Nonlocal Physics-Informed Neural Networks – A unified theoretical and computational framework for nonlocal models

Marta D'Elia[1]*, George E. Karniadakis[2], Guofei Pang[2], Michael L. Parks[1]
[1]Center for Computing Research, Sandia National Laboratories,
1450 Innovation Parkway SE, Albuquerque, NM, 87123
mdelia,mlparks@sandia.gov
[2] Applied Mathematics Department, Brown University,
170 Hope Street, Providence, RI 02912,
george_karniadakis,guofei_pang@brown.edu

Abstract

Nonlocal models provide an improved predictive capability thanks to their ability to capture effects that classical partial differential equations fail to capture. Among these effects we have multiscale behavior and anomalous behavior such as super- and sub-diffusion. These models have become incredibly popular for a broad range of applications, including mechanics, subsurface flow, turbulence, plasma dynamics, heat conduction and image processing. However, their improved accuracy comes at a price of many modeling and numerical challenges. In this work we focus on the estimation of model parameters, often unknown, or subject to noise. In particular, we address the problem of model identification in presence of sparse measurements. Our approach to this inverse problem is based on the combination of 1. Machine Learning and Physical Principles and 2. a Unified Nonlocal Vector Calculus and Versatile Surrogates such as neural networks (NN). The outcome is a flexible tool that allows us to learn existing and new nonlocal operators. We refer to our technique as nPINNs (nonlocal Physics-Informed Neural Networks); here, we model the nonlocal solution with a NN and we solve an optimization problem where we minimize the residual of the nonlocal equation and the misfit with measured data. The result of the optimization are the weights and biases of the NN and the set of unknown model parameters.

Challenges of nonlocal modeling

Nonlocal equations are model descriptions for which the state of a system at any point depends on the state in a neighborhood of points, i.e. every point in a domain interacts with a neighborhood of points. As such, interactions can occur at distance, without contact. These models are such that they can capture effects that traditional PDEs fail to capture; in fact, their solutions can be irregular: non-differentiable, singular, and discontinuous. Among those effects, we mention: 1) Multiscale behaviors and discontinuities such as cracks and fractures and 2) Anomalous behaviors such as super- and sub-diffusion. In case 1) we refer to nonlocal truncated operators where the neighborhood is a ball of radius δ (usually much smaller than the domain) surrounding any point. In case 2) refer to fractional operators where the interactions can be infinite ($\delta = \infty$); a standard representative of this class is the fractional Laplacian operator $(-\Delta)^s$.

As a consequence, nonlocal models provide an improved predictive capability for several scientific and engineering applications including fracture mechanics (Ha and Bobaru 2011; Littlewood 2010; Silling 2000), anomalous subsurface transport (Benson, Wheatcraft, and Meerschaert 2000; Schumer et al. 2003; 2001), phase transitions (Bates and Chmaj 1999; Delgoshaie et al. 2015; Fife 2003), image processing (A. Buades, Coll, and Morel 2010; Gilboa and Osher 2007; 2008; Lou et al. 2010), multiscale and multiphysics systems (Alali and Lipton 2012; Askari 2008), MHD (Schekochihin, Cowley, and Yousef 2008), and stochastic processes (Burch, D'Elia, and Lehoucq 2014; D'Elia et al. 2017; Meerschaert and Sikorskii 2012; Metzler and Klafter 2000).

In its simplest form, a nonlocal operator can be defined as

$$Lu(x) = \int_{B_\delta(x)} (u(y) - u(x)) k(x, y)\, dy, \tag{1}$$

where $B_\delta(x)$ is the ball of radius δ centered at x and where k is an application dependent *kernel* that determines the regularity properties of the solution. The integral form allows us to catch long-range forces and reduces the regularity requirements of the solution.

We consider nonlocal diffusion problems of the form

$$\begin{cases} -Lu = f & x \in \Omega \\ \quad u = g & x \in \Omega_I, \end{cases} \tag{2}$$

where $\Omega \subset \mathbb{R}^n$ is an open bounded domain and Ω_I is the *interaction domain*, a layer of thickness δ surrounding the

*Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. Report 2019-14015.

domain where nonlocal boundary conditions must be prescribed for the well-posedness of the problem.

Two very important concerns arise when addressing the solution of (2).

Q1 Is (1) general enough? How broad is the class of nonlocal operators that can be described by one single formula and analyzed through one unified calculus?

Q2 What is the "right" kernel for a given phenomenon? How can available data help determine the appropriate nonlocal model and its parameters? Can we design a unified data-driven tool for model identification and simulation of a broad class of nonlocal models?

The first concern arises from the fact that in the literature we have independent definitions, formulations and theory of nonlocal models. Similarities are evident, but they have not been rigorously proved. This is addressed in the next section.

