

Highlights

Similarity Equivariant Graph Neural Networks for Homogenization of Metamaterials

Fleur Hendriks, Vlado Menkovski, Martin Doškář, Marc G.D. Geers, Ondřej Rokoš

- Development of a Similarity-Equivariant Graph Neural Network (SimEGNN) for homogenization of metamaterials
- $E(n)$ -Equivariant Graph Neural Networks extended to higher-order tensors
- Incorporation of all relevant symmetry groups to achieve similarity in-/equivariance
- Formulation & implementation of Representative Volume Element in-/equivariance (periodic boundary conditions) in a graph neural network
- Efficient graph representation of the finite element mesh; speed-up and better scaling with mesh size

Similarity Equivariant Graph Neural Networks for Homogenization of Metamaterials

Fleur Hendriks^{a,d}, Vlado Menkovski^{b,d}, Martin Doškář^c, Marc G.D. Geers^{a,d}, Ondřej Rokoš^{a,d}

^a*Mechanics of Materials, Department of Mechanical Engineering, Eindhoven University of Technology, Postbus 513, 5600 MB, Eindhoven, The Netherlands*

^b*Data Mining, Department of Mathematics and Computer Science, Eindhoven University of Technology, Postbus 513, 5600 MB, Eindhoven, The Netherlands*

^c*Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29, Prague 6, Czech Republic*

^d*Eindhoven AI Systems Institute (EAISI), PO Box 513, 5600 MB, Eindhoven, The Netherlands*

Abstract

Soft, porous mechanical metamaterials exhibit pattern transformations that may have important applications in soft robotics, sound reduction and biomedicine. To design these innovative materials, it is important to be able to simulate them accurately and quickly, in order to tune their mechanical properties. Since conventional simulations using the finite element method entail a high computational cost, in this article we aim to develop a machine learning-based approach that scales favorably to serve as a surrogate model. To ensure that the model is also able to handle various microstructures, including those not encountered during training, we include the microstructure as part of the network input. Therefore, we introduce a graph neural network that predicts global quantities (energy, stress stiffness) as well as the pattern transformations that occur (the kinematics). To make our model as accurate and data-efficient as possible, various symmetries are incorporated into the model. The starting point is an $E(n)$ -equivariant graph neural network (which respects translation, rotation and reflection) that has periodic boundary conditions (i.e., it is in-/equivariant with respect to the choice of RVE), is scale in-/equivariant, can simulate large deformations, and can predict scalars, vectors as well as second and fourth order tensors (specifically energy, stress and stiffness). The incorporation of scale equivariance makes the model equivariant with respect to the similarities group, of which the Euclidean group $E(n)$ is a subgroup. We show that this network is more accurate and data-efficient than graph neural networks with fewer symmetries. To create an efficient graph representation of the finite element discretization, we use only the internal geometrical hole boundaries from the finite element mesh to achieve a better speed-up and scaling with the mesh size.

Keywords: Graph Neural Networks, Similarity Equivariance, $E(n)$ -equivariance, Periodicity, Computational Homogenization, Mechanical Metamaterials

1. Introduction

Metamaterials are materials with a microstructure designed to exhibit special properties. For instance, mechanical metamaterials have unusual mechanical properties [1, 2], such as a negative Poisson's ratio [3] or negative compressibility [4]. Here, we are specifically interested in flexible, porous mechanical metamaterials that have a tunable stiffness as a result of a pattern transformation [5, 6] that can be activated, for example, by mechanical loading, pneumatics [7] or magnetic fields [8]. Such pattern transformations can also change the acoustic properties, allowing the design of mechanically tunable acoustic metamaterials [9–12], or they can be used to control the shape of the material, with applications in soft robotics [7, 13, 14]. They also have applications in biomedicine [15].

These kinds of pattern-transforming materials provide a large design space to be explored, because there are a lot of possibilities regarding geometry (shape and number of holes and inclusions), electromagnetic, chemical and mechanical properties of the base material, and different means of loading and activation. Moreover, thanks to the recent advancements in 3D printing [16–21], the broad design space is no longer inherently restricted by manufacturing concerns. Consequently, being able to design mechanical metamaterials for a specific target response (i.e., having a prescribed stiffness and buckling exactly when needed) is highly appealing. This can be achieved with shape [22, 23] and/or topology optimization [24, 25] on the Representative Volume Element (RVE) [26]. Topology optimization can be used to design a microstructure for various homogenized properties [27, 28], including the buckling strength of materials [29–31]. Because topology and/or shape optimization are almost always iterative, being able to rapidly evaluate the performance of new designs using numerical simulations is critical for accelerating the design process.

However, the behavior of the metamaterials of interest is complex, highly nonlinear, and involves large deformations, which requires repeated evaluation of the constitutive law of the matrix material. The buckling behavior that enables pattern transformations makes the system even more complicated to solve (here avoided by using a perturbation). Moreover, to capture all fine microstructural details, the discretization needs to have a sufficiently high resolution. For these reasons, the numerical simulations are typically expensive to evaluate, which makes them too inefficient to explore the large design space. Therefore, the main objective of this article is to build an efficient, accurate and fast-to-evaluate machine learning-based surrogate model for the modeling of periodic elastomeric mechanical metamaterials with a highly non-linear response.

Given the goal is to develop a model enabling the design of new mechanical metamaterials, the model must have the potential to generalize to unseen, arbitrary microstructures, even including topology changes. This implies that the surrogate model should be able to take a description of the full geometry as input. This rules out other approaches, such as analytical (closed-form) homogenization methods (e.g. mixture methods [32] or methods based on Eshelby's [33] analytical result for an ellipsoidal inclusion in infinite matrix [34, 35]), reduced order models [36–44], data-driven approaches [45], self-consistent clustering analysis [46], relatively simple machine learning approaches that map deformation gradient (and optionally some microstructural parameters) to global quantities [47–49], and graph networks (not to

be confused with graph neural networks) [50], which model porous mechanical metamaterials as rigid crosses connected by neural network-modeled springs.

These methods are certainly valuable in their developed context, but only work very well in the linear-elastic regime or only for a fixed geometry and topology. One reason neural networks are promising is the fact that they are universal approximators [51–53], which means they can approximate any arbitrary function and are therefore not constrained to any specific form of macroscopic constitutive law. Specifically, we are interested in deep learning solutions, since deep learning excels at using high-dimensional data [54], which allows us to use a full description of the geometry of the RVE as an input and which also scales favorably (often linearly) with the size of the input and the number of parameters. Especially in 3D, the scaling advantage could be considerable.

However, creating an efficient deep learning model is not trivial. Specifically, because the following properties need to be incorporated into the architecture of the model:

- The model should be able to handle large deformations (to cover buckling), high strains and rotations.
- If the input is a mesh, the model should be in-/equivariant under permutation of nodes (the order of nodes should not matter). The concepts of invariance and equivariance are defined and explained in detail in Section 3.
- Similarity in-/equivariance ($E(n)$ in-/equivariance, i.e., rotation, translation, reflection symmetry), to satisfy material objectivity. These transformations correspond to choosing a different coordinate system. This is also one of the requirements of the principle of material frame indifference, also called objectivity [55, p.195].¹
- Scale in-/equivariance; since we limit ourselves to first-order computational homogenization of a hyperelastic material (Bertoldi-Boyce [5], Equation (8)), a change in scale of the RVE should not change its behavior.
- Periodicity, which implies in-/equivariance with respect to a shift of the RVE (i.e., RVE window translation) and with respect to merging multiple RVEs into a new, bigger one. This assumes the RVE is already big enough to capture all relevant buckling patterns. If not, a larger RVE could allow for a new buckling pattern, which breaks the in-/equivariance with respect to merging multiple RVEs.
- Prediction of scalars (energy), vectors (displacement), and higher-order tensors (stress and stiffness).

To the best of our knowledge, there are no existing machine-learning based models that satisfy all these constraints. See Table 1 for an overview of the constraints respected by different types of models.

¹Usually, when discussing material frame indifference, reflection symmetry is omitted. However, it still applies, because changing the orientation of the coordinate system does not change the behavior of the material, and that is why, for completeness, we include it here as well.

Table 1: Overview of GNN architectures and the constraints they address, along with the presently developed EGNN and similarity-equivariant GNN (SimEGNN).

Architecture	Updates node positions?	Invariances/equivariances					
		Translation	Rotation	Reflection	Scale	Periodicity	
						Shifted RVE	Extended RVE
CNN	✗	✓	✗	✗	✓	✓	✓
GNN	✓	✓	✗	✗	✗	✓	✓
GNN*	✗	✓	✓	✓	✗	✓	✓
Original EGNN	✓	✓	✓	✓	✗	✗	✗
This paper's GNN	✓	✓	✗	✗	✗	✓	✓
This paper's EGNN	✓	✓	✓	✓	✗	✓	✓
This paper's SimEGNN	✓	✓	✓	✓	✓	✓	✓

*assuming only $E(n)$ -invariant attributes such as distances and angles are used.

Convolutional neural networks applied to 2D or 3D images [56–62] respect translation in-/equivariance and can respect periodicity when used with circular/periodic padding.² They also respect scale in-/equivariance, because the input picture does not change in that case. They can be made to respect rotation and reflection in-/equivariance [63], but not for any arbitrary rotation or reflection; only with respect to rotations that are multiples of 90° and reflections that are horizontal or vertical. In addition, they can only predict vectors and higher-order tensors by treating them as an arbitrary list of numbers (that therefore do not necessarily transform as tensors), they cannot track the movement of nodes, and their geometry input is constrained to a square grid.

A promising alternative is a graph neural network (GNN): a type of neural network that takes a graph as input. The mesh representation used for the finite element simulations is well-suited for a graph representation. GNNs also respect permutation in-/equivariance, which is advantageous, because any description of the geometry of the material microstructure has to respect permutation in-/equivariance with respect to the geometrical features (i.e., the order in which the nodes, elements or entire holes are described should be irrelevant). GNNs have been used before [64–67] for the simulation of materials. However, these GNNs (i) only predict scalars (e.g. energy, mean pressure, deviatoric stress, fracture probability, damage), (ii) operate typically in the small-strain regime, and (iii) are neither $E(n)$ -in-/equivariant nor (iv) scale in-/equivariant. This makes them unsuitable for the goal of simulating the specific metamaterials of interest here. MeshGraphNets [67] using translation equivariant GNNs to recreate results from physical simulations *do* allow for large strains,

²A CNN with circular padding will not be exactly in-/equivariant with respect to shifting the RVE, because of the max pooling and the flattening that usually happens at the end of the convolutional layers, which is fed into the dense layers afterwards.

as demonstrated by a simulation of a body made of a hyperelastic material. However, these models are neither $E(n)$ -equivariant nor scale in-/equivariant.

$E(n)$ -in-/equivariance in GNNs can be achieved simply by using only $E(n)$ -invariant features such as distances or angles, as many GNNs for molecular property prediction do [68–73]. In addition, periodicity can be incorporated by merging nodes from one side of the RVE with their corresponding side (left to right, top to bottom), or equivalently, have ‘ghost nodes’ [74]. Unfortunately, these molecular GNNs tend to be suitable only for graphs with a small graph diameter, whereas a finite element (FE) mesh is usually much larger with a large graph diameter, requiring a high number of message passing steps, which are the iterations used by a GNN to update the graph, see Section 4. Moreover, using these $E(n)$ -invariant features precludes any dependence on a global coordinate system. Consequently, the node positions cannot be updated directly.

There exist other ways to update node positions in a global coordinate system such that the updates are independent of the global coordinate system, using, e.g., expansions in spherical harmonics. Relying on spherical harmonics allows for updating positions independently of a global coordinate system, albeit at increased computational cost [75, 76]. Alternatively, GVP-GNNs (Geometric Vector Perceptron Graph Neural Networks) [77], PaiNN (polarizable atom interaction neural network) [78] and GemNet (geometric message passing neural network) [79] can predict vector quantities as well, without the use of spherical harmonics, which makes them also suitable to predict deformation. Here, we choose another approach named $E(n)$ -equivariant graph neural networks (EGNNs) [80], due to their simplicity, which updates node positions by ‘pulling’ or ‘pushing’ them along their edges.

EGNNs currently do not respect periodicity, because the distances are recomputed from the new coordinates of the nodes after each message passing step. This approach does not work for edges that wrap around from one side of the RVE to another (i.e., for periodic edges). We show that it is relatively easy to adjust an EGNN such that it is periodic (referred to as ‘EGNN’ in this paper). EGNNs are also not scale invariant, because there is a dependence on the distances between nodes. We show this can also be fixed (SimEGNN).

The paper is structured as follows: Section 2 covers the essentials of first-order computational homogenization and describes the generation of the training dataset. Section 3 describes the concept of in-/equivariance, and how it applies here. Section 4 details the developed GNN architecture. Numerical results and discussion are provided in Section 5, while Section 6 concludes this paper. All the code can be found at: <https://github.com/FHendriks11/SimEGNN>, and all the data (including the trained networks) is available upon request.

On notation: scalars (except for the strain energy density \mathfrak{W}) are in italics (e.g. r_{ij}), abstract ‘vectors’ (neural network-generated lists of numbers without a physical meaning), which can have any dimension, are in bold (e.g. \mathbf{h}_i), physical/geometric vectors, which are always 2D or 3D, are in italics with an arrow (e.g. \vec{x}_i), geometric tensors are in bold capitals (e.g. \mathbf{P}), of which the 4th-order tensors have a superscript 4 in front (e.g. ${}^4\mathbf{D}$). The indices indicate how many nodes are relevant to that quantity, e.g., \mathbf{h}_i is an abstract vector specific to one node i , r_{ij} is a scalar quantity specific to an edge ij between node i and j . We use

the following definitions of the tensor product (also called open product or dyadic product)

$$\mathbf{A} = \vec{a} \otimes \vec{b} \iff A_{ij} = a_i b_j \quad (1)$$

$${}^4\mathbf{C} = \mathbf{A} \otimes \mathbf{B} \iff C_{ijkl} = A_{ij} B_{kl} \quad (2)$$

and the dot product

$$\vec{b} = \vec{a} \cdot \mathbf{A} \iff b_j = \sum_i a_i B_{ij} \quad (3)$$

$$\vec{b} = \mathbf{A} \cdot \vec{a} \iff b_i = \sum_j A_{ij} a_j \quad (4)$$

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{B} \iff C_{ik} = \sum_j A_{ij} B_{jk}. \quad (5)$$

2. Data Set Generation

To create a training data set, finite element simulations are exploited, in combination with the computational homogenization scheme to compute representative volume element (RVE) responses and their effective properties. Specifically, we construct a dataset of the following mapping from the microscopic RVE (geometry and material) and the macroscopic deformation gradient tensor \mathbf{F} to the microscopic deformation and the macroscopic target variables

$$\left. \begin{array}{c} \text{RVE geometry \& material} \\ \mathbf{F} \end{array} \right\} \rightarrow \begin{cases} \vec{x} \\ \mathfrak{W} \\ \mathbf{P} \\ {}^4\mathbf{D} \end{cases}, \quad (6)$$

where \mathfrak{W} is the effective (macroscopic) strain energy density, \mathbf{P} is the effective first Piola-Kirchhoff stress tensor and ${}^4\mathbf{D}$ is the effective stiffness tensor and \vec{x} is the deformed configuration, meaning the final position of each node. This stiffness is defined through the incremental stiffness relationship

$$\delta \mathbf{P} = {}^4\mathbf{D} : \delta \mathbf{F}. \quad (7)$$

Even though \vec{x} is not strictly necessary, we want to obtain it as well, to make the eventual machine learning model more interpretable, and to offer the possibility of observing the various pattern transformations directly. This dataset will later be used to train a machine learning model.

