbf{x}_{u_i},\boldsymbol{\theta})|^2) + \delta_1 \sum_{i=1}^N (|\kappa_i - \tilde{\kappa}(\mathbf{x}_{\kappa_i},\boldsymbol{\xi})|^2) + \delta_2 \|\mathcal{L}(\kappa(x,\boldsymbol{\xi}),\tilde{u}(x,\boldsymbol{\theta}))\|_{L_2}^2$, we solve the two-term optimization problem $\arg\min_{\boldsymbol{\xi},\boldsymbol{\theta}} \sum_{i=1}^M (|u_i - \tilde{u}(x_i,\boldsymbol{\theta})|^2) + \delta \|\mathcal{L}(\kappa(x,\boldsymbol{\xi}),\tilde{u}(x,\boldsymbol{\theta}))\|_{L_2}^2$, where $\mathcal{L}$ represents the partial differential operator.

This approach further reduces the desired number of observations to obtain acceptable accurate estimates. We observed that with the same number of $\kappa$ measurements and state observations this coupled training approach can achieve more accurate estimates for both $\kappa$ and $\tilde{u}$ than standard PINNs.

**Potential Flow Generator with $L_2$ Optimal Transport Regularity for Generative Models:** In [52], we proposed a potential flow generator with $L_2$ optimal transport regularity, which can be easily integrated into a wide range of generative models including different versions of GANs and flow-based models. We show the correctness and robustness of the potential flow generator in several 2D problems, and illustrate the concept of "proximity" due to the $L_2$ optimal transport regularity. Subsequently, we demonstrate the effectiveness of the potential flow generator in image translation tasks with unpaired training data from the MNIST dataset and the CelebA dataset.

**Bayesian PINNs for Forward and Inverse PDE problems with Noisy Data:** In [53] we proposed a *Bayesian PINN* (B-PINN) to solve both forward and inverse nonlinear problems described by PDEs and noisy data. In this Bayesian framework, the Bayesian NN combined with a PINN for PDEs serves as the prior, while the Hamiltonian Monte Carlo (HMC), or the variational inference (VI), could serve as an estimator of the posterior. B-PINNs make use of both physical laws *and* scattered noisy measurements to provide predictions and quantify the *aleatoric uncertainty* arising from the noisy data in the Bayesian framework. Compared with PINNs, in addition to uncertainty quantification, B-PINNs obtain more accurate predictions in scenarios with large noise due to their capability of avoiding overfitting. We conducted a systematic comparison between the two different approaches for the B-PINNs posterior estimation (i.e., HMC or VI), along with dropout used for quantifying uncertainty in DNNs. Our experiments showed that HMC is more suitable than VI with mean field Gaussian approximation for the B-PINNs posterior estimation, while dropout employed in PINNs can barely provide accurate predictions with reasonable uncertainty.

**GANs for Learning Stochastic Dynamical Systems:** P. Atzberger (UCSB), P. Stinis (PNNL), and C. Daskalakis (MIT) have been collaborating to investigate deep learning generative modeling approaches for learning models for stochastic dynamical systems when there is access to observables or partial information on system state evolution. To affect the generative modeling, we are employing GANs. We are utilizing recent results developed by C. Daskalakis on subadditive inequalities satisfied by many metrics and divergences that can be used in the GAN framework to accelerate training and improve the accuracy of the trained model. However, using GANs on physical systems also poses novel challenges to capture constraints such as energy and momentum conservation or statistical mechanical analogs in the stochastic setting (fluctuation-dissipation balance, detailed-balance). Thus, we are working to formulate GAN frameworks that utilize structure in order to achieve statistical gains from physical datasets, and to use related mathematical results to develop more reliable prediction models.

### 4.2 Training

**Fast Hybrid Optimizers for Training:** Block coordinate descent prescribes a broad class of optimizers employing splittings of relevant variables to extract easily solved subproblems. In [54] we present a scheme exploiting a recent trend in nonlinear approximation theory interpreting DNNs as an adaptive basis [55, 56] to identify an effective splitting. For regression and PINN loss functions the model reduces to solving a least squares problem at each iteration and has been shown to provide a $O(10-100\times)$ increase in accuracy using $O(10-100\times)$ fewer iterations. We have recently extended this splitting to consider classification examples [57]. A key feature of this work has been the ability to train networks, which realize a notion of convergence with width and depth, albeit only linear. Unpublished work concluding this fall involves a modification to develop an architecture, which demonstrates high-order hp-convergence while avoiding the curse of dimensionality.

**Min-Max Optimization:** Min-max optimization, used among other classical applications in GANs and in adversarial training more broadly, is empirically challenging. Understanding its computational complexity and/or developing empirically well-performing algorithms for even arriving at locally and approximately optimal solutions has been very challenging. C. Daskalakis (MIT) [58] sheds light on why min-max optimization is computationally intractable. We study the optimization problem $\min_x \max_y f(x,y)$ with $(x,y) \in S$, where $S \subseteq \mathbb{R}^d$ is a compact convex set and $f$ is $L$-smooth and $G$-Lipschitz, but not necessarily convex in $x$ and concave in $y$. In the absence of the latter, classical min-max/convex-programming duality theorems do not apply. A complexity-theoretic result established by Daskalakis et al. shows that the problem is PPAD-complete, which means that no algorithm (first-order, second-order, or whatever) can compute approximate local min-max equilibria in polynomial time unless P=PPAD. Thus, while min-max equilibria of two-player zero-sum games with convex-concave objectives are

tractable, even approximate and local min-max equilibria of nonconvex-nonconcave objectives are as intractable as Nash equilibria in general-sum normal-form games. The broader implications of the afore-described results are that, while the impressive recent progress of deep learning owes to the ability of gradient descent (GD) and its variants to discover approximately optimal local minima, which also generalize well for many interesting learning tasks, GANs, adversarial training, and broader applications wherein multiple learning agents interact cannot be founded on a similar expectation that GD or more general light-weight optimization methods will arrive at approximately optimal local min-max solutions.

**Plateau Phenomenon in GD Training of Rectified Linear Unit (ReLU) Networks:** In general, NNs are trained by GD type optimization methods, or a stochastic variant thereof. In practice, such methods result in the loss function rapidly decreasing at the beginning of training, but then, after a relatively small number of steps, it significantly slow down. The loss may even appear to stagnate over the period of a large number of epochs, only to then suddenly start to decrease fast again for no apparent reason. This so-called plateau phenomenon manifests itself in many learning tasks. In [59] we aim to identify and quantify the root causes of plateau phenomenon. No assumptions are made on the number of neurons relative to the number of training data, and the results hold for both the lazy and adaptive regimes. The main findings are: plateaux correspond to periods during which activation patterns remain constant, where activation pattern refers to the number of data points that activate a given neuron; quantification of convergence of the gradient flow dynamics; and, characterization of stationary points in terms of solutions of local least squares regression lines over subsets of the training data. Based on these conclusions, a new iterative training method is proposed, the Active Neuron Least Squares, characterized by the explicit adjustment of the activation pattern at each step, which is designed to enable a quick exit from a plateau.

**Sample Amplification and Data Augmentation:** G. Valiant (Stanford) investigated several different questions related to *data augmentation*, the process of taking a given dataset and somehow producing a larger dataset to train on. Even though data augmentation can impart significant boosts in performance, almost no theory exists to explain how and why certain augmentation strategies are better than others or to understand the possibilities and limitations of augmentation. In [60], we provide some rigorous understanding of why and how various augmentations work. We consider a family of linear transformations and study their effects on the ridge estimator in an over-parametrized linear regression setting. First, we show that transformations preserving the labels of the data can improve estimation by enlarging the span of the training data. Second, we show that transformations in which mix data can improve estimation by playing a regularization effect. Finally, we validate our theoretical insights on MNIST. Based on the insights, we propose an augmentation scheme that searches over the space of transformations by how uncertain the model is about the transformed data. We validate our proposed scheme on image and text datasets. For example, our method outperforms RandAugment by 1.24% on CIFAR-100 using Wide-ResNet-28-10. Furthermore, we achieve comparable accuracy to the SoTA Adversarial AutoAugment on CIFAR datasets. In a rather different vein, in [61], we address the extremely fundamental question of how/when "perfect" augmentation is possible: Given data drawn from an unknown distribution, $D$, to what extent is it possible to "amplify" this dataset and faithfully output an even larger set of samples that appear to have been drawn from $D$? We formalize this question as follows: an $(n, m)$ amplification procedure takes as input n independent draws from an unknown distribution $D$, and outputs a set of $m > n$ "samples" which must be indistinguishable from $m$ samples drawn iid from $D$. We consider this sample amplification problem in two fundamental settings: the case where $D$ is an arbitrary discrete distribution supported on k elements, and the case where $D$ is a $d$-dimensional Gaussian with unknown mean, and fixed covariance matrix. Perhaps surprisingly, we show a valid amplification procedure exists for both of these settings, even in the regime where the size of the input dataset, $n$, is significantly less than what would be necessary to learn distribution D to non-trivial accuracy. We also show that our procedures are optimal up to constant factors. Beyond these results, we describe potential applications of sample amplification and formalize a number of curious directions for future research.

**Locally Adaptive Activation Functions with Slope Recovery for Deep Learning and PINNs:** In [62], we propose two approaches of locally adaptive activation functions namely, layer-wise and neuron-wise locally adaptive activation functions, which improve the performance of deep and physics-informed NNs. The local adaptation of activation function is achieved by introducing a scalable parameter in each layer (layer-wise) and for every neuron (neuron-wise) separately, and then optimizing it using a variant of SGD algorithm. In order to further increase the training speed, an activation slope-based slope recovery term is added in the loss function, which further accelerates convergence, thereby reducing the training cost. On the theoretical side, we prove

that in the proposed method the GD algorithms are not attracted to sub-optimal critical points or local minima under practical conditions on the initialization and learning rate and that the gradient dynamics of the proposed method is not achievable by base methods with any (adaptive) learning rates. We further show that the adaptive activation methods accelerate the convergence by implicitly multiplying conditioning matrices to the gradient of the base method without any explicit computation of the conditioning matrix and the matrix–vector product. The different adaptive activation functions are shown to induce different implicit conditioning matrices.

