Data-driven learning of nonlocal models:
from high-fidelity simulations to constitutive laws

Huaiqian You,[1] Yue Yu,[1] Stewart Silling,[2] Marta D'Elia[3]

[1] Mathematics Department, Lehigh University, PA
[2] Center for Computing Research, Sandia National Laboratories, Albuquerque, NM
[3] Computational Science and Analysis, Sandia National Laboratories, Livermore, CA
huy316@lehigh.edu, yuy214@lehigh.edu, sasilli@sandia.gov, mdelia@sandia.gov.

Abstract

We show that machine learning can improve the accuracy of simulations of stress waves in one-dimensional composite materials. We propose a data-driven technique to learn nonlocal constitutive laws for stress wave propagation models. The method is an optimization-based technique in which the nonlocal kernel function is approximated via Bernstein polynomials. The kernel, including both its functional form and parameters, is derived so that when used in a nonlocal solver, it generates solutions that closely match high-fidelity data. The optimal kernel therefore acts as a homogenized nonlocal continuum model that accurately reproduces wave motion in a smaller-scale, more detailed model that can include multiple materials. We apply this technique to wave propagation within a heterogeneous bar with a periodic microstructure. Several one-dimensional numerical tests illustrate the accuracy of our algorithm. The optimal kernel is demonstrated to reproduce high-fidelity data for a composite material in applications that are substantially different from the problems used as training data.

Introduction

Nonlocal models use integral operators acting on a length-scale δ, known as horizon. This feature allows nonlocal models to capture long-range forces at small scales and multiscale behavior, and to reduce regularity requirements on the solutions, which are allowed to be discontinuous or even singular. In recent decades, nonlocal equations have been successfully used to model several engineering and scientific applications, including fracture mechanics (Silling 2000; Ha and Bobaru 2011; Trask et al. 2019), subsurface transport (Benson, Wheatcraft, and Meerschaert 2000; Schumer et al. 2003), image processing (D'Elia, De los Reyes, and Trujillo 2019; Gilboa and Osher 2007), multiscale and multiphysics systems (Alali and Lipton 2012; Askari et al. 2008; You, Yu, and Kamensky 2020), finance (Scalas, Gorenflo, and Mainardi 2000), and stochastic processes (D'Elia et al. 2017; Meerschaert and Sikorskii 2012).

However, it is often the case that *nonlocal kernels* defining nonlocal operators are justified a posteriori and it is not clear how to define such kernels to faithfully describe a physical system. The problem of learning an appropriate kernel for a specific application is one of the most challenging open problems in nonlocal modeling. The literature on techniques for learning kernel *parameters* for a given functional form is vast, see, e.g., (Burkovska, Glusa, and D'Elia 2020; D'Elia and Gunzburger 2016) for control-based approaches, and (Pang et al. 2020; Pang, Lu, and Karniadakis 2019) for machine-learning approaches. However, the use of machine learning to learn the *functional form* of the kernel is still in its infancy, (Xu and Foster 2020; Xu, D'Elia, and Foster 2020; You et al. 2021) being the only relevant works that we are aware of.

In this work we use an approach similar to the one developed in (You et al. 2021) to learn nonlocal kernels whose associated nonlocal wave equation is well posed by construction and can be used as an accurate surrogate for more detailed, high-fidelity wave propagation models. In particular, we present an application to wave propagation at the microscale in a heterogeneous solid. In this context, the machine-learned nonlocal kernel embeds the material constitutive behavior so that the material interfaces do not have to be treated explicitly and, more importantly, the material microstructure can be unknown. Furthermore, the corresponding nonlocal models allow for accurate simulations at scales that are much larger than the microstructure.

Our main contributions are:
• The design of an optimization technique that bridges micro and continuum scales by providing accurate and stable model surrogates for the simulation of wave propagation in heterogeneous materials.
• The illustration of this method via one-dimensional experiments that confirm the applicability of our technique and the improved accuracy compared with state-of-the-art results.
• The demonstration of generalization properties of our algorithm whose associated model surrogates are effective even on problem settings that are substantially different from the ones used for training in terms of loading and time scales.

Nonlocal kernel learning

We introduce the high-fidelity (HF) model that faithfully represents the system: for $\Omega \in \mathbb{R}^d$, the scalar function

$u(x, t)$ solves, for $(x, t) \in \Omega \times [0, T]$

$$\frac{\partial^2 u}{\partial t^2}(x, t) - \mathcal{L}_{\text{HF}}[u](x, t) = f(x, t), \qquad (1)$$

provided some boundary conditions on $\partial\Omega$ for $u(x, t)$ and initial conditions at $t = 0$ for u and $\partial u/\partial t$ are satisfied. Here, $\mathcal{L}_{\text{HF}}$ is the HF operator, which can either be a differential or integral operator, and f represents a forcing term.

