LSTC
Livermore Software
Technology Corp.

USNCCM15 Short Course:
Machine Learning Data-driven Discretization Theories, Modeling and Applications

Transfer learning in structure-property predictions

Zeliang Liu, Ph.D., Research Scientist
Livermore Software Technology Corporation (LSTC)
zlliu@lstc.com

2:40~3:10 PM, July 28, 2019

15th U.S. National Congress on Computational Mechanics
Austin, TX, USA

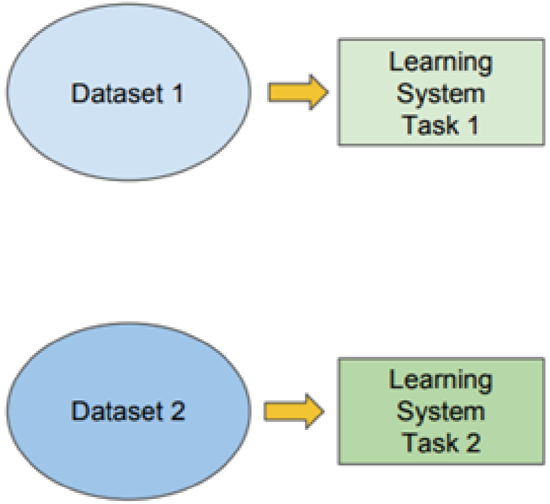
This class

- **Motivations:** Unified set of databases covering the full-range structure-property relationship, improvement of training convergence
- **Transfer-learning strategy:** Non-random initialization, database extrapolation and interpolation
- **Micromechanics:** Continuous structure-property relationship
- **A materials design problem:** Design of material toughness and ultimate tensile strength
- **Summary and future work**

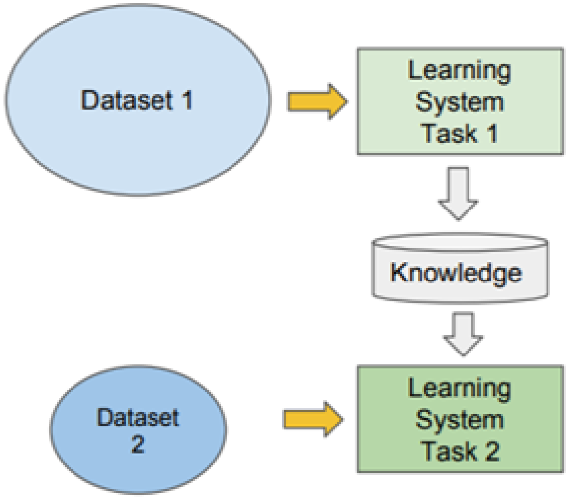
General idea of transfer learning

Traditional ML vs Transfer Learning

- Isolated, single task learning:
  - Knowledge is not retained or accumulated. Learning is performed w.o. considering past learned knowledge in other tasks

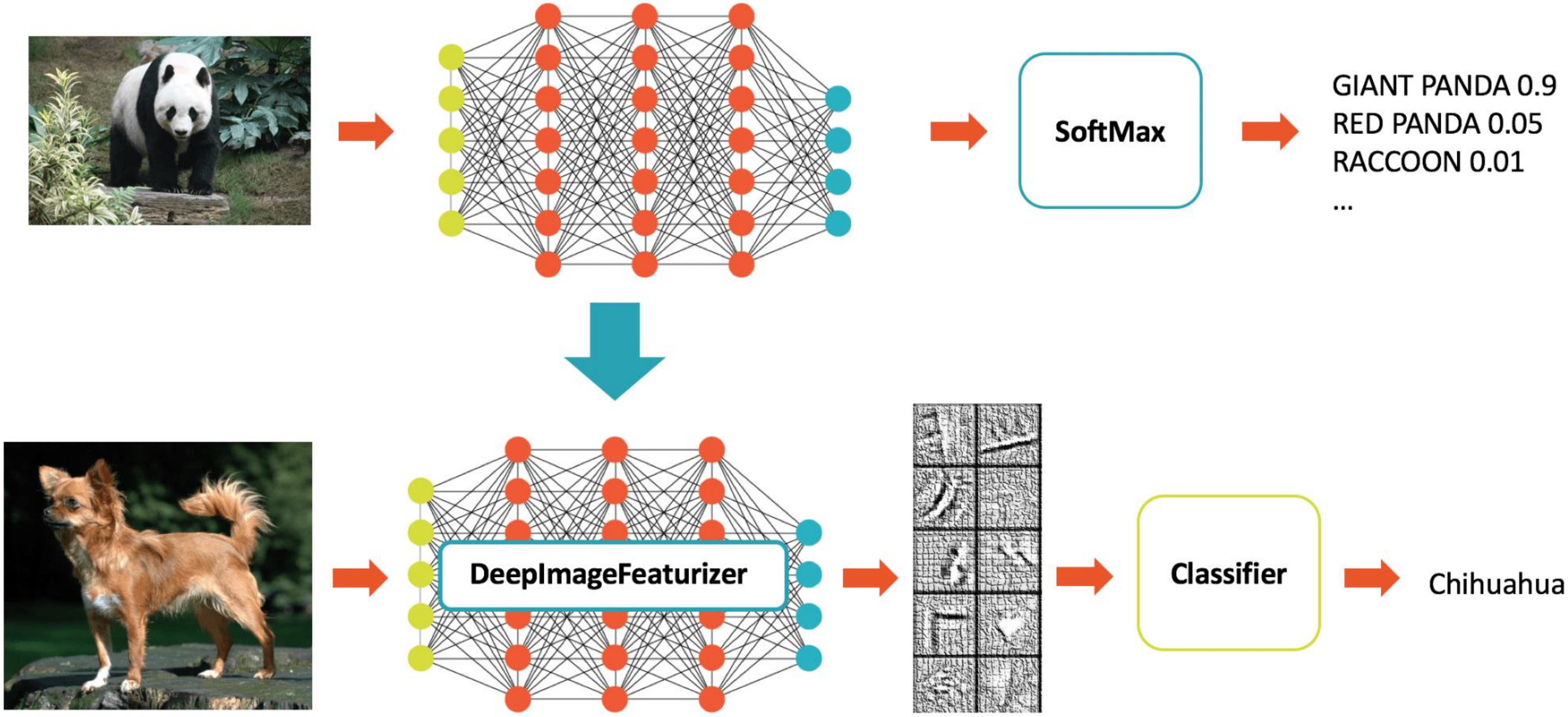


- Learning of a new tasks relies on the previous learned tasks:
  - Learning process can be faster, more accurate and/or need less training data



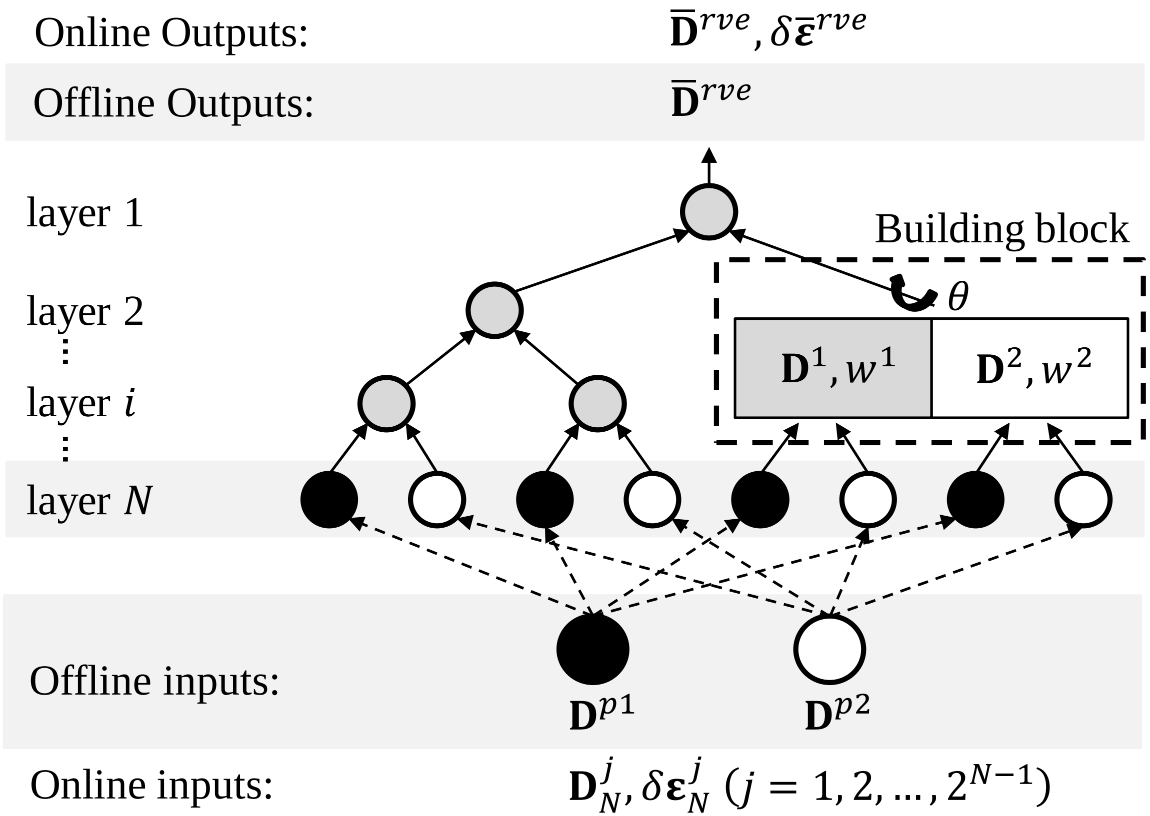
[Source: towardsdatascience.com]