### 4.3 Generalization

**Quantifying the Generalization Error in Deep Learning in Terms of Data Distribution and NN Smoothness:** The accuracy of deep learning (i.e., DNNs), can be characterized by dividing the total error into three main types: approximation error, optimization error, and generalization error. Whereas there are some satisfactory answers to the problems of approximation and optimization, much less is known about the theory of generalization. To derive a meaningful bound, we study the generalization error of NNs for classification problems in terms of data distribution and NN smoothness in [63]. We introduce the cover complexity to measure the difficulty of learning a data set and the inverse of modules of continuity to quantify NN smoothness. A quantitative bound for expected accuracy/error is derived by considering both the cover complexity and NN smoothness. We validate our theoretical results by several data sets of images. The numerical results verify that the expected error of trained networks scaled with the square root of the number of classes has a linear relationship with respect to the cover complexity. In addition, we observe a clear consistency between test loss and NN smoothness during the training process.

**On the Convergence and Generalization of PINNs:** In [3] we established a mathematical foundation of PINNs. As the number of data grows, PINNs generate a sequence of minimizers that correspond to a sequence of NNs. We posed the question: Does the sequence of minimizers converge to the solution to the PDE? This question is also related to the generalization of PINNs. We considered two classes of PDEs: elliptic and parabolic. By adapting the Schuader approach, we showed that the sequence of minimizers strongly converges to the PDE solution in $L_2$. Furthermore, we showed that if each minimizer satisfies the initial/boundary conditions the convergence mode can be improved to $H^1$. Computational examples were provided to illustrate our theoretical findings. To the best of our knowledge, this is the first theoretical work that shows the consistency of the PINNs methodology.

**Implicit Regularization for DNNs Driven by an Ornstein-Uhlenbeck-like Process:** This work [64] is motivated by the grand challenge of explaining—in a rigorous way—why deep learning performs as well as it does. Despite the explosion of interest in deep learning, driven by many practical successes across numerous domains, there are many basic mysteries regarding why it works so well. Why do networks with orders of magnitude more parameters than the dataset size, trained via SGD, often yield trained networks with small generalization error, despite the fact that such networks and training procedures are capable of fitting even randomly labeled training points [65]? Why do deeper networks tend to generalize better, as opposed to worse, as one might expect given their increased expressivity? Why does the test performance of deep networks often continue to improve after their training loss plateaus or reaches zero? In [64], we introduce a framework that sheds light on the above questions. Our analysis focuses on deep networks, trained via SGD, but where the gradient updates are computed with respect to noisy training labels. Specifically, for a SGD update for training data point $x$ and corresponding label $y$, the gradient is computed for the point $(x, y+Z)$ for some zero-mean, bounded random variable $Z$, chosen independently at each step of SGD. We analyze this specific form of SGD with independent label noise because such training dynamics seem to reliably produce "simple" models, independent of network initialization, even when trained on a small number of data points. This is not true for SGD without label noise, which has perhaps hindered attempts to rigorously formalize the sense in which training dynamics leads to "simple" models. We do, however, discuss the possibility that a variant of our analysis might apply to SGD without label noise, provided the training set is sufficiently large and complex that the randomness of SGD mimics the effects of the explicit label noise that we consider. Our main result characterizes the zero-training-error attractive fixed points of the dynamics of SGD with label noise and $\ell_2$ loss, in terms of the local optima of an implicit regularization term. This result holds for any network architecture, and any activation functions. We also illustrate the implications of this general characterization in three basic settings for which the implicit regularization term can be easily analyzed: matrix sensing as in [66], 2-layer ReLU networks trained on one-dimensional data, and 2-layer networks with logistic or tanh activations trained on a single labeled datapoint. In all three cases, empirically training via

SGD with label noise yields "simple" models, where training without label noise results in models that are not simple and that depend on the initialization. Our theory explains this, and we show that in these settings, being at a local optima of the implicit regularizer reduces the set of models to *only* "simple" ones.

## 5 Software Dissemination

**ADCME.jl:** (Stanford) The NN-based constitutive modeling has been developed using ADCME.jl, which is our dedicated software for inverse modeling problems in computational engineering. In the past year, this software has been heavily developed and is now equipped with several new functionalities. They include in particular new modules that support general finite element grids and discretizations, and distributed computing capabilities via MPI. This extends existing capabilities such as the ability to run in parallel on multicore processors and GPU processors. ADCME.jl is available to download at `https://github.com/kailaix/ADCME.jl` and `https://kailaix.github.io/ADCME.jl/dev/`.

**DeepXDE – A Deep Learning Library for Solving Differential Equations:** (Brown) In [67] we present an overview of PINNs, which embed a PDE into the loss of the NN using AD. The PINN algorithm is simple, and it can be applied to different types of PDEs, including integro-differential equations, fractional PDEs, and stochastic PDEs. Moreover, from the implementation point of view, PINNs solve inverse problems as easily as forward problems. We propose a new residual-based adaptive refinement method to improve the training efficiency of PINNs. For pedagogical reasons, we compare the PINN algorithm to a standard finite element method. We also present a Python library for PINNs, DeepXDE, which is designed to serve both as an educational tool to be used in the classroom as well as a research tool for solving problems in computational science and engineering. Specifically, DeepXDE can solve forward problems given initial and boundary conditions, as well as inverse problems given some extra measurements. DeepXDE supports complex-geometry domains based on the technique of constructive solid geometry and enables the user code to be compact, resembling closely the mathematical formulation (fig. 12). We introduce the usage of DeepXDE and its customizability, and we also demonstrate the capability of PINNs and the user-friendliness of DeepXDE for five different examples. More broadly, DeepXDE contributes to the more rapid development of the emerging Scientific ML field. DeepXDE has been uploaded into PyPI (`https://pypi.org/project/DeepXDE/` and Anaconda (`https://anaconda.org/conda-forge/deepxde`). As of September 2020 it has reached 56k downloads.

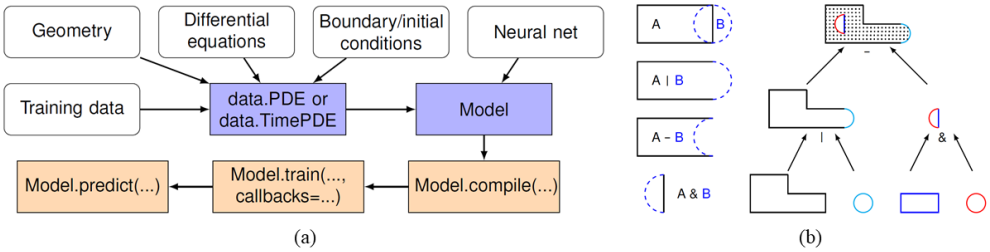



**Figure 12:** Schematic of DeepXDE. (a) Flowchart of DeepXDE. The white boxes define the PDE problem and the training hyperparameters. The blue boxes combine the PDE problem and training hyperparameters in the white boxes. The orange boxes are the three steps (from right to left) to solve the PDE. (b) Examples of constructive solid geometry in 2D. (left) A and B represent the rectangle and circle, respectively. The union $A|B$, difference $A - B$, and intersection $A\&B$ are constructed from A and B. (right) A complex geometry (top) is constructed from a polygon, a rectangle, and two circles (bottom) through the union, difference, and intersection operations. This capability is included in the module geometry of DeepXDE.

## 6 Integration and Outreach

Dissemination and outreach have been made more difficult due to the pandemic. For example, a PhILMs-led SciML-themed workshop in the planning stages was shelved back in March when it became clear such in-person activities no longer made sense. However, we have leveraged electronic means for dissemination as well as our existing broad professional networks to communicate the results of PhILMs and have many examples of external adoption of PhILMs capabilities by laboratories, universities, and industry as depicted in fig. 13.

**PhILMs Broad Outreach:** We have a regular webinar presenting forum for external speakers (listed on the PhILMs website) and internal discussions about strategic research directions. Publications and presentations have also served as a primary means of scientific outreach (see Appendix A). We will have a substantial presence at select future conferences, both online and in person, including SIAM CSE and the Mechanistic Machine Learning and Digital Twins for Computational Science, Engineering & Technology Conference.

**PhILMs Internal Collaborations:** PhILMs cross-institutional collaborations are embodied in the technical results. As a team, we have been working with and training postdocs and PhD students, who are fully engaged in all PhILMs activities, including PhILMs webinars, and relevant conferences and workshops. All academic PIs have visited both PNNL and SNL, as well as each other's universities, for seminars or PhILMs project meetings.

**PhILMs Intra- and Inter-Laboratory Collaborations:** National laboratory PIs have been active in engaging with ML-related projects and application owners to locate additional opportunities for impact. At SNL, PhILMs work has been adopted by mission applications funded by ASC, including Z-machine applications, plasma physics applications, and an NVBL effort providing a PINN prototype for fitting compartment models for COVID, as well as by LDRD- and DARPA-funded project work. At PNNL, A. Tartakovsky contributed to the Science-informed Machine Learning for Accelerating Real-Time Decisions in Subsurface Applications (SMART) Initiative (`https://edx.netl.doe.gov/smart`).PhILMs efforts have been adopted by projects at LANL, ANL, and ORNL through joint collaboration. An example is the collaboration of Karniadakis with the Los Alamos plasma team (X. Tang) on the DeepFusion project.

**PhILMs External University Collaborations:** PNNL has active collaborations with the University of Washington, Washington State University, and the University of Pennsylvania. Staff at SNL have joint PhILMs-related collaborations with faculty at Michigan State, the University of Tennessee, the University of Pittsburgh, Florida State, U. Trier, Lehigh University, the University of Texas at Austin, the University of Illinois, Tufts, and Boston University. Brown has collaborations with MIT, Johns Hopkins University, and University of Tokyo.

**PhILMs Industrial Collaborations and Technology Adoption:** PhILMs has active collaborations with NVIDIA, ANSYS, LaVision, and SIEMENS. NVIDIA has recruited a postdoc of Karniadakis (M. Raissi) and implemented a parallel version of PINNS, called SimNet [68]. ANSYS has recruited two researchers from the CRUNCH group of Karniadakis. We have collaborated also with LaVision who provided Schlieren video data to us for inferring fluid mechanics in the mask/no-mask application for COVID-19 and also for inferring the velocity and pressure fields over an espresso cup, which they validated with additional PIV experiments.