We assume that solutions to this HF problem may be approximated by solutions to a nonlocal problem of the form

$$\frac{\partial^2 u}{\partial t^2}(x, t) - \mathcal{L}_K[u](x, t) = f(x, t), \qquad (2)$$

for $(x, t) \in \Omega \times [0, T]$, augmented with nonlocal boundary conditions on Ω_δ (a layer of thickness δ that surrounds the domain) and the same initial conditions on the variable u and its derivative as in (1). The forcing f may coincide with the forcing term in (1) or it could be an appropriate representation of the same.

We seek $\mathcal{L}_K$ as a nonlocal operator of the form

$$\mathcal{L}_K[u](x, t) = \int_{\overline{\Omega}} K(|x - y|) \left(u(y, t) - u(x, t) \right) \, dy \quad (3)$$

where K is a radial, sign-changing, *kernel function*, compactly supported on the ball of radius δ centered at x, i.e., $B_\delta(x)$ and $\overline{\Omega} = \Omega \cup \Omega_\delta$.

The algorithm

To learn the kernel K, we assume that we are given N pairs of forcing terms and corresponding solutions to (1), normalized with respect to the L^2 norm of each solution over $\Omega \times [0, T_{\text{tr}}]$. These are denoted by

$$\mathcal{D}_{\text{tr}} = \left\{ (u_k(x, t), f_k(x, t)) \right\}_{k=1}^N, \qquad (4)$$

for $x \in \Omega$ and $t \in (0, T_{\text{tr}}]$. Similarly to (You et al. 2021), we represent K as a linear combination of Bernstein basis polynomials:

$$K\left(\frac{|y|}{\delta} \right) = \sum_{m=0}^M \frac{C_m}{\delta^{d+2}} B_{m,M}\left(\left| \frac{y}{\delta} \right| \right), \qquad (5)$$

where the Bernstein basis functions are defined as

$$B_{m,M}(x) = \binom{M}{m} x^m (1 - x)^{M-m} \quad \text{for } 0 \le x \le 1$$

and where $C_m \in \mathbb{R}$. Note that, by construction, this kernel guarantees that (2) is well-posed (Du, Tao, and Tian 2018).

We machine-learn the nonlocal model by finding optimal parameters $\{C_m\}$ such that solutions $\hat{u}_k$ to (2), for $f = f_k$ and the kernel function K associated to $\{C_m\}$, are as close as possible to the training variable u_k.

In this work we numerically approximate $\hat{u}_k$ by $\bar{u}_k$ using a central-differencing scheme in time with time step dt, i.e.

$$\begin{aligned}
\bar{u}_k^{n+1}(x_i) = \; & 2\bar{u}_k^n(x_i) - \bar{u}_k^{n-1}(x_i) \\
& + dt^2 \left(\mathcal{L}_{K,h}[\bar{u}_k^n](x_i) + f_k(x_i, t^n) \right),
\end{aligned} \qquad (6)$$

where $\bar{u}_k^{n+1}(x_i)$ represents the k-th approximate solution at time step t^{n+1} and at discretization point x_i, and $\mathcal{L}_{K,h}$ is an
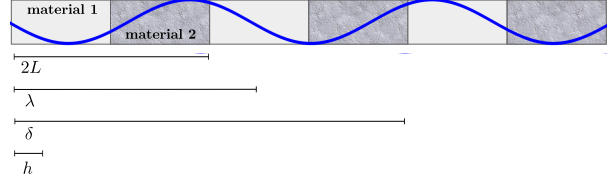

Figure 1: One-dimensional bar with ordered microstructure of period $2L$. Material 1 and 2 have the same density and Young modulus E_1 and E_2. The horizon δ, the wave length λ, and the discretization size, h, are reported for comparison.

approximation of $\mathcal{L}_K$ by Riemann sum with uniform grid spacing h. The optimal parameters are obtained by solving the following optimization problem.

$$\min_{C_m} \frac{T_{\text{tr}}}{dt^3 \, N} \sum_{k=1}^N \sum_{n=1}^{T_{\text{tr}}/dt} \left\| \bar{u}_k^{n+1} - u_k(t^{n+1}) \right\|_{\ell^2}^2 + \mathcal{R}(\{C_m\}),$$
$$(7)$$

$$\text{s.t.} \quad \bar{u}_k \text{ satisfies (6) and} \qquad (8)$$
$$\qquad K \text{ satisfies physics-based constraints.} \qquad (9)$$

Here, the ℓ^2 norm is taken over the space-discretization points x_i, $\mathcal{R}(\cdot)$ is a regularization term on the coefficients that improves the conditioning of the optimization problem, and (9) depends on the physics of the problem (as an example, it may correspond to enforcing that the surrogate model reproduces exactly a certain class of solutions).