2.1. Computational Homogenization

We focus on calculations pertinent to computational homogenization, which is a workhorse in shape or topology optimization methods of materials. It replaces an explicit constitutive law at the macroscale with a boundary value problem posed on an RVE at the microscale, loaded by a deformation gradient \mathbf{F} , which is extracted from the macroscale (for an overview

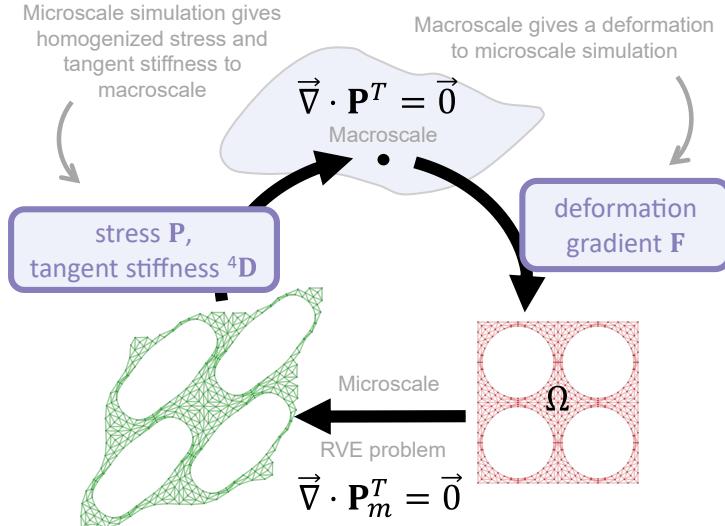


Figure 1: Schematic overview of the first-order computational homogenization procedure.

of the basics and extensions of computational homogenization, see [81]). The macroscale (global scale) is the scale over which the mechanical loading of the entire system varies, whereas the microscale is the scale of the heterogeneities in the microstructure of the underlying material. In this section, the microscopic quantities will be denoted by the subscript m . See Figure 1 for a schematic representation of the computational homogenization procedure.

We focus here on the microscale computation, which is a finite element (FE) simulation of an RVE. Since the microstructural features in our materials are quite fine, the microscale FE mesh must be adequately small as well, which makes the micro-scale simulations expensive.

Even though the nature of architected materials often calls for a generalized continuum at a macroscale (second-order [82, 83] or micromorphic [84–86]), we limit ourselves here to the standard first-order computational homogenization which leads to a classical Cauchy continuum at the macroscale, assuming separation of scales.

2.2. Microstructural RVE problem

The underlying pattern-transforming 2D metamaterial consists of a flexible, hyperelastic matrix with circular (or, in our case, almost circular) holes arranged in a square grid as shown in Figure 2a [84, 87]. This material exhibits auxetic behavior after buckling. We assume a hyperelastic, finite-strain model at the microscale. One RVE has four holes, which has been confirmed to be of a sufficient size to capture the pattern transformation when buckling (e.g., [5]). Each hole has a diameter 0.45 times the side length ℓ of the RVE.

The above-described metamaterial constitutes a challenging test case, because of the multiple buckling modes it shows, and the strong dependence of effective properties on the deformation state. Depending on the loading conditions, i.e., prescribed macroscopic \mathbf{F} , a bifurcation breaking the RVE’s symmetry can occur. There are two bifurcation modes: the

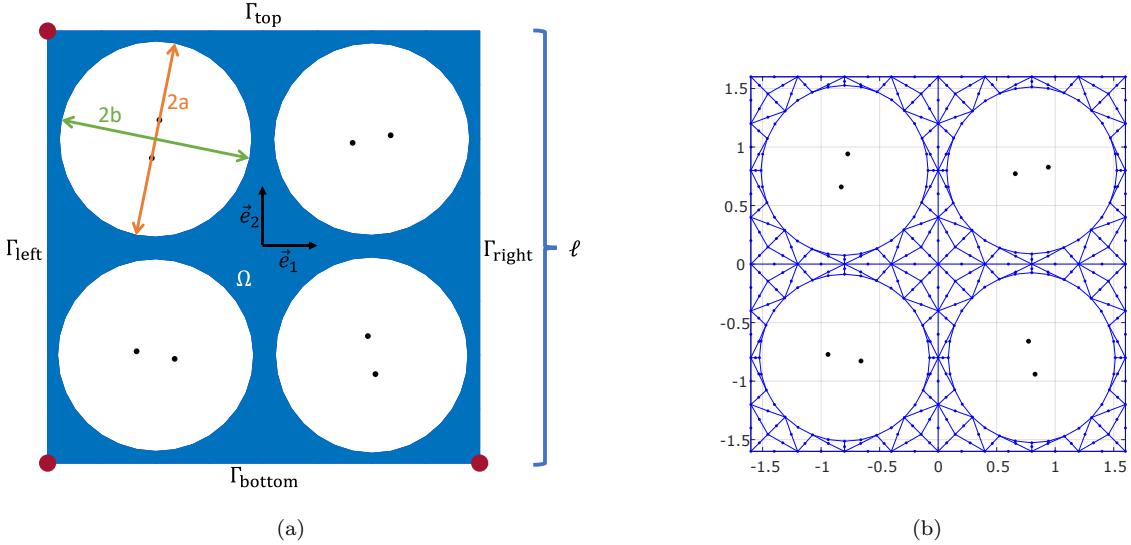


Figure 2: (a) The RVE geometry, with four holes, slightly flattened by 1% into ellipses, with $2a$ the major axis and $2b$ the minor axis. The periodic boundaries are indicated by Γ , the control points by red dots, and the foci of the ellipses by black dots. The length and width of the RVE is ℓ and the domain is Ω . The directions of the basis vectors are indicated by \vec{e}_1 and \vec{e}_2 . (b) The finite element discretization of this geometry.

microstructure can either (i) buckle clockwise or counterclockwise – referred to in this paper as ‘rotational bifurcation’ –, or (ii) bulge out left or right (or up or down if $F_{11} < F_{22}$) – ‘left/right bifurcation’,³ or even both situations can occur one after another along the loading path, see also Figure 3. Buckling significantly alters the homogenized stress and stiffness; Figure 4 shows how the normal stress P_{11} and stiffness component D_{1111} change under biaxial compression.

We use the Bertoldi-Boyce constitutive law for the polymer, given by the strain energy density function [5]

$$\mathfrak{W}_m(\mathbf{F}_m) = m_1(I_1 - 2) + m_2(I_1 - 2)^2 - 2m_1 \ln(J) + \frac{1}{2}\kappa(J - 1)^2, \quad (8)$$

where $J = \det \mathbf{F}_m$, $I_1 = \text{tr } \mathbf{C}_m$, \mathbf{C}_m is the microscopic right Cauchy-Green tensor $\mathbf{C}_m = \mathbf{F}_m^T \cdot \mathbf{F}_m$, and m_1 , m_2 and κ are constants, which take the values $m_1 = 0.55 \text{ MPa}$, $m_2 = 0.3 \text{ MPa}$, $\kappa = 55 \text{ MPa}$.⁴

The mechanical behavior is modeled using a standard Cauchy continuum. Assuming no body forces and neglecting inertia, this amounts to solving

$$\vec{\nabla}_0 \cdot \mathbf{P}_m(\mathbf{F}_m(\vec{u})) = \vec{0} \quad \forall \vec{x} \in \Omega \quad (9)$$

³This left/right buckling mode would actually correspond to a global buckling mode, and the RVE with 2×2 holes we use therefore does not capture it accurately. However, because here the system only serves as a test case for the GNN, we ignore this fact for now.

⁴In 3D, the first two terms would have $I_1 - 3$ instead of $I_1 - 2$.

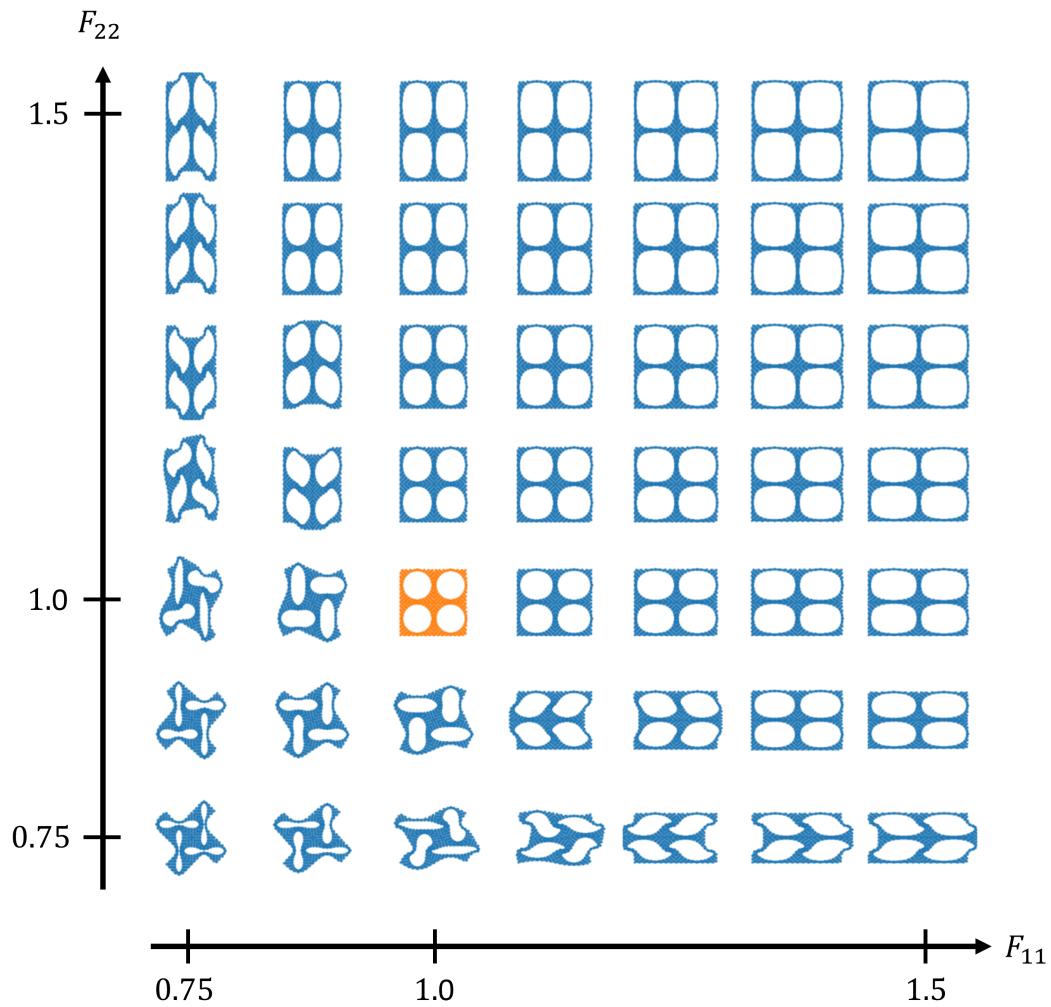


Figure 3: The behavior of the metamaterial RVE under various loading conditions. In this plot, U_{ij} are components of the right stretch tensor, see Section 2.3. The components U_{11} and U_{22} are varied, while U_{12} is constrained at zero. The reference configuration (i.e., the unloaded state) is plotted in orange.

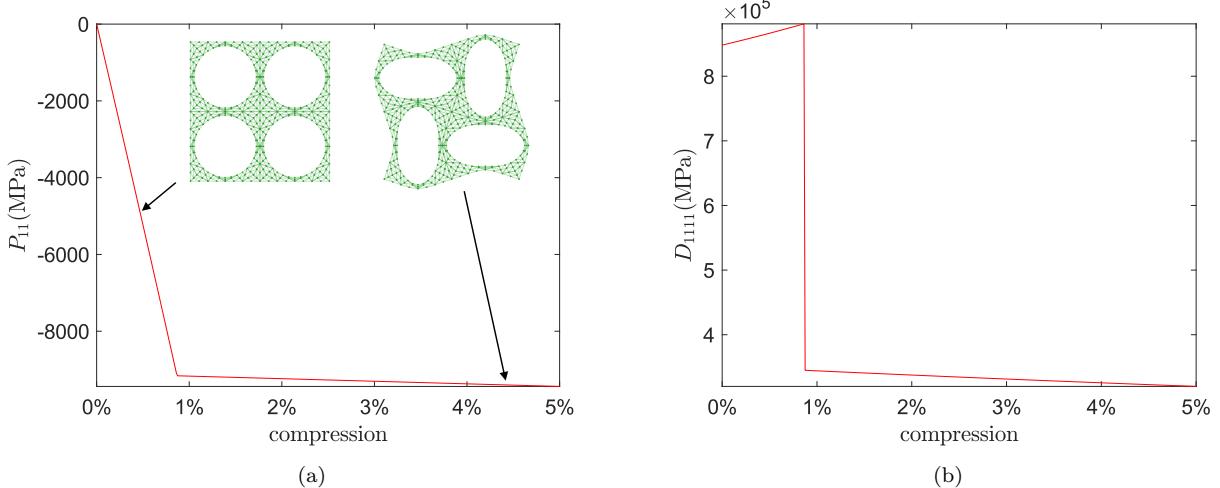


Figure 4: Change in (a) stress component P_{11} and (b) stiffness component D_{1111} as a function of the biaxial compression.

for the microscopic displacement field \vec{u} , with $\vec{\nabla}_0$ the gradient operator with respect to the *initial* microstructural position \vec{x}^{ref} , and \vec{x} the new position on the domain Ω . This equation and its solutions respect the in- and equivariances briefly mentioned in the Introduction and further discussed in Section 3. The new position \vec{x} is decomposed as

$$\vec{x} = \vec{x}^{\text{ref}} + \vec{u}(\vec{x}^{\text{ref}}) = \mathbf{F} \cdot \vec{x}^{\text{ref}} + \vec{w}(\vec{x}^{\text{ref}}), \quad (10)$$

where $\vec{u}(\vec{x})$ is the displacement field, $\mathbf{F} \cdot \vec{x}^{\text{ref}}$ is the affine part dictated by the prescribed macroscopic deformation gradient \mathbf{F} and \vec{w} denotes the fluctuation part of the solution.

From this solution, the homogenized quantities can be obtained as follows:

$$\mathfrak{W} = \frac{1}{\Omega} \int_{\Omega} \mathfrak{W}_m (\mathbf{F}_m (\vec{u})) d\vec{x}, \quad (11)$$

$$\mathbf{P} = \frac{\partial \mathfrak{W}}{\partial \mathbf{F}} = \frac{1}{\Omega} \int_{\Omega} \mathbf{P}_m (\mathbf{F}_m (\vec{u})) d\vec{x}, \quad (12)$$

$${}^4\mathbf{D} = \frac{\partial^2 \mathfrak{W}}{\partial \mathbf{F} \partial \mathbf{F}} = \frac{\partial \mathbf{P}}{\partial \mathbf{F}}. \quad (13)$$

We do not provide the full definition for ${}^4\mathbf{D}$, but it can be found in [55, p.227] and the method of computing it is adopted from [88, §2.4.4], see also [89].

The periodic boundary conditions enforce the top and bottom boundaries of the RVE to deform in the same way, and likewise for the left and right boundaries. This implies that the microfluctuation field \vec{w} (i.e., the displacement field \vec{u} minus its affine part) is equal for corresponding points on these boundaries [90], see Figure 2a:

$$\vec{w}(\Gamma_{\text{left}}) = \vec{w}(\Gamma_{\text{right}}), \quad (14)$$

$$\vec{w}(\Gamma_{\text{top}}) = \vec{w}(\Gamma_{\text{bottom}}). \quad (15)$$

The macroscale deformation gradient \mathbf{F} is imposed on this RVE by applying it as an affine transformation through the three corner control nodes (the fourth one is dependent on the others due to periodicity) in multiple adaptive load increments.

