This is the accepted manuscript made available via CHORUS. The article has been published as:

Interpreting Neural Operators: How Nonlinear Waves Propagate in Nonreciprocal Solids

Jonathan Colen, Alexis Poncet, Denis Bartolo, and Vincenzo Vitelli

Phys. Rev. Lett. **133**, 107301 — Published 3 September 2024

DOI: 10.1103/PhysRevLett.133.107301

Interpreting neural operators: how nonlinear waves propagate in non-reciprocal solids

Jonathan Colen,[1,2,3,*] Alexis Poncet,[4,*] Denis Bartolo,[4,†] and Vincenzo Vitelli[1,2,‡]

[1] *James Franck Institute, University of Chicago, Chicago, IL 60637*
[2] *Department of Physics, University of Chicago, Chicago, IL 60637*
[3] *Joint Institute on Advanced Computing for Environmental Studies, Old Dominion University, Norfolk, Virginia 23539*
[4] *Univ. Lyon, ENS de Lyon, Univ. Claude Bernard, CNRS, Laboratoire de Physique, F-69342, Lyon.*

(Dated: July 27, 2024)

We present a data-driven pipeline for model building that combines interpretable machine learning, hydrodynamic theories, and microscopic models. The goal is to uncover the underlying processes governing nonlinear dynamics experiments. We exemplify our method with data from microfluidic experiments where crystals of streaming droplets support the propagation of nonlinear waves absent in passive crystals. By combining physics-inspired neural networks, known as neural operators, with symbolic regression tools, we infer the solution, as well as the mathematical form, of a nonlinear dynamical system that accurately models the experimental data. Finally, we interpret this continuum model from fundamental physics principles. Informed by machine learning, we coarse grain a microscopic model of interacting droplets and discover that non-reciprocal hydrodynamic interactions stabilise and promote nonlinear wave propagation.

Introduction. — The goal of physical modeling is to build a set of rules which can predict and explain behaviors observed in experiments. In continuum theories, these rules are typically grounded in symmetries and conservation laws, which determine the relevant degrees of freedom and the equations they must satisfy. This approach faces challenges when applied to experiments that only reveal the dynamics of a finite subset of all the interacting degrees of freedom. This situation is not exceptional, in particular it is the norm when soft matter is driven out-of-equilibrium [1, 2]. In these systems, the particles, such as colloids, droplets, or bubbles continuously exchange energy and momentum with a solvent. As a result they exhibit a variety of behaviors, driven by effective interactions that violate microscopic reversibility constraints such as detailed balance [3, 4] or Newton's third law [5–15]. Such complications require new approaches to model building that can identify and predict the effects of these microscopic interactions in experiments [16–19].

Data-driven methods, which have shown much promise in learning physical models from observation data, have the potential to overcome these challenges. Deep neural networks [20, 21] can learn to accurately predict behavior for a variety of physical [22–30] and biological [31–35] systems. More recently, neural operators have been introduced to directly learn the solution operator for partial differential equations [36]. When incorporated within deep neural networks, they are capable of forecasting the behavior of complex dynamical systems such as the weather [36, 37]. While these methods are invaluable for efficiently and accurately forecasting dynamics from data, less is known about how to interpret what they have learned about a physical system.

In this case study, we demonstrate how interpreting neural operators can lead to data-driven model discovery by focusing on a paradigmatic out-of-equilibrium hydrodynamic problem: particles driven in a surrounding fluid [38]. Such systems exhibit a variety of phenomena: phonons in overdamped systems [39, 40], shock formation [41, 42], and symmetry-dependent melting of crystal phases [43]. We build on an experimental set-up in which water droplets advected inside a microfluidic channel form a propagating shock due to hydrodynamic interactions. Using a neural network with an internal neural operator layer [36], we transform the experimental movie into a set of variables whose dynamics are solved by an unknown partial differential equation. Using sparse regression [44–46], we then discover the equation that these machine-learned variables satisfy. We find that the discovered continuum model contains an essential term that had remained overlooked. Informed by our data-driven method, we then trace this term to the non-reciprocal nature of hydrodynamic interactions by solving suitable microscopic models. Beyond the specifics of our microfluidic experiments, we show how microscopic non-reciprocity generically results in the propagation of macroscopic solitons through non-equilibrium crystals. Our approach demonstrates how combining interpretable neural operators with theoretical modeling can provide previously-unknown insights into the physics governing complex non-linear systems.

Learning experimental dynamics. — To learn macroscopic physics from microscopic data, we consider an experimental movie [42] in which a one dimensional stream of water droplets (diameter 140 μm) are advected by an organic solvent inside a microfluidic channel, see [42] for details. The droplet stream develops a shock in response to a jam induced at $t = 0$ (Fig. 1a).