Dispersion in heterogeneous materials

We apply the learning algorithm described above to the propagation of waves in a one-dimensional heterogeneous bar, like the one reported in Figure 1, with an *ordered* microstructure, i.e. two materials with the same length alternate periodically. Our goal is to learn a nonlocal model able to reproduce wave propagation on distances that are much larger than the size of the microstructure without resolving the microscales. The high-fidelity model we rely on is the classical wave equation; the corresponding high-fidelity data used for training and validation are obtained with the solver described below.

High-fidelity data

For both training and validation purposes we generate data using high-fidelity simulations for the propagation of stress waves within the microstructure of the heterogeneous, linear elastic bar. This method, which will be referred to as Direct Numerical Solution (DNS), constructs an arbitrarily complex wave diagram (also called an x-t diagram), that treats the mutual interaction and superposition of many wavefronts moving in either direction. The bar is discretized into nodes such that it takes a constant amount of time Δt for a wave to travel between nodes γ and $\gamma + 1$, regardless of the elastic wave speed in the material between these two nodes. Therefore, in a heterogeneous medium, the spacing between nodes is not constant. Each node γ, at each time step n, has velocity v_γ^n (note that, in this case, the subscript refers to position,
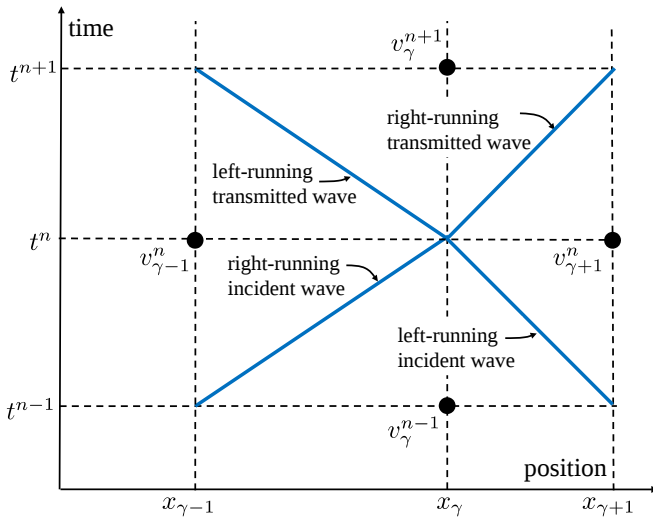
Figure 2: Interaction of two waves in the DNS method. Each node may or may not be located at a material interface.

as opposed to the previous section where it corresponds to a specific sample k). To compute the velocities in the next time step, it is assumed that two waves moving in opposite directions converge on the node γ at time step n (see Figure 2). The waves shown in the figure can have unequal slopes on the x-t diagram because the materials on either side of node γ can have different waves speeds c. The jump conditions for the waves are applied that relate the stress change $[\sigma]$ across a wave to the velocity change $[v]$. These jump conditions have the following form:

$$[\sigma] = \pm \rho c [v],$$

where ρ is the mass density, and where the $+$ and $-$ signs apply to right-running waves and left-running waves respectively. From these conditions, the velocity v_γ^{n+1} can be computed explicitly from the values at the adjacent nodes in time step $n-1$. Externally applied forces can also be included in a straightforward way. After v_γ^{n+1} is computed, the updated displacement is approximated by

$$u_\gamma^{n+1} = u_\gamma^n + \Delta t v_\gamma^{n+1}.$$

Details of the method can be found in (Silling 2020).

This DNS solver has the important advantage of not using an approximate representation of derivatives in space or time for the computation of the velocity, which is, therefore, free from truncation error and other sources of discretization error that are usually encountered with PDE solvers. This allows us to model the propagation of waves through many thousands of microstructural interfaces without the need to worry about what features of the velocity are real and what are numerical artifacts.

We consider four types of data and use the first two for training and the last two for validation of our algorithm. In all our experiments we set $L = 0.2$, $E_1 = 1$, $E_2 = 0.25$, $\rho = 1$, and the symmetric domain $\Omega = (-b, b)$. Discretization parameters for the DNS solver are set to $\Delta t = 0.01$ and $\max\{\Delta x\} = 0.01$.

1) Oscillating source. We set $b = 50$, $v(x, 0) = u(x, 0) = 0$,
$f(x, t) = e^{-\left(\frac{2x}{5kL}\right)^2} e^{-\left(\frac{t-t_0}{t_p}\right)^2} \cos^2\left(\frac{2\pi x}{kL}\right)$, $k = 1, 2, \ldots, 20$,
$t_0 = t_p = 0.8$.

