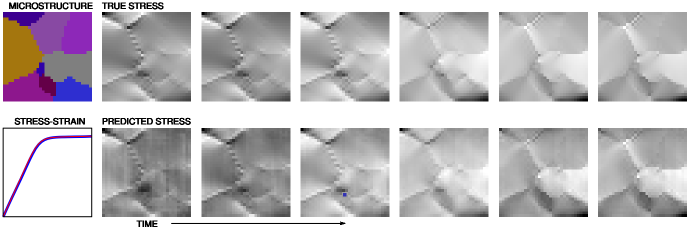
16th U.S. National Congress on Computational Mechanics (July 25-29, 2021)
308 Physics-Informed Learning and Data-Enabled Predictive Modeling and Discovery of Complex Systems
16:00–16:20 2021-07-26 Room 24-B-24

Machine Learning Constitutive Models

Reese E. Jones, A. Frankel

Sandia National Laboratories, Livermore, CA 94551, USA



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-NA-0003525.

boratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LL
Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE

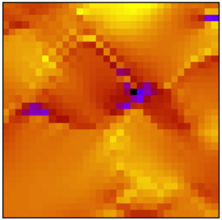
Overview

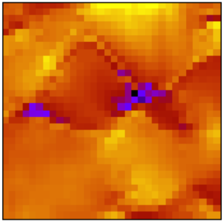
Goal: efficient accurate surrogates of material processes

Everyone is doing machine learning, it is easy and sometimes useful.

Anonymous - a paraphrase of George Box

Which one is the ML prediction?





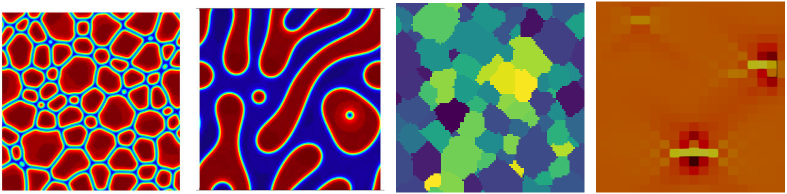
one is a "deep fake"

Outline

Please ask questions

Microstructural problems of interest

Premise: the state of each of these systems/processes can be encoded as an **image**/field with multiple **channels** $\phi(\mathsf{X})$.



bubbles multi-phase polycrystal pores/inclusions

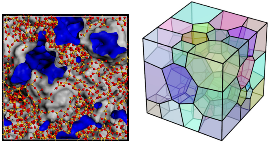
Classes of problems:

- **property estimation**: map initial image $\phi(\mathsf{X})$ to a static quantity ε, e.g. diffusivity
- **homogenization**: map initial image $\phi(\mathsf{X})$ and forcing $\epsilon(t)$ to evolving scalar quantity $\Psi(t)$, e.g. energy
- **field prediction**: map initial image $\phi(\mathsf{X})$ and forcing $\mathsf{f}(t)$ to an evolving field $\boldsymbol{\sigma}(\mathsf{X}, t)$, e.g. stress field

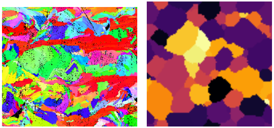
Applications: subgrid models, structure-property exploration/optimization, & material uncertainty quantification

Microstructural problems of interest: challenges

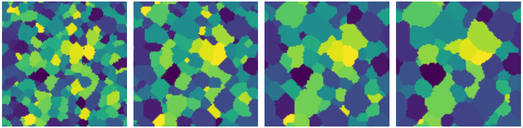
data representation: point clouds, graphs, grids, meshes



noise: real vs simulated

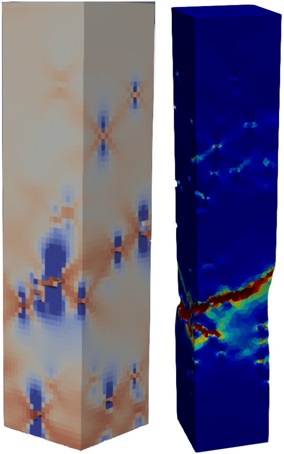


evolution: causal changes with time



… values and topology

regularity: sensitivity

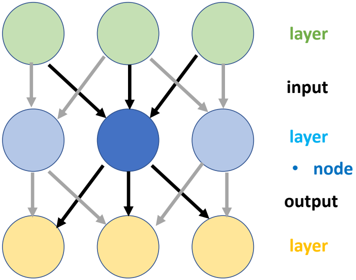


failure

Neural networks - basics

The simplest **neural network** (NN) is a multilayer perceptron (MLP), a directed graph of densely connected nodes organised in layers. Inputs are weighted, summed and transformed to outputs by *non-linear* ramp/switch-like activation functions.

$$y_j = f\underbrace{\left(\sum_i w_{ij} x_i + b_j\right)}_{\text{affine/linear}}$$



The parameters w, b are trained via backpropagation and stochastic descent i.e. regression.

