Learning Constitutive Models in Continuum Mechanics

Learning Homogenized Models in PDEs

Andrew Stuart

Computing and Mathematical Sciences
California Institute of Technology

AFOSR, ARL, NSF, ONR
DataSig, Oxford

January 27, 2022

Collaborators:

- Kaushik Bhattacharya
  - (Mechanial and Civil Engineering, Caltech)
- Burigede Liu
  - (Engineering, Cambridge)
- Margaret Trautner
  - (Computing and Mathematical Sciences, Caltech)

Liu, Kovachki, Li, Azizzadenesheli, Anandkumar, AMS, Bhattacharya '22 [11]

Bhattacharya, Liu, AMS, Trautner '22 [3]

Table of Contents

Talk Outline

Operator Learning

Learning Constitutive Models

Elasticity

Viscoelasticity

Plasticity

Conclusions

The Problem

Formulation

Function Spaces: $\mathcal{E}, \mathcal{S}: \ \mathbb{R}^d$ – valued over domain $D \subset \mathbb{R}^q$

Input-Output Map: $\Psi^\dagger : \mathcal{E} \to \mathcal{S}$

Data: $\{e_n, s_n\}_{n=1}^N, \ s_n \approx \Psi^\dagger(e_n), \ e_n \overset{i.i.d.}{\sim} \mu,$

Goal: Supervised Learning in Banach Space

Parameter Space $\Theta \subseteq \mathbb{R}^p$

Operator Class: $\Psi : \mathcal{E} \times \Theta \to \mathcal{S}$

Operator Approximation: $\Psi(\cdot\,; \theta^\star) \approx \Psi^\dagger$

Principal Components Analysis: PCA-NET

Architecture Bhattacharya, Hosseini, Kovachki and AMS '21 [2]

$$\Psi_{PCA}(e;\theta)(z) = \sum_{j=1}^{m} \alpha_j(Le;\theta)\psi_j(z), \quad \forall e \in \mathcal{E} \qquad z \in D.$$

Details

- L maps e to its PCA coefficients under μ.
- $\{\psi_j\}$ are PCA basis functions under $(\Psi^\dagger)^\sharp\mu$.
- $\{\alpha_j\}$ are finite dimensional neural networks.

Neural Operator: NOP-NET

Architecture Li, Kovachki, Azizzadenesheli, Liu, Bhattacharya, AMS and Anandkumar et al '20 [10]

$$\Psi_{NOP}(e;\theta)(z) = \mathcal{Q}_\theta \circ \mathcal{L}_L \circ \cdots \mathcal{L}_2 \circ \mathcal{L}_1 \circ \mathcal{R}_\theta(e)(z), \; \forall e \in \mathcal{E}, \; z \in D,$$

$$\mathcal{L}_l(v)(z;\theta) = \sigma\big(W_l \mathcal{P} v(z) + b_l 1(z) + (\mathcal{K}_{l,\theta}(\mathcal{P}v)(z)\big).$$

Details

- σ activation function, applied pointwise (Nemitskii operator).
- $\mathcal{Q}_\theta, \mathcal{R}_\theta$ pointwise NNs or linear transformations.
- W_l, b_l define pointwise affine transformations.
- $\mathcal{K}_{l,\theta}$ integral operator (convolution: FFT).

Recurrent Neural Operator: RNO-NET. $D = (0, T)$

Architecture Bhattacharya, Liu, AMS, Trautner '22 [3]

$$\Psi_{RNO}(e;\theta)(t) = F\Big(e(t), \frac{de}{dt}(t), r(t);\theta\Big), \quad \forall e \in \mathcal{E} \qquad t \in [0, T],$$
$$\frac{dr}{dt} = G(r, e;\theta), \quad \forall e \in \mathcal{E} \qquad t \in (0, T], \quad r(0) = 0.$$

Details

- F, G neural networks;
- Two-layer used in this talk.