Transfer learning in image classification



Source: databricks.com

Deep material network: Revisit



The fitting parameters are randomly initialized previously:

$$z_N^{j\,(0)} \sim U(0.2, 0.8) \quad \text{and} \quad \theta_i^{k(0)} \sim U(-\pi/2, \pi/2).$$

Z. Liu, C.T. Wu, M. Koishi. CMAME 345 (2019): 1138-1168.

Particle-reinforced RVEs with various volume fractions

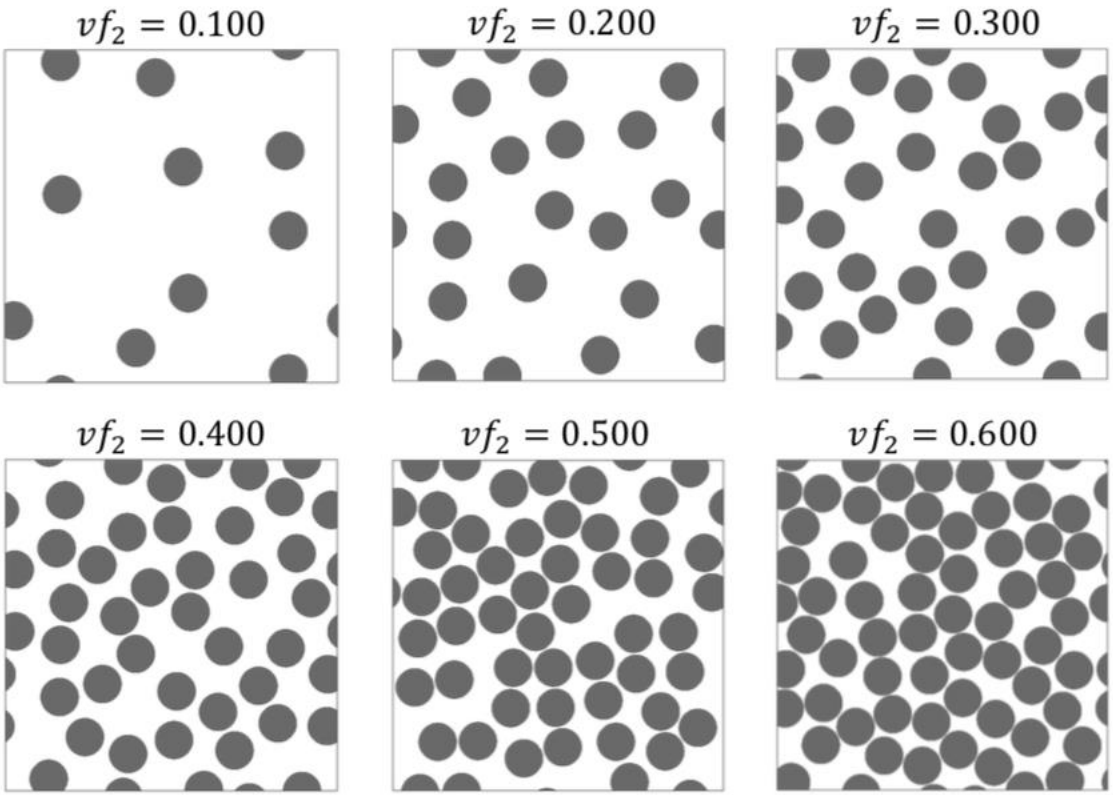
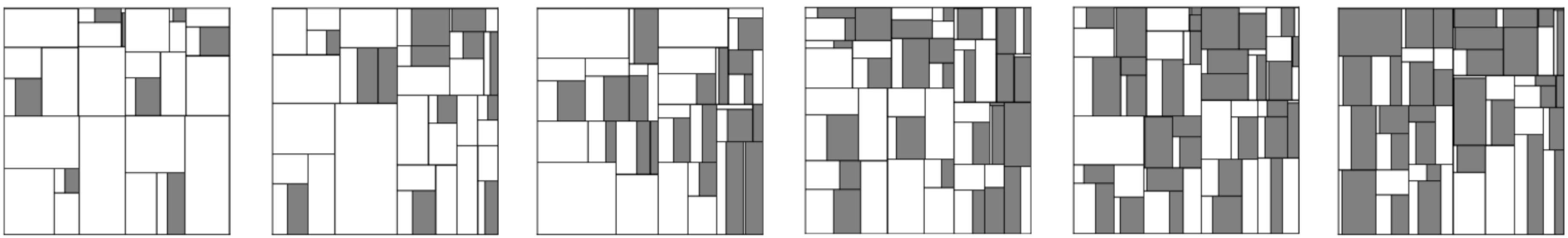


Fig. 2 DNS microstructures of particle-reinforced 2D RVEs with the volume fraction of particle phase vf_2 ranging from 0.1 to 0.6.

Motivation: Unify the databases of different RVEs

Random initialization of fitting parameters



$vf_2 = 0.096, N_a = 35$ | $vf_2 = 0.199, N_a = 44$ | $vf_2 = 0.301, N_a = 50$ | $vf_2 = 0.400, N_a = 69$ | $vf_2 = 0.502, N_a = 75$ | $vf_2 = 0.605, N_a = 72$

(c) Treemaps of DMN after 10000 epochs of training.

- With random initialization, the DMN databases trained for different RVEs are not analogous to each other in terms of the topological structure.
- Continuous migration between different database can not be realized through direct interpolation of the fitting parameters