A NN is basically a functional form to be fit. Like box of **LEGOS**™, layers with particular characteristics can be linked together to create a model.

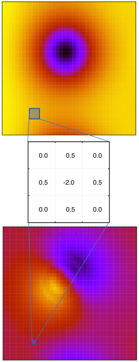
Deep learning: convolutional neural networks

All informative features may not be apparent. Given a microstructural image $\phi(X)$ is hard to see what determines the response $\Psi(t)$.

Convolution with a kernel is a standard technique in (time) signal and (spatial) image processing that has been adapted to ML.

For example, filters can:

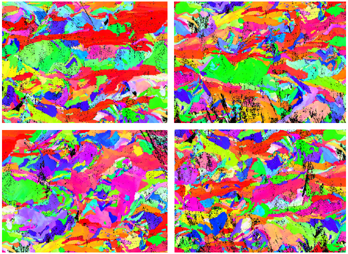
- **Smooth/filter noise:** convolving an image with a Gaussian kernel.
- **Average/coarsen:** multiplying with constant moving patch
- **Gradients and higher derivatives:** filter corresponding a finite difference stencil.
- **Features:** edge detection, clustering, segmentation, ...

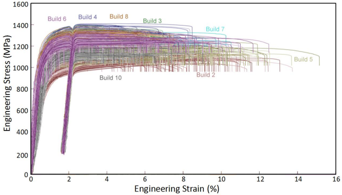


A convolutional NN trains the weights w_{ij} and bias b for a kernel (smaller than image MLP) and multiple filters can detect multiple hidden features.

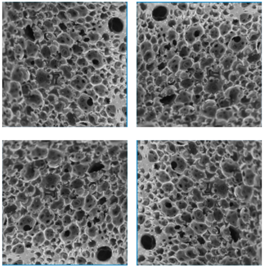
Exemplar microstructure problems

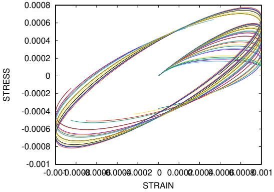
(A) polycrystal





(B) porosity



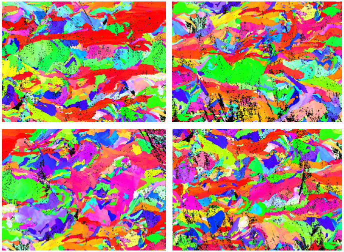


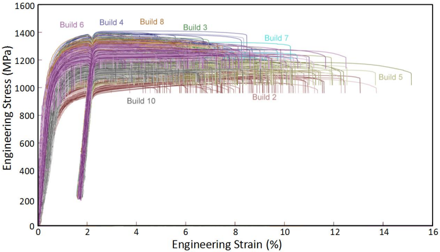
Which output is related to a particular input image?

Crystal plasticity

Predicting metal plasticity is still a hard problem. Polycrystalline plasticity is a good exemplar for other microstructural problems.

If we observe **initial** microstructures and mechanical tests:



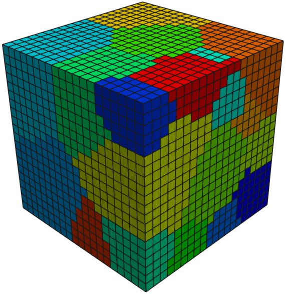


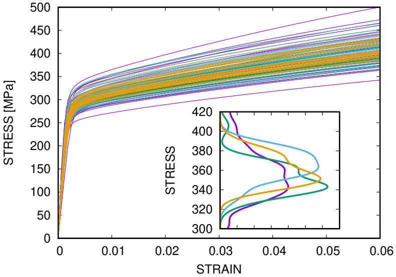
can we **predict** particular: (a) stress-strain averages or, (b) full field evolution of a polycrystalline material or a material with microstructure in general?