Universal Approximation

Theorems

For every $\epsilon > 0$ there is choice of parameters θ such that

- PCA-NET [2] $\mathbb{E}^{data}\|\Psi(\cdot;\theta) - \Psi^\dagger\|^2_{L^2_\mu(\mathcal{E};\mathcal{S})} < \epsilon.$
- NOP-NET [8] $\sup_{x \in compact} \|\Psi(x;\theta) - \Psi^\dagger(x)\|_{\mathcal{S}} < \epsilon.$

- These theorems:

  Bhattacharya, Hosseini, Kovachki and AMS '21 [2],

  Kovachki, Li, Liu, Azizzadenesheli, Bhattacharya, AMS and Anandkumar '21 [8];
- DEEP-ONET:

  Lanthaler, Mishra and Karniadakis '21 [9];
- Curse of dimensionality:

  Lanthaler, Mishra and Karniadakis '21 [9],

  Kovachki, Lanthaler and Mishra '21 [7];
- Two-layer neural networks on Banach space:

  Korolev '21 [6];
- RNO-NET – this talk.

Talk Outline

Operator Learning

Learning Constitutive Models

Elasticity

Viscoelasticity

Plasticity

Conclusions

Big Picture

Multiscale Problem Gonzalez and AMS '08 [5]

Displacement $u^\epsilon(x,t)$, stress $\sigma^\epsilon(x,t)$, $0 < \epsilon \ll 1$:

$$\rho\,\partial_t^2 u^\epsilon = \nabla \cdot (\sigma^\epsilon) + f,$$
$$\sigma^\epsilon = \Psi^\epsilon\left(\{\nabla u^\epsilon\}, \frac{x}{\epsilon}\right)$$

Homogenized Problem Bensoussan, Lions, Papanicolaou [1]

Approximate $u^\epsilon = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \dots$

Determine map Ψ, so that small scales are removed in u_0:

$$\rho\,\partial_t^2 u_0 = \nabla \cdot (\sigma) + f,$$
$$\sigma = \Psi(\{\nabla u_0\}).$$

Operator Learning

Approximate $\Psi \approx \Psi_{NN}$

Talk Outline

Multiscale Problem

Elasticity Multiscale Problem

$$
\begin{aligned}
-\nabla \cdot (A^{\epsilon} \nabla u^{\epsilon}) &= f, \quad x \in \Omega \\
u^{\epsilon} &= 0, \quad x \in \partial\Omega \\
A^{\epsilon}(x) &= A\left(\frac{x}{\epsilon}\right), \quad A : \mathbb{T}^d \to \mathbb{R}^{d \times d}
\end{aligned}
$$

Homogenization

Seek $u^{\epsilon} = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \ldots$ and determine A_0 such that

$$
\begin{aligned}
-\nabla \cdot (A_0 \nabla u_0) &= f, \quad x \in \Omega \\
u_0 &= 0, \quad x \in \partial\Omega
\end{aligned}
$$

Operator Learning

Constitutive Model Bensoussan, Lions, Papanicolaou '78 [1], Pavliotis and AMS '08 [12][Ch12]

A_0 determined by $\chi : \mathbb{T}^d \to \mathbb{R}^d$

$$-\nabla_y \cdot (\nabla_y \chi A^T) = \nabla_y \cdot A^T, \quad y \in \mathbb{T}^d,$$

$$A_0 = \int_{\mathbb{T}^d} \left(A(y) + A(y) \nabla \chi(y)^T \right) dy$$

Goal: Supervised Learning (NEU-NET)

Learn map $F : A(\cdot) \to A_0$ from function on torus to coefficient tensor:

$$A_0 = F(A)$$

such that

$$-\nabla \cdot (F(A) \nabla u_0) = f, \quad x \in \Omega,$$

$$u_0 = f, \quad x \in \partial\Omega$$

Talk Outline

Operator Learning

Learning Constitutive Models

Elasticity

Viscoelasticity

Plasticity

Conclusions

Multiscale Problem I Francfort and Suquet '86 [4]

Viscoelasticity Multiscale Problem

$$
\begin{aligned}
-\nabla \cdot (\sigma^\epsilon) &= f, \quad x \in \Omega \\
u^\epsilon &= 0, \quad x \in \partial\Omega \\
u^\epsilon|_{t=0} &= u_i, \quad x \in \Omega \\
\sigma^\epsilon &= \nu^\epsilon \partial_t \nabla u^\epsilon + E^\epsilon \nabla u^\epsilon \\
E^\epsilon(x) &= E\left(\frac{x}{\epsilon}\right), \quad E : \mathbb{T}^d \to \mathbb{R}, \quad \nu^\epsilon(x) = \nu\left(\frac{x}{\epsilon}\right), \quad \nu : \mathbb{T}^d \to \mathbb{R}.
\end{aligned}
$$