Transfer learning of deep material network

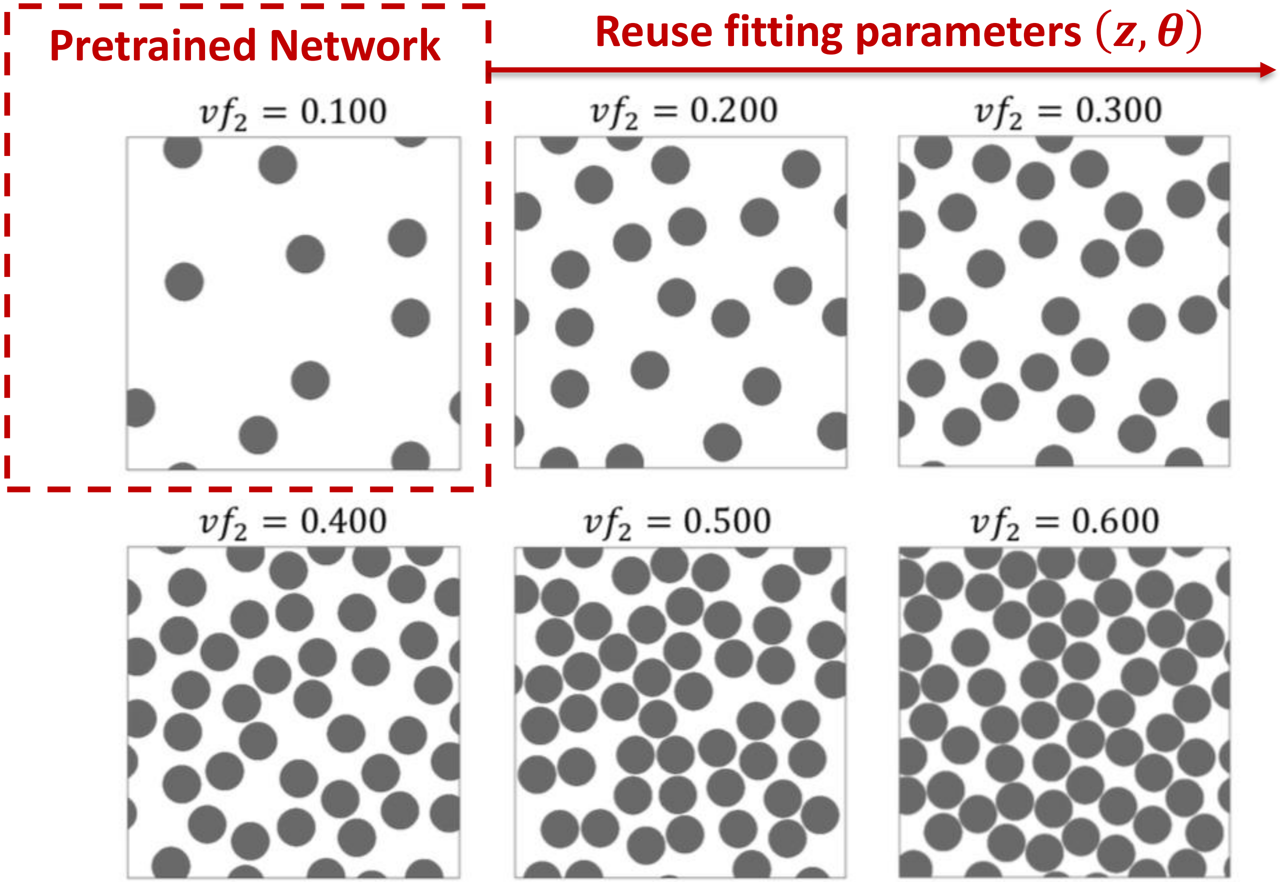
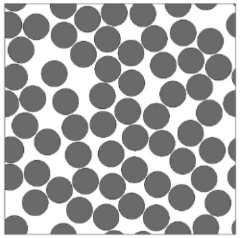


Fig. 2 DNS microstructures of particle-reinforced 2D RVEs with the volume fraction of particle phase vf_2 ranging from 0.1 to 0.6.

Z. Liu, C.T. Wu, M. Koishi. Computational Mechanics (2019)

Offline generation of training and test datasets

For each RVE: 

Generating orthotropic elastic samples:

- Inputs: $\mathbf{D}^{p1}, \mathbf{D}^{p2}$
- Outputs: $\overline{\mathbf{D}}^{dns}$

$$\mathbf{D}^{p1} = \begin{Bmatrix} 1/E_{11}^{p1} & -\nu_{12}^{p1}/E_{22}^{p1} & \\ & 1/E_{22}^{p1} & \\ & & 1/(2G_{12}^{p1}) \end{Bmatrix} \qquad \mathbf{D}^{p2} = \begin{Bmatrix} 1/E_{11}^{p2} & -\nu_{12}^{p2}/E_{22}^{p2} & \\ & 1/E_{22}^{p2} & \\ & & 1/(2G_{12}^{p2}) \end{Bmatrix}$$

Design of Experiments (DoE): 400 training / 100 test samples

Phase contrast: $E_{11}^{p1} E_{22}^{p1} = 1, \quad \log_{10}(E_{11}^{p2} E_{22}^{p2}) \in U[-6, 6]$

Anisotropy: $\log_{10}(E_{22}^{p1}/E_{11}^{p1}) \in U[-1, 1], \quad \log_{10}(E_{22}^{p2}/E_{11}^{p2}) \in U[-1, 1]$

Shear moduli: $\dfrac{G_{12}^{p1}}{\sqrt{E_{22}^{p1} E_{11}^{p1}}} \in U[0.25, 0.5], \quad \dfrac{G_{12}^{p2}}{\sqrt{E_{22}^{p2} E_{11}^{p2}}} \in U[0.25, 0.5]$

A naïve approach: Update the activations without training

Extrapolation of fitting parameters: Update the activations of the base network to match the target volume fraction, and all angles remains unaltered.

$$z^{j(t)} = \begin{cases} \dfrac{1 - vf_2^{(t)}}{1 - vf_2^{(b)}} a(z^{j(b)}) & \text{if } j \text{ is odd} \\ \dfrac{vf_2^{(t)}}{vf_2^{(b)}} a(z^{j(b)}) & \text{if } j \text{ is even,} \end{cases}$$

$$\theta_i^{k(t)} = \theta_i^{k(b)}$$

(t): Target; (b): Base;

a(x): ReLU activation function

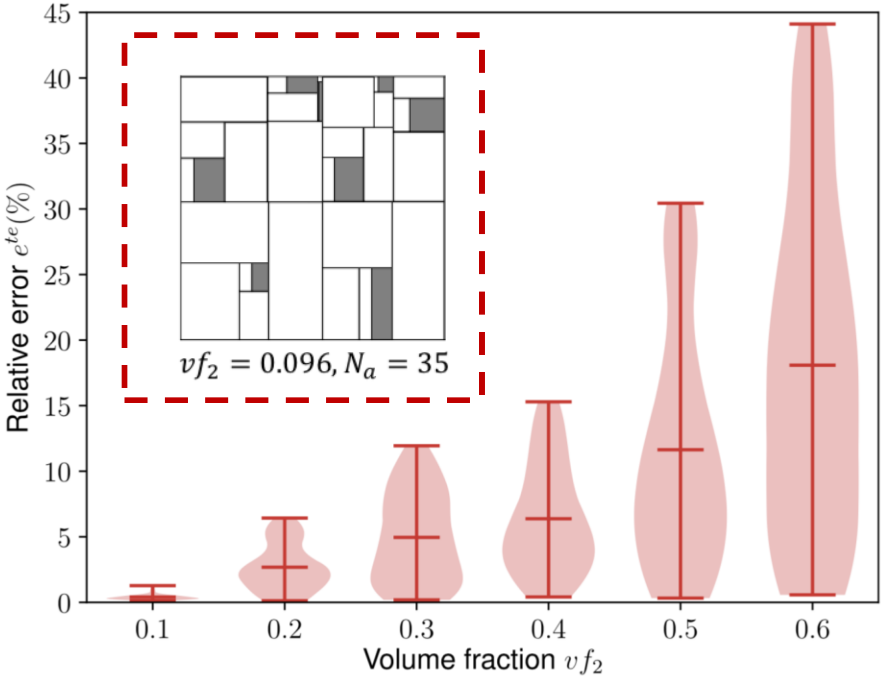
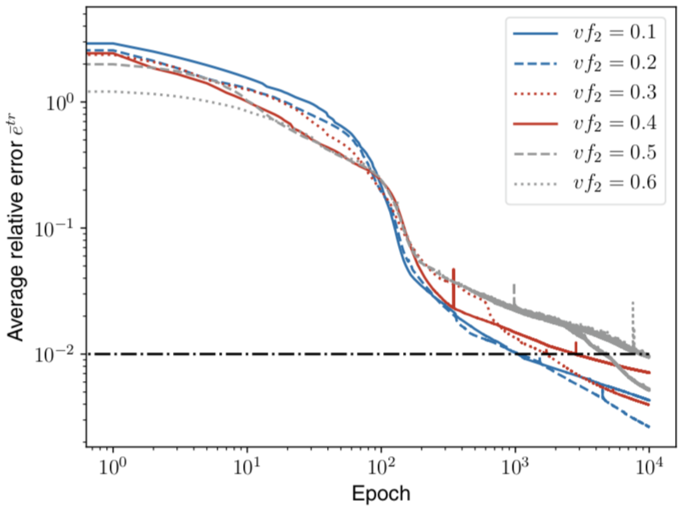


Fig. 4 Distributions of errors on the test datasets from the naive approach. The network for $vf_2 = 0.1$ is chosen as the base of migration

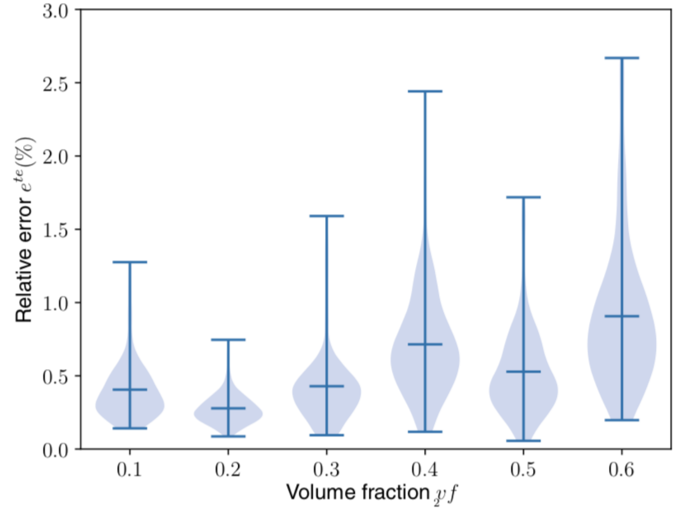
The naive approach is adopted mainly for the initialization of transfer learning.