2) Plane wave. For $b = 50$, $f(x, t) = 0$ and $u(x, 0) = 0$, we prescribe $v(-b, t) = \sin(\omega t)$ for $\omega = 0.35, 0.7, \cdots, 3.85$.

3) Wave packet. For $b = 133.3$, $f(x, t) = 0$ and $u(x, 0) = 0$, we prescribe $v(-b, t) = \sin(\omega t) \exp(-(t/5 - 3)^2)$ for $\omega = 2, 3.9, 5$.

4) Impact. For $b = 266.6$, $f(x, t) = 0$ and $u(x, 0) = 0$, we prescribe $v(x, 0) = 1$ for all $x \in [-b, -b + 1.6]$ and $v = 0$ outside of this interval. This initial condition represents an impactor hitting the bar at time zero, generating a velocity pulse of width roughly 3.2 that propagates into the interior of the bar. The pulse attenuates and changes shape as it encounters the many microstructural interfaces.

Training procedure

For the optimization problem (7) we choose a Tikhonov regularization of the form $\mathcal{R}(\{C_m\}) = \frac{\epsilon}{M+1} \sum_{m=0}^{M} C_m^2$, where the regularization weight ϵ is chosen empirically to guarantee accurate predictions, as we explain later on. The physics-based constraints in (9) are defined as follows and also discretized by Riemann sum; they are used to explicitly prescribe values of C_{M-1} and C_M:

$$\sum_{m=0}^{M} C_m \int_0^\delta \frac{y^2}{\delta^3} B_{m,M}\left(\frac{|y|}{\delta}\right) dy = \rho c_0^2,$$

$$\sum_{m=0}^{M} C_m \int_0^\delta \frac{y^4}{\delta^3} B_{m,M}\left(\frac{|y|}{\delta}\right) dy = -4\rho c_0^3 R, \tag{10}$$

where ρ is the density and c_0 is the effective wave speed for infinitely long wavelengths. For $\rho = 1$, it is given by $c_0 = (2/(1/E_1 + 1/E_2)^{\frac{1}{2}}$. R is the second derivative of the wave group velocity with respect to the frequency ω evaluated at $\omega = 0$. Both parameters are obtained by simulating a very low frequency plane wave propagating through the microstructure over a long distance using DNS (Silling 2020). These parameters primarily affect simulations at large times, $t > 10$. However, due to practical limitations on computer resources, our training simulations are restricted to $t \leq 2$. Therefore, we incorporate these parameters as constraints obtained from DNS as indicated in (10), rather than attempting to learn these through our algorithm. The first constraint in (10) is also used for similar purposes in (Xu, D'Elia, and Foster 2020) and prescribed in a weak sense by penalization.

Training is performed with DNS data of type 1) and 2). Parameters for the nonlocal solver and the optimization algorithm are set to $h = 0.05$, $dt = 0.02$, $T_{\text{tr}} = 2$, $\delta = 1.2$, $M = 24$ and $\epsilon = 0.01$. The optimization problem (7) is solved with L-BFGS. Note that we empirically choose δ and ϵ in such a way that the group velocity, defined below, corresponding to the optimal kernel is as close as possible to the one computed with DNS.

The optimal kernel, K_{opt}, is reported in Figure 3; as expected from the literature (Xu and Foster 2020; Xu, D'Elia,
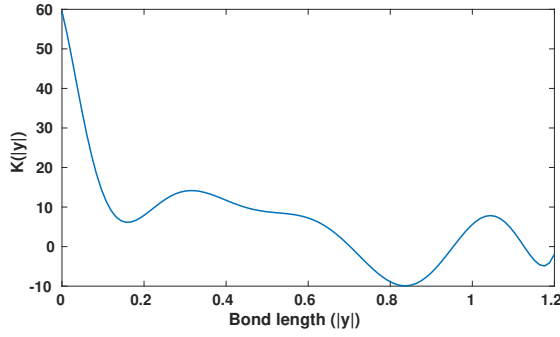
Figure 3: Optimal kernel K_{opt} as a function of distance.
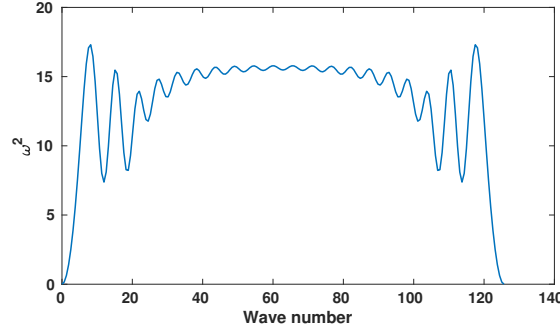

Figure 4: Dispersion curve associated with K_{opt}.