**Diversity and Inclusion:** The PhILMs team, as well as each institution involved[1], is committed to diversity and inclusion activities to strengthen the ML and data science community. Karniadakis participated in a unique University of Texas at El Paso conference and townhall meeting (09-08-2020) on strengthening the national foundation for advanced modeling and simulation at HBCUs/MIs as a panelist (only one of three university professors), along with other DOE, NSF, and AFOSR representatives, and he proposed specific plans of close collaboration of Brown and PhILMs with MIs, their students, and faculty.

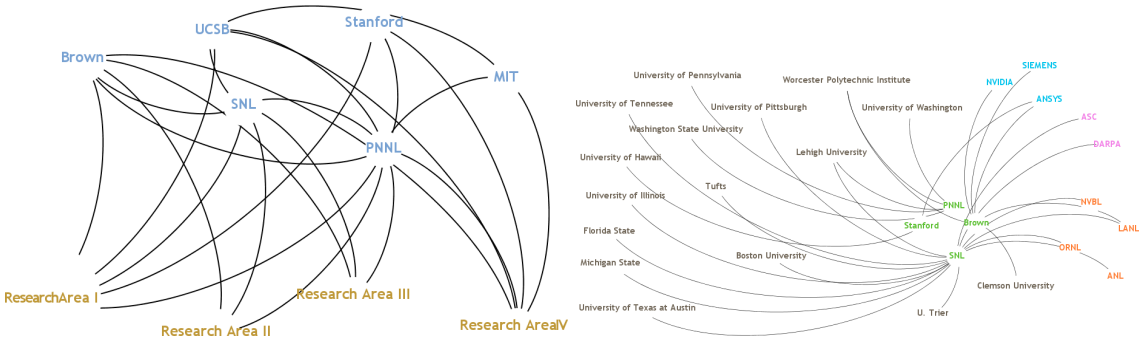



**Figure 13:** PhILMs collaboration (left) and outreach (right) graphs.

[1] `https://www.pnnl.gov/diversity-inclusion`, `https://www.sandia.gov/about/diversity/`, `https://www.brown.edu/about/administration/institutional-diversity/`, `https://equity.stanford.edu/`, `https://diversity.mit.edu/`, `https://diversity.ucsb.edu/`

## References


1. Ainsworth, M. & Dong, J. Galerkin Neural Networks: A Framework for Approximating Variational Equations with Error Control. *SIAM Journal on Scientific Computing (submitted)* (2020).
2. Kharazmi, E., Zhang, Z. & Karniadakis, G. E. Variational physics-informed neural networks for solving partial differential equations. *arXiv preprint arXiv:1912.00873* (2019).
3. Shin, Y., Darbon, J. & Karniadakis, G. E. On the convergence and generalization of physics informed neural networks. *arXiv preprint arXiv:2004.01806* (2020).
4. Bochev, P. B. & Gunzburger, M. D. *Least-squares finite element methods* (Springer Science & Business Media, 2009).
5. Chang, C.-L. & Gunzburger, M. D. A finite element method for first order elliptic systems in three dimensions. *Applied Mathematics and Computation* **23,** 171–184 (1987).
6. Bochev, P. B. & Hyman, J. M. in *Compatible spatial discretizations* 89–119 (Springer, 2006).
7. Dodziuk, J. Combinatorial and continuous Hodge theories. *Bulletin of the American Mathematical Society* **80,** 1014–1016 (1974).
8. Gao, X., Huang, A., Trask, N. & Reza, S. *Physics-Informed Graph Neural Network for Circuit Compact Model Development* tech. rep. (Sandia National Laboratories, SAND2020-9413, 2018).
9. Murray, C. *et al. Basic Research Needs for Microelectronics: Report of the Office of Science Workshop on Basic Research Needs for Microelectronics, October 23–25, 2018* tech. rep. (USDOE Office of Science (SC)(United States), 2018).
10. Gross, B., Trask, N., Kuberry, P. & Atzberger, P. Meshfree methods on manifolds for hydrodynamic flows on curved surfaces: A Generalized Moving Least-Squares (GMLS) approach. *Journal of Computational Physics* **409,** 109340. ISSN: 0021-9991. `http://www.sciencedirect.com/science/article/pii/S0021999120301145` (2020).
11. Trask, N., Patel, R., Atzberger, P. & Gross, B. GMLS-Nets: A machine learning framework for unstructured data. *Proceedings of AAAI-MLPS.* `http://ceur-ws.org/Vol-2587/article_9.pdf` (2020).
12. Trask, N., Patel, R. G., Gross, B. J. & Atzberger, P. J. GMLS-Nets: A framework for learning from unstructured data. *arXiv:1909.05371.* `https://arxiv.org/abs/1909.05371` (Sept. 2019).
13. Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367,** 1026–1030 (2020).
14. Lu, L. *et al.* Extraction of mechanical properties of materials through deep learning from instrumented indentation. *Proceedings of the National Academy of Sciences* **117,** 7052–7062 (2020).
15. Chen, Y., Lu, L., Karniadakis, G. E. & Dal Negro, L. Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Optics Express* **28,** 11618–11633 (2020).
16. D'Elia, M., Gulian, M., Olson, H. & Karniadakis, G. E. A Unified Theory of Fractional, Nonlocal, and Weighted Nonlocal Vector Calculus. *arXiv preprint arXiv:2005.07686* (2020).
17. Gulian, M., D'Elia, M., Mengesha, T. & Scott, J. M. *A generalized nonlocal calculus for vector-valued functions: theory and applications* In progress. 2020.
18. Olson, H., Gulian, M. & D'Elia, M. *On the equivalence of tempered fractional and truncated fractional operators* CSRI summer proceedings 2020, Sandia National Laboratories. In progress. 2020.
19. Pang, G., D'Elia, M., Parks, M. & Karniadakis, G. nPINNs: nonlocal Physics-Informed Neural Networks for a parametrized nonlocal universal Laplacian operator. Algorithms and Applications. *Journal of Computational Physics* **422,** 109760 (2020).
20. Pang, G., D'Elia, M., Parks, M. & Karniadakis, G. *Nonlocal Physics-Informed Neural Networks - A unified theoretical and computational framework for nonlocal models.* in *AAAI Spring Symposium: MLPS* (2020).
21. Trask, N., Patel, R. G., Gross, B. J. & Atzberger, P. J. GMLS-Nets: A framework for learning from unstructured data. *arXiv preprint arXiv:1909.05371* (2019).

22. Trask, N., Patel, R. G., Atzberger, P. J. & Gross, B. J. *GMLS-Nets: A Machine Learning Framework for Unstructured Data*. in *AAAI Spring Symposium: MLPS* (2020).

23. Gross, B. J., Trask, N., Kuberry, P. & Atzberger, P. J. Meshfree methods on manifolds for hydrodynamic flows on curved surfaces: a generalized moving least-squares (GMLS) approach. *Journal of Computational Physics* **409,** 109340 (2020).

24. Trask, N., Bochev, P. & Perego, M. A conservative, consistent, and scalable meshfree mimetic method. *Journal of Computational Physics* **409,** 109187 (2020).

25. You, H., Lu, X., Trask, N. & Yu, Y. An asymptotically compatible approach for Neumann-type boundary condition on nonlocal problems. *ESAIM: Mathematical Modelling and Numerical Analysis* **54,** 1373–1413 (2020).

26. Trask, N. & Kuberry, P. Compatible meshfree discretization of surface PDEs. *Computational Particle Mechanics* **7,** 271–277 (2020).

27. Pang, G., Lu, L. & Karniadakis, G. E. fPINNs: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing* **41,** A2603–A2626 (2019).

28. Avsarkisov, V., Hoyas, S., Oberlack, M. & Garcia-Galache, J. P. Turbulent plane Couette flow at moderately high Reynolds number. *Journal of Fluid Mechanics* **751** (2014).

29. Li, Y. *et al. Johns Hopkins Turbulence Database (JHTDB)* 2008.

30. Chen, J. H. *et al.* Terascale direct numerical simulations of turbulent combustion using S3D. *Computational Science & Discovery* **2,** 015001 (2009).

31. Lu, L., Jin, P. & Karniadakis, G. E. DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv preprint arXiv:1910.03193* (2019).

32. Patel, R. G., Trask, N. A., Wood, M. A. & Cyr, E. C. A physics-informed operator regression framework for extracting data-driven continuum models.

33. You, H., Yu, Y., Trask, N., Gulian, M. & D'Elia, M. *Data-driven learning of robust nonlocal physics from high-fidelity synthetic data* arXiv:2005.10076. 2020.

34. Weckner, O. & Silling, S. A. Determination of the constitutive model in peridynamics from experimental dispersion data. *International Journal of Multiscale Computational Engineering, under review* (2011).

35. McKelvey, J. P. *Solid State and Semiconductor Physics* (R.E. Krieger Publishing Co., Malabar, FL, 1986).

36. Hochbruck, M. & Ostermann, A. Exponential integrators. *Acta Numerica* **19,** 209–286 (2010).

37. Hanson, J., Bochev, P. & Paskaleva, B. Learning Compact Physics-Aware Photocurrent Models Using Dynamic Mode Decomposition. *Statistical Analysis and Data Mining* **Submitted** (2020).

38. Yang, L., Daskalakis, C. & Karniadakis, G. E. Generative Ensemble-Regression: Learning Stochastic Dynamics from Discrete Particle Ensemble Observations. *arXiv preprint arXiv:2008.01915* (2020).

39. Chen, X., Yang, L., Duan, J. & Karniadakis, G. E. Solving Inverse Stochastic Problems from Discrete Particle Observations Using the Fokker-Planck Equation and Physics-informed Neural Networks. *arXiv preprint arXiv:2008.10653* (2020).

40. Gao, P., Yang, X. & Tartakovsky, A. M. Learning Coarse-Grained Potentials for Binary Fluids. *Journal of Chemical Information and Modeling* **60,** 3731–3745 (2020).

41. Zhao, L. *et al.* Active- and transfer-learning applied to microscale-macroscale coupling to simulate viscoelastic flows. *arXiv preprint arXiv:2005.04382* (2020).

42. Jagtap, A. D., Kharazmi, E. & Karniadakis, G. E. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering* **365,** 113028 (2020).