The RVE’s geometry shown in Figure 2a is discretized with quadratic triangular elements, using Gmsh (<https://gmsh.info/>) [91] to create the mesh shown in Figure 2b. This mesh is fairly coarse, which means the results are not as accurate as desired for FE², but good enough to capture the overall behavior and to quickly generate a large data set to test the proposed method. The FE simulations were performed using an in-house Matlab code [85].

Because the neural network cannot handle the ambiguity in buckling (i.e., left vs right, clockwise vs counterclockwise), we ensure that the material always buckles in the same direction, by steering the bifurcation through a minor perturbation of the holes’ initial geometry. To this end, the circles are flattened into ellipses by 1% (see Figure 2b for the resulting finite element mesh, where the foci of the four ellipses are depicted as black dots). Moreover, the ellipses are tilted such that both lateral and rotational bifurcations are triggered consistently. A stability analysis was performed to verify if this adequately controls the bifurcation ambiguity, which was indeed the case.

2.3. Sampling and Data Augmentation

For the purpose of training GNNs, a training dataset was generated by uniformly sampling the RVE’s response to a prescribed macroscopic deformation gradient \mathbf{F} . Using the polar decomposition, \mathbf{F} can be decomposed as $\mathbf{F} = \mathbf{R} \cdot \mathbf{U}$ into the symmetric right-stretch tensor \mathbf{U} and a macroscopic rotation tensor \mathbf{R} . Because \mathbf{R} encodes only a rigid-body rotation, which does not meaningfully affect the results, we only sampled $\mathbf{F} = \mathbf{U}$. This simplification left us with only three components U_{11} , U_{12} and U_{22} .

Sampling was not uniform, because each simulation starts from $\mathbf{U} = \mathbf{I}$, and all intermediate configurations on the loading path to the prescribed final \mathbf{U} were also saved and used for training, which is why sampling included only cases of \mathbf{U} with at least one component at its extreme. To sample different values of \mathbf{U} , in turn one of the three components was kept at its extreme value (minimum or maximum) while the other two were uniformly sampled from a specified range. The values of U_{11} and U_{22} were sampled between 0.75 and 1.5 in steps of 0.05, and $U_{12} = U_{21}$ had values between 0 and 0.5, also with steps of 0.05. This is a wide enough range to capture the different types of behavior. Because of the symmetry of the RVE (disregarding the small perturbation), we only consider \mathbf{U} where $U_{11} > U_{22}$, and $U_{12} \geq 0$.

Whenever a computed response featured overlapping elements, indicating the presence of contact (not considered in the numerical model), the corresponding \mathbf{U} was removed from the dataset. This resulted in a data set of 8451 distinct data points.

In the results Section 5, the newly developed machine learning model that respects all symmetries is compared to two models that have fewer symmetries incorporated. To train models with less symmetries to a higher degree of accuracy, a method called data augmentation can be used. Here, one trains the model on transformed versions of the data to promote learning of these symmetries (e.g., teach it rotation by including rotated versions of each data point). Therefore, to make the comparison more fair, we augmented the dataset

with rotated, reflected and scaled versions of the data points. To this end, we transformed each load case, using Latin hypercube sampling for the rotation and scaling, and random booleans for the x -reflection and y -reflection. The bounds on rotation were set between 0 and 2π radians and the scaling factor was chosen between 0.1 and 10 (sampling as 10^x , with linear sampling of $x \in [-1, 1]$).

3. In- and Equivariance

Inductive biases are a set of assumptions about the data, built into machine learning models, that do not need to be learned from data. Incorporating domain knowledge into a model improves the model’s efficiency with respect to the amount of training data needed and improves its generalization capabilities. The symmetries (invariances and equivariances) mentioned in the introduction are examples of such inductive biases [92].

Invariance with respect to a certain transformation T means that when this transformation is applied to the input \mathbf{x} , the output \mathbf{y} of the model ϕ stays the same

$$\phi(\mathbf{x}) = \phi(T\mathbf{x}). \quad (16)$$

Equivariance with respect to a transformation means that there is a corresponding transformation T' that can be performed on the model’s output to get the same result

$$T'\phi(\mathbf{x}) = \phi(T\mathbf{x}). \quad (17)$$

Often $T = T'$, but not always. For example, if the input configuration is rotated (because the coordinate system is chosen differently, or because \mathbf{F} contains a rotational part), then the energy density \mathfrak{W} stays the same (invariance), while the output configuration \vec{x} is also rotated (equivariance), and the homogenized stress and stiffness tensors are transformed as tensors under a rotation of the basis vectors (also equivariance). See Figure 5 for a visual explanation of these concepts. So, rather than giving many examples of such transformations in the training data and letting the model learn these invariances and equivariances, we aim to incorporate these symmetries directly into the model architecture and guarantee the model’s performance under such transformations.

As mentioned in the introduction, in addition to the node label permutation in-/equivariance that GNNs already respect, we consider similarity in-/equivariance (translation, rotation, reflection, scaling) and RVE in-/equivariance (periodicity). These in-/equivariances result directly from the governing balance equation of the system, Equation (9).

Figure 6 illustrates these equivariances for the deformation, and Table 2 indicates which of the two (invariance or equivariance) applies for each type for each target variable. These are the node positions \vec{x} , the macroscopic strain energy density \mathfrak{W} , the first Piola-Kirchhoff stress tensor \mathbf{P} and the stiffness ${}^4\mathbf{D}$.

Incorporating more in-/equivariances makes it easier for the model to construct and use features in a general way [92]. Otherwise, it will have to learn them from the data, which can be encouraged by using data augmentation (see section 2.3).

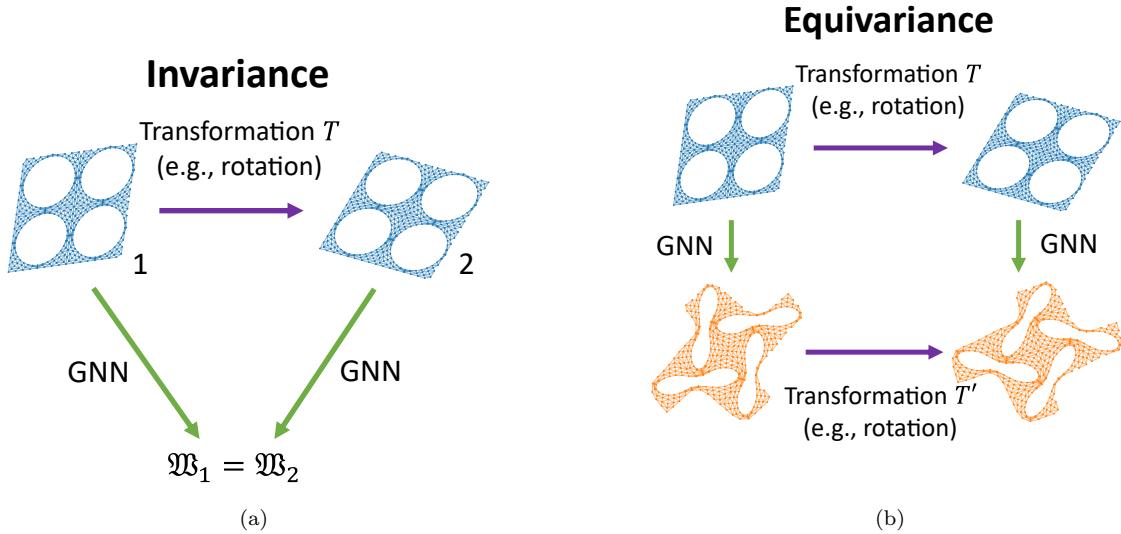


Figure 5: The concepts of (a) invariance and (b) equivariance illustrated, using a GNN predicting the strain energy density and deformation as examples, respectively.

Table 2: Overview of the invariance (**I**) or equivariance (**E**) of target variables with respect to different types of transformations

Transformation	\vec{x}	\mathfrak{W}	P	${}^4\mathbf{D}$
Translation	E	I	I	I
Rotation	E	I	E	E
Reflection	E	I	E	E
Scaling	E	I	I	I
Shifting RVE	E	I	I	I
Extending RVE	E	I	I	I

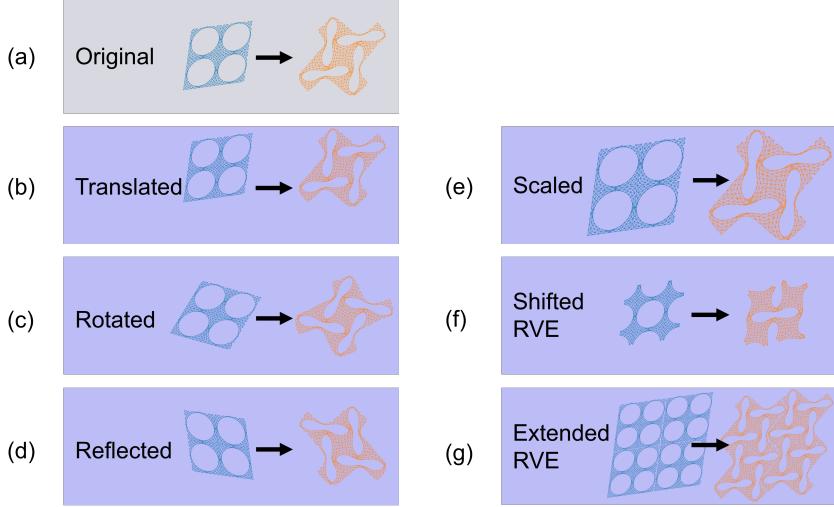


Figure 6: Different types of equivariance that should be respected by the model. Each case shows that the output configuration (orange) is transformed similarly to the input (blue), which is the meaning of equivariance.

4. GNN Approach

To achieve the goals set out in the introduction, we have developed a graph neural network (GNN)-based approach as detailed in this section. The final model – the SimEGNN – respects all the in-/equivariances discussed in Table 3. Respecting these symmetries is important for data and parameter efficiency and generalizability [92].

We first discuss the data structure encoding, i.e., graph encoding, and then present and compare three GNN architectures that were created specifically for metamaterial simulations. Each successive network has more symmetries built in, to quantify the added value of incorporating these symmetries. The first model, discussed in Subsection 4.3, is a base GNN model that can work with the data presentation, and respects the periodicity (i.e., the ‘Shifted RVE’ and ‘Extended RVE’ in-/equivariance in Table 2), which is explained in Subsection 4.2 as well as translation in-/equivariance. The second architecture, discussed in Subsection 4.4, extends this base GNN to also incorporate $E(n)$ -in-/equivariance. Finally, the third architecture, discussed in Subsection 4.5, is our newly proposed SimEGNN, which also adds scale in-/equivariance. This means it respects all in-/equivariances listed in Table 2.

Details on the implementation, including the training process, can be found in Appendix A.

4.1. Graph Encoding

A graph consists of a set of nodes \mathcal{V} and a set of edges \mathcal{E} connecting pairs of the nodes. The element edges and nodes of a FE mesh (e.g., the one used for data generation) are an obvious option for the input graph; however, one can also choose a different graph. Our starting point in creating a graph representation of the metamaterial’s microstructure is

the finite element discretization in Figure 2b, but we only keep the nodes and edges at the boundary of the holes (i.e., we remove all the bulk nodes). We do this because incorporating fewer nodes in the input graph leads to a reduced computational cost, and a smaller diameter (the maximum degree of separation of any two nodes), which results in fewer message passing steps needed to propagate information throughout the entire graph. Because the deformation of the boundary nodes of the mesh is enough to see how the material deforms, and the rest of the quantities of interest are global (\mathfrak{W} , \mathbf{P} , ${}^4\mathbf{D}$), we do not require any local fields, which makes the bulk nodes unnecessary in the first place. Although it is possible to use the full FE mesh, our numerical experiments showed that it is a slower and less accurate option, presumably due to the higher number of message passing steps needed, leading to problems such as oversmoothing [93].⁵

When constructing the new graph, we make no distinction between nodes at the corner of the quadratic triangular elements and the mid-edge nodes. The elimination of the bulk nodes leaves us with four disconnected components, each corresponding to one of the four holes. We join these components by connecting each node to the closest node on every other hole boundary (i.e., adding three new edges for each node, adding the reverse of these edges as well, and then deduplicating the set of edges). Due to the RVE’s periodicity, an edge may start on one side of the RVE and ‘wrap around’ to the other side of the RVE (i.e., a Pac-Man World geometry). The resulting graph, derived from the FE mesh shown in Figure 2b, is shown in Figure 7. The original mesh contained 413 nodes and 1056 edges, while the generated input graph has 128 nodes and 480 edges.

In the constructed input graph, each node i has a position \vec{x}_i (a 2D or 3D vector) and an abstract embedding \mathbf{h}_i associated with it. In addition, every edge ij has a vector \vec{r}_{ij} pointing along it and an abstract embedding \mathbf{e}_{ij} ; see Figure 8 for a visual explanation of these graph quantities. Initially, the edge embedding \mathbf{e}_{ij}^0 is -1 if at a hole boundary and +1 else. The initial node embedding \mathbf{h}_i^0 is empty.

4.2. General approach

GNNs incorporate permutation symmetry by updating the node and edge attributes by message passing, which is a way of propagating information throughout the graph [94]. For all three models we compare, the message passing steps update the node positions \vec{x}_i and generate abstract node and edge embeddings \mathbf{h}_i and \mathbf{e}_{ij} . The node positions after the last message passing steps are the predicted positions.⁶

Message passing involves constructing a vector \mathbf{m}_{ij} (called a ‘message’) for each edge $ij \in \mathcal{E}$, and then aggregating at each node $i \in \mathcal{V}$ the ‘incoming’ messages (the messages associated with edges that connect to the node) to update the node embedding \mathbf{h}_i . In its

⁵When fully detailed local fields *are* of interest, it may be possible to limit the number of necessary message passing steps by using graph pooling strategies.

⁶Because any predicted displacement that contains a rigid body translation is still a valid solution, we compare the final GNN prediction for \vec{x} to the FEM ground truth after setting the mean position such that it matches the FEM ground truth.

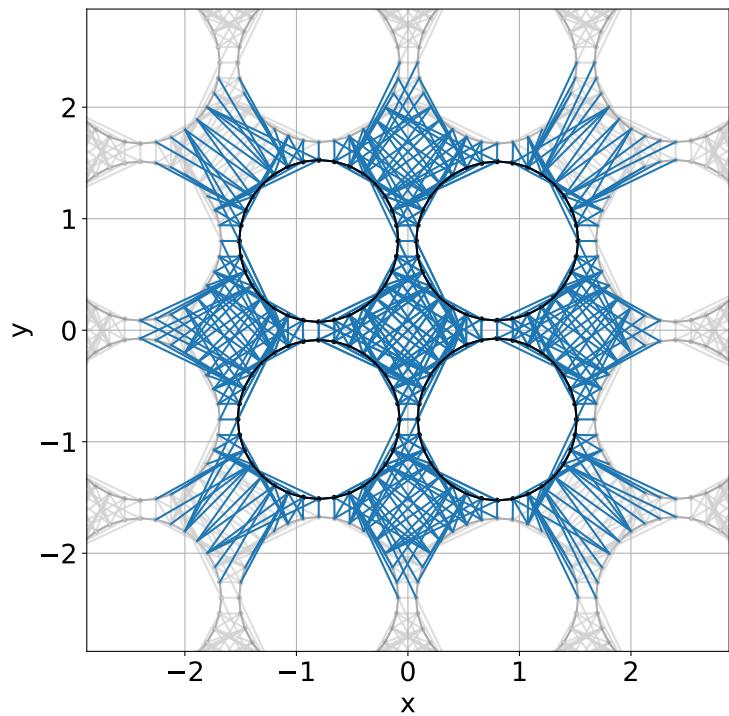


Figure 7: The input graph used for all GNNs considered in this paper, shown in blue and black. Periodic copies are shown in gray. The graph includes only nodes at the boundaries of the holes and edges connecting the nodes to the nearest nodes at the boundaries of the other holes.