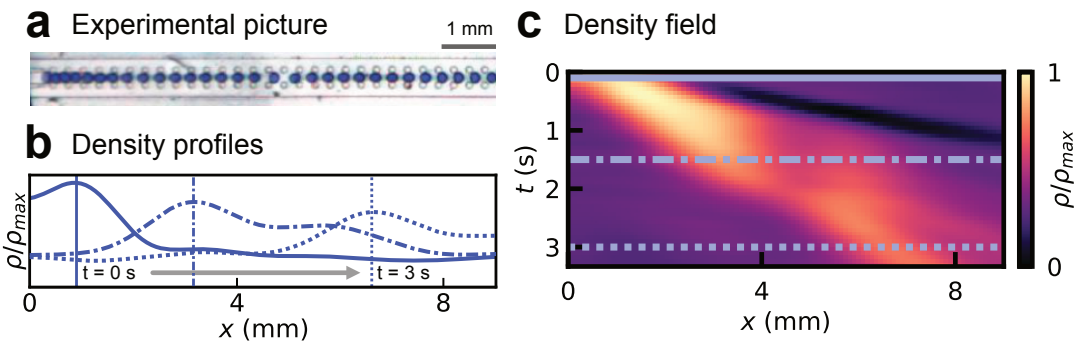


FIG. 1. **Coarse-graining microfluidic particle experiments.** (**a**) Experimental picture of a microfluidic particle stream [42].(**b**) Experimental pictures are coarse-grained into a density profile and normalized by the maximum density ρ_{max}. (**c**) Spatio-temporal evolution of the density field shows formation of a shock front in response to an initial jam.

Coarse-graining the droplet positions yields a density field $\rho(x,t)$ which shows the shock deforming and propagating through space and time (Fig. 1). The system behavior is driven by microscopic interactions between the droplets, which may include a simple nearest-neighbor repulsive interaction as well as hydrodynamic backflows. Previous studies modeled this experiment at the continuum level to leading order in derivatives and nonlinearities. The resulting model reduced to a Burgers' equation obeyed by the droplet density $\rho(x,t)$ [42]

$$\partial_t \rho + (c - \alpha\rho)\partial_x \rho = 0, \tag{1}$$

where c is a basic advection speed, and α is a nonlinearity that creates shocks. Here, without making any assumption about the relevant interactions, we learn to predict the dynamics of this movie using a deep neural network. To train the network, we segmented the experimental movie into overlapping $3\,\mathrm{s} \times 9\,\mathrm{mm}$ training examples. For each example, the network receives the one-dimensional slice $\rho(x) = \rho_0(x)$ and predicts the full $\rho(x,t)$ response to that initial condition (Fig. 2a). The network architecture has three components: (i) an input convolutional block which transforms the experimental density field $\rho_0(x)$ into a latent variable $\phi_0(x)$, (ii) a single-layer Fourier neural operator which evolves ϕ in time according to an unknown *linear* partial differential equation, and (iii) an output convolutional block which transforms the predicted $\phi(x,t)$ field into a predicted density field $\rho(x,t)$. Both the input and output blocks used 1-D convolutional layers which only use spatial information to construct features. The full architecture is shown in Fig. 2a. We note that the dynamics were solely predicted by the central neural operator layer. However, while the neural operator itself is linear, the input and output transformations enable the network to predict nonlinear dynamics [47]. Once trained, the neural network is capable of predicting experimental dynamics including the propagation of the initial jam in space and time from past to present (Fig. 2a-c) [48].

Interpreting the neural operator. — The single-layer linear Fourier neural operator can be interpreted as a Green's function for the internal variable ϕ. This learned Green's function predicts how the response to an initial perturbation will propagate in space over time (Fig. 2a-b). In the co-moving frame, defined by the spatial coordinate $x - ct$ where c is the speed of the wave, we observe that the response decreases in magnitude while the width stays nearly constant (Fig. 2c-e). The encoder/decoder architecture transforms the nonlinear dynamics of the density ρ into a linear problem that can be solved using this Green's function [47].

A similar approach can exactly solve the Burgers' equation (1) previously proposed for this system. Upon adding a small diffusion term $D\,\partial_x^2\rho$ and effecting a Cole-Hopf transformation, we obtain a linear partial differential equation whose Green's function is that of the standard diffusion equation (see SI). In Figure 2f-h, we plot the dynamics predicted by Burgers' equation from experimental initial conditions. As a benchmark for our machine-learned solution, we trained an identical neural network on a synthetic movie of an exact solution to Burgers' equation (SI Fig. S3). While this model learned an internal variable transformation that was distinct from Cole-Hopf, the Green's function is qualitatively similar to the diffusion kernel with a response whose amplitude decays and width grows over time.