Laplace Transform

Let $A^\epsilon = s\nu^\epsilon + E^\epsilon$, then

$$
\begin{aligned}
-\nabla \cdot \left(A^\epsilon \left(\nabla \widehat{u^\epsilon}\right)\right) &= \widehat{f}, \quad x \in \Omega, \\
u^\epsilon &= 0, \quad x \in \partial\Omega.
\end{aligned}
$$

Multiscale Problem II Bhattacharya, Liu, AMS, Trautner '22 [3]

Theorem (Piecewise-Constant Approximation)

Approximating E^ϵ, ν^ϵ in $L^\infty(\Omega)$ by piecewise constants $E^\epsilon_{PC}, \nu^\epsilon_{PC}$ gives

$$\sup_{0\le t\le T} \|u^\epsilon - u^\epsilon_{PC}\|_{H^1_0} \le C\Big(\|E - E_{PC}\|_{L^\infty} + \|\nu - \nu_{PC}\|_{L^\infty}\Big).$$

Theorem (Piecewise-Constant Homogenization)

In piecewise-constant case homogenized equation for u_0 is Markovian:

$$\begin{aligned}
-\nabla\cdot(\sigma) &= f, \quad x\in\Omega,\\
u_0 &= 0, \quad x\in\partial\Omega\\
u_0\big|_{t=0} &= u_i, \quad x\in\Omega\\
\sigma &= \nu'\partial_t\nabla u_0 + E'\nabla u_0 + \langle \mathbb{1}, r\rangle\\
\partial_t r_\ell &= -\alpha_\ell r_\ell + \beta_\ell \nabla u_0, \quad \ell\in\{1,2,\cdots,L\},
\end{aligned}$$

for some choice of $E'\in\mathbb{R}_+$, $\nu'\in\mathbb{R}_+$, $\alpha\in\mathbb{R}^L_+$, $\beta\in\mathbb{R}^L$, $L\in\mathbb{Z}_+$.

Operator Learning

True Solution Map

Let $\Psi : \{\nabla u_0(\tau)\}_{\tau=0}^{t} \to \sigma|_{\tau=t}$ be the map such that the homogenized constitutive relation is

$$\sigma = \Psi(\{\nabla u_0\}).$$

Goal: Supervised Learning (RNO-NET)

Learn map $\Psi_{RNO} : \{\nabla u_0(\tau)\}_{\tau=0}^{t} \to \sigma|_{\tau=t}$ approximating Ψ with the form

$$\begin{aligned} \sigma &= F\left(\nabla u_0, \partial_t \nabla u_0, r\right) \\ \partial_t r &= G(r, \nabla u_0), \quad r(0) = 0. \end{aligned}$$

Effect of Resolution on Test Error

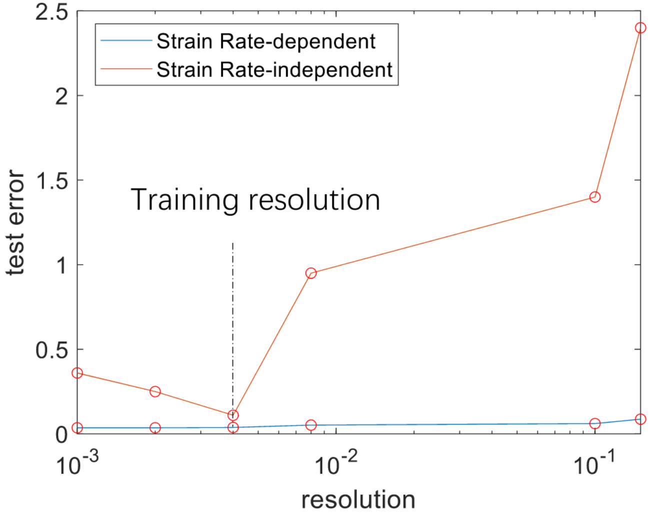

Figure: Viscoelasticity: Old architecture includes the strain rate as an input while new architecture does not include the strain rate