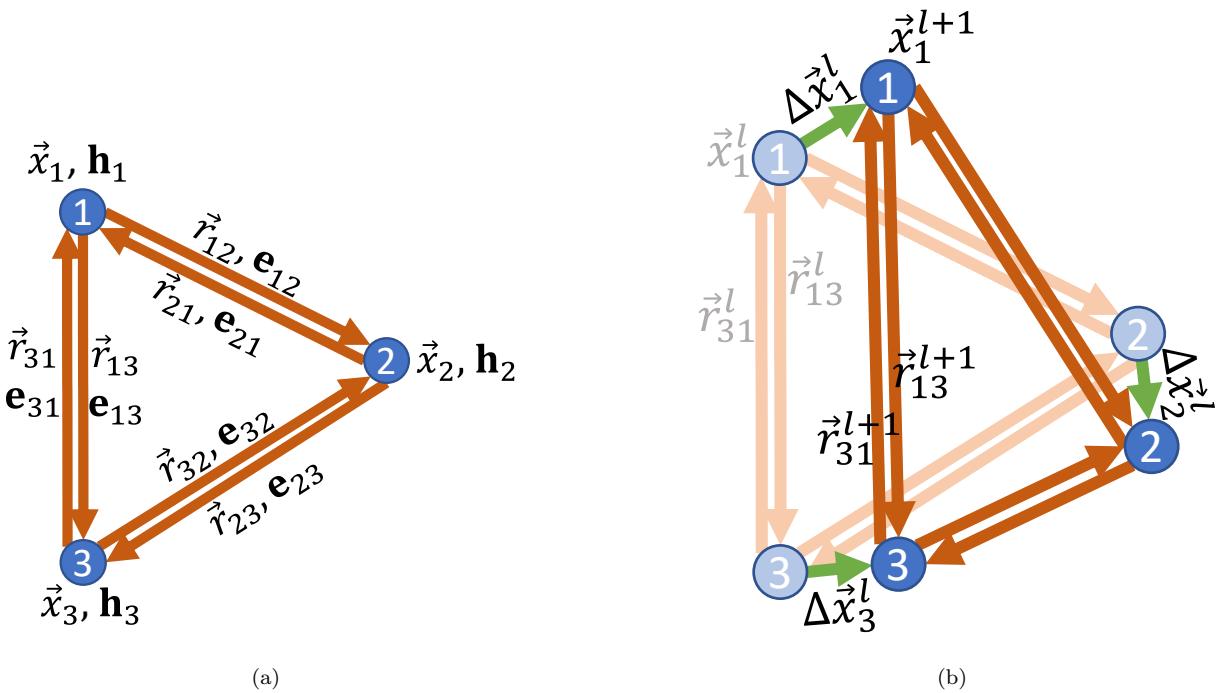


Figure 8: Definition of the graph quantities: (a) each node $i \in \mathcal{V}$ has a position \vec{x}_i (a 2D or 3D vector) and an abstract embedding \mathbf{h}_i associated with it. Each edge ij has a vector \vec{r}_{ij} pointing along it and an abstract embedding \mathbf{e}_{ij} . (b) During a message passing step l , a shift $\Delta\vec{x}_i^l$ is applied to each node, which moves the node from its old position \vec{x}_i^l to the new position $\vec{x}_i^{l+1} = \vec{x}_i^l + \Delta\vec{x}_i^l$. The edge vectors are then updated correspondingly.

most general form, this looks like

$$\mathbf{m}_i^l = \square_{j \in \mathcal{N}(i)} \phi_m^l (\mathbf{h}_i^l, \mathbf{h}_j^l, \mathbf{e}_{ij}), \quad (18)$$

$$\mathbf{h}_i^{l+1} = \phi_h^l (\mathbf{h}_i^l, \mathbf{m}_i^l), \quad (19)$$

where \square is the aggregation function (e.g. sum or average), \mathbf{m}_{ij}^l is the message from node j to i in message passing step l , $\mathcal{N}(i)$ is the set of neighbors of node i , and ϕ_m^l and ϕ_h^l are learnable functions (in our case, one linear layer plus an activation function). We also update the edge embeddings from the messages:

$$\mathbf{e}_{ij}^{l+1} = \phi_e^l (\mathbf{m}_{ij}^l), \quad (20)$$

where ϕ_e^l is another learnable function. A graph neural network generally consists of multiple message passing steps. The learnable functions are different for each message passing layer, although each layer can be repeated multiple times. In our case, this message passing scheme needs to be modified to also take into account geometric quantities like the distances r_{ij} between nodes. The exact form of the resulting message passing scheme differs for each of the three models we developed, but all have in common that they update the position \vec{x}_i^l of the nodes in each message passing step l , by computing for each node a shift in position $\Delta\vec{x}_i^l$

$$\vec{x}_i^{l+1} = \vec{x}_i^l + \Delta\vec{x}_i^l \quad (21)$$

The way $\Delta\vec{x}_i^l$ is computed differs per model.

At the end of the messages passing steps, still only local abstract quantities are available: node and edge embeddings, and the messages from the last message passing step. The way these quantities are converted into global predictions of the strain energy density \mathfrak{W} , stress \mathbf{P} and stiffness ${}^4\mathbf{D}$ also differs per model, but they all have in common that some local quantity (either an existing one or a newly created one) is being averaged. We use averaging, because it is permutation invariant with respect to the order of the nodes and edges, and is also the most common approach to homogenize local quantities to global ones in computational homogenization.

The mean square error in the prediction of each of the four target quantities ($\vec{w}, \mathfrak{W}, \mathbf{P}, {}^4\mathbf{D}$) is used in the loss. Details about training and balancing these loss terms are given in Appendix A.

Periodic Boundary Conditions

In order to properly implement periodic boundary conditions as described in Section 2, we need to construct and process the graph in such a way that the ‘Shifted RVE’ and ‘Extended RVE’ in-/equivariance shown in Figure 6 are respected. For this purpose, when constructing the input graph, we modify the computation of the edge vectors \vec{r}_{ij} that are associated with the “wraparound” edges. Instead of pointing from \vec{x}_i to \vec{x}_j , \vec{r}_{ij} then points from \vec{x}_i to the periodic image \vec{x}_j on the correct side. Because we use only \vec{r}_{ij} and the distance r_{ij} , which are independent of the actual location of the nodes, in updating the node embeddings, node positions and edge embeddings, the graph is updated independently of how the RVE is chosen.

However, these wraparound edges entail a problem if not handled carefully; their edge vectors \vec{r}_{ij} cannot be directly recalculated from the node positions \vec{x}_i . These positions \vec{x}_i are updated multiple times during the message passing steps, and in the standard implementations of the MeshGraphNets and EGNNS, the updated \vec{r}_{ij} are then recalculated from \vec{x}_i as $\vec{r}_{ij} = \vec{x}_j - \vec{x}_i$, which will be inaccurate for the wraparound edges. In order to avoid using periodic images all the time, we only calculate \vec{r} once in the beginning, after that, we keep \vec{x}_i and \vec{r}_{ij} separate, and update \vec{r}_{ij} directly using

$$\vec{r}_{ij}^{l+1} = \vec{r}_{ij}^l - \Delta\vec{x}^l + \Delta\vec{x}_j, \quad (22)$$

where $\Delta\vec{x}_i$ is the shift in node position $\vec{x}_i^{l+1} = \vec{x}_i^l + \Delta\vec{x}_i$ in message passing step l . This ensures that the updates of \vec{r}_{ij} are still independent of the absolute positions.

To ensure that in-/equivariance with respect to the transformations ‘Extended RVE’ and ‘Shifted RVE’ in Figure 6 is still preserved after applying \mathbf{F} , we apply \mathbf{F} as an affine transformation to all reference node positions \vec{x}_i^{ref} and edge vectors $\vec{r}_{ij}^{\text{ref}}$ at the same time. This provides the initial positions and edge vectors fed into the GNN:

$$\vec{x}_i^0 = \mathbf{F} \cdot \vec{x}_i^{\text{ref}}, \quad (23)$$

$$\vec{r}_{ij}^0 = \mathbf{F} \cdot \vec{r}_{ij}^{\text{ref}}. \quad (24)$$

Figure 9 shows the general structure of all three models, including the preprocessing. In the following paragraphs, the ‘message passing’ and ‘read-out’ blocks of this diagram will be described per model.

4.3. Base GNN

For the base GNN used as a baseline in this paper, we use an approach similar to MeshGraphNets [67]. For that, we use the following message computation:

$$\mathbf{m}_{ij}^l = \phi_m^l (\mathbf{h}_i^l, \mathbf{h}_j^l, \vec{r}_{ij}^{\text{ref}}, r_{ij}^{\text{ref}}, \vec{r}_{ij}^l, r_{ij}^l, \mathbf{e}_{ij}^l), \quad (25)$$

with ϕ_m^l a learnable function (in our case, one linear layer plus an activation function). This means the message computation for edge ij uses both the reference edge length $r_{ij}^{\text{ref}} = \|\vec{r}_{ij}^{\text{ref}}\|_2$ and all components of the edge vector $\vec{r}_{ij}^{\text{ref}}$ (before any deformation) and the current edge length $r_{ij}^l = \|\vec{r}_{ij}^l\|_2$ and edge vector \vec{r}_{ij}^l (after affine deformation and the message passing steps up to l), as well as the current node embeddings $\mathbf{h}_i^l, \mathbf{h}_j^l$ of both nodes, and the edge embedding \mathbf{e}_{ij}^l .

The messages \mathbf{m}_{ij}^l are then used to update the node and edge embeddings

$$\mathbf{h}_i^{l+1} = \phi_h^l \left(\mathbf{h}_i^l, \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{m}_{ij}^l \right), \quad (26)$$

$$\mathbf{e}_{ij}^{l+1} = \phi_e^l(\mathbf{m}_{ij}^l), \quad (27)$$

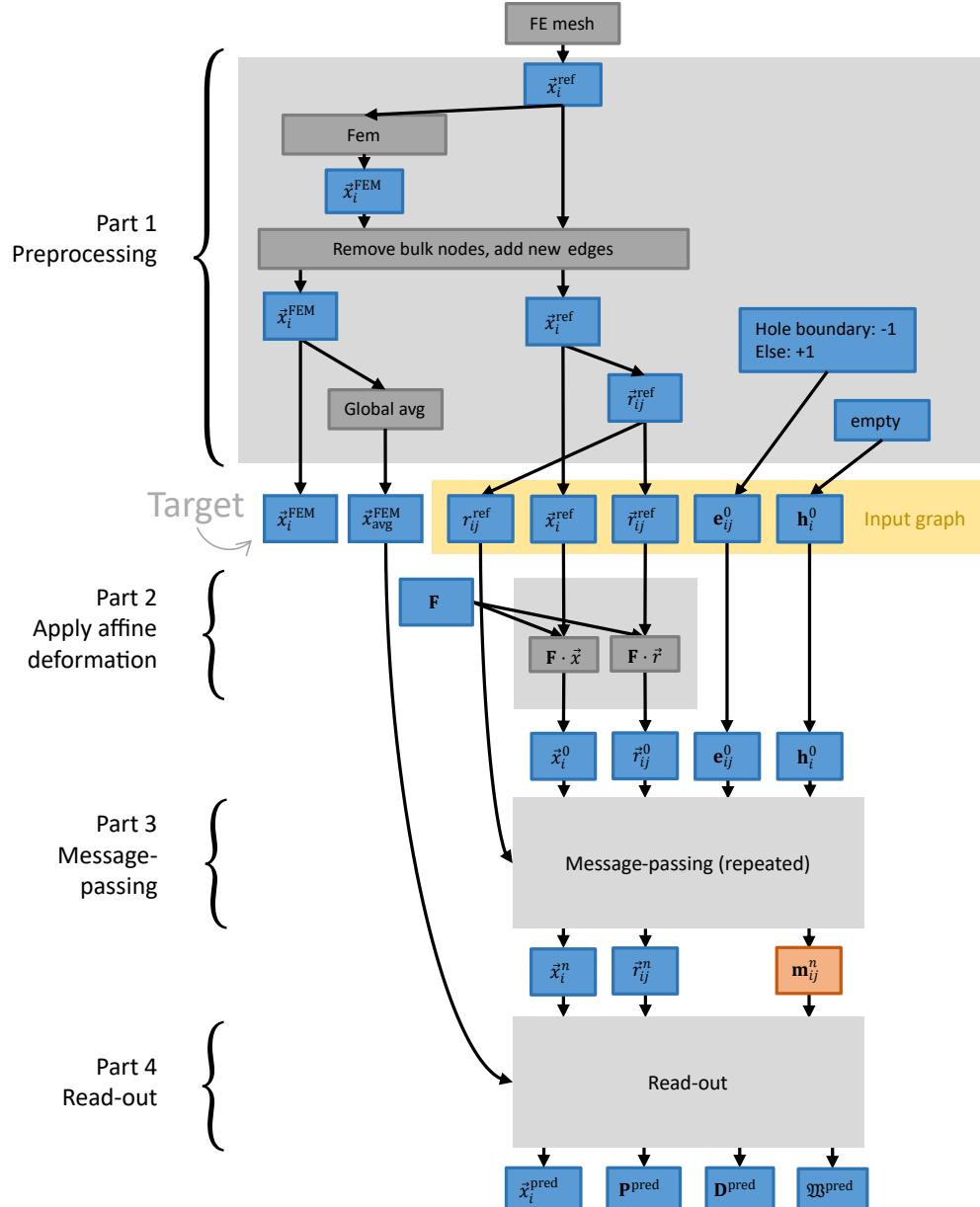


Figure 9: Diagram of the structure of all models, including the preprocessing. All blue quantities have a physical meaning, the orange quantity is abstract and generated by the neural network. The dark gray boxes indicate operations.

where $|\mathcal{N}(i)|$ is the number of neighbors of node i and ϕ_e^l and ϕ_h^l are, again, learnable functions.

From the new node embedding \mathbf{h}_i^{l+1} , during every message passing step the shift $\Delta\vec{x}_i$ in the position \vec{x}_i^l , as illustrated in Figure 8b, is calculated as

$$\Delta\vec{x}_i^l = \phi_x^l(\mathbf{h}_i^{l+1}), \quad (28)$$

with another learnable function ϕ_x^l . The edge vectors are then updated correspondingly according to Equation (22). Figure 10a provides a flowchart of the message passing scheme of this base GNN.

To predict the global quantities \mathfrak{W} , \mathbf{P} and ${}^4\mathbf{D}$, we aggregate the final messages \mathbf{m}_{ij}^n (after all n message passing steps) into a global abstract quantity, to which we then apply three final learnable functions $\phi_{\mathfrak{W}}$, ϕ_P , ϕ_D . Therefore, we predict \mathfrak{W} , \mathbf{P} and ${}^4\mathbf{D}$ with

$$\mathfrak{W} = \phi_{\mathfrak{W}} \left(\frac{1}{E} \sum_{i,j} \mathbf{m}_{ij}^n \right), \quad (29)$$

$$\mathbf{P} = \phi_P \left(\frac{1}{E} \sum_{i,j} \mathbf{m}_{ij}^n \right), \quad (30)$$

$${}^4\mathbf{D} = \phi_D \left(\frac{1}{E} \sum_{i,j} \mathbf{m}_{ij}^n \right), \quad (31)$$

where E is the total number of edges and $\phi_{\mathfrak{W}}$, ϕ_P and ϕ_D are learnable functions. See Figure 10b for a flowchart of this last part of the GNN.