The neural network (Figure 2c-e) more accurately predicts the experimental movie than an exact solution to Burgers' equation (Figure 2f-h) and uses a qualitatively distinct Green's function to do so. We stress that while both Green's functions decay in amplitude over time, the diffusion kernel has a response width that grows over time while the neural operator solution remains constant (Fig. 2e,h). As a result, the neural network shock remains sharp unlike the Burgers' shock that disperses (Fig. 2a,c,f). The distinct behavior of the neural operator trained on experiments from both the exact and benchmark solutions to Burgers' equation suggests that

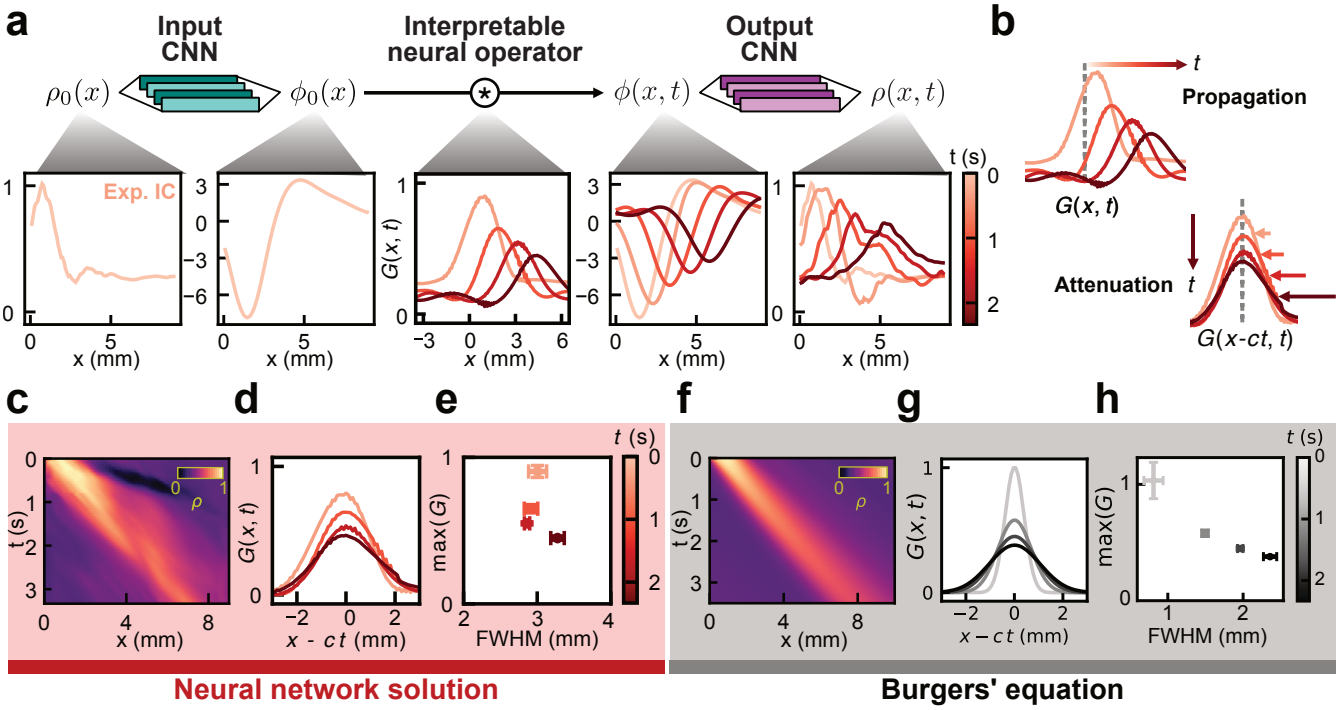


FIG. 2. Learning the dynamics of shocks in microfluidic particles experiments. (**a**) A neural network forecasts the particle density field $\rho(x)$ from initial conditions. The network uses a 1-D convolutional block to map the density field to an internal variable $\phi(x)$. A single-layer linear Fourier neural operator forecasts the evolution of ϕ. A second 1-D convolutional block converts $\phi(x,t)$ into a predicted density field $\rho(x,t)$. (**b**) The neural operator is interpretable as a Green's function for an unknown linear PDE which defines the response to an perturbation. The learned Green's function predicts a response which propagates in space and attenuates over time. (**c**) Neural network predictions for the density field $\rho(x,t)$. (**d**) Plot of the machine-learned Green's function in the co-moving frame. (**e**) Comparing the magnitude and width of the kernel $G(x,t)$ shows that the response to a perturbation will decrease in magnitude over time, but not disperse. (**f-h**) Exact solution to Burgers' equation from experimental initial conditions. $G(x,t)$ is a diffusion kernel applied after a Cole-Hopf transformation of the density ρ. The neural network learns distinct physics from Burgers', which exhibits spreading due to viscous dissipation.

the non-linear waves seen in experiments obey physics beyond Burgers' equation.