Comparison: Training and test errors

Random initialization

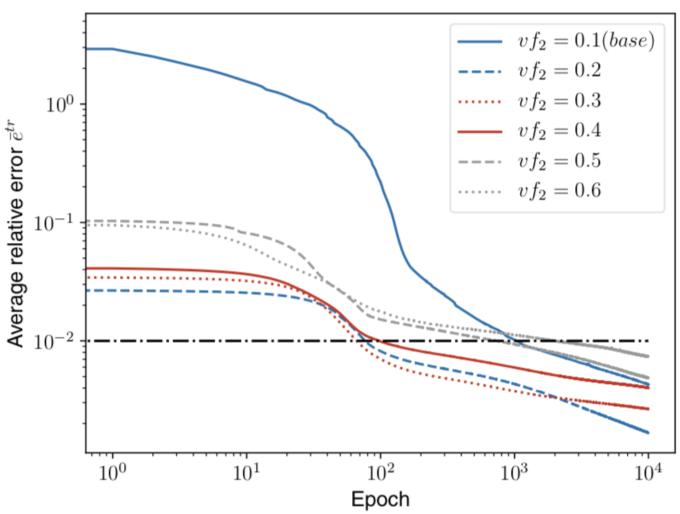


(a) Histories of average training error $\bar{e}^{tr}$.

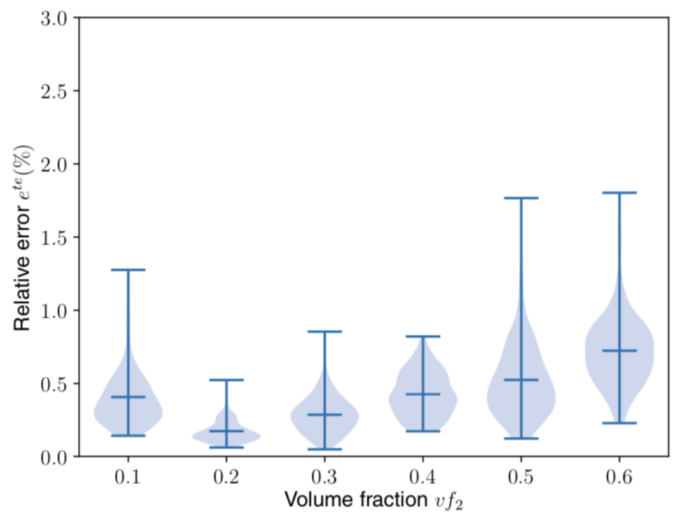


(b) Distributions of test error e^{te}.

Transfer learning



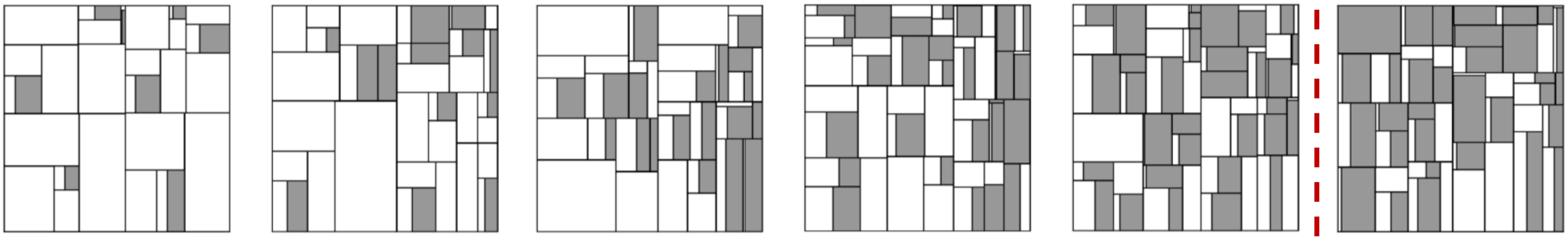
(a) Histories of average training error $\bar{e}^{tr}$.



(b) Distributions of test error e^{te}.

Comparison: Topology structure of DMN

Random initialization

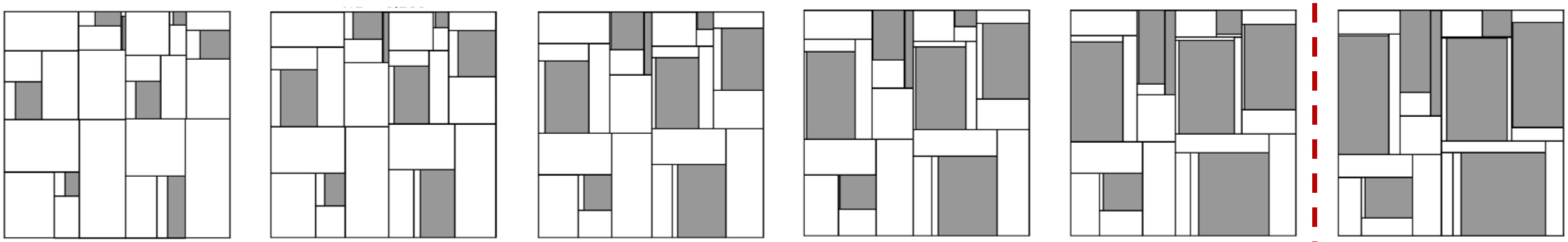


$vf_2 = 0.096, N_a = 35$ | $vf_2 = 0.199, N_a = 44$ | $vf_2 = 0.301, N_a = 50$ | $vf_2 = 0.400, N_a = 69$ | $vf_2 = 0.502, N_a = 75$ | $vf_2 = 0.605, N_a = 72$

(c) Treemaps of DMN after 10000 epochs of training.

- Discontinuous topology structure
- Redundancy in the network

Transfer learning

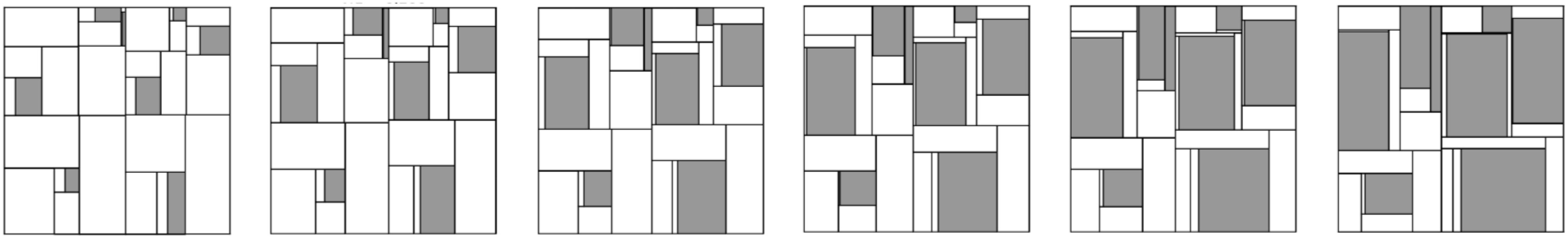


$vf_2 = 0.096, N_a = 35$ | $vf_2 = 0.200, N_a = 34$ | $vf_2 = 0.302, N_a = 34$ | $vf_2 = 0.402, N_a = 34$ | $vf_2 = 0.502, N_a = 31$ | $vf_2 = 0.609, N_a = 31$