In Equation (25), the model treats different components of the edge vectors simply as different inputs. In Equations (28), (30) and (31) different components of $\Delta\vec{x}$, \mathbf{P} and ${}^4\mathbf{D}$ are treated as different outputs. This implies that the network treats these components as unrelated, which means the model does not respect rotation, reflection or scaling in-equivariance. It respects only translation in-equivariance, because then all the inputs and outputs stay the same, as well as RVE in-equivariance because of the periodic boundary conditions. Hence, this model will struggle with generalizing over rotations, reflections and scaling.

4.4. EGNN

Even though all positions must be expressed in a coordinate system, the way they evolve is independent of this coordinate system. Therefore, for our second GNN, we aim to update the positions in the chosen coordinate system, in a way that is independent of that coordinate system. To this end, we use $E(n)$ -equivariant graph neural networks (EGNNs) by Satorras et al [80].⁷ This is achieved by pulling or pushing the nodes along their edges, i.e., the shift

⁷The original paper [80] also includes velocities \vec{v}_i . However, these are not relevant for our purposes, because we only consider a quasi-static system.

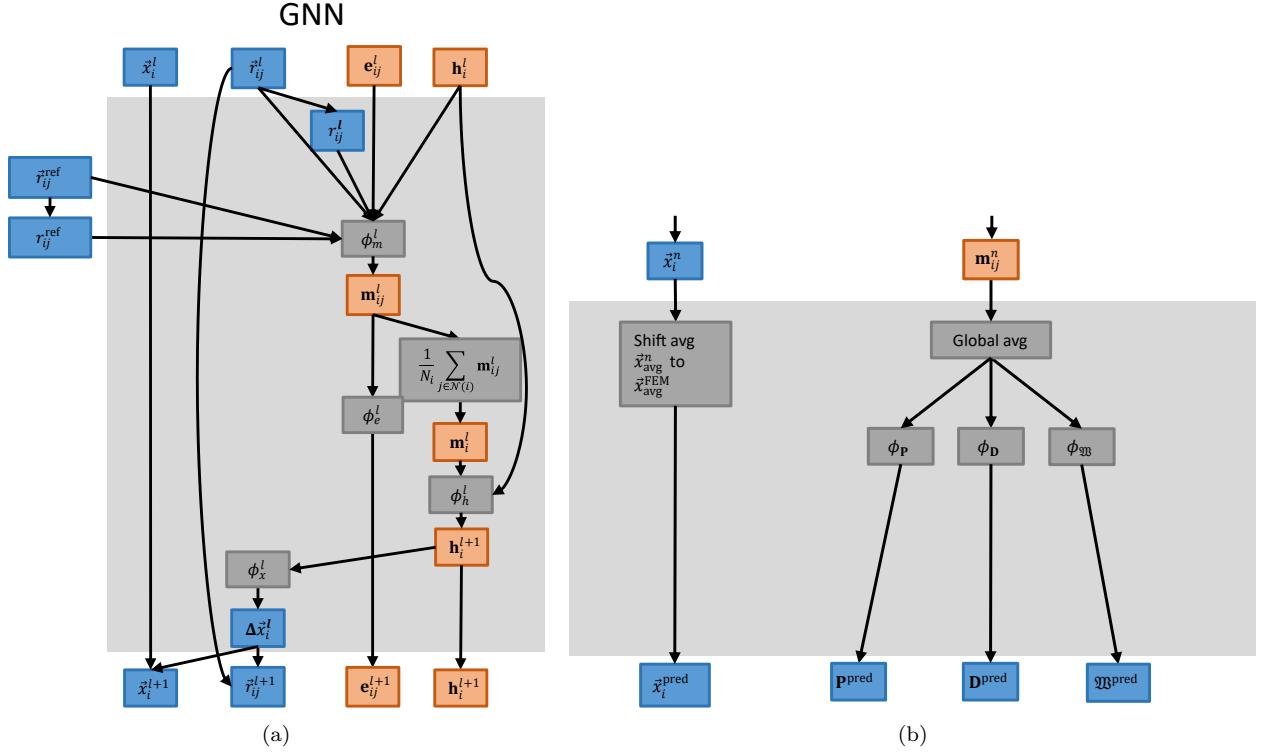


Figure 10: Diagrams of the base GNN as applied to metamaterial simulations. (a) One message passing step. (b) The read-out part, which computes final predictions for all quantities. The blue quantities have a physical meaning whereas the orange quantities are abstract and generated by the neural network (except for the initial edge and node embeddings \mathbf{e}_{ij}^0 and \mathbf{h}_i^0 , which may also have a physical meaning since they are part of the input). The dark gray boxes indicate operations, either explicitly stated or through an underlying neural network.

$\Delta\vec{x}_i$ in the node position \vec{x}_i is computed by using vectors \vec{r}_{ij} as an overcomplete basis set.⁸ The message passing scheme is described by the following relations:

$$\mathbf{m}_{ij}^l = \phi_m(\mathbf{h}_i^l, \mathbf{h}_j^l, r_{ij}^l, \mathbf{e}_{ij}), \quad (32)$$

$$\Delta\vec{x}_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \vec{r}_{ij}^l \tanh(\phi_x(\mathbf{m}_{ij}^l)), \quad (33)$$

$$\mathbf{m}_i^l = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{m}_{ij}^l, \quad (34)$$

$$\mathbf{e}_{ij}^{l+1} = \phi_e(\mathbf{h}_i^l, \mathbf{m}_i^l), \quad (35)$$

$$\mathbf{h}_i^{l+1} = \phi_h(\mathbf{h}_i^l, \mathbf{m}_i^l), \quad (36)$$

and summarized in the flowchart of Figure 11a. In the expressions above, \mathbf{m}_i^l denotes the aggregated message that node i receives. The hyperbolic tangent in Equation (33) prevents the network from shifting the nodes over large distances, which could otherwise cause an exponential blow-up. In this approach, when there is a change in the coordinate system, the edge vectors \vec{r}_{ij}^l are changed correspondingly. Consequently, since the predicted shifts in nodal positions are linear combinations of these edge vectors and the scalars $\tanh(\phi_x^l(\mathbf{m}_{ij}^l))$ are independent of the coordinate system, the shifts $\Delta\vec{x}_i^l$ also change with the coordinate system, resulting in the desired $E(n)$ -equivariance. For a proof of this, see [80].

To predict second- and fourth-order tensors (stress and stiffness, respectively) in an $E(n)$ -equivariant way, we take inspiration from the calculation of $\Delta\vec{x}_i^l$ and use the edge vectors \vec{r}_{ij} as an overcomplete basis set again, although the final ones \vec{r}_{ij}^n this time. A second-order tensor for each node i can be constructed using dyadic products of \vec{r}_{ij}^n , i.e.,

$$\mathbf{A}_i = \frac{1}{|\mathcal{N}(i)|^2} \sum_{j \in \mathcal{N}(i)} \sum_{k \in \mathcal{N}(i)} c_{i,jk} \vec{r}_{ij} \otimes \vec{r}_{ik}, \quad (37)$$

$$c_{i,jk} = \phi_A(\mathbf{m}_{ij}^n, \mathbf{m}_{ik}^n), \quad (38)$$

where $c_{i,jk}$ is a scalar, computed with a learnable layer ϕ_A , for each triplet i, j, k where node j and k are from the neighborhood $\mathcal{N}(i)$ of node i . For the maximum number of neighbors of the graph in Figure 7 $|\mathcal{N}(i)| = 8$, Equation (37) has $|\mathcal{N}(i)|^2 = 64$ terms, which is reasonable.

Theoretically, we can extend this approach to predict a fourth-order tensor as follows

$${}^4\mathbf{B}_i = \frac{1}{|\mathcal{N}(i)|^4} \sum_{j \in \mathcal{N}(i)} \sum_{k \in \mathcal{N}(i)} \sum_{l \in \mathcal{N}(i)} \sum_{m \in \mathcal{N}(i)} c_{i,jklm} \vec{r}_{ij} \otimes \vec{r}_{ik} \otimes \vec{r}_{il} \otimes \vec{r}_{im}, \quad (39)$$

$$c_{i,jklm} = \phi_B(\mathbf{m}_{ij}^n, \mathbf{m}_{ik}^n, \mathbf{m}_{il}^n, \mathbf{m}_{im}^n) \quad (40)$$

where $c_{i,jklm}$ is a scalar, computed with a learnable layer ϕ_B , for each combination of edge vectors. However, for the maximum number of neighbors $|\mathcal{N}(i)| = 8$ this results in $|\mathcal{N}(i)|^4 =$

⁸Meaning, \vec{x}_i is computed as a linear combination of \vec{r}_{ij} . This basis set is overcomplete because the number of edge vectors \vec{r}_{ij} connected to node i will always be larger than the number of dimensions.

4096 terms, which becomes computationally too expensive. For that reason, we first predict a second-order tensor \mathbf{A}'_i for each node i using the approach in Equations (37) and (38), and then use these tensors of neighboring nodes as a basis to predict a fourth-order tensor

$${}^4\mathbf{B}_i = \frac{1}{|\mathcal{N}(i)|^2} \sum_{j \in \mathcal{N}(i)} \sum_{k \in \mathcal{N}(i)} c'_{i,jk} \mathbf{A}'_j \otimes \mathbf{A}'_k, \quad (41)$$

$$c'_{i,jk} = \phi_B(\mathbf{m}_{ij}^n, \mathbf{m}_{ik}^n). \quad (42)$$

In this approach, because the predicted second and fourth-order tensors are linear combinations of dyadic/tetradic products of the edge vectors and the scalars $c_{i,jk}$ are independent of the coordinate system, the \mathbf{A}_i and ${}^4\mathbf{B}_i$ are transformed as tensors under a change in the coordinate system, resulting in the desired $E(n)$ -equivariance for these tensors as well.

For the final prediction of the global \mathbf{P} and ${}^4\mathbf{D}$, we take a global average of the per-node second- and fourth-order tensors, such that

$$\mathbf{P}_{\text{pred}} = \frac{1}{N} \sum_i {}^4\mathbf{B}_i \quad (43)$$

$${}^4\mathbf{D}_{\text{pred}} = \frac{1}{N} \sum_i {}^4\mathbf{D}_i. \quad (44)$$

The whole read-out part of the EGNN is illustrated with the flowchart in Figure 11b.

All the predictions of this model respect $E(n)$ -in-/equivariance, as well as RVE in-/equivariance because of the periodic boundary conditions. Only scale in-/equivariance is missing. Hence, this model will struggle with generalizing over scaling.

Since we are modeling a conservative system, an obvious alternative approach to predict tensors would be using Sobolev training [95] (i.e., training not just for a certain output, but also for one or more of its derivatives). This means the network would predict only the predicted energy density \mathfrak{W} . The macroscopic first Piola-Kirchhoff stress tensor and the stiffness tensor would then be calculated from \mathfrak{W} according to Equations (12) and (13). This means calculating $\mathbf{P} = \frac{\partial \mathfrak{W}}{\partial \mathbf{F}}$ and ${}^4\mathbf{D} = \frac{\partial^2 \mathfrak{W}}{\partial \mathbf{F}^2}$ using the autodifferentiation that any package for neural network training such as PyTorch is capable of, and then training for \mathfrak{W} as well as \mathbf{P} and ${}^4\mathbf{D}$. Many molecular GNNs similarly train for the energy as well as the forces [70, 73]. This approach would also guarantee that the differentiability relations in Equations 12 and 13 are exactly satisfied, which is not the case with the current approach. However, we found that for the second derivative ${}^4\mathbf{D}$ this is prohibitively slow and unstable. Presumably this is due to the repeated differentiation, which – especially over a deep network – means a small change in \mathfrak{W} can lead to a large change in ${}^4\mathbf{D}$. For the first derivative \mathbf{P} this approach is nevertheless feasible, but still not as fast or accurate as predicting \mathbf{P} directly.

4.5. Similarity-Equivariant GNN (SimEGNN)

Finally, the last architecture developed and tested in this work is SimEGNN, possessing all required in-/equivariances of Table 2. The Euclidean group $E(n)$ appearing in the

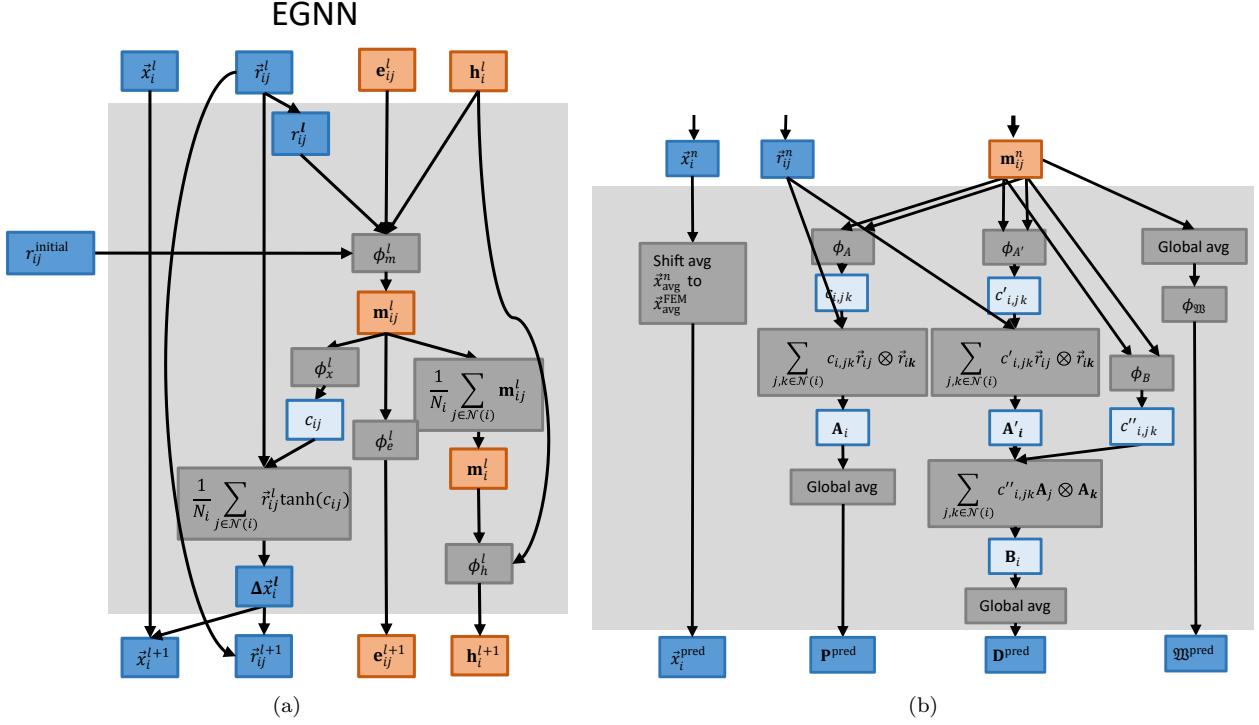


Figure 11: Diagrams of the EGNN as applied to metamaterial simulations. (a) One message passing step, which iteratively updates the node positions and node and edge embeddings, (b) the read-out part, which computes final predictions for all quantities. Blue quantities have a physical meaning, and the orange quantities are abstract quantities generated by the neural network (except for the initial edge and node embeddings e_{ij}^0 and h_i^0 , which may also have physical meaning as they are part of the input); finally, light blue quantities are not abstract but do not have a direct physical interpretation either. The dark gray boxes indicate operations, either explicitly stated or through an underlying neural network.