To learn an equation governing the internal variable $\phi(x,t)$ (Fig. 3a), we resort to the SINDy method [44], which has proven to effectively learn non-linear dynamical equations from experimental data [19, 28] . As the linear neural operator ensures ϕ obeys a linear PDE, we used a library of ϕ and its derivatives and found the following equation

$$\partial_t\phi - 1.86\,\partial_x\phi = -0.31\,\phi + 0.003\,\partial_x^2\phi - 0.008\,\partial_x^3\phi \quad (2)$$

Simulating this equation for $\phi(x,t)$ yields excellent correspondence with the neural network solution (Fig. 3a, $R^2 = 0.96$), indicating that it is indeed an excellent model for the dynamics solved by the neural operator. We applied the neural network's output CNN to this predicted ϕ field to achieve an accurate prediction of the density dynamics (SI Fig. S2). This approach of learning the dynamics of the internal variable predicted by the neural operator yielded better predictions than applying SINDy directly to the density field.

The learned equation (2) resembles a linearized KdV-Burgers' equation, containing a dispersive $\partial_x^3\phi$ term characteristic of the Korteweg de-Vries equation. The $\partial_x\phi$ term captures both the basic advection term and a linearization of the $\rho\partial_x\rho$ term in (1). We do note that the model has learned a damping term whose importance to the learned dynamics is weaker than the dispersive $\partial_x^3\phi$ term but stronger than the diffusive $\partial_x^2\phi$ term (SI Table S1). The new KdV term in the dynamics of ϕ, as well as the qualitatively distinct Green's function learned by the neural operator, corroborates our conclusion that the dynamics of the experimental density field is governed by physics beyond Burgers' equation. Hence, we consider the more general nonlinear KdV-Burgers' equation:

$$\partial_t\rho + (c - \alpha\,\rho)\partial_x\rho = D\partial_x^2\rho + \beta\,\partial_x^3\rho \quad (3)$$

We used the experimental movie to directly fit coefficients to a Burgers' equation ($\beta = 0$), a KdV equation ($D = 0$), and a KdV-Burgers' equation (3). We plot the predictions and compare accuracy in Figure 3b-c. Both Burgers' and KdV-Burgers' achieve similar error rates on the

entire field, although neither matches the performance of the neural network. The difference between the two is more apparent near the shock front. The KdV-Burgers' solution maintains a more pronounced shock over time (Fig. 3b) and achieves a lower error rate than Burgers' in the vicinity of this shock front (Fig. 3c).