A unified nonlocal calculus

The purpose of a unified nonlocal notation and theory is to

- Connect the nonlocal and fractional communities that would benefit from each other's research;
- Include as special cases the well-known classical differential calculus at the limit of vanishing interactions and the fractional calculus at the limit of infinite interactions;
- Provide the groundwork for *new model discovery* thanks to the broad class of operators that it describes;
- Describe intrinsically nonlocal phenomena that have not been analyzed or used due to the lack of theory.
- Guide algorithm/discretization/solver design.

In this work we introduce a generalized nonlocal operator, in the spirit of a unified calculus, that bridges local, truncated nonlocal and fractional diffusion operators:

$$L_{\delta,s}u(x) = C_{\delta,s}\int_{B_\delta(x)} \frac{u(x)-u(y)}{|x-y|^{n+2s}}\,dy \qquad (3)$$

where $C_{s,\delta}$ is such that the corresponding solutions span a broad range of nonlocal diffusion processes including local and fractional diffusion at the limit of vanishing and increasing nonlocality, i.e.

$$\lim_{\delta\to 0} L_{\delta,s}u = \Delta u \;\text{ and }\; \lim_{\delta\to\infty} -L_{\delta,s}u = (-\Delta)^s u.$$

A unified computational framework

The unified nonlocal vector calculus, and more specifically the operator in (3) provides us with a universal definition of parametrized nonlocal operators that describe both well-known nonlocal phenomena and may describe new intrinsically nonlocal phenomena not yet analyzed and used due to lack of theory. However, the universal nature of these new mathematical models and the abundance of data raise important questions.

Q3 What are the true model parameters δ and s?

Q4 How can we deal with data sparsity and noise (the forcing term f and the nonlocal boundary condition g in (2) may be sparse or subject to noise)?

We propose a new approach to model learning that is in stark contrast with previously developed UQ and PDE-constrained-like optimization techniques. The game changer is the combination of 1) *Machine Learning* and *Physical Principles*, and 2) *Unified Calculus* and *Versatile Surrogates*, such as neural networks. The outcome is a Data-Driven Physics-Informed tool for learning new complex nonlocal phenomena.

We refer to our strategy as *nPINNs* (nonlocal Physics-Informed Neural Networks); this is an extension of PINNs (Raissi, Perdikaris, and Karniadakis 2018) and fPINNs (Pang, Lu, and Karniadakis 2018) designed for PDEs and fractional operators respectively. More specifically nPINNs includes the methods above as special instances. In the next section we describe our strategy and its main properties.

Nonlocal Physics-Informed Neural Networks

The nPINNs algorithm consists of three simple steps.

1 Collect observations of solution and data in training sets: $f_m(x_i)$, $x_i \in \mathcal{T}_f$, and $u_m(x_j)$, $x_j \in \mathcal{T}_u$;

2 Approximate the solution with a Neural Network: $u(x) = u_{NN}(x)$;

3 Minimize the loss function

$$\min_{u;\delta,s} \mathcal{L}oss(u;\delta,s) = \frac{1}{2}\sum_{x_i\in\mathcal{T}_f}\left(L_{\delta,s}u_{NN}(x_i) - f_m(x_i)\right)^2 + \frac{\beta}{2}\sum_{x_i\in\mathcal{T}_u}\left(u_{NN}(x_j) - u_m(x_j)\right)^2,$$

where the minimization with respect to u must be regarded as minimization with respect to the weights and biases of the NN. The two, distinct, training sets in **1** depend only on data availability and are not necessarily associated with quadrature points. Note that $\mathcal{L}oss$ has a physics-driven and a data-driven component: the first term controls the residual of the nonlocal equation, whereas the second the mismatch between solution and data. The outcome of the optimization are the weights and biases of the NN and the model parameters. This strategy

- Is as *accurate* as any other discretization method for the forward problem. As an example, numerical tests show that it has the same convergence rate, as the number of training points increases, of fPINNs and of a standard Finite Difference discretization. However, due to the increased computation cost, nPINNs is not yet recommended for the solution of forward problems.
- Is *not tied* to any discretization method.
- Requires *minimal implementation effort:* available solvers can be used as black boxes.
- Easily handles *sparsity*.

We tested this method on one-dimensional forward and inverse problem (to illustrate our theoretical findings and learn model parameters) and on two- and three- dimensional forward problems (to show applicability in higher dimensions). Also, we applied nPINNs to the solution of turbulent Couette flow for the estimation of the dispersion rate s and the characteristic length δ. Computational results are

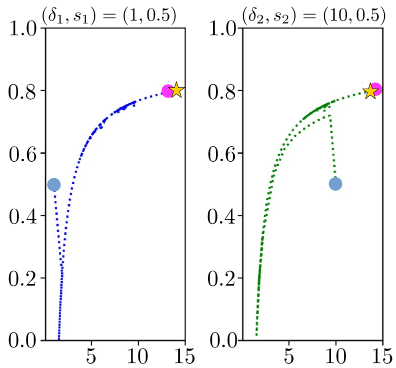

Figure 1: Trajectories of the optimization algorithm for the initial guesses (δ_1, s_1)=$(1, 0.5)$, left, and (δ_2, s_2)=$(10, 0.5)$, right. The blue dot indicates the initial guess, the pink dot the optimal value and the yellow star the true value.

promising and show that the versatility of NN allows one to describe complex phenomena, to identify model parameters and to handle data sparsity.