43. Swiler, L., Gulian, M., Frankel, A., Safta, C. & Jakeman, J. A Survey of Constrained Gaussian Process Regression: Approaches and Implementation Challenges. *arXiv preprint arXiv:2006.09319* (2020).

44. Capodaglio, G., D'Elia, M., Bochev, P. & Gunzburger, M. An energy-based coupling approach to nonlocal interface problems. *Computers and Fluids.* To appear (2020).

45. Burkovska, O., Glusa, C. & D'Elia, M. *An optimization-based approach to parameter learning for fractional type nonlocal models* 2020.

46. D'Elia, M. & Glusa, C. *A fractional model for anomalous diffusion with increased variability. Analysis, algorithms and applications to interface problems* In progress. 2020.

47. Capodaglio, G. *et al. A general framework for substructuring-based domain decomposition methods for models having nonlocal interactions* arXiv:2008.11780. 2020.

48. Reyes, B., Howard, A. A., Perdikaris, P. & Tartakovsky, A. M. Learning Unknown Physics of non-Newtonian Fluids. *arXiv preprint arXiv:2009.01658* (2020).

49. Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B. A. & Vesselinov, V. V. *Inverse modeling of subsurface flow and transport properties: A review with new developments* 2008.

50. Rajib, A. I., Assumaning, G. A., Chang, S.-Y. & Addai, E. B. Use of Multiple Data Assimilation Techniques in Groundwater Contaminant Transport Modeling. *Water Environment Research* **89,** 1952–1960. ISSN: 1061-4303 (2017).

51. He, Q., Barajas-Solano, D., Tartakovsky, G. & Tartakovsky, A. M. Physics-informed neural networks for multiphysics data assimilation with application to subsurface transport. *Advances in Water Resources,* 103610 (2020).

52. Yang, L. & Karniadakis, G. E. Potential Flow Generator with $L_2$ Optimal Transport Regularity for Generative Models. *arXiv preprint arXiv:1908.11462* (2019).

53. Yang, L., Meng, X. & Karniadakis, G. E. B-PINNs: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *arXiv preprint arXiv:2003.06097* (2020).

54. Cyr, E. C., Gulian, M. A., Patel, R. G., Perego, M. & Trask, N. A. *Robust training and initialization of deep neural networks: An adaptive basis viewpoint* in *Mathematical and Scientific Machine Learning* (2020), 512–536.

55. He, J., Li, L., Xu, J. & Zheng, C. ReLU deep neural networks and linear finite elements. *arXiv preprint arXiv:1807.03973* (2018).

56. Opschoor, J. A., Petersen, P. C. & Schwab, C. Deep ReLU networks and high-order finite element methods. *SAM, ETH Zürich* (2019).

57. Patel, R. G., Trask, N. A., Gulian, M. A. & Cyr, E. C. A block coordinate descent optimizer for classification problems exploiting convexity. *arXiv preprint arXiv:2006.10123* (2020).

58. Daskalakis, C., Skoulakis, S. & Zampetakis, M. The Complexity of Constrained Min-Max Optimization. *arXiv preprint arXiv:2009.09623* (2020).

59. Ainsworth, M. & Shin, Y. Plateau Phenomenon in Gradient Descent Training of ReLU networks: Explanation, Quantification and Avoidance. *arXiv preprint arXiv:2007.07213* (2020).

60. Wu, S., Zhang, H., Valiant, G. & Re, C. *On the Generalization Effects of Linear Transformations in Data Augmentation* in *International Conference on Machine Learning (ICML)* (2020).

61. Axelrod, B., Garg, S., Sharan, V. & Valiant, G. *Sample Amplification: Increasing dataset size even when learning is impossible* in *International Conference on Machine Learning (ICML)* (2020).

62. Jagtap, A. D., Kawaguchi, K. & Karniadakis, G. E. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. *Proceedings of the Royal Society A* **476,** 20200334 (2020).

63. Jin, P., Lu, L., Tang, Y. & Karniadakis, G. E. Quantifying the generalization error in deep learning in terms of data distribution and neural network smoothness. *Neural Networks* **130,** 85–99 (2020).

64. Blanc, G., Gupta, N., Valiant, G. & Valiant, P. *Implicit regularization for deep neural networks driven by an Ornstein-Uhlenbeck like process* in *Conference on Learning Theory (COLT)* (2020).

65. Zhang, C., Bengio, S., Hardt, M., Recht, B. & Vinyals, O. Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530* (2016).

66. Li, Y., Ma, T. & Zhang, H. *Algorithmic regularization in over-parameterized matrix sensing and neural networks with quadratic activations* in *Conference on Learning Theory (COLT)* (2018).

67. Lu, L., Meng, X., Mao, Z. & Karniadakis, G. E. DeepXDE: A deep learning library for solving differential equations. *arXiv preprint arXiv:1907.04502* (2019).

68. *NVIDIA SimNet: AI-Accelerated Simulation Toolkit,* `https://developer.nvidia.com/simnet`. 2020.

# A Presentations and publications

The PhILMs team measures its success, in part, on its fundamental contributions to the scientific literature, focusing primarily on mathematics and materials science, as well as the generation and dissemination of algorithms and open-source software. A Publications list also is available online at `https://www.pnnl.gov/computing/philms/Publications.stm`.

## A.1 Publications


1. Aadithya, K., Kuberry, P., Paskaleva, B., Bochev, P., Leeson, K., Mar, A., Mei, T. & Keiter, E. *Data-driven Compact Models for Circuit Design and Analysis* in (eds Lu, J. & Ward, R.) **107** (PMLR, Princeton University, Princeton, NJ, USA, 2020), 555–569. `http://proceedings.mlr.press/v107/aadithya20a.html`

2. Aadithya, K., Kuberry, P., Paskaleva, B., Bochev, P., Leeson, K., Mar, A., Mei, T. & Keiter, E. *Development, Demonstration, and Validation of Data-driven Compact Diode Models for Circuit Simulation and Analysis* 2020. arXiv: `2001.01699 [cs.LG]`

3. Ainsworth, M. & Mao, Z. Phase field crystal based prediction of temperature and density dependence of elastic constants through a structural phase transition. *Physical Review B* **100,** 104101 (2019)

4. Blumers, A. L., Li, Z. & Karniadakis, G. E. Supervised parallel-in-time algorithm for long-time Lagrangian simulations of stochastic dynamics: Application to hydrodynamics. *Journal of Computational Physics* **393,** 214–228 (2019)

5. Burkovska, O., Glusa, C. & D'Elia, M. *An optimization-based approach to parameter learning for fractional type nonlocal models* 2020

6. Capodaglio, G., D'Elia, M., Bochev, P. & Gunzburger, M. An energy-based coupling approach to nonlocal interface problems. *Computers and Fluids*. To appear (2020)

7. Capodaglio, G., D'Elia, M., Gunzburger, M., Bochev, P., Klar, M. & Vollmann, C. *A general framework for substructuring-based domain decomposition methods for models having nonlocal interactions* arXiv:2008.11780. 2020

8. Chen, X., Duan, J. & Karniadakis, G. E. Learning and meta-learning of stochastic advection-diffusion-reaction systems from sparse measurements. *arXiv preprint arXiv:1910.09098* (2019)

9. Chen, Y., Lu, L., Karniadakis, G. E. & Dal Negro, L. Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Optics Express* **28,** 11618–11633 (2020)

10. Dagan, Y., Daskalakis, C., Dikkala, N. & Kandiros, A. V. Estimating Ising Models from One Sample. *arXiv preprint arXiv:2004.09370* (2020)

11. Daskalakis, C., Skoulakis, S. & Zampetakis, M. The Complexity of Constrained Min-Max Optimization. *arXiv preprint arXiv:2009.09623* (2020)

12. D'Elia, M., Gulian, M., Olson, H. & Karniadakis, G. E. A Unified Theory of Fractional, Nonlocal, and Weighted Nonlocal Vector Calculus. *arXiv preprint arXiv:2005.07686* (2020)

13. D'Elia, M. & Bochev, P. Formulation, analysis and computation of an optimization-based local-to-nonlocal coupling method. *arXiv preprint arXiv:1910.11214* (2019)

14. D'Elia, M. & Glusa, C. *A fractional model for anomalous diffusion with increased variability. Analysis, algorithms and applications to interface problems* In progress. 2020

15. D'Elia, M., Gulian, M., Olson, H. & Karniadakis, G. E. A Unified Theory of Fractional, Nonlocal, and Weighted Nonlocal Vector Calculus. *arXiv preprint arXiv:2005.07686* (2020)

16. Ding, M., Daskalakis, C. & Feizi, S. When Do Local Discriminators Work? On Subadditivity of Probability Divergences. *arXiv preprint arXiv:2003.00652* (2020)

17. Fan, T., Xu, K., Pathak, J. & Darve, E. Solving Inverse Problems in Steady State Navier-Stokes Equations using Deep Neural Networks. *arXiv preprint arXiv:2008.13074* (2020)

18. Fan, D., Yang, L., Triantafyllou, M. S. & Karniadakis, G. E. Reinforcement Learning for Active Flow Control in Experiments. *arXiv preprint arXiv:2003.03419* (2020)

19. Gao, P., Yang, X. & Tartakovsky, A. M. Learning Coarse-Grained Potentials for Binary Fluids. *Journal of Chemical Information and Modeling* **60,** 3731–3745 (2020)

20. Gross, B., Trask, N., Kuberry, P. & Atzberger, P. Meshfree methods on manifolds for hydrodynamic flows on curved surfaces: A Generalized Moving Least-Squares (GMLS) approach. *Journal of Computational Physics* **409,** 109340. ISSN: 0021-9991. `http://www.sciencedirect.com/science/article/pii/S0021999120301145` (2020)

21. Gross, B. J., Trask, N., Kuberry, P. & Atzberger, P. J. Meshfree methods on manifolds for hydrodynamic flows on curved surfaces: a generalized moving least-squares (GMLS) approach. *Journal of Computational Physics* **409,** 109340 (2020)

22. Gulian, M., D'Elia, M., Mengesha, T. & Scott, J. M. *A generalized nonlocal calculus for vector-valued functions: theory and applications* In progress. 2020