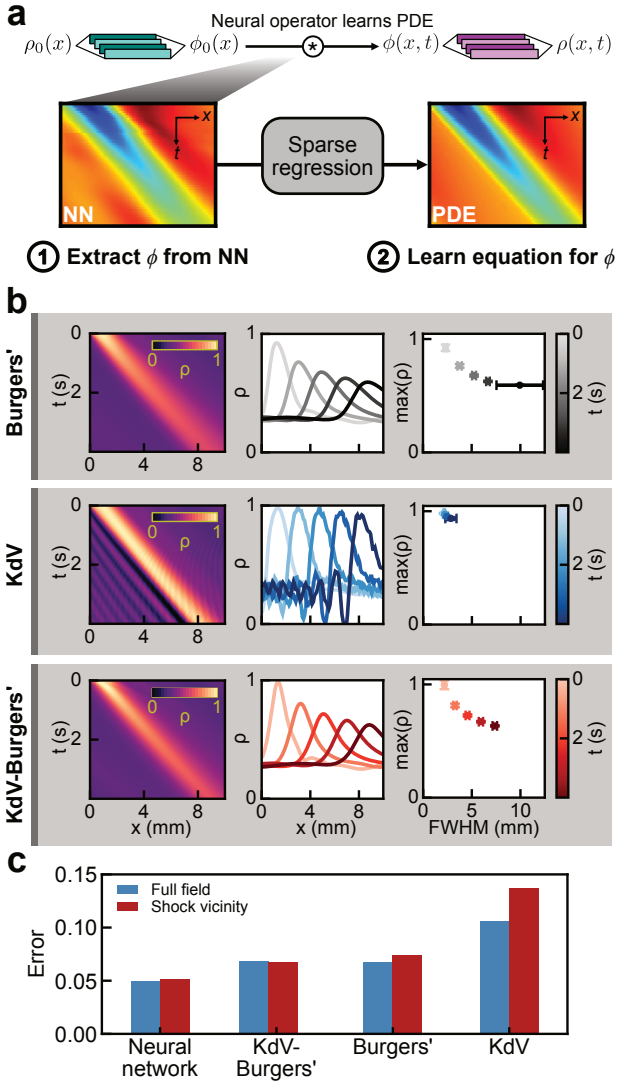


FIG. 3. Interpreting the neural operator. (**a**) The neural operator evolves the internal variable $\phi(x,t)$ according to a partial differential equation. Using SINDy, we learn the PDE governing $\phi(x,t)$. The resulting equation (2) includes a novel ∂_x^3 KdV term [42]. (**b**) Simulations of Burgers' equation (first row), KdV equation (second row), and KdV-Burgers' equation (third row) from experimental initial conditions. For each equation, we plot the density field (left) and traces at fixed time points (middle). To characterize the shock front, we plot the height vs. width at different time points (right). KdV-Burgers' preserves the sharpness of the shock and shows less spreading over time compared to Burgers'. (**c**) Error rates for neural network (Fig. 2) and analytic solutions (Fig. 3). Blue bars are the mean absolute error $\langle|\rho-\rho_0|\rangle$ for the full field, while red is the error within 1 mm of the shock front. Numerical values in SI Table S2.

Interpretation of learned non-linear dynamics from microscopics. — The neural operator discovered a KdV-Burgers dynamics but provided no first-principles interpretation for it. We now seek to understand its physical origin. To make progress, we build a minimal model where the dynamics of the droplets is determined by the competition between contact and hydrodynamic interactions. The equations of motion for the droplet positions R_n are

$$\zeta(\partial_t R_n - v_0) = \sum_{m\neq n} \left(f^H_{m\to n} + f^E_{m\to n}\right), \tag{4}$$

where ζ a friction coefficient and v_0 the advection speed of an isolated droplet. f^E is an elastic force that models the repulsion between the soft droplets which we approximate by linear springs $f^E_{n+1\to n} = k(R_{n+1} - R_n - a)$ (a is the lattice spacing and k the elastic stiffness). f^H models the hydrodynamic force experienced by a droplet in response to the backflows induced by the motion of its neighbors, see Fig. 4a. Unlike f^E, hydrodynamic forces do not obey Newton's third law [6]. They are non-reciprocal: $f^H_{m\to n} = f^H_{m\to n} = f_H(R_n - R_m)$. In practice, we use the standard form $f_H(r) = A\,\mathrm{csch}^2(\pi r/W)$, where W is the channel width [39].

We perform numerical simulations of Eq. 4 with periodic boundary conditions, $\zeta = a = 1$, $W = 2$, and $A = -(\pi/W)^2$. To probe the relevance of KdV-Burgers physics, we choose initial conditions that correspond to a 2-soliton solution of the KdV equation (Fig. 4b-e, see also SI for different initial conditions). When ignoring the soft contact interactions ($k = 0$), the particle dynamics is perfectly captured by a continuum model having the form of a KdV equation. Two solitons cross each other and propagate freely (Fig. 4c). In the limit of strong elastic forces ($k = 0.2$), the initial shape is not preserved, shocks develop and their dynamics is well described by a Burgers equation (Fig. 4e). In the intermediate regime, where both contact and hydrodynamic forces compete on equal footing ($k = 0.02$), we find that the propagation of the initial density fluctuation is quantitatively predicted by the KdV-Burgers equation Eq. (3), (Fig. 4d).

We are now equipped to interpret the KdV-Burgers dynamics revealed by our neural operator. We focus on large distances ($\tilde{x} = \epsilon(n - ct)$ with $\epsilon \to 0$) and large times ($\tilde{t} = \epsilon^3 t$), we then can recast Eq. (4) into a PDE obeyed by the density field $\rho(x,t)$. Informed by our Neural operator, we do not restrain our expansion to leading order in gradients and find that the dynamics of $\rho(x,t)$ is ruled by a KdV-Burgers equation (3). We provide a detailed derivation of our continuum model in SI and generalize it beyond the specifics of hydrodynamic interactions [49].