previous architecture includes transformations that preserve Euclidean distances (i.e., rotations, reflections, translations). However, in order to apply it to first-order homogenization problems, we also require scale in-/equivariance. The group equivariance we need is thus with respect to the similarities group S , which also includes isotropic scaling and preserves distance ratios. To achieve this, instead of the distance between nodes r_{ij}^l in the message computation of Equation (32), we use the strain measure defined as $\varepsilon_{ij}^l = (r_{ij}^l - r_{ij}^{\text{ref}})/r_{ij}^{\text{ref}}$, i.e.,

$$\mathbf{m}_{ij}^l = \phi_m^l(\mathbf{h}_i^l, \mathbf{h}_j^l, \varepsilon_{ij}^l, \mathbf{e}_{ij}^l). \quad (45)$$

Strains are also used in the computation of the shift in Equation (33), i.e.,

$$\Delta\vec{x}_i^l = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \vec{r}_{ij}^l \tanh(\varepsilon_{ij}^l \phi_x^l(\mathbf{m}_{ij}^l)), \quad (46)$$

which guarantees that there is no position shift for $\mathbf{F} = \mathbf{I}$, because then all strains ε_{ij}^l will be zero, which will make all $\Delta\vec{x}_i^l$ equal to zero. The resulting message passing scheme is shown in the flowchart of Figure 12a. Because \mathfrak{W} , \mathbf{P} and ${}^4\mathbf{D}$ are scale *invariant*, we normalize the final edge vectors \vec{r}_{ij}^n used as an overcomplete basis set to predict node-wise tensors \mathbf{A}_i (recall Equation (37)):

$$\mathbf{A}_i = \frac{1}{|\mathcal{N}(i)|^2} \sum_{j \in \mathcal{N}(i)} \sum_{k \in \mathcal{N}(i)} c_{ijk} \hat{\vec{r}}_{ij} \otimes \hat{\vec{r}}_{ik}, \quad (47)$$

where $\hat{\vec{r}}_{ij} = \vec{r}_{ij}/r_{ij}$ are the normalized edge vectors. This read-out part of the SimEGNN is illustrated with the flowchart in Figure 12b.

5. Results

The model evaluation consists of testing 7 different versions of the 3 graph neural networks described in Section 4, from the smallest to the largest number of symmetries, on 6 different test cases, for 4 target quantities. We test them on the homogenization problem introduced in Section 2. In addition to the three types of GNN architectures, for the GNN and EGNN, we tested augmenting the training data as described in Section 2.3, as a different way to enforce the symmetries. The models trained on augmented data are marked ‘DA’ for ‘Data Augmentation’. Here, ‘DA $\times 1$ ’ means the training data consisted of one transformed version of each data point, and ‘DA $\times 2$ ’ means it consisted of two transformed versions of each data point (i.e., meaning the data set was effectively twice as large). This resulted in 7 models in total: 3 GNNs, 3 EGNNs and 1 SimEGNN. The results cover 6 different test cases:

1. the validation data (labeled ‘untransformed’, where none of the configurations have been rotated, reflected, translated or scaled),
2. the validation data reflected in the y -axis (‘reflected’),
3. rotated by $\pi/4$ radians (‘rotated’),

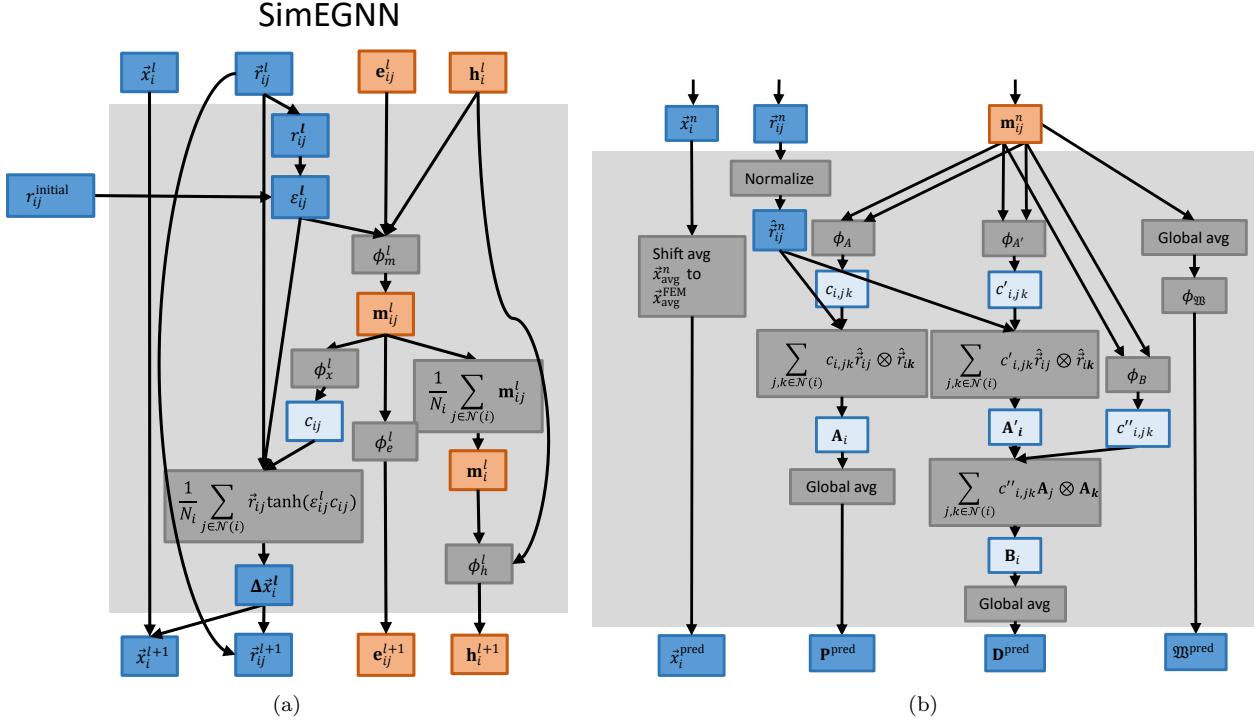


Figure 12: Diagrams of the SimEGNN as applied to metamaterial simulations. (a) One message passing step, which iteratively updates the node positions and node and edge embeddings, (b) the read-out part, which computes final predictions for all quantities. Blue quantities have a physical meaning, and the orange quantities are abstract quantities generated by the neural network (except for the initial edge and node embeddings \mathbf{e}_{ij}^0 and \mathbf{h}_i^0 , which may also have a physical meaning since they are part of the input); finally, light blue quantities are not abstract but do not have a direct physical interpretation either. The dark gray boxes indicate operations, either explicitly stated or through an underlying neural network.

4. with a shifted RVE, which means the material is shifted up and left by one quarter the size of the RVE, such that the RVE is centered on one hole (labeled ‘shifted RVE’),
5. with a larger RVE that consists of four original RVEs merged together (‘extended RVE’),
6. and an RVE which has been scaled by a factor 1.5 (‘scaled’).

These test cases represent the various in-/equivariances of Figure 6 and Table 2 that we built into the model, and were chosen to evaluate the effect of embedding these in-/equivariances. The 4 different target quantities are the microfluctuation field $\vec{w} = \vec{x} - \mathbf{F} \cdot \vec{x}^{\text{ref}}$ (see Equation 10), homogenized strain energy density \mathfrak{W} , homogenized stress \mathbf{P} and tangent stiffness tensor ${}^4\mathbf{D}$.

The spider web charts in Figure 13 show the fraction of variance unexplained (FVU) – i.e., the variance in the errors of quantities predicted by the models divided by the variance in the data of the target quantity⁹ – obtained by each neural network architecture. Each chart corresponds to one target quantity: \vec{w} , \mathfrak{W} , \mathbf{P} or \mathbf{D} . The investigated models are shown in different line colors and styles, where the color indicates the type of neural network (GNN, EGNN or SimEGNN) and the line style (solid, dashed or dotted) indicates the amount of data augmentation. The different spokes of the spider plots correspond to the six test cases (reference, reflected, rotated, shifted RVE, larger RVE, scaled).

Additionally, for four of these six test cases, we show a comparison of the predicted hole deformations from three of the networks against FEM outputs – considered the ground truth in this work – for biaxial compression in Figure B.16. We exclude the ‘Shifted RVE’ and ‘Extended RVE’ cases, because all models respect periodicity. The Figure only shows the networks trained without data augmentation. Tables with the data behind Figure 13 are provided in Appendix B, as well as tables with the relative error. This appendix also contains figures similar to Figure B.16 for different \mathbf{F} , such that all bifurcation patterns are included.

The results of Figure 13 show that the periodicity is correctly implemented in all models, because all of them can predict the ‘shifted RVE’ and ‘larger RVE’ cases with an FVU equal to the untransformed case. By construction, the EGNN indeed additionally respects $E(n)$ -in-/equivariance (‘reflected’ and ‘rotated’, translated not shown in Figure 13), and the SimEGNN additionally respects scale in-/equivariance.¹⁰ The figure also shows that the microfluctuation \vec{w} and stiffness ${}^4\mathbf{D}$ are generally more difficult to predict than the strain energy density \mathfrak{W} and the stress \mathbf{P} ; in the plots for \vec{w} and ${}^4\mathbf{D}$, the FVU stays above 10^{-4} , whereas in the plots for \mathfrak{W} and \mathbf{P} the FVU gets close to 10^{-6} .

For GNNs (the blue lines in Figure 13), augmenting the data with reflected, rotated and scaled cases improves the predictions for those cases, i.e., it makes the result more similar to the ‘untransformed’ result. However, it worsens the ‘untransformed’ result at the same

⁹i.e., $1 - R^2$, where R^2 is the coefficient of determination

¹⁰The EGNN and SimEGNN still show a small deviation in the FVU for the rotated and reflected cases for the stress \mathbf{P} and stiffness ${}^4\mathbf{D}$; this is because those transformations slightly change the variance in the ground truth data. The mean square error does remain the same, however.

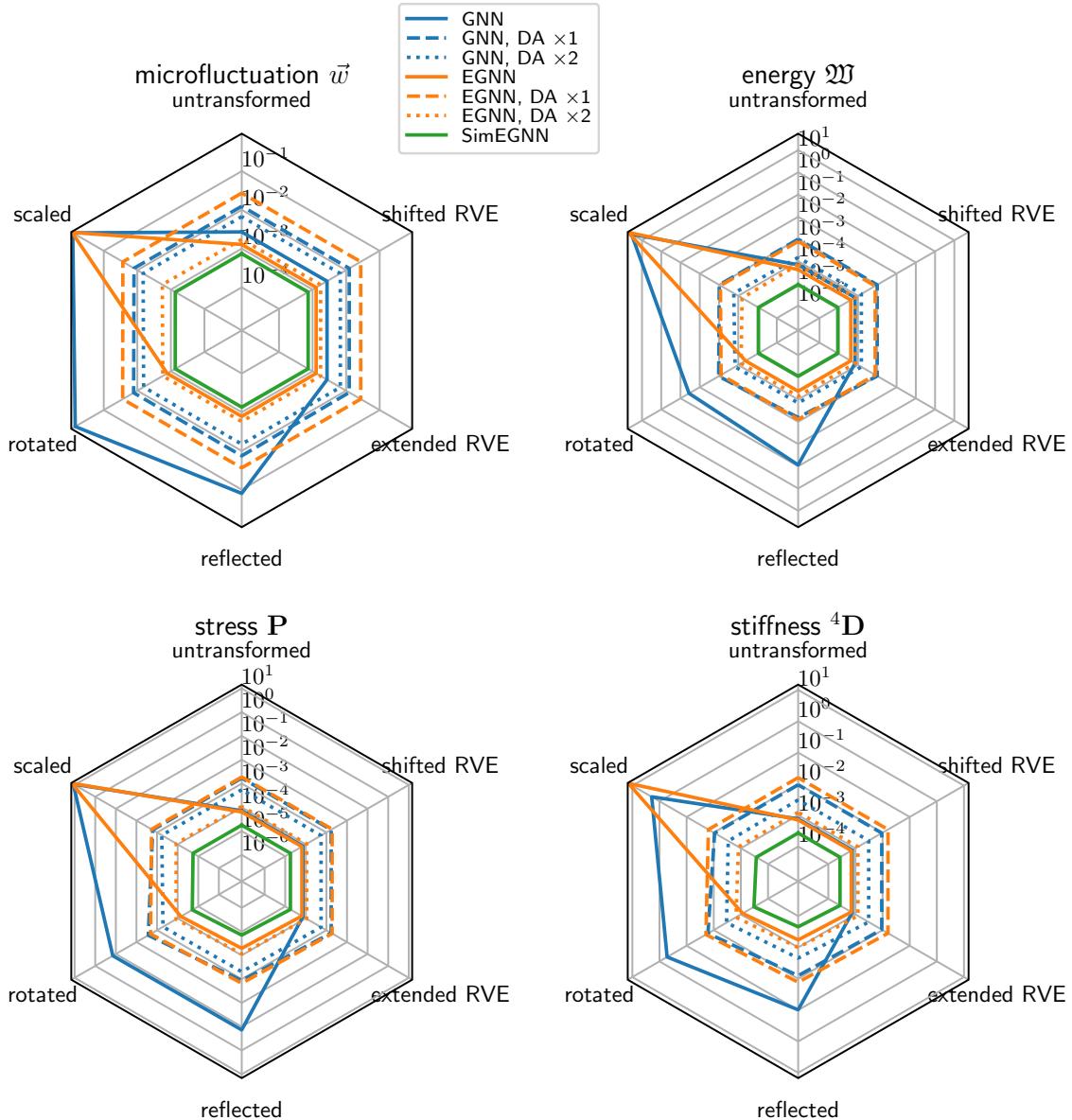


Figure 13: Fraction of variance unexplained of the four target quantities ($\vec{w}, \mathfrak{W}, \mathbf{P}, {}^4\mathbf{D}$) achieved by the investigated models (shown in different line colors and styles) for the six test cases (untransformed, reflected, rotated, shifted RVE, larger RVE, scaled). The SimEGNN is the only model that respects all symmetries, and it outperforms the other models.