Effect of Different L on Test Error

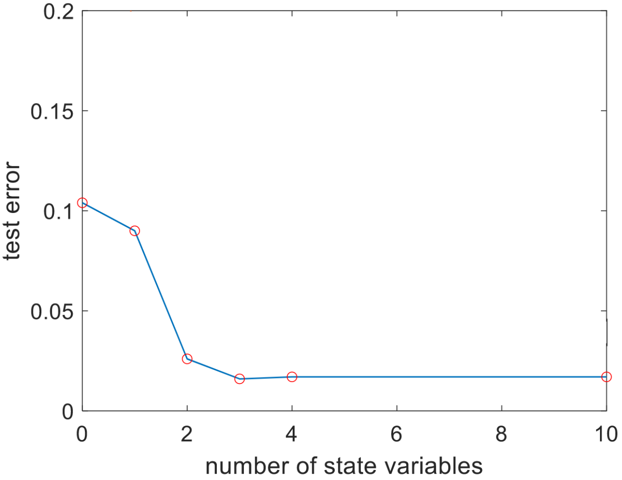

Figure: Viscoelasticity: Test error vs. L, the number of constant pieces

Talk Outline

Multiscale Problem 1 Liu, Kovachki, Li, Azizzadenesheli, Anandkumar, AMS, Bhattacharya '22 [11]

Plasticity Multiscale Problem

$$
\begin{aligned}
\rho\,\partial_t^2 u^\epsilon &= \nabla \cdot \sigma^\epsilon + f, \qquad x \in \Omega \\
\partial_t \xi^\epsilon &= K(\xi^\epsilon, \nabla u^\epsilon), \qquad x \in \Omega \\
\sigma^\epsilon &= \Psi^\epsilon\left(\nabla u^\epsilon, \xi^\epsilon, \frac{x}{\epsilon}\right) \\
u^\epsilon|_{t=0} &= u_i, \quad \partial_t u^\epsilon|_{t=0} = \nu_i, \quad \xi^\epsilon|_{t=0} = \xi_i, \quad x \in \Omega \\
u^\epsilon &= u^* \quad x \in \partial_1\Omega, \quad \sigma^\epsilon n = s^* \quad x \in \partial_2\Omega
\end{aligned}
$$

Multiscale Problem 2

Homogenized Plasticity Problem

$$
\begin{aligned}
\rho\,\partial_t^2 u_0 &= \nabla \cdot \sigma_0 + f, \qquad x \in \Omega \\
\sigma_0 &= \Psi\Big(\{\nabla u_0\}\Big) \\
u_0|_{t=0} &= u_i, \quad \partial_t u_0|_{t=0} = \nu_i, \quad x \in \Omega \\
u_0 &= u^* \quad x \in \partial_1\Omega, \quad \sigma_0 n = s^* \quad x \in \partial_2\Omega
\end{aligned}
$$

Operator Learning

Goal: Supervised Learning (PCA-NET)

Learn map $\Psi_{PCA} : \{\nabla u_0(\tau)\}_{\tau=0}^{t} \to \{\sigma(\tau)\}_{\tau=0}^{t}$ approximating Ψ. In particular causality must be learned.

Goal: Supervised Learning (RNO-NET)

Learn map $\Psi_{RNO} : \{\nabla u_0(\tau)\}_{\tau=0}^{t} \to \sigma|_{\tau=t}$ approximating Ψ with the form

$$\sigma = F\left(\nabla u_0, \partial_t \nabla u_0, r\right)$$
$$\partial_t r = G(r, \nabla u_0), \quad r(0) = 0.$$