Training data

Even with high-throughput tests we cannot current generate more than $\approx 10^2$ tests, we need a dataset with $\approx 10^4$ samples for a NN.

So, we resort to high-fidelity simulation data ☹

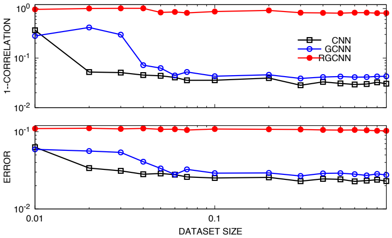




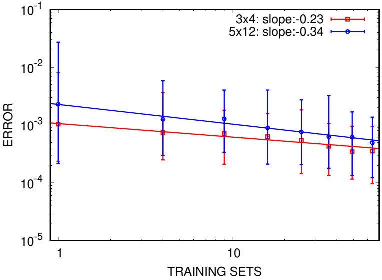
We generated realizations of oligocrystals with different textures (crystal orientations) and run crystal plasticity simulations with a variety of loadings.

Training data

Sampling loading modes, microstructures, etc to obtain sufficient data is expensive



Error vs (microstructure) sample size for a fixed size test set



Error vs random sampling of modes for homogeneous material

Steep descrease until number of samples ≈ number of parameters, then slow improvement

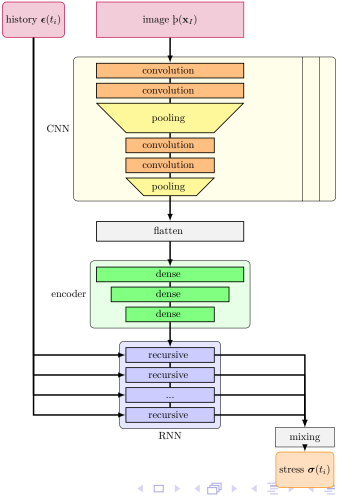
A hybrid RNN-CNN network for time evolution

To predict the evolution of the average stress we augment the output of a CNN that processes the initial microstructure with the loading/time dependence. The CNN output is only correlated with the observable through a RNN.

A recurrent NN (RNN) uses a causal time filter to process history information. An RNN for time is an analog to the CNN for spatial data.

How much to reduce the image to n_{features}?

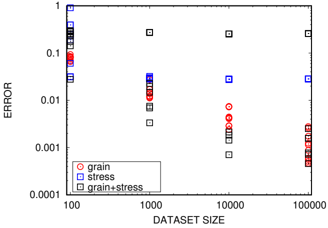
We are also exploring alternatives to RNN such as more traditional time integrators.



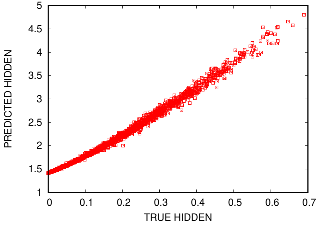
Predicting the response due to “hidden” features

Does the deep NN discover the hidden features?

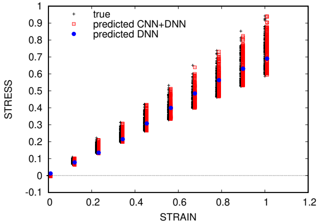
A **test problem** where we *know* what “hidden” microstructural features the observable stress depends on, e.g. average misorientation



Training NN on stress-strain alone stalls, but given initial **microstructures** continues to learn



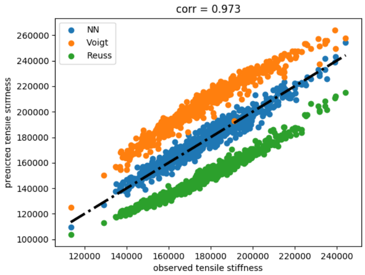
True and learned hidden feature are **highly correlated** - but not identical



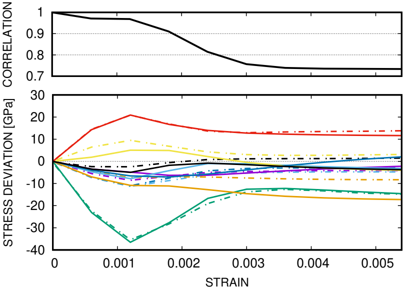
Observing the microstructure enables prediction of microstructure variations

Predicting the particular response to microstructure

Using data from the ensemble of polycrystals, we can make predictions of the crystal plastic mechanical response that are significantly better than traditional homogenization theory.