23. Hanson, J., Bochev, P. & Paskaleva, B. Learning Compact Physics-Aware Photocurrent Models Using Dynamic Mode Decomposition. *Statistical Analysis and Data Mining* **Submitted** (2020)

24. He, Q., Barajas-Solano, D., Tartakovsky, G. & Tartakovsky, A. M. Physics-informed neural networks for multiphysics data assimilation with application to subsurface transport. *Advances in Water Resources,* 103610 (2020)

25. Huang, D. Z., Xu, K., Farhat, C. & Darve, E. Predictive modeling with learned constitutive laws from indirect observations. *arXiv preprint arXiv:1905.12530* (2019)

26. Huang, D. Z., Xu, K., Farhat, C. & Darve, E. Learning constitutive relations from indirect observations using deep neural networks. *Journal of Computational Physics,* 109491 (2020)

27. Jagtap, A. D., Kawaguchi, K. & Karniadakis, G. E. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. *Proceedings of the Royal Society A* **476,** 20200334 (2020)

28. Jagtap, A. D., Kawaguchi, K. & Karniadakis, G. E. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics* **404,** 109136 (2020)

29. Jagtap, A. D., Kharazmi, E. & Karniadakis, G. E. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering* **365,** 113028 (2020)

30. Jin, P., Zhang, Z., Zhu, A., Tang, Y. & Karniadakis, G. E. SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems. *Neural Networks* (2020)

31. Jin, P., Lu, L., Tang, Y. & Karniadakis, G. E. Quantifying the generalization error in deep learning in terms of data distribution and neural network smoothness. *Neural Networks* **130,** 85–99 (2020)

32. Kharazmi, E., Zhang, Z. & Karniadakis, G. E. hp-VPINNs: Variational Physics-Informed Neural Networks With Domain Decomposition. *arXiv preprint arXiv:2003.05385* (2020)

33. Kharazmi, E., Zhang, Z. & Karniadakis, G. E. Variational physics-informed neural networks for solving partial differential equations. *arXiv preprint arXiv:1912.00873* (2019)

34. Li, D., Xu, K., Harris, J. M. & Darve, E. Time-lapse full waveform inversion for subsurface flow problems with intelligent automatic differentiation. *arXiv preprint arXiv:1912.07552* (2019)

35. Li, D., Xu, K., Harris, J. M. & Darve, E. F. Time-lapse Full-waveform Inversion of Dynamic Geophysical Processes with Examples from Multi-phase Flow. *AGUFM* **2019,** H32A–05 (2019)

36. Li, D., Xu, K., Harris, J. M. & Darve, E. Coupled Time-Lapse Full-Waveform Inversion for Subsurface Flow Problems Using Intrusive Automatic Differentiation. *Water Resources Research* **56,** e2019WR027032 (2020)

37. Lu, L., Dao, M., Kumar, P., Ramamurty, U., Karniadakis, G. E. & Suresh, S. Extraction of mechanical properties of materials through deep learning from instrumented indentation. *Proceedings of the National Academy of Sciences* **117,** 7052–7062 (2020)

38. Lu, L., Jin, P. & Karniadakis, G. E. DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv preprint arXiv:1910.03193* (2019)

39. Meng, X., Li, Z., Zhang, D. & Karniadakis, G. E. PPINN: Parareal physics-informed neural network for time-dependent PDEs. *Computer Methods in Applied Mechanics and Engineering* **370,** 113250 (2020)

40. Meng, X. & Karniadakis, G. E. A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems. *Journal of Computational Physics* **401,** 109020 (2020)

41. Olson, H., Gulian, M. & D'Elia, M. *On the equivalence of tempered fractional and truncated fractional operators* CSRI summer proceedings 2020, Sandia National Laboratories. In progress. 2020

42. Pang, G., D'Elia, M., Parks, M. & Karniadakis, G. nPINNs: nonlocal Physics-Informed Neural Networks for a parametrized nonlocal universal Laplacian operator. Algorithms and Applications. *Journal of Computational Physics* **422,** 109760 (2020)

43. Pang, G., Lu, L. & Karniadakis, G. E. fPINNs: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing* **41,** A2603–A2626 (2019)

44. Patel, R. G., Trask, N. A., Gulian, M. A. & Cyr, E. C. A block coordinate descent optimizer for classification problems exploiting convexity. *arXiv preprint arXiv:2006.10123* (2020)

45. Patel, R. G., Trask, N. A., Wood, M. A. & Cyr, E. C. A physics-informed operator regression framework for extracting data-driven continuum models. *CMAME* **Submitted** (2020)

46. Pazner, W., Trask, N. & Atzberger, P. J. Stochastic Discontinuous Galerkin Methods (SDGM) based on fluctuation-dissipation balance. *Results in Applied Mathematics* **4,** 100068 (2019)

47. Pourzanjani, A., Jiang, R. M., Mitchell, B., Atzberger, P. J. & Petzold, L. General Bayesian Inference over the Stiefel Manifold via the Givens Transform. *Bayesian Analysis* (July 2020)

48. Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367,** 1026–1030 (2020)

49. Reyes, B., Howard, A. A., Perdikaris, P. & Tartakovsky, A. M. Learning Unknown Physics of non-Newtonian Fluids. *arXiv preprint arXiv:2009.01658* (2020)

50. Shin, Y., Darbon, J. & Karniadakis, G. E. On the convergence and generalization of physics informed neural networks. *arXiv preprint arXiv:2004.01806* (2020)

51. Shin, Y. & Karniadakis, G. E. Trainability of ReLU Networks and data-dependent initialization. *Journal of Machine Learning for Modeling and Computing* **1** (2020)

52. Trask, N., Bochev, P. & Perego, M. A conservative, consistent, and scalable meshfree mimetic method. *Journal of Computational Physics* **409,** 109187 (2020)

53. Trask, N. & Kuberry, P. Compatible meshfree discretization of surface PDEs. *Computational Particle Mechanics* **7,** 271–277 (2020)

54. Trask, N., Patel, R. G., Gross, B. J. & Atzberger, P. J. GMLS-Nets: A framework for learning from unstructured data. *arXiv:1909.05371.* `https://arxiv.org/abs/1909.05371` (Sept. 2019)

55. Wang, Y., Li, Z., Ouyang, J. & Karniadakis, G. E. Controlled release of entrapped nanoparticles from thermoresponsive hydrogels with tunable network characteristics. *Soft Matter* **16,** 4756–4766 (2020)

56. Wang, Y., Li, Z., Xu, J., Yang, C. & Karniadakis, G. E. Concurrent coupling of atomistic simulation and mesoscopic hydrodynamics for flows over soft multi-functional surfaces. *Soft matter* **15,** 1747–1757 (2019)

57. Xu, K. & Darve, E. Adversarial Numerical Analysis for Inverse Problems. *arXiv preprint arXiv:1910.06936* (2019)

58. Xu, K., Li, D., Darve, E. & Harris, J. M. Learning Hidden Dynamics using Intelligent Automatic Differentiation. *arXiv preprint arXiv:1912.07547* (2019)

59. Xu, K. & Darve, E. The neural network approach to inverse problems in differential equations. *arXiv preprint arXiv:1901.07758* (2019)

60. Xu, K. & Darve, E. *Calibrating multivariate Lévy processes with neural networks* in *Mathematical and Scientific Machine Learning* (2020), 207–220

61. Xu, K., Tartakovsky, A. M., Burghardt, J. & Darve, E. Inverse Modeling of Viscoelasticity Materials using Physics Constrained Learning. *arXiv preprint arXiv:2005.04384* (2020)

62. Xu, K. & Darve, E. Isogeometric collocation method for the fractional Laplacian in the 2D bounded domain. *Computer Methods in Applied Mechanics and Engineering* **364,** 112936 (2020)

63. Xu, K., Huang, D. Z. & Darve, E. Learning constitutive relations using symmetric positive definite neural networks. *arXiv preprint arXiv:2004.00265* (2020)

64. Xu, K. & Darve, E. Physics constrained learning for data-driven inverse modeling from sparse observations. *arXiv preprint arXiv:2002.10521* (2020)

65. Yang, L., Daskalakis, C. & Karniadakis, G. E. Generative Ensemble-Regression: Learning Stochastic Dynamics from Discrete Particle Ensemble Observations. *arXiv preprint arXiv:2008.01915* (2020)

66. Yang, L. & Karniadakis, G. E. Potential Flow Generator with $L_2$ Optimal Transport Regularity for Generative Models. *arXiv preprint arXiv:1908.11462* (2019)

67. Yang, L., Zhang, D. & Karniadakis, G. E. Physics-informed generative adversarial networks for stochastic differential equations. *SIAM Journal on Scientific Computing* **42,** A292–A317 (2020)

68. Yang, L., Meng, X. & Karniadakis, G. E. B-PINNs: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *arXiv preprint arXiv:2003.06097* (2020)

69. You, H., Lu, X., Trask, N. & Yu, Y. An asymptotically compatible approach for Neumann-type boundary condition on nonlocal problems. *ESAIM: Mathematical Modelling and Numerical Analysis* **54,** 1373–1413 (2020)

70. You, H., Yu, Y., Trask, N., Gulian, M. & D'Elia, M. *Data-driven learning of robust nonlocal physics from high-fidelity synthetic data* arXiv:2005.10076. 2020

71. Zhang, D., Guo, L. & Karniadakis, G. E. Learning in modal space: Solving time-dependent stochastic PDEs using physics-informed neural networks. *SIAM Journal on Scientific Computing* **42,** A639–A665 (2020)

72. Zhang, K., Li, Z., Maxey, M., Chen, S. & Karniadakis, G. E. Self-cleaning of hydrophobic rough surfaces by coalescence-induced wetting transition. *Langmuir* **35,** 2431–2442 (2019)

73. Zhao, L., Li, Z., Wang, Z., Caswell, B., Ouyang, J. & Karniadakis, G. E. Active- and transfer-learning applied to microscale-macroscale coupling to simulate viscoelastic flows. *arXiv preprint arXiv:2005.04382* (2020)

74. Zheng, Q., Zeng, L. & Karniadakis, G. E. Physics-informed semantic inpainting: Application to geostatistical modeling. *Journal of Computational Physics* **419,** 109676 (2020)