This procedure makes it possible to understand the physical origin of all terms in Eq. (3). The damping (D) is solely controlled by the elastic interactions f^E. Conversely, the advection, the non-linear and the

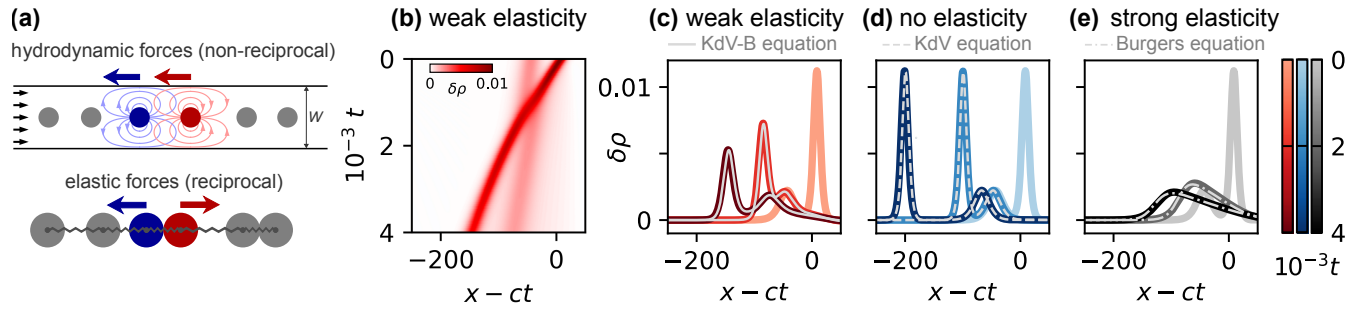


FIG. 4. Microscopic modeling of the experiments and confirmation of the KdV-Burger behavior. (a) We model the droplets as an overdamped chain of beads that interact by two kinds of interactions: hydrodynamic forces which are non-reciprocal ($f^H_{m\to n} = f^H_{n\to m}$) and elastic forces which correspond to usual springs ($f^E_{m\to n} = -f^E_{n\to m}$). (b) We run numerical simulations starting from small density fluctuations $\delta\rho = \rho - a^{-1}$ corresponding to a two-soliton solution of the KdV equation. At weak elasticity ($k = 0.02$) the solitons propagate by crossing one-another, and are damped by the elasticity. (c-e) The colored lines show the time-evolution in the simulations at weak elasticity ($k = 0.02$, corresponding to panel b), no elasticity ($k = 0$), and strong elasticity ($k = 0.2$). In the three cases, this evolution is in excellent agreement with the numerical integration of the KdV-Burgers equation (shown in panel (c) with thin gray lines, see SI for measurements of the error). Dashed lines in panels (d) and (e) are respectively the analytical solutions of the KdV equation and the Burgers equation, showing the two limit cases.

dispersive terms (c, α and β) originate from the hydrodynamic forces ($c = v_0 - 2\sum_{m=1}^{\infty} m f'_H(ma)$, $\alpha = -2\sum_{m=1}^{\infty} m^2 f''_H(ma)$, and $\beta = \sum_{m=1}^{\infty} m f'_H(ma)/3$). Unlike in [50], the dispersive and non-linear terms rooted in hydrodynamic interactions conspire to stabilize and propel solitary waves without relying on inertia.

In this case study, we used interpretable machine learning and microscopic modeling to learn rules governing a microfluidic experiment. Using an interpretable neural network architecture built around a linear neural operator, we learned to predict experimental behavior from data. Next, we examined the behavior of the neural operator and extracted a dynamical rule for our machine learned variable which contained a previously overlooked KdV term. By considering suitable microscopic models, we traced the origin of this novel term to the non-reciprocal nature of hydrodynamic interactions between the droplets.

Our first principles modeling and machine learning algorithms play a mutually beneficial role. While the neural operator can suggest the presence of a KdV term, it could not provide a mechanism justifying its existence. Our simulations and theoretical argument help rationalizing these findings which had been hitherto overlooked. They provide insight on the origin of the KdV-like physics: beyond the specifics of our driven emulsions, non-reciprocal interactions generically stabilize and propagate solitary waves in overdamped one-dimensional systems. Neural networks' role as universal approximators provides an alternative route to characterizing experiments. In learning to predict dynamics, they build a "maximal model" which can then serve as a target for more principled theoretical investigations.

This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101019141) (D.B. and A.P.). V.V. acknowledges support from the National Science Foundation through the Physics Frontier Center for Living Systems (PHY-2317138). J.C. acknowledges support from the Hampton Roads Biomedical Research Consortium, United States, as part of the efforts associated with the Old Dominion University-Thomas Jefferson National Accelerator Facility Joint Institute for Advanced Computing on Environmental Studies.