Figure 5: Comparison of the group velocity. Upper: group velocity corresponding to K_{opt}, K_{const}, and DNS data. Bottom: group velocity corresponding to K_{opt} for different pairs (δ, ϵ).

and Foster 2020; You et al. 2021; Weckner and Silling 2011), we observe a sign-changing behavior. We also compute the corresponding dispersion $\omega(k)$ and group velocity $v_g(\omega) = d\omega/dk$. For a given kernel K and different frequencies $k_i = 0, \frac{2\pi}{200h}, \cdots, \frac{2\pi}{h}$, the corresponding angular frequency $\omega(k_i)$ and group velocity $v_g(\omega(k_i))$ are approximated by

$$\omega(k_i)^2 \approx \frac{1}{\rho} \sum_q K(|y_q|)(1 - \cos(k_i y_q))h,$$

$$v_g(\omega(k_i)) \approx \frac{\omega(k_{i+1}) - \omega(k_{i-1})}{k_{i+1} - k_{i-1}},$$

where y_q belong to a uniform grid of size h in $(-\delta, \delta)$.

The dispersion curve is reported in Figure 4, its positivity indicates that K_{opt} corresponds to a physically stable material model. The group velocity is reported in the upper plot of Figure 5 in comparison with the curve computed with DNS by observing the speed of a wave packet of a given frequency as it moves through the microstructure. We also display the group velocity associated with an alternative kernel obtained for the same material by a completely different method (Silling 2020). This alternative kernel is a constant, specifically, we have that $C_m = K_{\mathrm{const}} = 0.7714$, for $M = 3$ and $\delta = 0.15$.

It is well known that layered, periodic elastic media have a band structure for wave propagation, see (Bedford and Drumheller 1994, pages 121-122). In the present study, because it is not possible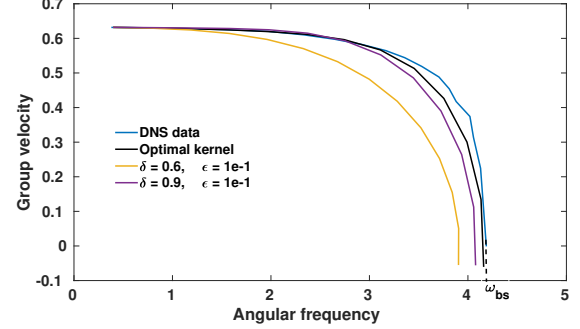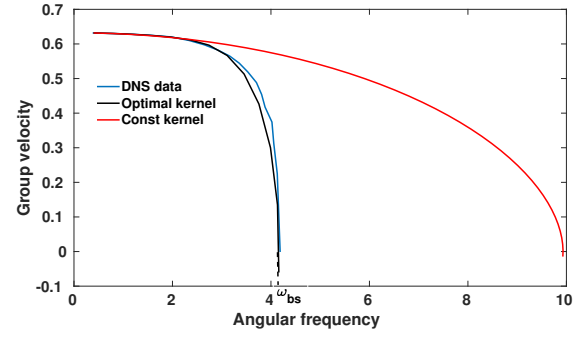 to reproduce the higher-frequency pass bands with the coarse discretization that is used in fitting our nonlocal kernel, we address only the first, low-frequency pass band, i.e. $\omega \in (0, \omega_{\mathrm{bs}})$, where "bs" stands for band stop. Hence, the optimal kernel is suitable only for wavelengths that are bigger than the microstructure; this is enough to reproduce the physically most important features of wave propagation in layered media for typical applications.

The profile of the group velocity shows the improved accuracy of our optimal kernel that not only matches the behavior for low values of ω, but also catches the behavior at $\omega = \omega_{\mathrm{bs}} \approx 4$. This fact has important consequences on the ability to reproduce wave propagation for values of ω bigger than ω_{bs}. In order to justify the statement above on the optimality of the parameters δ and ϵ, we report the group velocity profile in correspondence of different pairs; it is clear from the profiles in Figure 5 that $(\delta, \epsilon) = (1.2, 0.01)$ provides the best match both in terms of curvature at $\omega = 0$ and identification of the band stop.

Numerical validation

We test the performance of the optimal kernel K_{opt} on data sets of type 3) and 4), i.e. the problem setting considered for validation has different model parameters, including the domain, than the one used for training and, hence, these tests serve as an indicator of the generalization properties of our algorithm.