75. Zhu, W., Xu, K., Darve, E. & Beroza, G. C. A General Approach to Seismic Inversion with Automatic Differentiation. *arXiv preprint arXiv:2003.06027* (2020)

### A.1.1 Peer-reviewed conference papers

1. Axelrod, B., Garg, S., Sharan, V. & Valiant, G. *Sample Amplification: Increasing dataset size even when learning is impossible* in *International Conference on Machine Learning (ICML)* (2020)

2. Blanc, G., Gupta, N., Valiant, G. & Valiant, P. *Implicit regularization for deep neural networks driven by an Ornstein-Uhlenbeck like process* in *Conference on Learning Theory (COLT)* (2020)

3. Brustle, J., Cai, Y. & Daskalakis, C. *Multi-Item Mechanisms without Item-Independence: Learnability via Robustness* in *Proceedings of the 21st ACM Conference on Economics and Computation (EC)* (2020)

4. Daskalakis, C., Foster, D. & Golowich, N. *Decoupled Policy Gradient Methods for Competitive Reinforcement Learning* in *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS)* (2020)

5. Daskalakis, C. & Zampetakis, M. *More Revenue from Two Samples via Factor Revealing SDPs* in *Proceedings of the 21st ACM Conference on Economics and Computation (EC)* (2020)

6. Daskalakis, C., Fishelson, M., Lucier, B., Syrgkanis, V. & Velusamy, S. *Simple, Credible, and Approximately-Optimal Auctions* in *Proceedings of the 21st ACM Conference on Economics and Computation (EC)* (2020)

7. Daskalakis, C., Dikkala, N. & Panageas, I. *Logistic regression with peer-group effects via inference in higher-order Ising models* in *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)* (2020)

8. Daskalakis, C., Rohatgi, D. & Zampetakis, M. *Constant-Expansion Suffices for Compressed Sensing with Generative Priors* in *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS)* (2020)

9. Daskalakis, C., Rohatgi, D. & Zampetakis, M. *Truncated Linear Regression in High Dimensions* in *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS)* (2020)

10. Golowich, N., Pattathil, S., Daskalakis, C. & Ozdaglar, A. E. *Last Iterate is Slower than Averaged Iterate in Smooth Convex-Concave Saddle Point Problems* in *Proceedings of the 33nd Annual Conference on Learning Theory (COLT)* (2020)

11. Golowich, N., Pattathil, S. & Daskalakis, C. *Tight last-iterate convergence rates for no-regret learning in multi-player games* in *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS)* (2020)

12. Ilyas, A., Zampetakis, E. & Daskalakis, C. *A Theoretical and Practical Framework for Regression and Classification from Truncated Samples* in *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)* (2020)

13. Pang, G., D'Elia, M., Parks, M. & Karniadakis, G. *Nonlocal Physics-Informed Neural Networks - A unified theoretical and computational framework for nonlocal models.* in *AAAI Spring Symposium: MLPS* (2020)

14. Stinis, P. Enforcing constraints for time series prediction in supervised, unsupervised and reinforcement learning. *Proceedings of AAAI-MLPS.* `http://ceur-ws.org/Vol-2587/article_5.pdf` (2020)

15. Trask, N., Patel, R., Atzberger, P. & Gross, B. GMLS-Nets: A machine learning framework for unstructured data. *Proceedings of AAAI-MLPS.* `http://ceur-ws.org/Vol-2587/article_9.pdf` (2020)

16. Wu, S., Zhang, H., Valiant, G. & Re, C. *On the Generalization Effects of Linear Transformations in Data Augmentation* in *International Conference on Machine Learning (ICML)* (2020)

### A.2 Conferences and workshops

#### A.2.1 Invited presentations

Atzberger P.J. "Geometric Approaches for Machine Learning in the Sciences and Engineering(Offsite link)", Invited to give seminar at University of California, Davis, May 2020.

Bochev, P.B. "Mimetic meshfree methods or how to be compatible without a mesh", Conference on Computational Mathematics and Applications", October 2019, Las Vegas, NV. Invited.

Bochev, P.B. "Development of data-driven models for radiation-induced photocurrent effects", *MLDL Workshop*, Sandia National Laboratories, August 2020.

Bochev, P.B. "Data driven exponential integrators for parabolic PDEs", *Eringen Medal Symposium at SES-2020*. Virtual. September 2020.

Darve E. "Deep Neural Networks for Inverse Modeling," *MIT Distinguished Seminar Series in Computational Science and Engineering,* November 2019 (Distinguished Seminar).

Daskalakis, C. "Game Theory and Computation", *Fundraising Gala for Institut des Hautes Études Scientifiques (IHES)*, November 2019, Harvard Club of NYC, New York, NY. (Public Lecture).

Daskalakis, C. "Learning from Censored and Dependent Data", *Brown University Kanellakis Lecture*, December 2019, Providence, RI. (Distinguished Lecture).

Daskalakis, C. "The Promise and Threat of Artificial Intelligence", *Concert Hall of Athens*, December 2019, Athens, Greece. (Public Lecture).

Daskalakis, C. "The Promise and Threat of Artificial Intelligence", *Eugenides Foundation*, January 2020, Athens, Greece. (Public Lecture).

Daskalakis, C. "Statistical Inference from Dependent Observations". *National Technical University of Athens*, January 2020, Athens, Greece. (Invited Talk).

Daskalakis, C. "How Computer Science is Changing the World", *Democritus University of Thrace*, January 2020, Xanthi, Greece. (Public Lecture).

Daskalakis, C. "Statistical Inference from Dependent Observations", *Institute for Advanced Studies Computer Science/Discrete Mathematics Seminar*, March 2020, Princeton, NJ. (Invited Talk).

Daskalakis, C. "Min-Max Optimization and Deep Learning", *Institute for Advanced Studies Special Year in Optimization, Statistics, and Theoretical Machine Learning Seminar*, March 2020, Princeton, NJ. (Invited Talk).

Daskalakis, C. "Robust Learning from Censored Data", *MIT-Microsoft Research Trustworthy and Robust AI Collaboration Workshop*, June 2020, Cambridge, MA. (Invited Talk, virtually).

Daskalakis, C. "Game Theory and Machine Learning", *Max Planck Institute for Intelligent Systems Machine Learning Summer School*, July 2020, Tuebingen, Germany. (Summer Course, virtually).

Daskalakis, C. "Learning from Biased Data", *National Centre of Scientific Research "Demokritos" Summer School*, July 2020, Athens, Greece. (Public Lecture, virtually).

Daskalakis, C. "The Complexity of Min-Max Optimization", *Université de Montreal Machine Learning-Optimization Seminar*, July 2020, Montreal, Canada. (Invited Talk, virtually).

Daskalakis, C. "Learning from Biased Data", *MIT Brains, Minds, and Machines Summer Course*, August 2020, Cambridge, MA. (Invited Lecture, virtually).

Daskalakis, C. "How does Machine Learning fail, and what to do about it?", *ERC organized session on "Artificial Intelligence: A blessing or a threat for society?" at EuroScience Open Forum (ESOF)*, September 2020, Trieste, Italy. (Invited talk and Panel, virtually).

Daskalakis, C. "Three ways Machine Learning fails and what to do about them", *Columbia University Computer Science Distinguished Lecture*, September 2020, NYC, NY. (Distinguished Lecture, virtually).

D'Elia, M. "Part A: Nonlocal models in computational science and engineering: theory and challenges". "Part B: Nonlocal models in computational Science and Engineering: treatment of interfaces in heterogeneous materials and media, image processing, and model learning". *Summer School at University of Roma, La Sapienza*, September 2020, Rome, Italy, virtually.

D'Elia, M. "A Unified Theoretical and Computational Nonlocal Framework: Generalized Nonlocal Vector Calculus and Physics-informed Neural Networks". *CMAI Colloquium*, August 2020, George Mason University, Fairfax, VA, virtually.

D'Elia, M. "A Unified Theory of Fractional and Nonlocal Vector Calculus". *Invited lecture*, August 2020, Brown University, virtually.

D'Elia, M. "A Unified Theoretical and Computational Nonlocal Framework: Generalized Nonlocal Vector Calculus and Physics-informed Neural Networks", *Recent progress in Nonlocal Modeling, Analysis and Computation*, June 2020, Beijing, China, virtually.

D'Elia, M. "Nonlocal Physics-Informed Neural Networks: a unified theoretical and computational framework". *AAAI MLPS*, March 2020, Stanford University, Palo Alto, CA, virtually.

D'Elia, M. "Nonlocal models in computational Science and Engineering". *GA Scientific Computing Symposium 2020*, February 2020, Emory University, Atlanta, GA. (Plenary).

D'Elia M. "Nonlocal models in computational Science and Engineering". *Invited lecture*, February 2020, University of New Mexico, Albuquerque, NM.

D'Elia M. "Nonlocal models in computational Science and Engineering: challenges and applications". *Invited lecture*, November 2019, University of California at Berkeley, Berkely, CA.

He, Q. "Machine Learning Enhanced Computational Mechanics", *SE Special Seminar in Computational Mechanics, Department of Structural Engineering at University of California San Diego*, March 2020, La Jolla, California. (Invited).

He, Q. "Machine Learning Enhanced Computational Mechanics: Reduced-Order Modeling and Physics-Informed Data-Driven Computing", *Sonny Astani Civil and Environmental Engineering Seminar, University of Southern California*, November 2019, Los Angeles, California. (Invited).

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *University of Pennsylvania*, December 2019.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Nanyang Technological University*, January 2020, Singapore.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *National Science Foundation/ SMU workshop*, February 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Applied Mathematics, South Methodist University*, February 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *ANSYS, Inc.*, April 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Applied Mathematics, ETH*, May 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Politecnico Di Milano, MOX*, June 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Siemens Corporation*, August 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *South Eastern University*, August 2020, Nanjing, China.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *KDD2020, Earth Day*, August 2020.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Center for Brains, Mind and Machines (CBMM), Massachusetts Institute of Technology*, September 2020, Massachusetts.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *Indian Institute of Technology, Roorkee (IITR)*, September 2020, India.

Karniadakis, G.E. "Physics-Informed Neural Networks PINNs and Applications", *The Oden Institute, University of Texas at Austin*, September 2020, Austin, Texas.