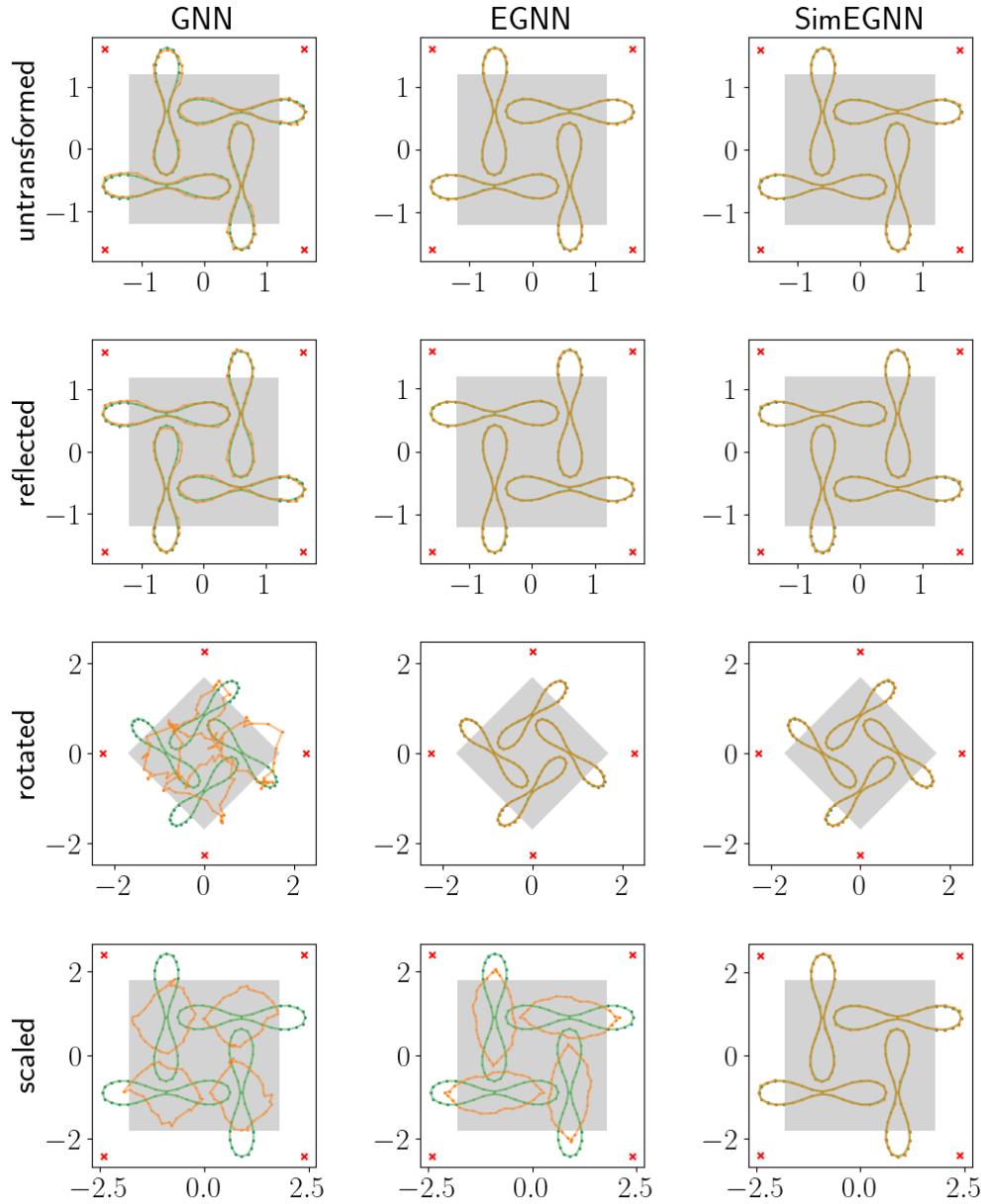


Figure 14: Predicted deformation, using the three models without data augmentation, of the hole boundaries (orange) compared to the FEM ground truth (green), for $\mathbf{F} = \begin{bmatrix} 0.75 & 0 \\ 0 & 0.75 \end{bmatrix}$ (biaxial compression, resulting in a rotational pattern).

time. This indicates that the model is trying to learn all the equivalent cases separately, which effectively reduces the capacity of the network. A similar effect can be seen for the EGNN (orange lines in Figure 13) and the ‘scaled’ case.

Figure 13 shows that with all symmetries incorporated directly into the network’s architecture, the FVU improves by about an order of magnitude or more, compared to the GNN with $2\times$ data augmentation, on all target quantities. This holds for all test cases, even for the untransformed case, where none of the configurations have been rotated, reflected, translated or scaled. Incorporating more symmetries thus makes it easier for the model to construct and use features in a general way. For example, it can more easily transfer learned features from one edge to predict the behavior of another longer edge with scale in/equivariance. Embedding the symmetries thus allows the network to focus on constructing actually meaningful features, because it does not first have to filter out the aspects that are irrelevant (scale, rotation, etc.). Consequently, the newly proposed SimEGNN outperforms the other models in all aspects.

Efficiency and Scaling

The SimEGNN trains just as fast as the GNN and EGNN and is approximately 8 times faster than the finite element method, if both are implemented on a single-threaded CPU. Both the finite element method and the GNN can be made faster using a GPU (and in fact, we did train our GNN on a GPU), but since our finite element code is not GPU-compatible yet, we cannot make a fair comparison. However, in general, machine learning methods tend to be very GPU-friendly. It is likely that our SimEGNN approach can also be made faster in the most obvious way: by lowering our standards. The current model is remarkably accurate. If we require a maximum relative error of about 5% in the homogenized target quantities (instead of the current 0.31% – 0.90%), a much smaller model will probably also suffice. We also expect our method to scale much better when applied to more complicated or 3D geometries, because we only use the boundary nodes. This means the number of used nodes scales linearly with the mesh size, while the number of nodes in the finite element mesh scales quadratically (in 2D). In 3D the number of boundary nodes scales quadratically and the number of bulk nodes scales to the power 3, hence we expect the SimEGNN to scale much better than the finite element method. It is also possible to perform the FE training simulations on a fine mesh, but approximate those using a coarser mesh for the training and prediction with the GNNs, to exploit the accuracy of a fine mesh while limiting the computational cost. The main limitation of the presented approach is the need for at least as many message passing steps as the graph diameter, although this can possibly be addressed by, e.g., using pooling strategies and adding long-distance edges to the graph.

6. Conclusion

For the development of new hyperelastic, flexible, porous, 2D mechanical metamaterials with complex microstructures, fast simulations of the homogenized response are needed. Conventional finite element simulations are often too slow for this purpose. In this paper, we present a graph neural network architecture called a Similarity-Equivariant Graph Neural

Network (SimEGNN) to speed up simulations of these materials. The model is trained on finite element simulations and incorporates all relevant symmetries. The network predicts the final node positions \vec{x} by repositioning the nodes during message passing. From the final graph, the homogenized strain energy density \mathfrak{W} , homogenized stress \mathbf{P} and tangent stiffness tensor ${}^4\mathbf{D}$ are predicted, which serve as fast surrogates in computational homogenization schemes.

All necessary in-/equivariances pertinent to computational homogenization are incorporated. To achieve this, we use $E(n)$ -equivariant graph neural networks [80] as the initial stepping stone and extend them to also (i) respect periodicity (i.e., RVE in-/equivariance), which we achieve by using a periodic graph (i.e., a Pac-Man world geometry) and coordinate-independent updates, (ii) respect scale in-/equivariance, for which we use strains instead of distances and normalize the edge vectors \vec{r}_{ij} where necessary, and (iii) equivariantly output higher-order tensors, by using normalized edge vectors $\hat{\vec{r}}_{ij}$ as an overcomplete basis for predicting \mathbf{P} and ${}^4\mathbf{D}$.

We demonstrated that the SimEGNN approximates the results of material simulations (both micro- and macroscopic response) in first-order homogenization with very high accuracy (i.e. $FVU < 10^{-3}$, or $R^2 > 0.999$ in all cases), even for cases with large deformations. Our numerical experiments also confirm that SimEGNN outperforms the architectures with fewer built-in in-/equivariances (base GNN and EGNN). Even though some symmetries can also be trained for in the GNN and EGNN without explicitly incorporating them (using data augmentation), training for them is achieved at the expense of a decreased performance and a bigger dataset, which also leads to slower training.

The fact that this network takes the geometry as an input means it should be able to generalize with respect to geometry, i.e., it can predict the response of unseen geometries. Our initial studies confirm our expectations; although more detailed investigations are needed, which are left for future study. The SimEGNN’s high accuracy and potential to generalize with respect to geometry make this model highly promising for the future development of new soft porous mechanical metamaterials.

In the future, we aim to test the newly developed SimEGNN on a wider variety of microstructures, to investigate its generalizability, such that we can use one model for a wide range of microstructures, and also predict on new, unseen microstructures. We will also address the main limitation of our current approach, which is the need for at least as many message passing steps as the graph diameter. The current model can only use one type of vector as input (the edge vectors). To extend the model such that it can handle more vectors and/or tensors as input, we investigate ways to increase model expressiveness, by allowing the features to be similarity-equivariant instead of similarity-invariant (similar to [77, 79]). See [96] for a more in-depth look at graph neural network expressiveness. We also investigate approaches based on generative models to predict all possible bifurcations.

7. Acknowledgements

This project has received funding from the Eindhoven Artificial Intelligence Institute (E AISI). MD acknowledges the support of the Czech Science Foundation through Project

No. 19-26143X. The authors furthermore acknowledge the High Performance Computing Lab of Eindhoven University of Technology for their support related to numerical experiments, and the NWO grant (number EINF-5845) for compute resources on the Dutch National Supercomputer Snellius.

Appendix A. Implementation

All GNN architectures involve message passing steps; recall Figures 10a, 11a and 12a. One can reuse the parameters (i.e., the ones parametrizing ϕ_m , ϕ_e , ϕ_h and ϕ_x in aforementioned Figures) of one message passing step for another step. In other words, one message passing ‘layer’ can be used for multiple message passing steps. In this work, each network has five distinct message passing layers, out of which the middle three layers are repeated three times each. This provides an adequate balance between the number of parameters and the expressiveness of the network. This results in eleven message passing steps in total. This is close to the graph diameter, which is nine for our input graph. The message size is set to 64 in all networks and the embedding sizes (edge and node embedding) at 32. ϕ_m , ϕ_e , ϕ_h and ϕ_x each only have one layer and use the softplus activation function. This setup resulted in the following number of learnable parameters: 76351 for base GNN, 91465 parameters for EGNN, and 91145 for SimEGNN.

The networks are implemented in PyTorch (<https://pytorch.org/>) and PyTorch Geometric (<https://pytorch-geometric.readthedocs.io/en/latest/>). For the EGNN and SimEGNN, each message passing step, described by Equations (32) to (36) and Equations (45) and (46), can be implemented as two sequential message passing steps. In the first message passing step, node embeddings \mathbf{h}_i are updated using computed messages \mathbf{m}_{ij} , which are aggregated into \mathbf{m}_i . In the second message passing step, \vec{x}_i^l is updated using ‘messages’ $\vec{r}_{ij}^l \tanh(\phi_x^l(\mathbf{m}_{ij}^l))$ or $\vec{r}_{ij}^l \tanh(\phi_x^l(\varepsilon_{ij}\mathbf{m}_{ij}^l))$ that are aggregated into $\Delta\vec{x}_i^l$. Accordingly, the networks are implemented such that each message passing layer contains two different PyTorch Geometric message passing layers that are alternated. (For the GNN, one type of message-passing layer is sufficient.)

For training all learnable parameters, we use the Adam optimizer, with a schedule that reduces the learning rate in steps from 2.5×10^{-4} to 2.5×10^{-6} . In total, we train for 1620 epochs, with a batch size of 12, which takes 6 to 10 hours per model on an NVIDIA A100-SXM4-40GB GPU. All outputs except the position of the nodes are scaled such that their target values have an average squared value of 1. The mean square error of each of the four outputs (position \vec{w} , \mathfrak{W} , \mathbf{P} and ${}^4\mathbf{D}$) is used in the loss. The loss terms are weighted proportionally to the inverse of the square value of the target values (which for all loss terms except \vec{w} was 1, because of the aforementioned scaling). See Figure A.15 for a plot of the validation loss during training for each model.

Appendix B. Full results

Tables B.3, B.4, B.5, B.6 show, for each target quantity respectively, the fraction of variance unexplained (FVU) obtained by each neural network architecture for each test case.

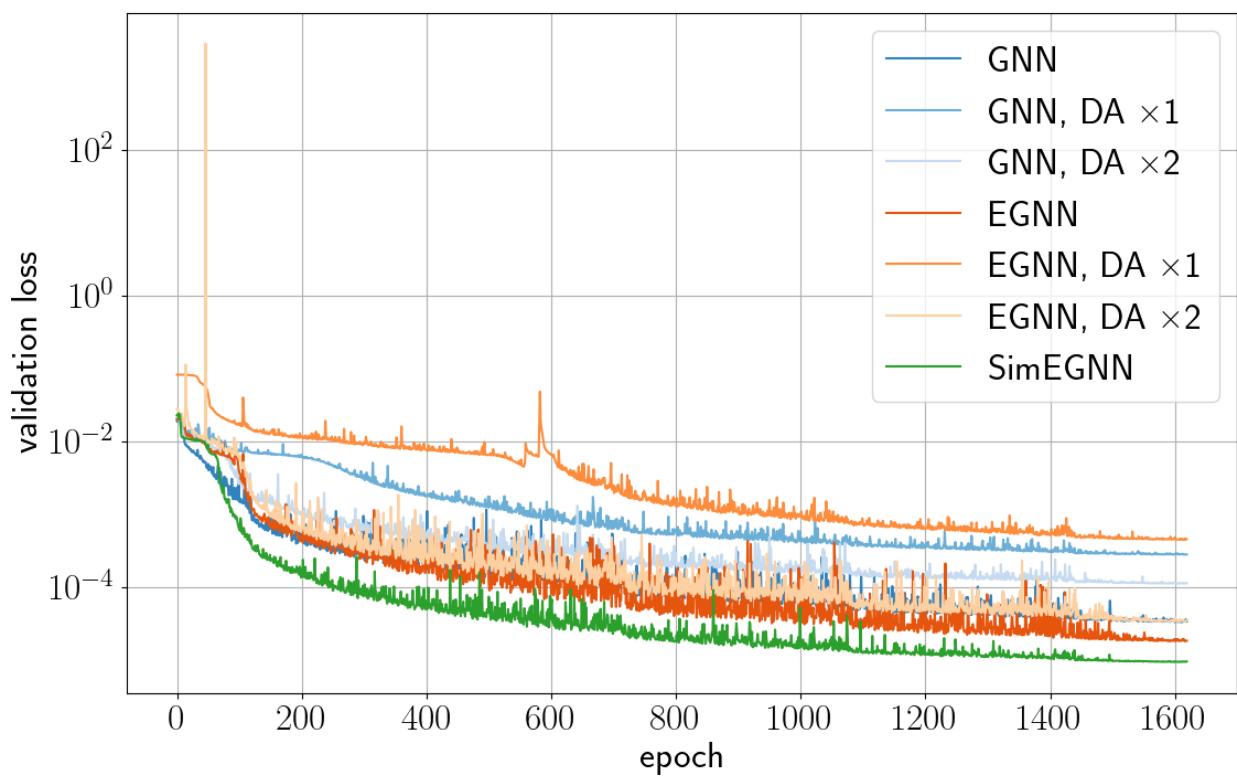


Figure A.15: The validation loss during training for each model.

Tables B.7, B.8, B.9, B.10 show the relative error, which we define as the mean Frobenius norm of the error divided by the mean Frobenius norm of the target

$$\frac{\frac{1}{N} \sum_i^N \|\mathbf{A}_{\text{pred},i} - \mathbf{A}_{\text{target},i}\|_F}{\frac{1}{N} \sum_i^N \|\mathbf{A}_{\text{target},i}\|_F} \times 100\%, \quad (\text{B.1})$$

where $\mathbf{A}_{\text{pred},i}$ is some predicted tensor, $\mathbf{A}_{\text{target},i}$ is the ground truth value of this tensor, $\|\dots\|_F$ is the Frobenius norm and N is the number of targets in the validation data, which for the global quantities \mathfrak{W} , \mathbf{P} , ${}^4\mathbf{D}$ is equal to the number of graphs, but for the local quantity \vec{w} is equal to the total number of nodes (number of graphs \times number of nodes per graph). For a scalar quantity such as \mathfrak{W} , this reduces to the root mean squared error divided by the root mean square target value.