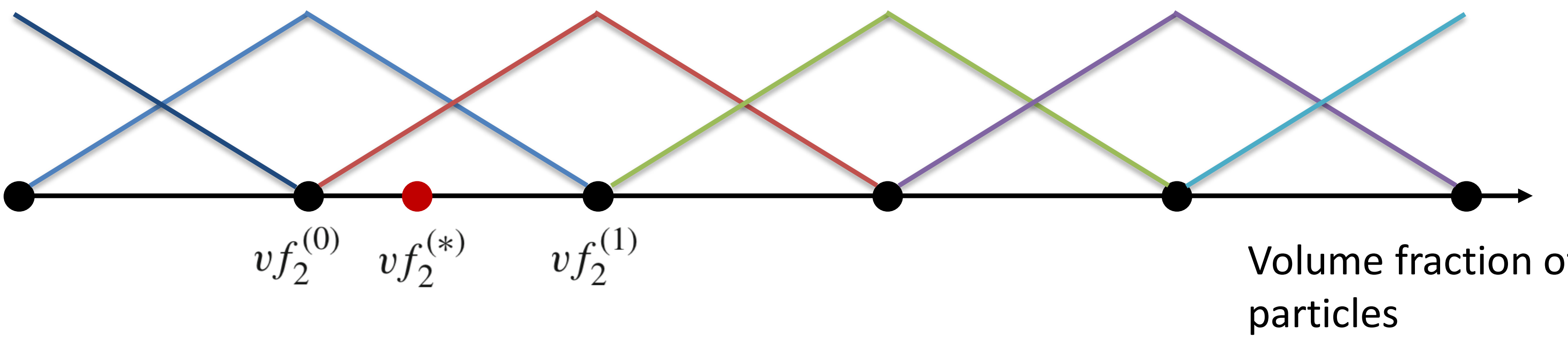
(c) Treemaps of DMN after 10000 epochs of training.

- Analogous topology structure, enabling interpolation of fitting parameters
- More compressed network, less number of DOFs

Interpolation of Database from Transfer Learning



$vf_2 = 0.096, N_a = 35$ $vf_2 = 0.200, N_a = 34$ $vf_2 = 0.302, N_a = 34$ $vf_2 = 0.402, N_a = 34$ $vf_2 = 0.502, N_a = 31$ $vf_2 = 0.609, N_a = 31$



The parameters for the interpolated database at

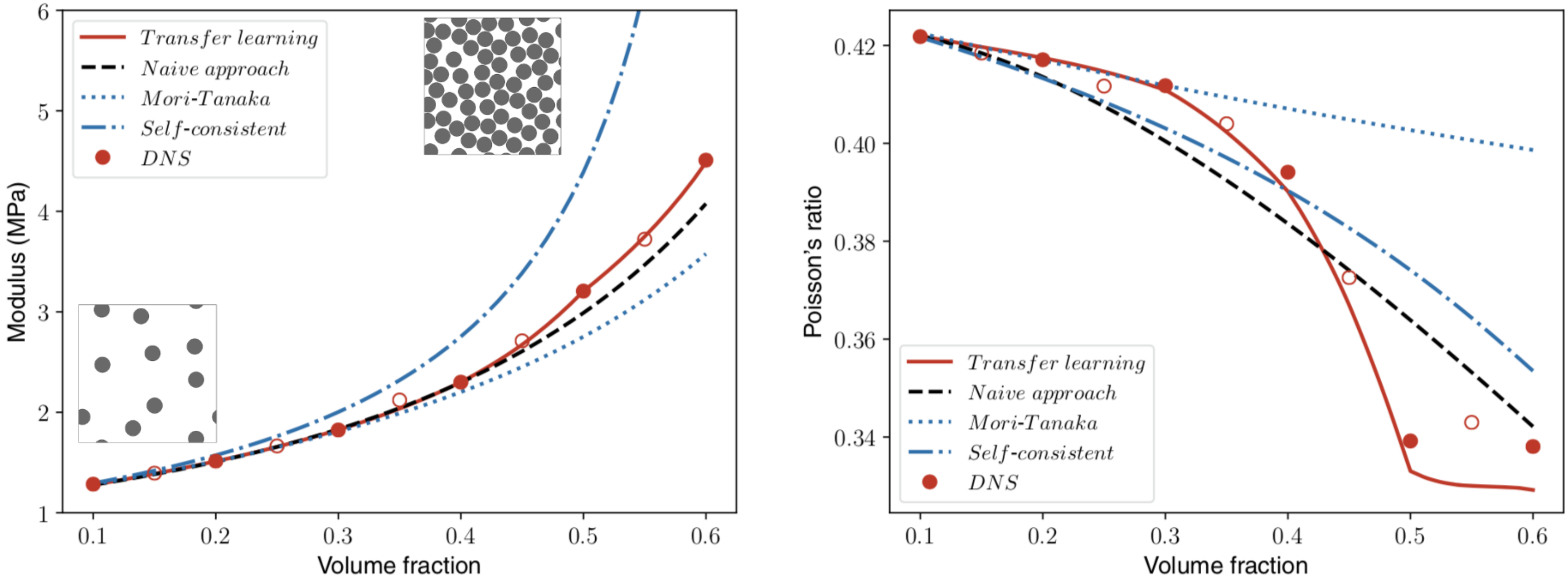
$$\tilde{z}^{j(*)} = N_0 a(\tilde{z}^{j(0)}) + (1 - N_0) a(\tilde{z}^{j(1)}) \qquad \theta_i^{k(*)} = N_0 \theta_i^{k(0)} + (1 - N_0) \theta_i^{k(1)}$$

The shape function is given by $$N_0 = \frac{vf_2^{(1)} - vf_2^{(*)}}{vf_2^{(1)} - vf_2^{(0)}}$$

Micromechanics I: Continuous structure-property relationship

Matrix (phase 1): $E_1 = 1\ \text{MPa}, \quad \nu_1 = 0.3.$

Hard particle (phase 2): $E_2 = 1000\ \text{MPa}, \quad \nu_2 = 0.3.$

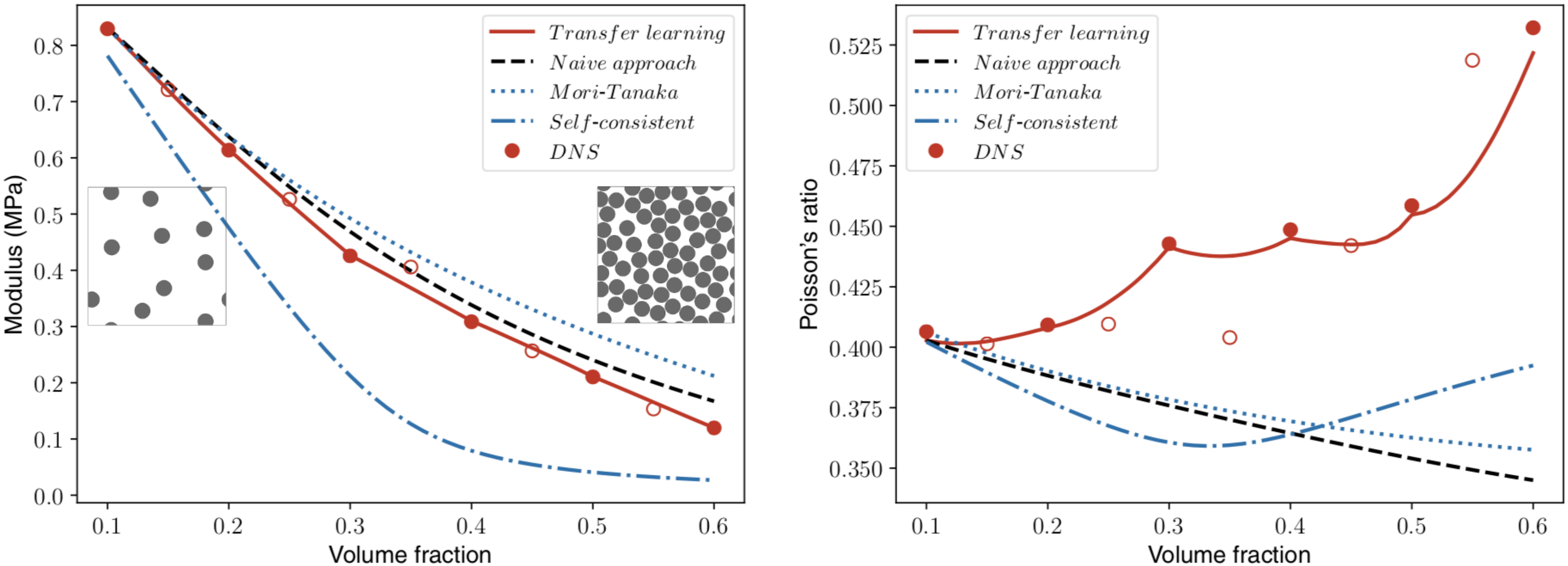


(a) Transverse Young's modulus. **(b)** Transverse Poisson's ratio.