Parks, M.L. "On Neumann-type Boundary Conditions for Nonlocal Models", *The 5th Annual Meeting of SIAM Central States Section*, October 2019, Ames, Iowa,

Qian, Y. H., Ghorbanidehno, M. Forghani, M. Farthing, T. Hesser, P.K. Kitanidis, and E.F. Darve. "Surf-zone Topography-informed Deep Learning Techniques to Nearshore Bathymetry with Sparse Measurements," *AAAI-MLPS, Association for the Advancement of Artificial Intelligence,* Spring Symposium Series Symposium, March 2020.

Qian, Y. H., J.H. Lee, M. Forghani, M. Farthing, T. Hesser, P.K. Kitanidis, and E.F. Darve. 2020. "Deep Learning Based Spatial Interpolation Methods for Nearshore Bathymetry with Sparse Measurements," *CMWR, Computational Methods in Water Resources,* June 2020 at Stanford University, Stanford, CA.

Stinis, P. "Enforcing constraints for time series prediction in supervised, unsupervised and reinforcement learning ", *AAAI 2020 Spring Symposium on Combining Artificial Intelligence and Machine Learning with Physical Sciences*, March 2020, Stanford, California. (Invited).

Stinis, P. "Enforcing constraints for time series prediction in supervised, unsupervised and reinforcement learning ", *SIAM Conference on Mathematics of Data Science (MDS20)*, June 2020, Cleveland, Ohio. (Invited).

Stinis, P. "Enforcing constraints for time series prediction in supervised, unsupervised and reinforcement learning ", *SIAM/CAIMS Annual Meeting (AN20)*, July 2020, Toronto, Canada. (Invited).

Trask, N. "Robust Training and Initialization of Deep Neural Networks An Adaptive Basis Viewpoint". *MSML2020 - Mathematical and Scientific Machine Learning Conference*, July 2020, Princeton, NJ, virtually.

Trask, N. "Physics-informed graph neural nets: A unification of NN architectures with mimetic PDE discretization". *SIAM Annual meeting*, July 2020, virtually.

Trask, N. "GMLS-Nets: A machine learning framework for unstructured data". *2020 AAAI Spring Symposium*, March 2020, Stanford, CA, virtually.

Trask, N. "Compatible meshfree discretization". *Invited presentation: UIUC civil engineering colloquium*, February 2020, Urbana-Champaign, IL.

Trask, N. "GMLS-Nets: A machine learning framework for unstructured data". *Neural Information Processing Systems*, December 2019, Vancouver, Canada.

Trask, N. "Compatible meshfree discretization". *Invited presentation: Tufts university applied mathematics colloquium*, December 2019, Medford, MA.

Trask N., Patel R.G., Gross B.J., and Atzberger P.J. "GMLS-Nets: Scientific Machine Learning Methods for Unstructured Data" AAAI-MLPS, Stanford, March 2020.

Valiant, G. "Constrained Learning (Plenary Session)", *Information Theory and Applications (ITA)*, February 2020, San Diego, CA. (Plenary Session).

Valiant, G. "Randomly Collected, Worst Case Data", *Workshop on Local Algorithms (WOLA)*, July 2020, (Plenary Talk, virtually).

Valiant, G. "New Problems and Perspectives on Learning, Sampling, and Memory, in the Small Data Regime", *Princeton Theory Seminar*, October 2019, Princeton NJ.

Valiant, G. "How bad is worst-case data if you understand where it comes from?", *MIT MIFODS Workshop on Learning Under Complex Structure*, January 2020, Cambridge MA.

Valiant, G. "Sample Amplification", *Workshop on data augmentation and equivariance*, September 2020, virtually.

Xu K. 2020. "Learning a Neural Network Monte Carlo Sampler for Stochastic Inverse Problems," *Combining Artificial Intelligence and Machine Learning with Physics Sciences Symposium,* March 2020, Stanford University.

Xu K. and E. Darve. 2020. "Data-driven Constitutive Relation Modeling using Deep Neural Networks," at the *Thermal & Fluid Sciences Industrial Affiliates and Sponsors Conference,* Stanford University, February 2020, Stanford, California.

Xu, K. and E. Darve. 2020. "Data-driven Inverse Modeling for Subsurface Properties with Physics-Based Machine Learning," *the SIAM Imaging Science Minisymposium on UQ and ML for the Subsurface,* July 2020, Toronto, Ontario, Canada.

### A.2.2 Organized conferences and workshops

Organized Symposium: Darve, E. and Jonghyun, HL. *AAAI 2020 Spring Symposium Series,* "Combining Artificial Intelligence and Machine Learning with Physical Sciences", March 22–24, 2021, Palo Alto, CA.

Co-organizer and scientific committee member: Darve, E. *1st IACM conference for machine learning, and digital twins for computational science and engineering,* San Diego, CA October 2021.

Organized Minisymposium: Howard, AA. and Rosenthal, W. Steven. "Mathematical Modeling for Multiphase Flow", *SIAM Pacific Northwest Section Meeting*, October 2019, Seattle, WA.

Organized Minisymposium: Gross BJ and Atzberger PJ. "Machine Learning Approaches for the Sciences and Engineering: Recent Developments," *SIAM Conference on Computational Science and Engineering (CSE)*, February 2019, Spokane, Washington.

Organized Minisymposium: Atzberger, PJ and Stinis, P., "Developments in Machine Learning: Foundations and Applications - Parts I - III," *SIAM Annual Meeting*, July 2020, virtual.

Organized international conference: D'Elia M. *et al.* "RAMSES: Reduced order models, Approximation theory, Machine Learning, Surrogates, Emulators, and Simulators", June 2021, SISSA, Trieste, Italy.

Organized international conference: D'Elia M. *et al.* "Optimal control and optimization for nonlocal models", part of *Semester on optimization*, October 2019, RICAM, Linz, Austria.

Organized Minisymposium: D'Elia M. and Glusa C. "Model Learning and Optimization for Nonlocal and Fractional Equations", *SIAM Computer Science and Engineering Conference*, March 2021, Fort Worth, TX.

Organized invited special session: D'Elia M. "Nonlocal interface problems for the simulation of heterogeneous materials and media", *Coupled Problems 2021*, June 2021, Chia Laguna, Italy.

Organized Minisymposium: Trask N, Nelson NH and Patel RG. "Learning Operators From Data", *SIAM Conference on Computational Science and Engineering (CSE)*, March 2021, Fort Worth, TX.

### A.2.3 Contributed presentations and posters

D'Elia, M., Pang, G., Karniadakis, G.E., and Parks, M. "nPINNs: Nonlocal Physics-Informed Neural Networks", poster at *CODA 2020*, February 2020, Santa Fe, NM.

Forghani, J., H.L.M. Farthing, T. Hesser, P. Kitanidis, E. Darve. 2019. "Deep learning techniques for riverine bathymetry and flow velocity estimation," *AGU Meeting,* December 2019. San Francisco, CA.

Ghorbanidehno, J. H.L.M. Farthing, T. Hesser, P. Kitanidis, E. Darve, M. Forghani. 2020. "Deep learning techniques for nearshore and riverine bathymetry estimation using water-surface observations," *2020 Oceans Sciences Meeting, AGU,* February 2020, San Diego, CA

He, Q., Tartakovsky, G., Barajas-Solano, D., Tartakovsky, A. "Physics-Informed Deep Neural Networks for Multiphysics Data Assimilation in Subsurface Transport Problems", *American Geophysical Union (AGU) Fall Meeting 2019*, December 2019, San Francisco, CA.

Trask N., Patel R.G., Gross B.J., and Atzberger P.J. "GMLS-Nets: Scientific Machine Learning Methods for Unstructured Data", poster at *NeurIPs 2019: Workshop on Machine Learning and the Physical Sciences*, December 2019, Vancouver, Canada.

Paskaleva, B. and Bochev, P. "Data driven compact device models", poster at *CODA 2020*, February 2020, Santa Fe, NM.

## A.3 Software

`ADCME.jl` Automatic Differentiation Software for Computational and Mathematical Engineering. Available at `https://github.com/kailaix/ADCME.jl`

`ADSeismic.jl` An Open Source High Performance Package for General Seismic Inversion. Available at `https://github.com/kailaix/ADSeismic.jl/`

`BBFMM` generalized fast multipole method. Available at `https://github.com/ruoxi-wang/BBFMM3D`

`DeepXDE` A deep learning library for solving differential equations. Available at `https://github.com/lululxvi/deepxde`

`FwiFlow.jl` Wave and Flow Inversion with Intrusive Automatic Differentiation. Available at `https://github.com/lidongzh/FwiFlow.jl`

`GMLS-Nets` Generalized Moving Least Squares (GMLS) Neural Networks for Scattered Datasets, P. Atzberger (UCSB), N. Trask (Sandia), B. Gross (UCSB), R. Patel (Sandia). Available for PyTorch at `https://github.com/atzberg/gmls-nets` and TensorFlow at `https://github.com/rgp62/gmls-nets`.