Correlation of elastic response (NN, Voigt and Reuss predictions), NN on par with Hill average.



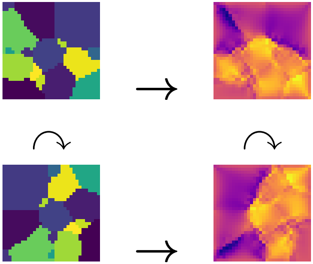
Trajectories of discrepancy from mean: solid lines data, dashed: NN prediction. Trajectories drift with accumulated error. Plastic response is better than Sachs or Taylor estimates.

Physical symmetries

Satisfaction of physical constraints and symmetries is expected in physical models and is necessary for conservation, stability, etc

How do we learn/impose physical constraints?

- **Augment** the dataset with many examples of what should happen, e.g. rotate the inputs and outputs (soft and inefficient)
- **Penalize** loss / training objective function (introduces a meta parameter and can be hard to formulate)
- **Embed** the symmetry in the NN architecture so that the response exactly preserves the symmetry (can be hard to formulate)



Objectivity and representation theory

We want to **embed symmetries** in the NN structure – so that they are exact/not learned. Let's go back to classical theory...

Material frame indifference for constitutive function M(A)

$$\mathsf{G}\mathsf{M}(\mathsf{A})\mathsf{G}^T = \mathsf{M}(\mathsf{G}\mathsf{A}\mathsf{G}^T) \ ,$$

for every member G of the orthogonal group.
Based on the spectral $\mathsf{A} = \sum_{i=1}^{3} \lambda_i \mathsf{a}_i \otimes \mathsf{a}_i$, and Cayley-Hamilton theorems

$$\mathsf{A}^3 - \operatorname{tr}(\mathsf{A})\mathsf{A}^2 + \frac{1}{2}\left(\operatorname{tr}^2 \mathsf{A} - \operatorname{tr} \mathsf{A}^2\right)\mathsf{A} - \det(\mathsf{A})\mathsf{I} = 0$$

one can obtain a compact **general representation**/model form:

$$\boxed{\mathsf{M}(\mathsf{A}) = c_0(\mathcal{I})\mathsf{I} + c_1(\mathcal{I})\mathsf{A} + c_2(\mathcal{I})\mathsf{A}^2 = \sum_i c_i(\mathcal{I})\mathsf{A}^i}$$

in form of **unknown** coefficient functions of invariants and a **known** tensor basis. Inputs: scalar invariants $\mathcal{I}$ & tensor basis $\mathcal{B}$.

A tensor basis neural network

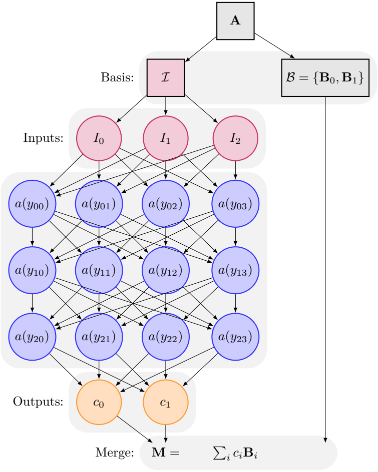
A *tensor basis* neural network is an NN implementation of this representation [Ling *JCP* 2016],

where the coefficients are unknown scalar functions of the invariants $\mathcal{I} = \{I_0, I_1, \ldots\}$

$$\mathrm{M} = \sum_i c_i(\mathcal{I})\mathrm{B}_i$$

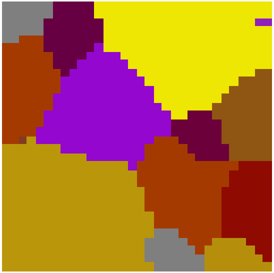
and a final merge/sum layer associates c_i with the **tensor basis** $\mathcal{B} = \{\mathrm{A}^0, \mathrm{A}^1, \ldots\}$.

It is adept at representing the response with exact invariance / avoiding the need for data augmentation for symmetry.