Micromechanics II: Continuous structure-property relationship

Matrix (phase 1): $E_1 = 1\ \text{MPa}, \quad \nu_1 = 0.3$

Soft particle (phase 2): $E_2 = 0.01\ \text{MPa}, \quad \nu_2 = 0.3$



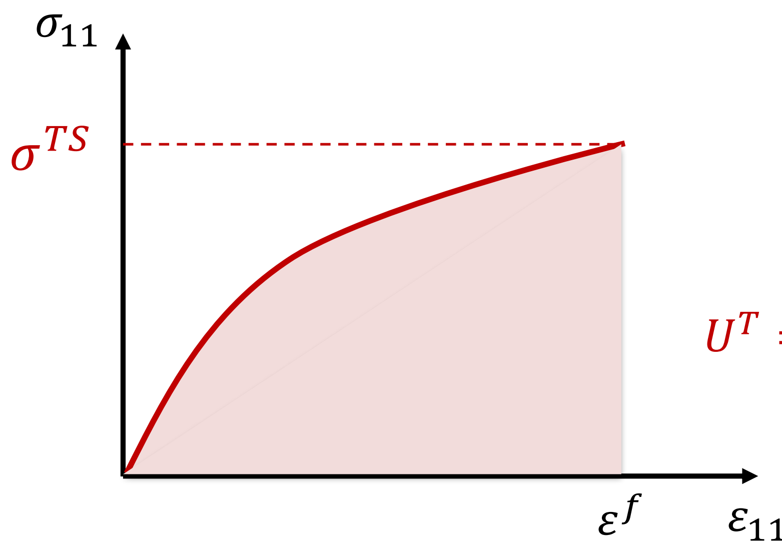
(a) Transverse Young's modulus. **(b)** Transverse Poisson's ratio.

A materials design problem

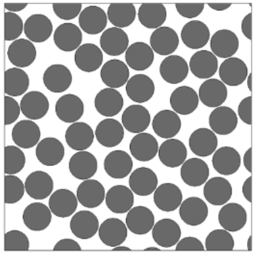
Design parameter: Young's modulus of particle E_2, volume fraction of particles vf_2

Objectives: Ultimate tensile strength σ^{TS}, material toughness U^T

Other settings: The matrix is assumed to be elasto-plastic with constant properties



$$U^T = \int_0^{\varepsilon^f} \sigma_{11} d\varepsilon_{11}$$



Definition of failure: The composite RVE fails if any of the two conditions is meet:

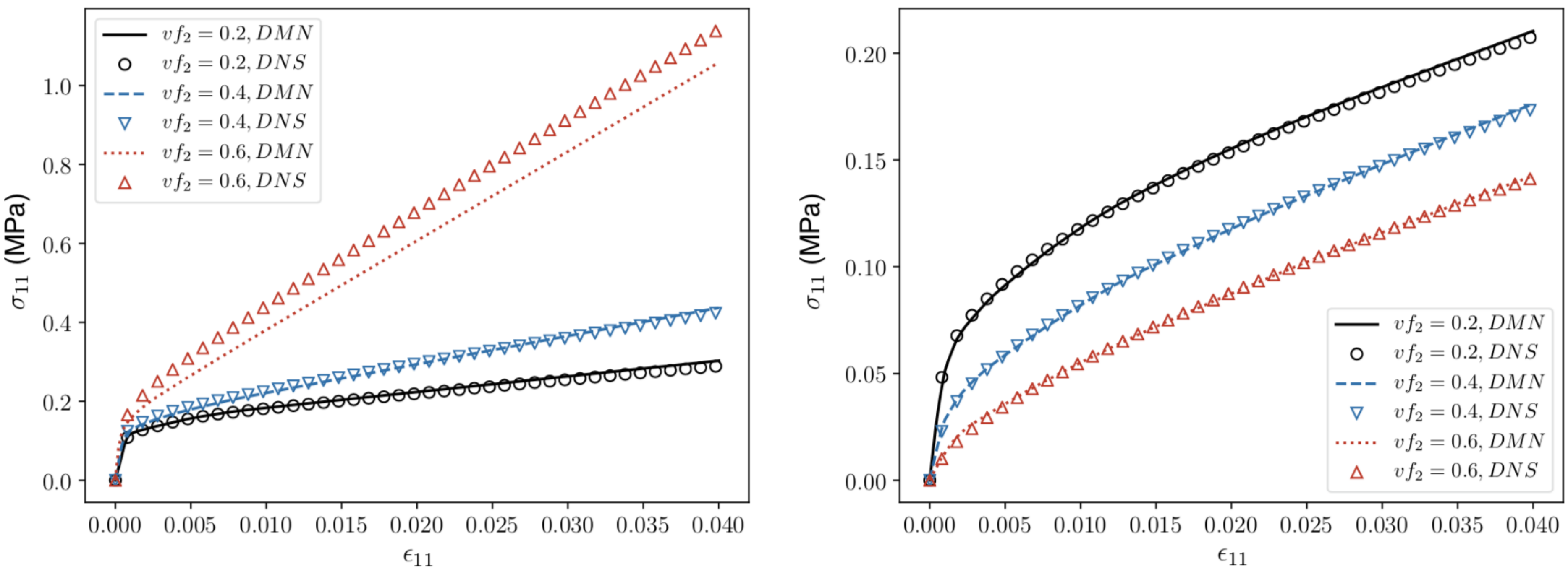
- 10% of the matrix phase has an effective plastic strain ε_1^p above 0.07;
- The mean effective plastic strain in the matrix phase $\bar{\varepsilon}_1^p$ is above 0.05.

DMN predictions on global stress-strain relations

DMN extrapolation

Offline stage: Both matrix and particle phases are orthotropic linear elastic

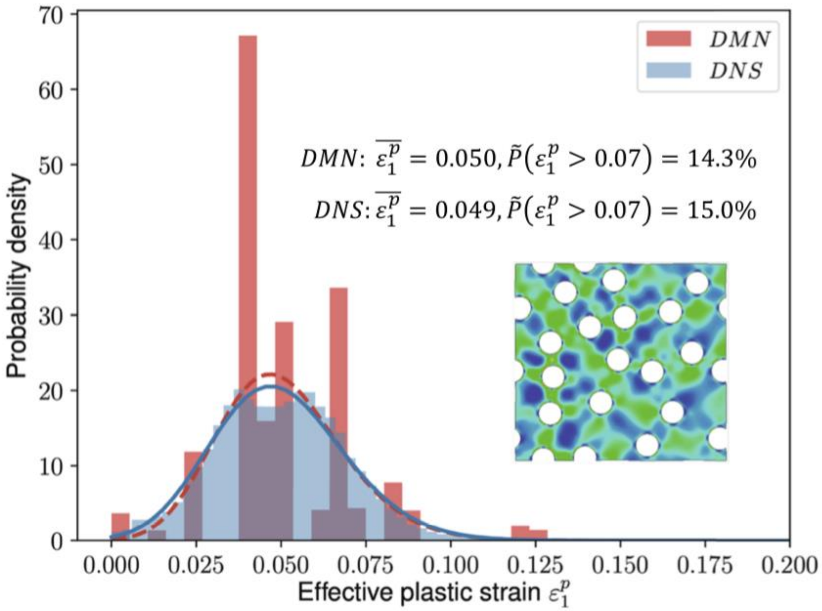
Online stage: The matrix is elasto-plastic with piece-wise linear hardening



(a) Hard particles: $E_2 = 5000$ MPa.