One-dimensional example In Figure 1 we report the outcome of our algorithm (steps **1**–**3**) for the estimation of δ and s. For Ω=$(0, 1)$ and $\Omega \cup \Omega_I$=$(-\delta, 1 + \delta)$, we consider the nonlocal diffusion problem (2) with g=0, f=$\sin(2\pi x)$ and L defined as in (3). The training data u_m are generated via accurate solution of (2) with parameters (δ^*, s^*)=$(14, 0.8)$; we refer to these values as *true values* and represent them with a yellow star in the plot. The training points are 100 uniformly spaced points in $\Omega \cup \Omega_I$. We run the algorithm for two initial guesses, represented by the blue dots and report their trajectories. Both of them, see pink dots in both plots, converge to the true values. The optimal u_{NN} corresponding to the estimated parameters are accurate for both initial guesses; in fact, their relative errors are of the order of 10^{-4}.

References

A. Buades, A.; Coll, B.; and Morel, J. 2010. Image denoising methods. a new nonlocal principle. *SIAM Review* 52:113–147.

Alali, B., and Lipton, R. 2012. Multiscale dynamics of heterogeneous media in the peridynamic formulation. *Journal of Elasticity* 106(1):71–103.

Askari, E. 2008. Peridynamics for multiscale materials modeling. *Journal of Physics: Conference Series, IOP Publishing* 125(1):649–654.

Bates, P., and Chmaj, A. 1999. An integrodifferential model for phase transitions: Stationary solutions in higher space dimensions. *J. Statist. Phys.* 95:1119–1139.

Benson, D.; Wheatcraft, S.; and Meerschaert, M. 2000. Application of a fractional advection-dispersion equation. *Water Resources Research* 36(6):1403–1412.

Burch, N.; D'Elia, M.; and Lehoucq, R. 2014. The exit-time problem for a markov jump process. *The European Physical Journal Special Topics* 223:3257–3271.

Delgoshaie, A.; Meyer, D.; Jenny, P.; and Tchelepi, H. 2015. Non-local formulation for multiscale flow in porous media. *Journal of Hydrology* 531(1):649–654.

D'Elia, M.; Du, Q.; Gunzburger, M.; and Lehoucq, R. 2017. Nonlocal convection-diffusion problems on bounded domains and finite-range jump processes. *Computational Methods in Applied Mathematics* 29:71–103.

Fife, P. 2003. *Some nonclassical trends in parabolic and parabolic-like evolutions*. Springer-Verlag, New York. chapter Vehicular Ad Hoc Networks, 153–191.

Gilboa, G., and Osher, S. 2007. Nonlocal linear image regularization and supervised segmentation. *Multiscale Model. Simul.* 6:595–630.

Gilboa, G., and Osher, S. 2008. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* 7:1005–1028.

Ha, Y. D., and Bobaru, F. 2011. Characteristics of dynamic brittle fracture captured with peridynamics. *Engineering Fracture Mechanics* 78(6):1156–1168.

Littlewood, D. 2010. Simulation of dynamic fracture using peridynamics, finite element modeling, and contact. In *Proceedings of the ASME 2010 International Mechanical Engineering Congress and Exposition, Vancouver, British Columbia, Canada.*

Lou, Y.; Zhang, X.; Osher, S.; and Bertozzi, A. 2010. Image recovery via nonlocal operators. *Journal of Scientific Computing* 42:185–197.

Meerschaert, M., and Sikorskii, A. 2012. *Stochastic models for fractional calculus*. Studies in mathematics, Gruyter.

Metzler, R., and Klafter, J. 2000. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports* 339(1):1–77.

Pang, G.; Lu, L.; and Karniadakis, G. 2018. Fractional physics-informed neural networks. Technical report. ArXiv:1811.08967.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. 2018. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*.

Schekochihin, A.; Cowley, S.; and Yousef, T. 2008. Mhd turbulence: Nonlocal, anisotropic, nonuniversal? In *In IUTAM Symposium on computational physics and new perspectives in turbulence*, 347–354. Springer, Dordrecht.

Schumer, R.; Benson, D.; Meerschaert, M.; and Wheatcraft, S. 2001. Eulerian derivation of the fractional advection-dispersion equation. *Journal of Contaminant Hydrology* 48:69–88.

Schumer, R.; Benson, D.; Meerschaert, M.; and Baeumer, B. 2003. Multiscaling fractional advection-dispersion equations and their solutions. *Water Resources Research* 39(1):1022–1032.

Silling, S. 2000. Reformulation of elasticity theory for discontinuities and long-range forces. *Journal of the Mechanics and Physics of Solids* 48:175–209.