Graph-based convolutional neural networks

CNNs work great for structured grid/rastered image but the **need interpolation** for mesh-based fields and do not inherently **satisfy invariance** $\mathsf{G}\boldsymbol{\sigma}(\epsilon, \phi)\mathsf{G}^T = \boldsymbol{\sigma}(\mathsf{G}\epsilon\mathsf{G}^T, \mathsf{G}\phi\mathsf{G}^T)$ where ϕ is the initial microstructure. E.g. a polycrystal has obvious segmentation that leads directly to graph.



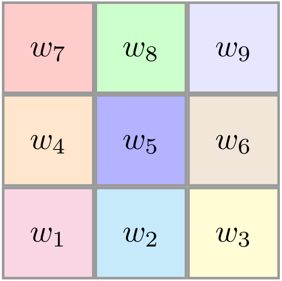
Reducing the grains to nodes and shared interfaces to edges has been shown effective [Vlassis *CMAME* 2020]

However this approach loses information (eg the details of the grain and interface geometry) and hence **requires featurization**.

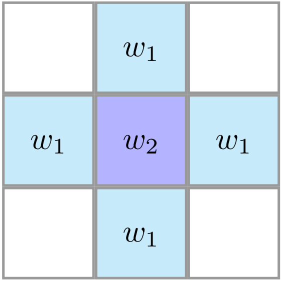
We propose applying the graph convolutions directly the mesh topology. This approach does not require featurization but can benefit from it. It does not increase the number of parameters since the same kernels are being employed.

Graph-based convolutional neural networks

Graph based convolution layers/filters [Kipf & Welling 2016] can be applied directly to the graph based on the mesh topology : elements are graph nodes and shared faces are graph edges.



CNN filter

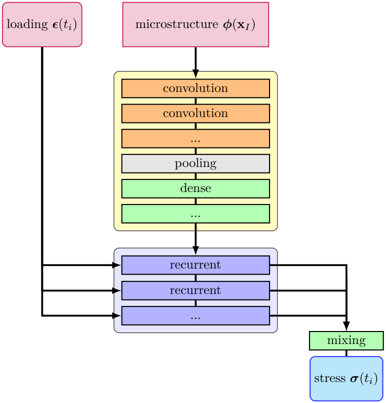


GCNN filter

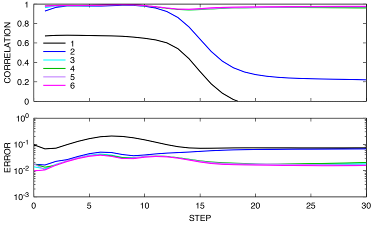
The GCNN filter uses the same weights for all the neighbors, hence it produces the same output when the image is rotated.

Graph-based convolutional neural networks

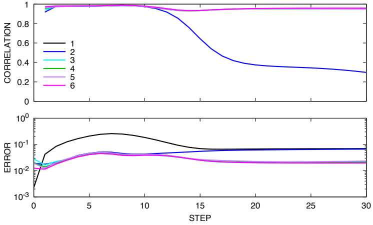
GCNNs have similar performance to CNN with fewer parameters and inherent invariance.



Convergence with number of filters



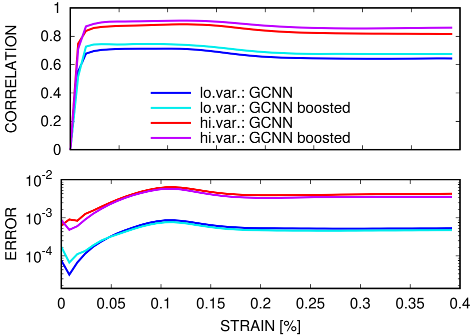
CNN



GCNN

Graph-based convolutional neural networks

GCNNs (and CNNs) can be boosted by embedding obvious features into the image (or further down the RNN-CNN pipeline)



adding node volumes to image of orientation angles

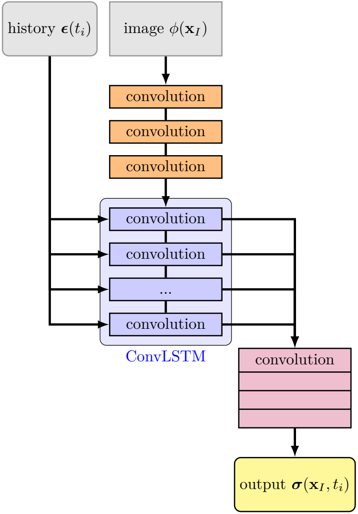
The improvement is marginal but distinct for a NN that is already fairly accurate.