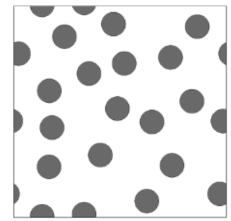
(b) Soft particles: $E_2 = 2$ MPa.

DMN predictions on local failure properties

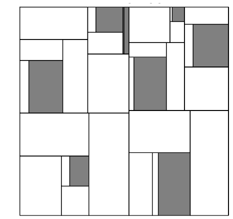


(a) RVE with $vf_2 = 0.2$ loaded at $\varepsilon_{11} = 0.036$.

DNS: 198212 elements



DMN: 34 nodes



Fit the DMN pdf to a log-norm distribution:

$$f_\varepsilon(\varepsilon, s, b, \eta) = \frac{1}{sy\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln y}{s}\right)^2\right)$$

Table 1 Ultimate tensile strength σ^{TS} (MPa) for RVEs with different particle moduli E_2 and volume fractions vf_2

vf_2	$E_2 = 5000$ MPa			$E_2 = 2$ MPa		
	0.2	0.4	0.6	0.2	0.4	0.6
DNS	0.275	0.298	0.417	0.244	0.228	0.207
DMN	0.278	0.302	0.381	0.246	0.230	0.198
Error	+1.1%	+1.3%	−8.6%	+0.8%	+0.9%	−4.3%

Table 2 Material toughness U^T (KJ/m^3) for RVEs with different particle moduli E_2 and volume fractions vf_2

vf_2	$E_2 = 5000$ MPa			$E_2 = 2$ MPa		
	0.2	0.4	0.6	0.2	0.4	0.6
DNS	6.81	4.43	2.61	9.22	9.01	8.26
DMN	6.92	4.65	2.58	9.01	8.83	7.27
Error	+1.6%	+5.0%	−1.1%	−2.3%	−2.0%	−12.2%

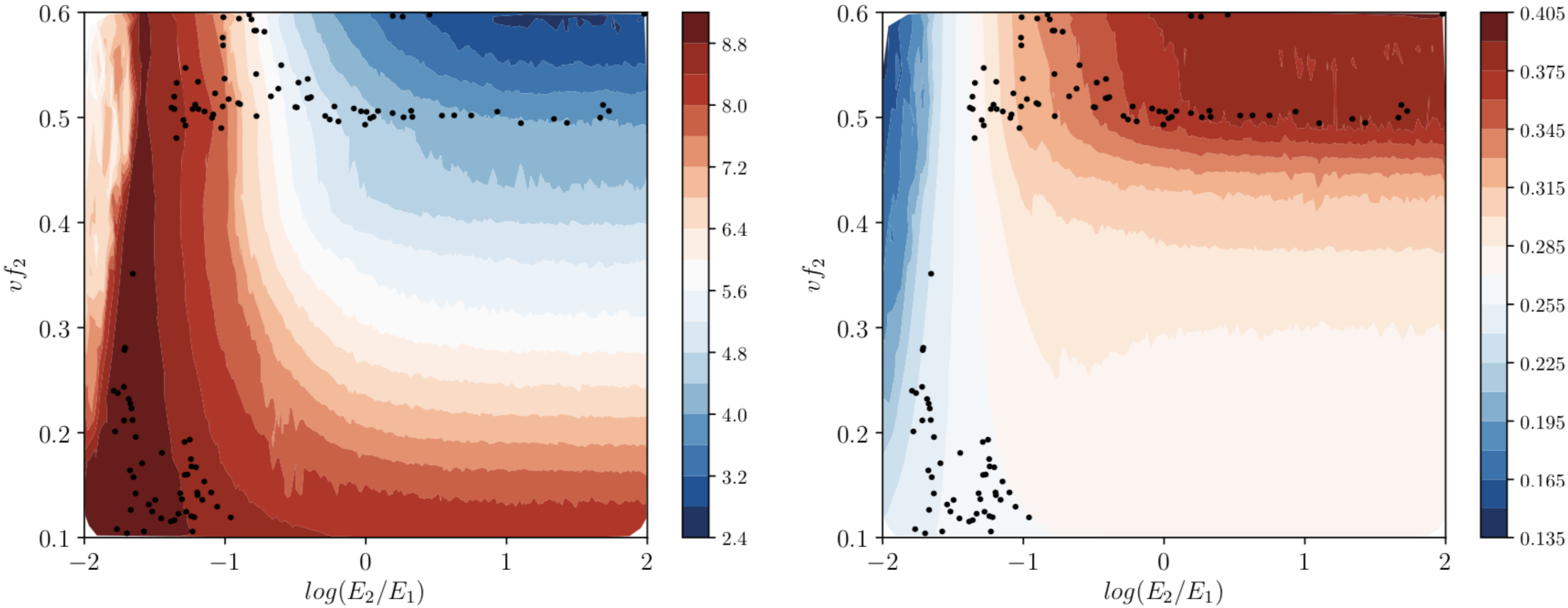
Design of Experiments on E_2 and vf_2

Matrix phase (elastoplastic):

$$E_1 = 100 \text{ MPa}, \quad \nu_1 = 0.3 \qquad \sigma_1^Y(\varepsilon_1^p) = \begin{cases} 0.1 + 5\varepsilon_1^p & \varepsilon_1^p \in [0, 0.008) \\ 0.124 + 2\varepsilon_1^p & \varepsilon_1^p \in [0.008, \infty) \end{cases}$$

Particle phase (elastic):

$$E_2 = [1, 10{,}000] \text{ MPa}, \quad vf_2 = [0.10, 0.60]. \quad \nu_2 = 0.3$$



(a) Material toughness KJ/m^3.

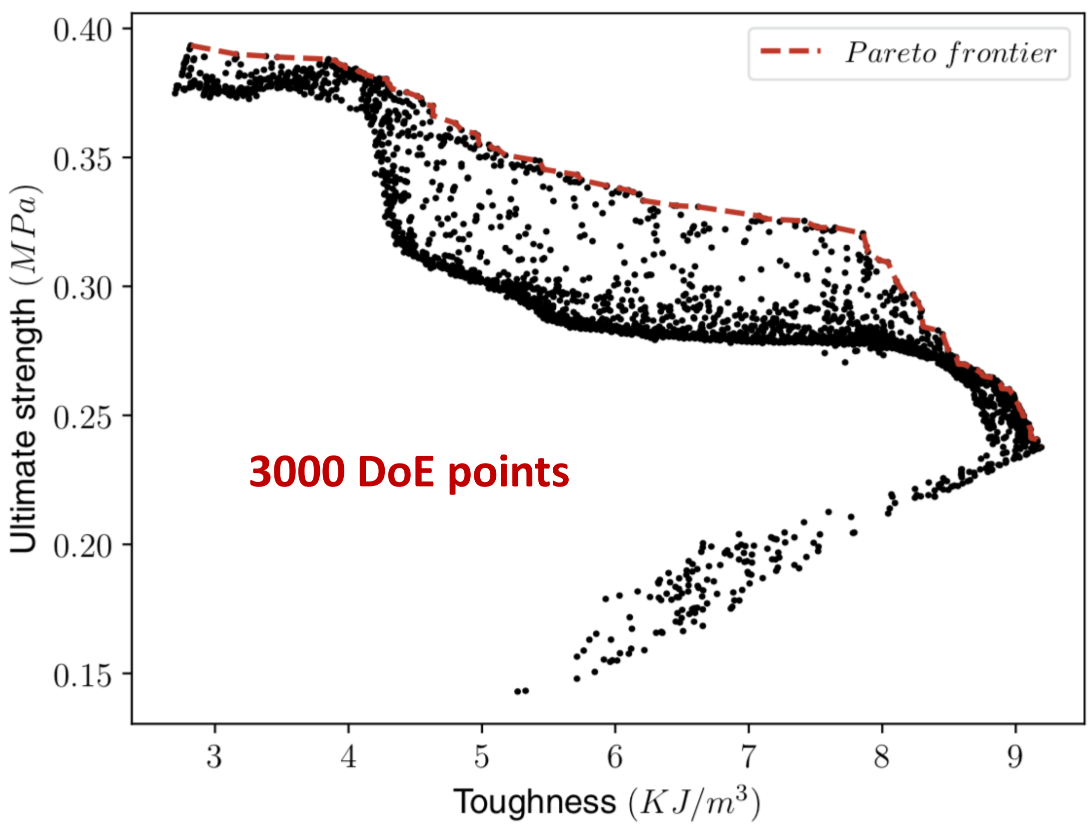
(b) Ultimate tensile strength (MPa).