`MOR-Physics` A physics-informed operator regression framework. Available at `https://github.com/rgp62/MOR-Physics`

`NNFEM.jl` Neural Network Approach for Data-Driven Constitutive Modeling. Available at `https://github.com/kailaix/NNFEM.jl/`

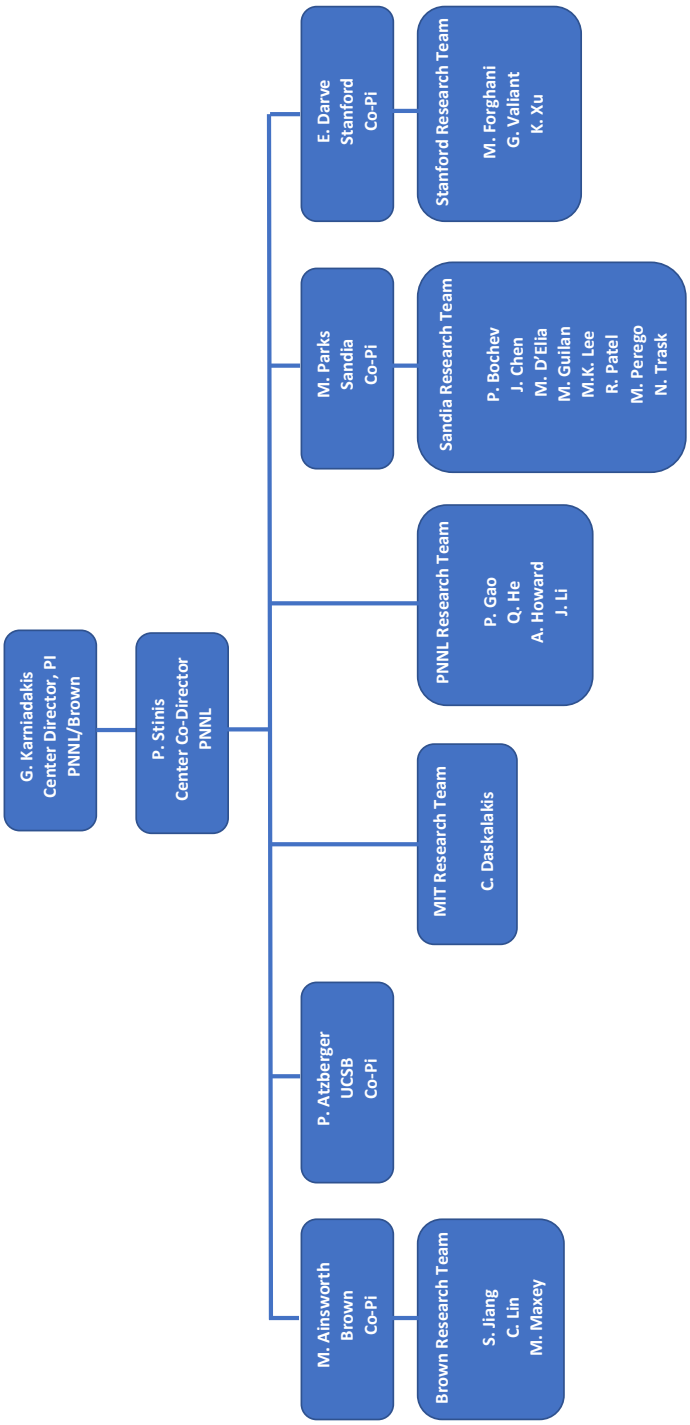
`PoreFlow.jl` Inverse modeling for geomechanics and fluid mechanics. Documentation available at `https://kailaix.github.io/PoreFlow.jl/dev/`. Code is currently private.

`pyPCGA` python library for principal component geostatistical approach. Available at `https://github.com/jonghyunharrylee/pyPCGA`

`SpaND` Fast Linear Solver for Large-Scale Sparse Linear system. Available at `https://github.com/leopoldcambier/spaND_public`

`TaskTorrent` parallel runtime library for executing concurrent directed acyclic graphs of computational tasks with a focus on a very low-overhead when executing micro-tasks. Available at `https://github.com/leopoldcambier/tasktorrent`

## B Org. chart


G. Karniadakis
Center Director, PI
PNNL/Brown
P. Stinis
Center Co-Director
PNNL
E. Darve
Stanford
Co-Pi
Stanford Research Team
M. Forghani
G. Valiant
K. Xu
M. Parks
Sandia
Co-Pi
Sandia Research Team
P. Bochev
J. Chen
M. D'Elia
M. Guilan
M.K. Lee
R. Patel
M. Perego
N. Trask
PNNL Research Team
P. Gao
Q. He
A. Howard
J. Li
MIT Research Team
C. Daskalakis
P. Atzberger
UCSB
Co-Pi
M. Ainsworth
Brown
Co-Pi
Brown Research Team
S. Jiang
C. Lin
M. Maxey

## C Work responsibilities and timelines

**Research Area I:** PDE-Based Modeling of Macroscales (Lead: Ainsworth-Brown; co-PIs: Parks, D'Elia, Trask, Bochev-SNL; Karniadakis, Stinis, Li-PNNL; C. Li, Maxey-Brown; Atzberger-UCSB; Darve-Stanford; Daskalakis-MIT)
Year 3: Continue developments and publish papers in the following areas: scalable physics-informed neural networks (PINNs, cvPINNs, and pigNNs) with domain decomposition that learn from multi-fidelity data for forward and inverse problems; unified nonlocal vector calculus, theory, and computation of nonlocal models: nonlocal PINNs (nPINNs) applied to turbulence modeling in flows with adverse pressure gradients; inference of constitutive laws of complex hard materials and of polymers from synthetic data using different types of PINNs; learning the hidden fluid mechanics and hidden fractional dynamics in seismology using PINNs machine learning methods; nonlinear functional and operator regression using extensions of DeepOnet and GMLS-nets.
Year 4: Integrate the advances in RA-III and RA-IV. Develop diverse machine learning frameworks for multiscale and multiphysics problems in exemplar applications, specifically subsurface reactive transport and ice sheets. Combine nonlocal and fractional operators with PINNs and their extensions to discover (stochastic) closures in the exemplar applications. Compare the approximation properties of extensions of scalable PINNs and meshfree high-order methods for multiphysics problems at the macroscale. Study the ability of Bayesian PINNs to quantify uncertainty in long-term predictions in geophysical applications and validate it using historical data.

**Research Area II:** Stochastic Modeling of Mesoscales (Lead: Stinis-PNNL; co-PIs: Atzberger-UCSB; Stinis, Howard, Gao, Li-PNNL; Trask, Parks-SNL; Kharazmi, Meng, Maxey-Brown; Darve-Stanford)
Year 3: Continue developments and publish papers in the following areas: learning effective CG potentials from DFT for multi-phase systems; learning the nucleation of nano-bubbles and micro-bubbles using many-body Dissipative Particle Dynamics (mDPD) and DNS data using DeepOnet; learn new PDEs from MD simulation data for color noise; connections between ML and model reduction with applications to unsupervised (GANS) and reinforcement learning; parallel-in-time PINNs for long-time integration of mesoscale stochastic equations.
Year 4: Integrate the advances in RA-I and RA-IV; synthesize and scale up peptoids to achieve desired properties and functionality at the macroscale. Employ the MZ-derived mDPD to study nucleation in soft materials and obtain phase diagrams. Evaluate the ability of RNNs to deal with multiple timescales in realistic soft material applications. In collaboration with RA-I, use fractional operators to represent nonlocal interactions (due to aggressive coarse-graining) in scaled-up functional materials and other systems.

**Research Area III:** Bridging Methods to Connect the Scales (Lead: Bochev-SNL; co-PIs: Chen, Trask, D'Elia, Perego, Parks-SNL; Ainsworth, Kharazmi, Meng-Brown; Howard, Gao, He, Karniadakis-PNNL; Atzberger-UCSB; Daskalakis-MIT; Darve, Valiant-Stanford)
Year 3: Continue developments and publish papers in the following areas: nPINNs for modeling heterogeneous materials; multiscale modeling in combustion; active and transfer learning of constitutive laws for multiscale modeling of non-Newtonian fluids using DPD data; SPH-SPH interface for viscoelastic media using the Multiscale Universal Interface (MUI); domain decomposition for PINNs for porous media with largely disparate conductivities; learning the kernel in nonlocal models for surface tension based on MD data; learning surrogate models for turbulent mixing and ignition (DFT, DSMC, DNS).
Year 4: Integrate the advances in RA-I, RA-II, and RA-IV. Apply active learning and upscaling to peptoids and DNS in combustion and examine accuracy and cost; validation using data from PNNL and SNL, respectively. Apply active learning and domain decomposition to subsurface reactive transport and ice sheets and examine feasibility and scaling up from laboratory scales to field scales. Validation with existing (classical) solvers and partial data available at PNNL and SNL.

**Research Area IV:** Statistical Learning (Lead: Darve-Stanford; co-PIs: Valiant-Stanford; Daskalakis-MIT; Karniadakis, Stinis, He-PNNL; Parks, Trask, Bochev-SNL; Atzberger-UCSB; Ainsworth, Kharazmi, Meng- Brown)
Year 3: Continue developments and publish papers in the following areas: investigate generalization of PINNs and their extensions; multifidelity training of B-PINNs; generalized existing techniques for estimating learnability of the best classifier in a specified class in the data regime in which there is insufficient data to learn even an approximation of such a classifier; explore approaches for integrating certain classes of invariances within a convolutional DNN architecture; developed Generative Adversarial Networks for dynamical systems, and for the optimal transport problem for solving high-dimensional stochastic PDEs (e.g., 10,000 dimensions in a porous media model for the Hanford site); designed stable optimization methods (optimistic GD) for DNNs and GANs; explore new data augmentation techniques and how they work
Year 4: Develop the next generation of PINNs and DeepOnets for continuum- and molecular-based physical systems, including multiscale and multiphysics NN frameworks. Systematically study the findings of RA-I–III and incorporate lessons learned into the new deep learning architectures and SGD and HNN algorithms. Develop and finalize information-theoretic approaches for designing a priori PhILMs with a specific number of layers and neurons and document accuracy bounds. Provide practical guidelines on learnability and generalization in DNNs.

## D Abbreviations

- AD: automatic differentiation
- B-PINN: Bayesian PINN
- CG: coarse-grained
- cPINN: conservative PINN
- cvPINN: control volume PINN
- DOE: Department of Energy
- DNN: deep neural network
- DPD: dissipative particle dynamics
- ETI: exponential time integrators
- FF: force field
- fPINN: fractional PINN
- GAN: generative adversarial network
- GD: gradient descent
- GMLS: generalized moving least squares
- HFM: hidden fluid mechanics
- HMC: Hamiltonian Monte Carlo
- MD: Molecular dynamics
- MIT: Massachusetts Institute of Technology
- ML: machine learning
- NLM: nonlocal model
- NN: neural network
- nPINN: nonlocal PINN
- PCL: Physics Constrained Learning
- PDE: Partial differential equation
- PDF: Probability density function
- PhILMs: Physics-Informed Learning Machines for Multiscale and Multiphysics Problems
- PI: principal investigator
- PINN: physics-informed neural network
- PNNL: Pacific Northwest National Laboratory
- RA: Research Area
- ReLU: rectified linear unit
- SGD: stochastic gradient descent

- SNL: Sandia National Laboratories
- SODE: stochastic ordinary differential equation
- UCSB: University of California, Santa Barbara
- VI: variational inference