PCA-net: 2D FFT

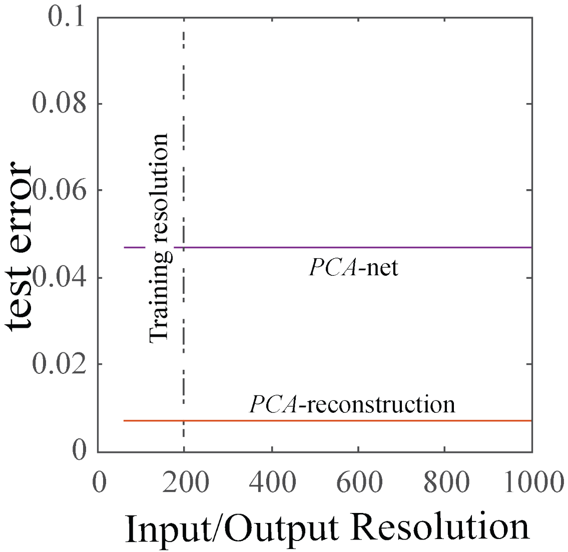

Figure: Crystal plasticity PCA-net

RNO-net 2D FFT

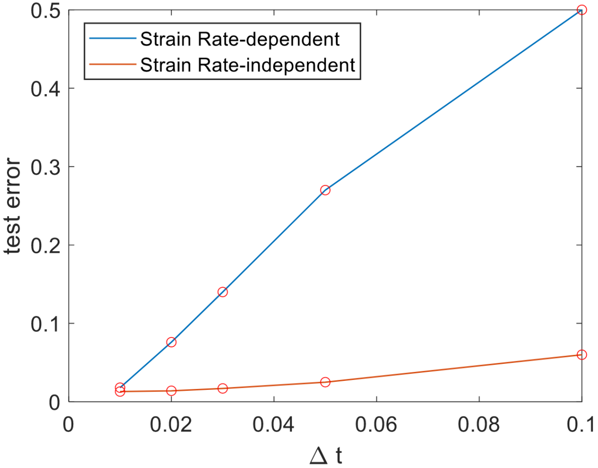

Figure: Crystal plasticity with $dt = 0.01$, 400 data

Talk Outline

Conclusions

- Supervised learning in Banach space valuable for:
  - Surrogate Modeling
  - Model Discovery
- Various homogenization problems can be framed this way
- Learning constitutive models can hence be framed this way
  - Elasticity (material properties to stress)
  - Viscoelasticity (history of strain to stress)
  - Plasticity (history of strain to stress)
- Mesh-independence useful to validate model-form hypothesis

References I

[1] A. Bensoussan, J.-L. Lions, and G. Papanicolaou.
Asymptotic analysis for periodic structures, volume 374.
American Mathematical Soc., 2011.

[2] K. Bhattacharya, B. Hosseini, N. B. Kovachki, and A. M. Stuart.
Model reduction and neural networks for parametric pdes.
SMAI Journal of Computational Mathematics, 7:121–157, 2021, arXiv:2005.03180.

[3] K. Bhattacharya, B. Liu, A. M. Stuart, and M. Trautner.
Learning Markovian homogenized models in viscoelasticity.
arXiv, 2022.

[4] G. A. Francfort and P. M. Suquet.
Homogenization and mechanical dissipation in thermoviscoelasticity.
Archive for Rational Mechanics and Analysis, 96(3):265–293, 1986.

[5] O. Gonzalez and A. M. Stuart.
A first course in continuum mechanics.
Cambridge University Press, 2008.

[6] Y. Korolev.
Two-layer neural networks with values in a Banach space.
arXiv preprint arXiv:2105.02095, 2021.

[7] N. Kovachki, S. Lanthaler, and S. Mishra.
On universal approximation and error bounds for fourier neural operators.
arXiv preprint arXiv:2107.07562, 2021.

[8] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar.
Neural operator: Learning maps between function spaces.
arXiv preprint arXiv:2108.08481, 2021.

References II

[9] S. Lanthaler, S. Mishra, and G. E. Karniadakis.
Error estimates for deeponets: A deep learning framework in infinite dimensions.
arXiv preprint arXiv:2102.09618, 2021.

[10] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar.
Neural operator: Graph kernel network for pdes.
arXiv preprint arXiv:2003.03485.

[11] B. Liu, N. Kovachki, Z. Li, K. Azizzadenesheli, A. Anandkumar, A. M. Stuart, and K. Bhattacharya.
A learning-based multiscale method and its application to inelastic impact problems.
Journal of the Mechanics and Physics of Solids, 158:104668, 2022.

[12] G. Pavliotis and A. Stuart.
Multiscale methods: averaging and homogenization.
Springer Science & Business Media, 2008.