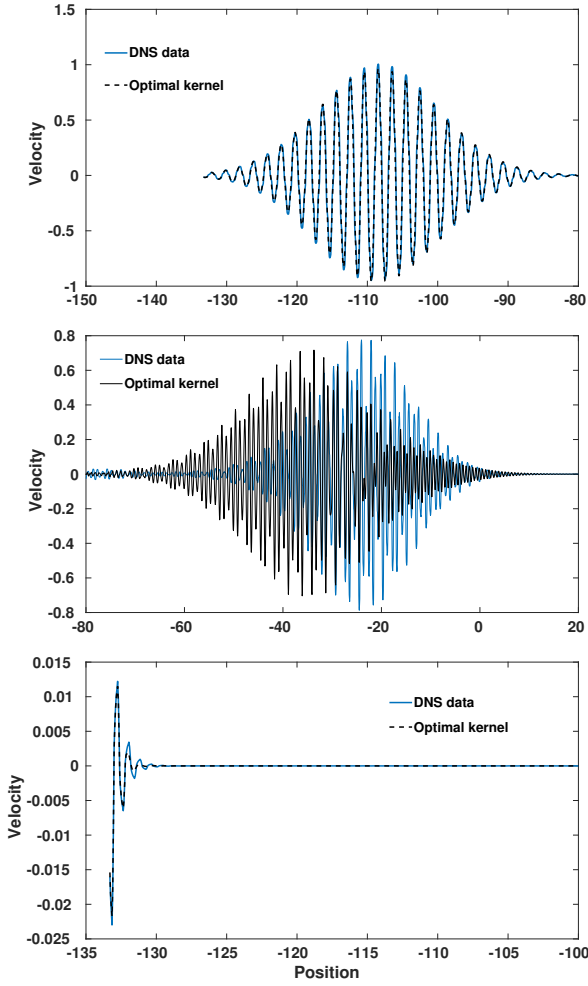*Wave packet.* For data type 3) we numerically compute so-

Figure 6: Velocity computed with K_{opt}. Plots from top to bottom correspond to: 1. $\omega_1 = 2$ at $t = 100$; 2. $\omega_2 = 3.9$ at $t = 320$; 3. $\omega_3 = 5$ at $t = 100$.

Figure 7: Velocity computed with K_{const}. Plots from top to bottom correspond to: 1. $\omega_1 = 2$ at $t = 100$; 2. $\omega_2 = 3.9$ at $t = 320$; 3. $\omega_3 = 5$ at $t = 100$. The optimal kernel K_{opt} obtained by machine learning (Figure 6) clearly performs better than K_{const} for the second and third cases (ω_2 and ω_3).

lutions to (2) using K_{opt} and DNS data as nonlocal boundary conditions. We consider solutions corresponding to three values of ω: $\omega_1 = 2 < \omega_{\text{bs}}$, $\omega_2 = 3.9 \approx \omega_{\text{bs}}$ and $\omega_3 = 5 > \omega_{\text{bs}}$. Note that the latter value is beyond the band stop and, as such, corresponds to a zero group velocity, i.e. the wave does not travel in time. In Figure 6 we report the velocity corresponding to the computed displacement $\bar{u}$ at time $t = 100$, $t = 320$, and $t = 100$ for ω_1, ω_2, and ω_3 respectively; as a reference, we also report the exact DNS velocity. Our results indicate that our kernel can accurately reproduce solutions of type 3) at times larger than T_{tr} and for all values of ω, even larger than ω_{bs}. This is possible because the group velocity corresponding to K_{opt} reproduces the true group velocity very accurately, see Figure 5. In particular, detecting the presence of a band stop allows us to accurately predict the wave propagation for values of $\omega > \omega_{\text{bs}}$. Due to the poor accuracy of the group velocity associated to K_{const}, corresponding solutions are not as accurate for ω in the proximity of ω_{bs} and beyond. To illustrate this phenomenon, we re-
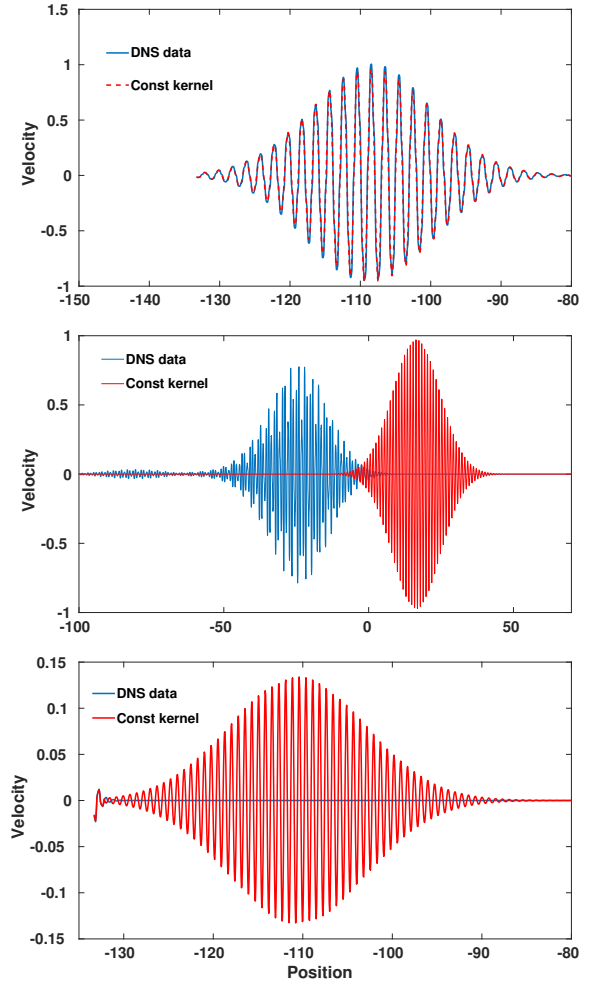
port in Figure 7 the behavior of the velocity corresponding to K_{const} at time $t = 100$, $t = 320$ and 100, respectively for ω_1, ω_2 and ω_3. Comparison with DNS data shows that, for ω_2 the wave associated with K_{const} is traveling faster than the exact one and, for ω_3, it keeps traveling while the exact wave is not propagating.