Pareto frontier of toughness and ultimate tensile strength

Multi-objective optimization: A data point $X' = \{E_2', vf_2'\}$ is defined to be Pareto efficient if there does not exist such a point $X'' = \{E_2'', vf_2''\}$ $(X'' \neq X')$ that the conditions

$$\sigma^{TS}(X'') > \sigma^{TS}(X') \text{ and } U^T(X'') > U^T(X')$$

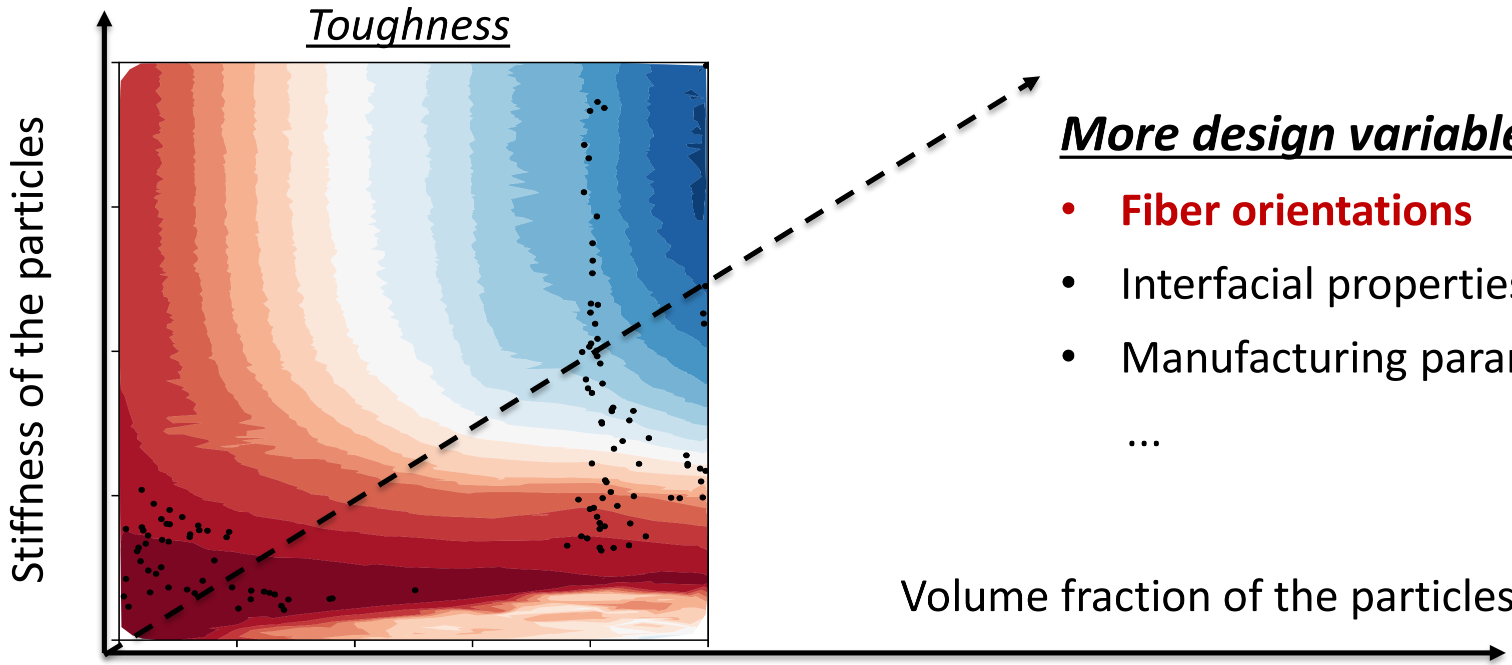
are both satisfied.



Computational time in offline and online stages

Table 3 Computational times for DMN with transfer learning and DNS on 10 Intel® Xeon® E5-2640 CPUs.

	Offline stage for 6 RVEs		Prediction stage (3000 DoE points)
	Sampling (3000 data points)	Training (60000 epochs)	
DNS	0	0	$\approx$ 3000 h (N/A)
DMN	71 h 30 min	12 h 40 min	38 min



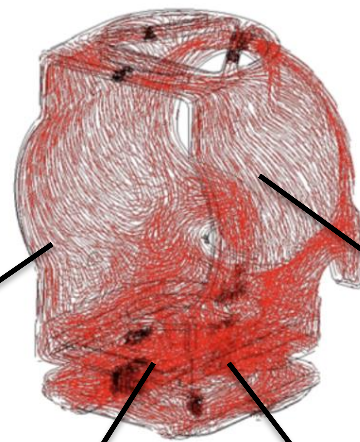
More design variables:

- Fiber orientations
- Interfacial properties
- Manufacturing parameters

...

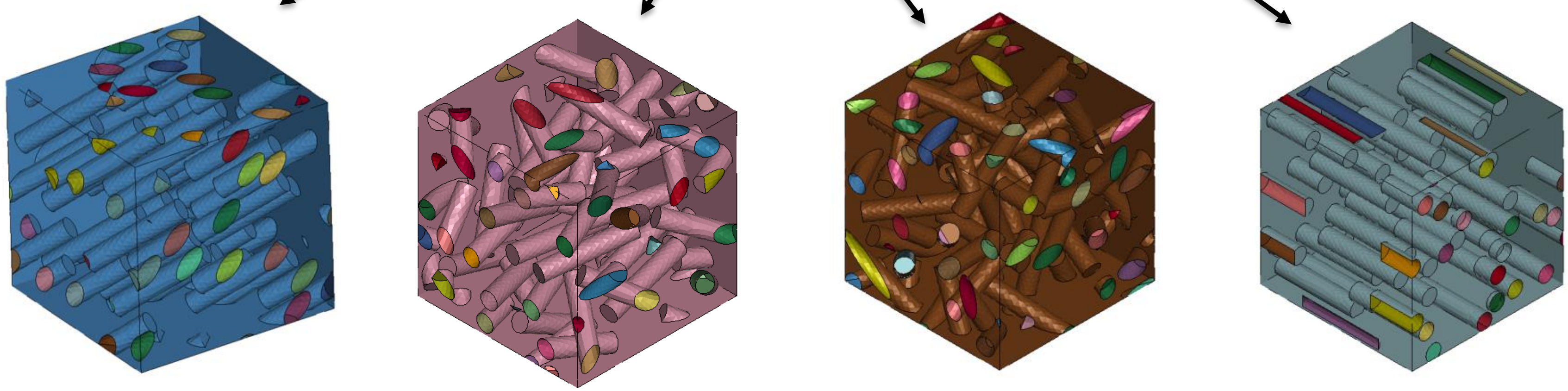
Real-world application: Short fiber composite materials

Injection molding simulation:



Orientation tensor:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$



$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ & 1/3 & 1/3 \\ sym. & & 1/3 \end{bmatrix} \quad \begin{bmatrix} 1/3 & 0 & 0 \\ & 1/3 & 0 \\ sym. & & 1/3 \end{bmatrix} \quad \begin{bmatrix} 0.7 & 0 & 0 \\ & 0.2 & 0 \\ sym. & & 0.1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ & 1 & 0 \\ sym. & & 0 \end{bmatrix}$$

Microstructural variations induced by Manufacturing Process

Summary

- **Transfer learning of deep material network**:
  - Initial database migration using a naïve approach
  - Faster convergence of training
  - Generation of analogous networks with the same base structure
- **A unified set of databases** are constructed by interpolating the fitting parameters
  - Continuous structure-property relationship
  - Encouraging micromechanical results of predicting the volume fraction effect on elastic properties
- **Materials design** enabled by the efficiency and accuracy of DMN extrapolation
  - Multi-objective optimization of material toughness and ultimate tensile strength
  - Failure prediction based on local distribution of effective plastic strain.

Future opportunities

- More material systems: rubber composite, short fiber composite, polycrystals…
- Database interpolation for more design variables
- Process-structure-property relationship
- Uncertainty quantification

Thank you!