* These two authors contributed equally
† denis.bartolo@ens-lyon.fr
‡ vitelli@uchicago.edu

[1] S. Ramaswamy, Advances in Physics **50**, 297 (2001).
[2] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Reviews of Modern Physics **85**, 1143 (2013).
[3] M. Fruchart, C. Scheibner, and V. Vitelli, Annual Review of Condensed Matter Physics **14**, 471 (2023).
[4] C. R. Scheibner, A. Souslov, D. Banerjee, P. Surowka, W. T. Irvine, and V. Vitelli, Nature Physics **16**, 475 (2020).
[5] A. V. Ivlev, J. Bartnick, M. Heinen, C.-R. Du, V. Nosenko, and H. Löwen, Phys. Rev. X **5**, 011035 (2015).
[6] A. Poncet and D. Bartolo, Physical Review Letters **128**, 048002 (2022).
[7] F. Brauns and M. C. Marchetti, "Non-reciprocal pattern formation of conserved fields," (2023), arXiv:2306.08868 [nlin.PS].
[8] S. Saha and R. Golestanian, "Effervescent waves in a binary mixture with non-reciprocal couplings," (2022), arXiv:2208.14985 [cond-mat.stat-mech].
[9] S. Saha, J. Agudo-Canalejo, and R. Golestanian, Physical Review X **10**, 041009 (2020).
[10] Z. You, A. Baskaran, and M. C. Marchetti, Proceedings

of the National Academy of Sciences **117**, 19767 (2020).
[11] R. Mandal, S. S. Jaramillo, and P. Sollich, "Robustness of travelling states in generic non-reciprocal mixtures," (2022), arXiv:2212.05618 [cond-mat.stat-mech].
[12] Y. Avni, M. Fruchart, D. Martin, D. Seara, and V. Vitelli, "The non-reciprocal ising model," (2023), arXiv:2311.05471 [cond-mat.stat-mech].
[13] C. Weis, M. Fruchart, R. Hanai, K. Kawagoe, P. B. Littlewood, and V. Vitelli, "Exceptional points in nonlinear and stochastic dynamics," (2023), arXiv:2207.11667 [nlin.CD].
[14] D. Martin, D. Seara, Y. Avni, M. Fruchart, and V. Vitelli, "The transition to collective motion in nonreciprocal active matter: coarse graining agent-based models into fluctuating hydrodynamics," (2024), arXiv:2307.08251 [cond-mat.stat-mech].
[15] M. Fruchart, R. Hanai, P. B. Littlewood, and V. Vitelli, Nature **592**, 363 (2021).
[16] H. Li, X.-q. Shi, M. Huang, X. Chen, M. Xiao, C. Liu, H. Chaté, and H. P. Zhang, Proceedings of the National Academy of Sciences **116**, 777 (2019).
[17] W. Van Saarloos, V. Vitelli, and Z. Zeravcic, *Soft Matter: Concepts, Phenomena, and Applications* (Princeton University Press, 2024).
[18] D. S. Seara, J. Colen, M. Fruchart, Y. Avni, D. Martin, and V. Vitelli, "Sociohydrodynamics: data-driven modelling of social behavior," (2024), arXiv:2312.17627 [cond-mat.soft].
[19] R. Supekar, B. Song, A. Hastewell, G. P. T. Choi, A. Mietke, and J. Dunkel, Proceedings of the National Academy of Sciences **120**, e2206994120 (2023).
[20] Y. Lecun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
[21] J. Schmidhuber, Neural Networks **61**, 85 (2015).
[22] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Journal of Computational Physics **378**, 686 (2019).
[23] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Reviews of Modern Physics **91**, 045002 (2019).
[24] F. Cichos, K. Gustavsson, B. Mehlig, and G. Volpe, Nature Machine Intelligence **2**, 94 (2020).
[25] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Nature Reviews Physics **3**, 422 (2021).
[26] J. Carrasquilla and R. G. Melko, Nature Physics **13**, 431 (2017).
[27] V. Bapst, T. Keck, A. Grabska-Barwińska, C. Donner, E. D. Cubuk, S. Schoenholz, A. Obika, A. Nelson, T. Back, D. Hassabis, and P. Kohli, Nature Physics **16**, 448 (2020).
[28] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, Annual review of fluid mechanics **52**, 477 (2020).
[29] Z. Zhou, C. Joshi, R. Liu, M. M. Norton, L. Lemma, Z. Dogic, M. F. Hagan, S. Fraden, and P. Hong, Soft Matter **17**, 738 (2021).
[30] J. Colen, M. Han, R. Zhang, S. A. Redford, L. M. Lemma, L. Morgan, P. V. Ruijgrok, R. Adkins, Z. Bryant, Z. Dogic, M. L. Gardel, J. J. de Pablo, and V. Vitelli, Proceedings of the National Academy of Sciences **118**, e2016708118 (2021).
[31] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, Nature **596**, 583–589 (2021).
[32] C. J. Soelistyo, G. Vallardi, G. Charras, and A. R. Lowe, Nature Machine Intelligence **4**, 636 (2022).
[33] A. Zaritsky, A. R. Jamieson, E. S. Welf, A. Nevarez, J. Cillay, U. Eskiocak, B. L. Cantarel, and G. Danuser, Cell Syst **12**, 733 (2021).
[34] M. S. Schmitt, J. Colen, S. Sala, J. Devany, S. Seetharaman, A. Caillier, M. L. Gardel, P. W. Oakes, and V. Vitelli, Cell **187**, 481 (2024).
[35] M. Lefebvre, J. Colen, N. Claussen, F. Brauns, M. Raich, N. Mitchell, M. Fruchart, V. Vitelli, and S. J. Streichan, "Learning a conserved mechanism for early neuroectoderm morphogenesis," (2023), bioRxiv 2023.12.22.573058.
[36] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Fourier neural operator for parametric partial differential equations," (2021), arXiv:2010.08895 [cs.LG].
[37] J. Pathak, S. Subramanian, P. Harrington, S. Raja, A. Chattopadhyay, M. Mardani, T. Kurth, D. Hall, Z. Li, K. Azizzadenesheli, P. Hassanzadeh, K. Kashinath, and A. Anandkumar, "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators," (2022), arXiv:2202.11214 [physics.ao-ph].
[38] T. Beatus, R. H. Bar-Ziv, and T. Tlusty, Physics Reports **516**, 103 (2012).
[39] T. Beatus, R. Bar-Ziv, and T. Tlusty, Physical Review Letters **99**, 124502 (2007).
[40] N. Desreumaux, J.-B. Caussin, R. Jeanneret, E. Lauga, and D. Bartolo, Physical Review Letters **111**, 118301 (2013).
[41] T. Beatus, T. Tlusty, and R. Bar-Ziv, Physical Review Letters **103**, 114502 (2009).
[42] N. Champagne, E. Lauga, and D. Bartolo, Soft Matter **7**, 11082 (2011).
[43] I. Saeed, H. K. Pak, and T. Tlusty, Nature Physics **19**, 536 (2023).
[44] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Proceedings of the National Academy of Sciences **113**, 3932 (2016).
[45] K. Kaheman, J. N. Kutz, and S. L. Brunton, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **476**, 20200279 (2020).
[46] K. Champion, P. Zheng, A. Y. Aravkin, S. L. Brunton, and J. N. Kutz, IEEE Access **8**, 169259 (2020).
[47] C. R. Gin, D. E. Shea, S. L. Brunton, and J. N. Kutz, Scientific Reports **11**, 21614 (2021).
[48] Neural network and simulation code is available at `https://github.com/jcolen/microfluidics_neural_operators`.
[49] See Supplemental Material for additional information on neural network training and analysis, simulation details, and the derivation of our continuum model. SI includes Refs. [51–61].
[50] J. Veenstra, O. Gamayun, X. Guo, A. Sarvi, C. V. Meinersen, and C. Coulais, Nature **627**, 528 (2024).
[51] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner,