Impact. We use the optimal kernel to compute solutions corresponding to data type 4). In Figure 8 we report the velocity profile at different time steps in correspondence of K_{opt} and DNS data, displayed for comparison. Figure 9 displays the same results in correspondence of K_{const}. These results indicate that our optimal kernel can accurately predict the short- and long-time wave propagation, as opposed to the constant kernel that successfully predicts the long-time behavior only. We also point out that for values of (δ, ϵ) for which the group
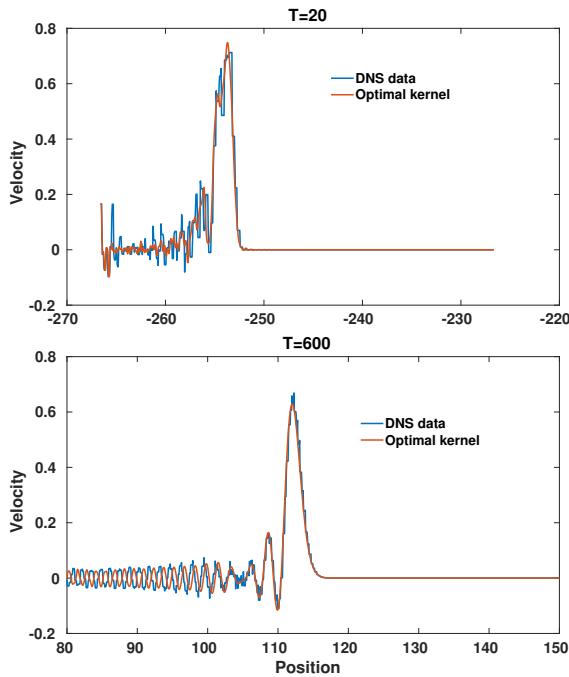
Figure 8: Velocity profile for the *Impact* problem at $T = 20$ and $T = 600$ with K_{opt}.

Figure 9: Velocity profile for the *Impact* problem at $T = 20$ and $T = 600$ with K_{const}. The optimal kernel K_{opt} obtained by machine learning (Figure 8) provides better agreement with the DNS data than K_{const} at the earlier time ($T = 20$) by reducing the size of the oscillations that trail the main pulse. For large T, the solution is dominated by the low frequency components of the pulse, for which the two kernels behave similarly.

velocity is not accurate, the predicted velocity and displacement exhibit non-physical oscillations, which disappear in correspondence of pairs that guarantee an accurate group velocity profile.

Conclusion

We introduced a new data-driven, optimization-based algorithm for the identification of nonlocal kernels in the context of wave propagation through material featuring heterogeneities at the microscale. The corresponding nonlocal model is well-posed by construction and allows for accurate simulations at a larger scale than the microstructure. We stress the fact that our algorithm does not require a priori knowledge of the microstructure (often unknown and/or hard to model), but only requires high-fidelity measurements of the displacements or the velocity. We also point out that our algorithm has excellent generalization properties as the optimal kernel performs well at much larger times than the time instants used for training and on problem settings that are substantially different from the training data set.