Full field predictions

An architecture similar to the one we used to predict system-level evolution can be used to predict full-field (element/pixel level) evolution.

Inputs: pairs of

- image of initial microstructure
- system level strain history



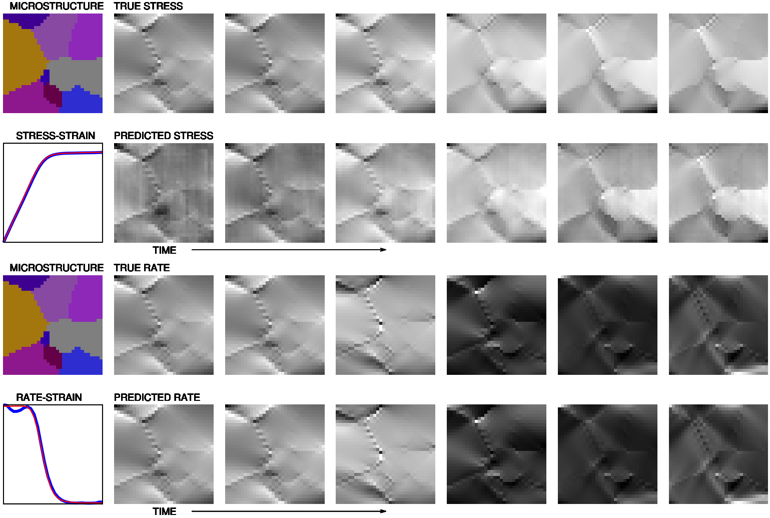
The image is fed to a *convolutional* neural network to process its latent features but not reduce them to a list of scalars. Perhaps the optimal number of filters is related to the dimensionality of the latent space of the microstructure in this context.

This is combined with the strain history in a *recurrent* neural network, specifically a convLSTM, to produce the output.

Output: full field stress evolution

Full field predictions

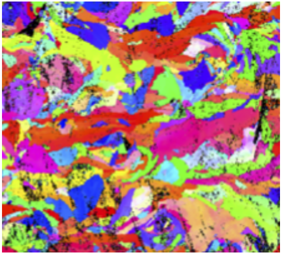
A convLSTM combines the RNN (time) and CNN (space) into PDE-like model



Conclusion

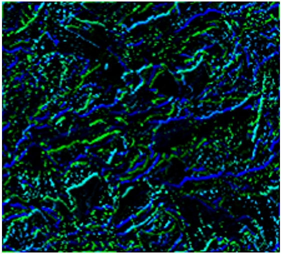
Applications:

- subgrid / multiscale surrogate models
- structure-property exploration / material optimization
- material uncertainty quantification



Open/current issues:

- meta parameter / architecture optimization
- interpretability (latent space / low dimensional manifold)
- training burden / multifidelity (experimental+simulation) data



rjones@sandia.gov

References: rjones@sandia.gov

- J. Ling, R.E. Jones, J.A. Templeton. *Machine learning strategies for systems with invariance properties.* Journal of Computational Physics (2016).
- R.E. Jones, J.A. Templeton, C.M. Sanders, J.T. Ostien. *Machine learning models of plastic flow based on representation theory.* Computer Modeling in Engineering & Sciences, (2018)
- A.L. Frankel, R.E. Jones, C. Alleman, and J.A. Templeton. *Predicting the mechanical response of oligocrystals with deep learning.* Computational Materials Science (2019).
- A.L. Frankel, K. Tachida, R.E. Jones. *Prediction of the evolution of the stress field of polycrystals undergoing elastic-plastic deformation with a hybrid neural network model.* Machine Learning: Science and Technology, (2020).
- A.L. Frankel, R.E. Jones, L. Swiler. *Tensor Basis Gaussian Process Models of Hyperelastic Materials.* Journal of Machine Learning for Modeling and Computing (2020).
- A.L. Frankel, C. Safta, C. Alleman, R.E. Jones, *Mesh-based graph convolutional neural network models of processes with complex initial states* arXiv, (2021)