L. Fang, J. Bai, and S. Chintala, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," (2019), arXiv:1912.01703 [cs.LG].

[52] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2022) pp. 11966–11976.

[53] O. Ronneberger, P. Fischer, and T. Brox, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, edited by N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi (Springer International Publishing, Cham, 2015) pp. 234–241.

[54] B. d. Silva, K. Champion, M. Quade, J.-C. Loiseau, J. Kutz, and S. Brunton, Journal of Open Source Software **5**, 2104 (2020), publisher: The Open Journal.

[55] A. A. Kaptanoglu, B. M. d. Silva, U. Fasel, K. Kaheman, A. J. Goldschmidt, J. Callaham, C. B. Delahunt, Z. G. Nicolaou, K. Champion, J.-C. Loiseau, J. N. Kutz, and S. L. Brunton, Journal of Open Source Software **7**, 3994 (2022).

[56] T. Dauxois and M. Peyrard, *Physics of Solitons* (Cambridge University Press, 2006).

[57] A. D. Polyanin and V. F. Zaitsev, *Handbook of Nonlinear Partial Differential Equations: Exact Solutions, Methods, and Problems* (Chapman and Hall/CRC, 2003).

[58] V. E. Adler and A. B. Shabat, Theoretical and Mathematical Physics **201**, 1442 (2019).

[59] M. Toda, Journal of the Physical Society of Japan **22**, 431 (1967).

[60] C. Rackauckas and Q. Nie, Journal of Open Research Software **5** (2017), 10.5334/jors.151.

[61] K. J. Burns, G. M. Vasil, J. S. Oishi, D. Lecoanet, and B. P. Brown, Physical Review Research **2**, 023068 (2020).