One of the most important findings in this work is the key role of the group velocity in the accuracy of the predictions; in fact, our criterion for the choice of the horizon δ and the regularization weight ϵ is the accurate prediction of the group velocity profile. Given the critical role of such quantity, our future work includes the identification of the optimal horizon by, possibly, embedding constraints on the group velocity to the training procedure. Another natural follow-up work is the illustration of the efficiency of our algorithm on two- and three-dimensional test cases.
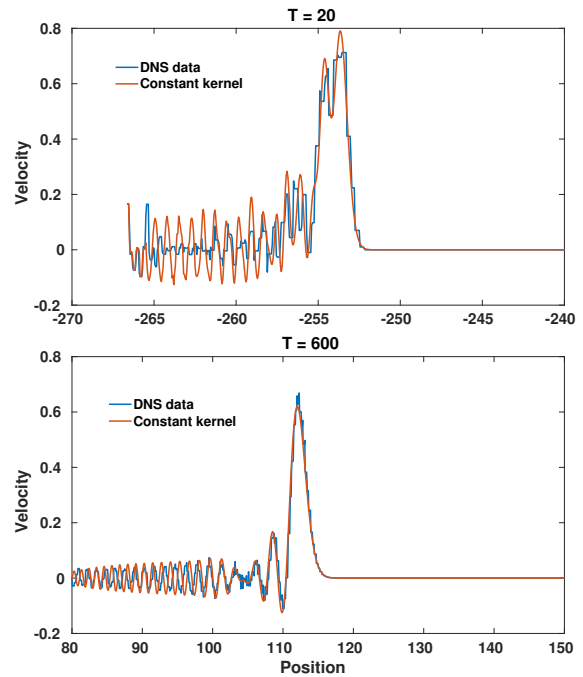
Acknowledgments

References

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

Askari, E.; Bobaru, F.; Lehoucq, R.; Parks, M. L.; Silling, S. A.; and Weckner, O. 2008. Peridynamics for multiscale materials modeling. In *Journal of Physics: Conference Series*, volume 125, 012078. IOP Publishing.

Bedford, A.; and Drumheller, D. S. 1994. *Introduction to Elastic Wave Propagation*. Wiley.

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

Burkovska, O.; Glusa, C.; and D'Elia, M. 2020. An optimization-based approach to parameter learning for fractional type nonlocal models. ArXiv:2010.03666.

D'Elia, M.; De los Reyes, J.; and Trujillo, A. M. 2019. Bilevel parameter optimization for nonlocal image denoising models. ArXiv1912.02347.

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.

Du, Q.; Tao, Y.; and Tian, X. 2018. A peridynamic model of fracture mechanics with bond-breaking. *Journal of Elasticity* 132(2): 197–218.

D'Elia, M.; and Gunzburger, M. 2016. Identification of the diffusion parameter in nonlocal steady diffusion problems. *Applied Mathematics & Optimization* 73(2): 227–249.

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

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

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

Pang, G.; D'Elia, M.; Parks, M.; and Karniadakis, G. E. 2020. nPINNs: nonlocal Physics-Informed Neural Networks for a parametrized nonlocal universal Laplacian operator. Algorithms and Applications. To appear in Journal of Computational Physics.

Pang, G.; Lu, L.; and Karniadakis, G. E. 2019. fPINNs: Fractional Physics-Informed Neural Networks. *SIAM Journal on Scientific Computing* 41: A2603–A2626.

Scalas, E.; Gorenflo, R.; and Mainardi, F. 2000. Fractional calculus and continuous time finance. *Physica A* 284: 376–384.

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. 2020. Propagation of a Stress Pulse in a Heterogeneous Elastic Bar. Sandia Report SAND2020-8197, Sandia National Laboratories.

Silling, S. A. 2000. Reformulation of elasticity theory for discontinuities and long-range forces. *Journal of the Mechanics and Physics of Solids* 48(1): 175–209.

Trask, N.; You, H.; Yu, Y.; and Parks, M. L. 2019. An asymptotically compatible meshfree quadrature rule for nonlocal problems with applications to peridynamics. *Computer Methods in Applied Mechanics and Engineering* 343: 151–165.

Weckner, O.; and Silling, S. A. 2011. Determination of nonlocal constitutive equations from phonon dispersion relations. *International Journal for Multiscale Computational Engineering* 9(6).

Xu, X.; D'Elia, M.; and Foster, J. 2020. Bond-Based Peridynamic Kernel Learning with Energy Constraint. ArXiv:2101.01095.

Xu, X.; and Foster, J. 2020. Deriving peridynamic influence functions for one-dimensional elastic materials with periodic microstructure. ArXiv:2003.05520.

You, H.; Yu, Y.; and Kamensky, D. 2020. An asymptotically compatible formulation for local-to-nonlocal coupling problems without overlapping regions. *Computer Methods in Applied Mechanics and Engineering* 366: 113038.

You, H.; Yu, Y.; Trask, N.; Gulian, M.; and D'Elia, M. 2021. Data-driven learning of robust nonlocal physics from high-fidelity synthetic data. *Computer Methods in Applied Mechanics and Engineering* 374: 113553.