In addition to Figure 14 in Section 5, Figures B.16, B.17, B.18, B.19 and B.20 show for all test cases a comparison of predicted deformations from all of the networks against FEM outputs. Each figure shows a different loading \mathbf{F} , such that all bifurcation patterns are included.

Table B.3: FVU of the microfluctuation \vec{w}

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	0.0027	0.012	0.0067	0.0013	0.027	0.0017	7.4×10^{-4}
noisy distances	0.0027	0.012	0.0067	0.0013	0.027	0.0017	7.4×10^{-4}
shifted RVE	0.0027	0.012	0.0067	0.0013	0.027	0.0017	7.4×10^{-4}
extended RVE	0.0027	0.012	0.0067	0.0013	0.027	0.0017	7.4×10^{-4}
reflected	0.12	0.013	0.0067	0.0013	0.027	0.0017	7.4×10^{-4}
rotated	0.71	0.013	0.0065	0.0013	0.027	0.0017	7.4×10^{-4}
scaled	0.84	0.013	0.0067	0.88	0.027	0.0018	7.4×10^{-4}

 Table B.4: FVU of the strain energy density \mathfrak{W}

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	7.0×10^{-5}	9.9×10^{-4}	1.5×10^{-4}	4.2×10^{-5}	7.8×10^{-4}	7.1×10^{-5}	8.9×10^{-6}
noisy distances	7.0×10^{-5}	9.9×10^{-4}	1.5×10^{-4}	4.2×10^{-5}	7.8×10^{-4}	7.1×10^{-5}	8.9×10^{-6}
shifted RVE	7.0×10^{-5}	9.9×10^{-4}	1.5×10^{-4}	4.2×10^{-5}	7.8×10^{-4}	7.1×10^{-5}	8.9×10^{-6}
extended RVE	7.0×10^{-5}	9.9×10^{-4}	1.5×10^{-4}	4.2×10^{-5}	7.8×10^{-4}	7.1×10^{-5}	8.9×10^{-6}
reflected	0.087	6.4×10^{-4}	1.4×10^{-4}	4.2×10^{-5}	7.8×10^{-4}	7.1×10^{-5}	8.9×10^{-6}
rotated	0.036	0.0010	1.5×10^{-4}	4.2×10^{-5}	7.8×10^{-4}	7.1×10^{-5}	8.9×10^{-6}
scaled	35.	9.7×10^{-4}	1.7×10^{-4}	52.	8.5×10^{-4}	6.4×10^{-5}	8.9×10^{-6}

 Table B.5: FVU of the first Piola-Kirchhoff stress tensor \mathbf{P}

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	7.2×10^{-5}	0.0017	5.5×10^{-4}	6.4×10^{-5}	0.0019	1.1×10^{-4}	1.8×10^{-5}
noisy distances	7.2×10^{-5}	0.0017	5.5×10^{-4}	6.4×10^{-5}	0.0019	1.1×10^{-4}	1.8×10^{-5}
shifted RVE	7.2×10^{-5}	0.0017	5.5×10^{-4}	6.4×10^{-5}	0.0019	1.1×10^{-4}	1.8×10^{-5}
extended RVE	7.2×10^{-5}	0.0017	5.5×10^{-4}	6.4×10^{-5}	0.0019	1.1×10^{-4}	1.8×10^{-5}
reflected	0.14	0.0011	4.9×10^{-4}	5.1×10^{-5}	0.0015	8.7×10^{-5}	1.4×10^{-5}
rotated	0.14	0.0027	5.1×10^{-4}	7.0×10^{-5}	0.0021	1.2×10^{-4}	1.9×10^{-5}
scaled	12.	0.0016	5.9×10^{-4}	14.	0.0018	1.2×10^{-4}	1.8×10^{-5}

 Table B.6: FVU of the stiffness tensor ${}^4\mathbf{D}$

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	7.6×10^{-4}	0.0093	0.0031	7.1×10^{-4}	0.016	0.0012	2.7×10^{-4}
noisy distances	7.6×10^{-4}	0.0093	0.0031	7.1×10^{-4}	0.016	0.0012	2.7×10^{-4}
shifted RVE	7.6×10^{-4}	0.0093	0.0031	7.1×10^{-4}	0.016	0.0012	2.7×10^{-4}
extended RVE	7.6×10^{-4}	0.0093	0.0031	7.1×10^{-4}	0.016	0.0012	2.7×10^{-4}
reflected	0.098	0.008	0.0022	5.7×10^{-4}	0.013	0.0010	2.2×10^{-4}
rotated	0.52	0.016	0.0034	8.6×10^{-4}	0.019	0.0015	3.3×10^{-4}
scaled	1.9	0.0093	0.0029	14.	0.016	0.0012	2.7×10^{-4}

Table B.7: Relative error of the microfluctuation \vec{w}

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	5.1%	12.%	8.1%	3.5%	17.%	4.2%	2.6%
shifted RVE	5.1%	12.%	8.1%	3.5%	17.%	4.2%	2.6%
extended RVE	5.1%	12.%	8.1%	3.5%	17.%	4.2%	2.6%
reflected	31.%	12.%	8.1%	3.5%	17.%	4.2%	2.6%
rotated	93.%	12.%	8.3%	3.5%	17.%	4.2%	2.6%
scaled	100.%	12.%	8.1%	94.%	16.%	4.2%	2.6%

Table B.8: Relative error of the strain energy density \mathfrak{W}

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	0.76%	3.0%	1.2%	0.65%	2.9%	0.79%	0.31%
shifted RVE	0.76%	3.0%	1.2%	0.65%	2.9%	0.79%	0.31%
extended RVE	0.76%	3.0%	1.2%	0.65%	2.9%	0.79%	0.31%
reflected	35.%	2.4%	1.1%	0.65%	2.9%	0.79%	0.31%
rotated	18.%	2.9%	1.2%	0.65%	2.9%	0.79%	0.31%
scaled	88.%	3.0%	1.2%	89.%	2.8%	0.76%	0.31%

Table B.9: Relative error of the first Piola-Kirchhoff stress tensor \mathbf{P}

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	0.83%	4.3%	2.4%	0.82%	4.5%	1.1%	0.39%
shifted RVE	0.83%	4.3%	2.4%	0.82%	4.5%	1.1%	0.39%
extended RVE	0.83%	4.3%	2.4%	0.82%	4.5%	1.1%	0.39%
reflected	33.%	4.2%	2.5%	0.82%	4.5%	1.1%	0.39%
rotated	35.%	5.1%	2.2%	0.82%	4.5%	1.1%	0.39%
scaled	82.%	4.2%	2.5%	84.%	4.1%	1.1%	0.39%

Table B.10: Relative error of the stiffness tensor ${}^4\mathbf{D}$

	GNN	GNN, DA $\times 1$	GNN, DA $\times 2$	EGNN	EGNN, DA $\times 1$	EGNN, DA $\times 2$	SimEGNN
untransformed	1.5%	6.7%	3.8%	1.5%	8.5%	2.2%	0.90%
shifted RVE	1.5%	6.7%	3.8%	1.5%	8.5%	2.2%	0.90%
extended RVE	1.5%	6.7%	3.8%	1.5%	8.5%	2.2%	0.90%
reflected	21.%	7.1%	3.6%	1.5%	8.5%	2.2%	0.90%
rotated	55.%	8.3%	3.6%	1.5%	8.5%	2.2%	0.90%
scaled	58.%	6.8%	3.8%	80.%	8.6%	2.2%	0.90%

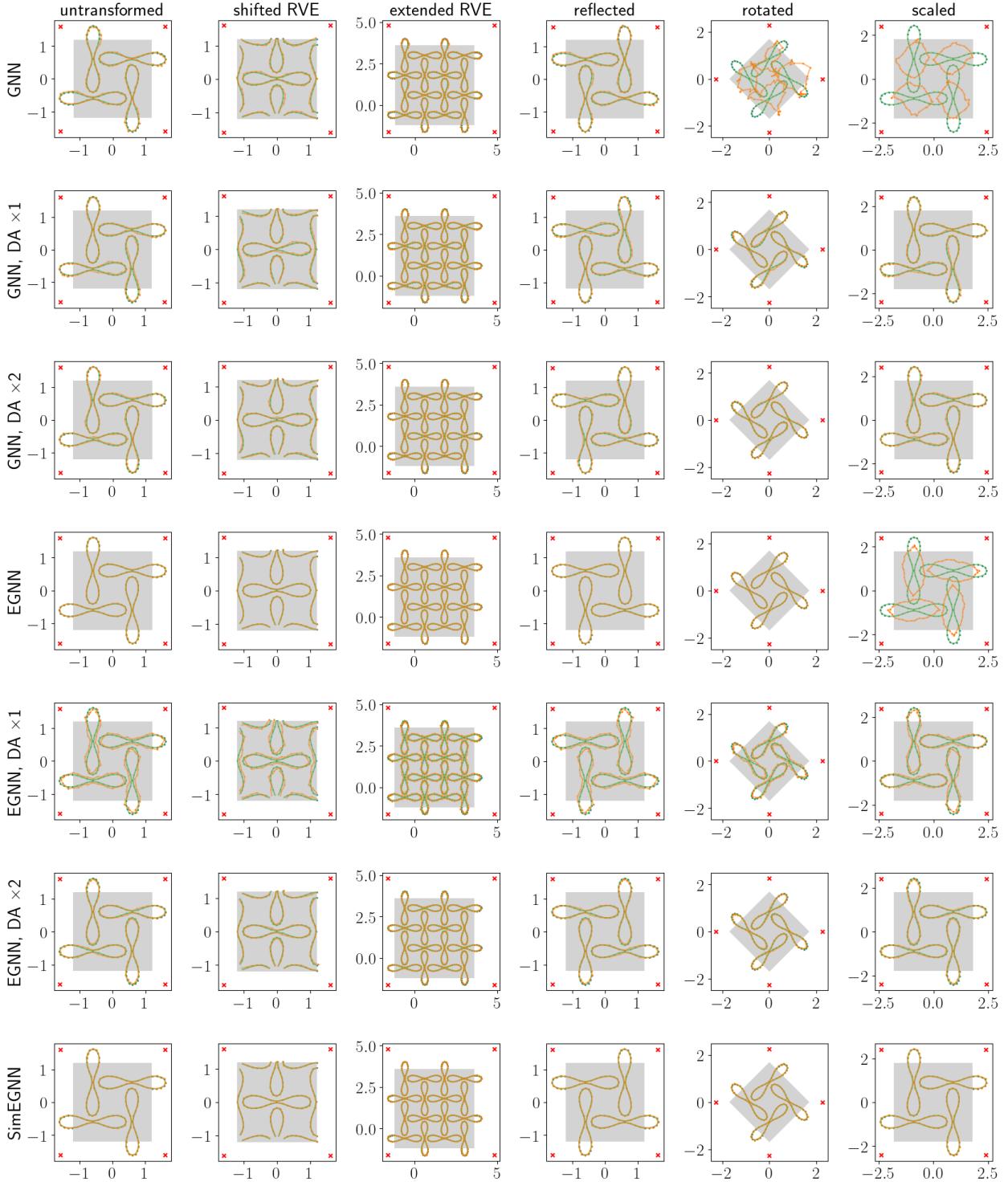


Figure B.16: Predicted deformation of the hole boundaries (orange) compared to the FEM ground truth (green), for $\mathbf{F} = \begin{bmatrix} 0.75 & 0 \\ 0 & 0.75 \end{bmatrix}$ (biaxial compression, resulting in a rotational pattern). The grey square indicates how \mathbf{F} deforms the square RVE through an affine transformation. The red crosses indicate the original positions of the corners.

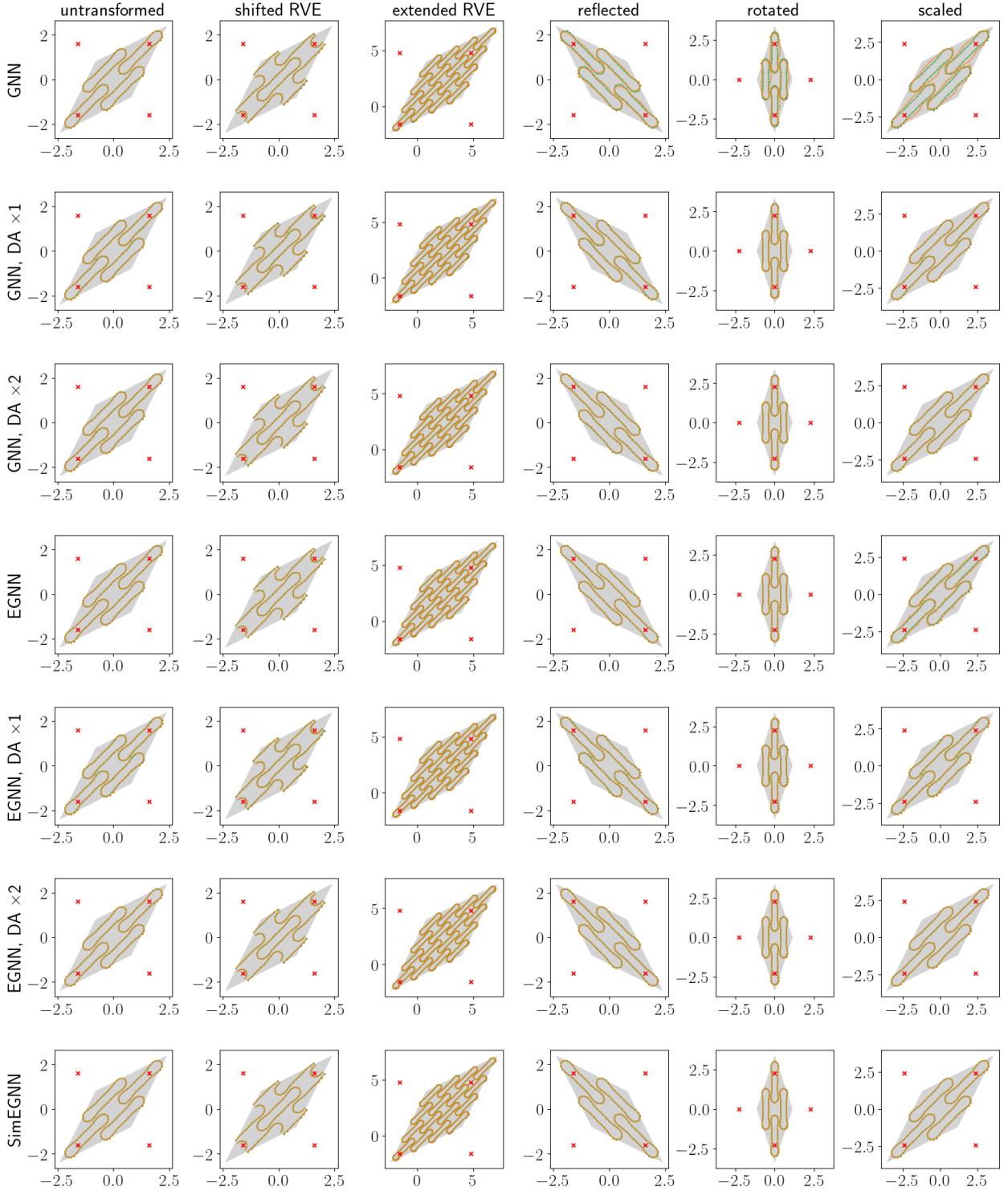


Figure B.17: Model-predicted deformation (orange) compared to the FEM ground truth (green), for $\mathbf{F} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ (pure shear). The grey square indicates how \mathbf{F} deforms the square RVE through an affine transformation. The red crosses indicate the original positions of the corners.

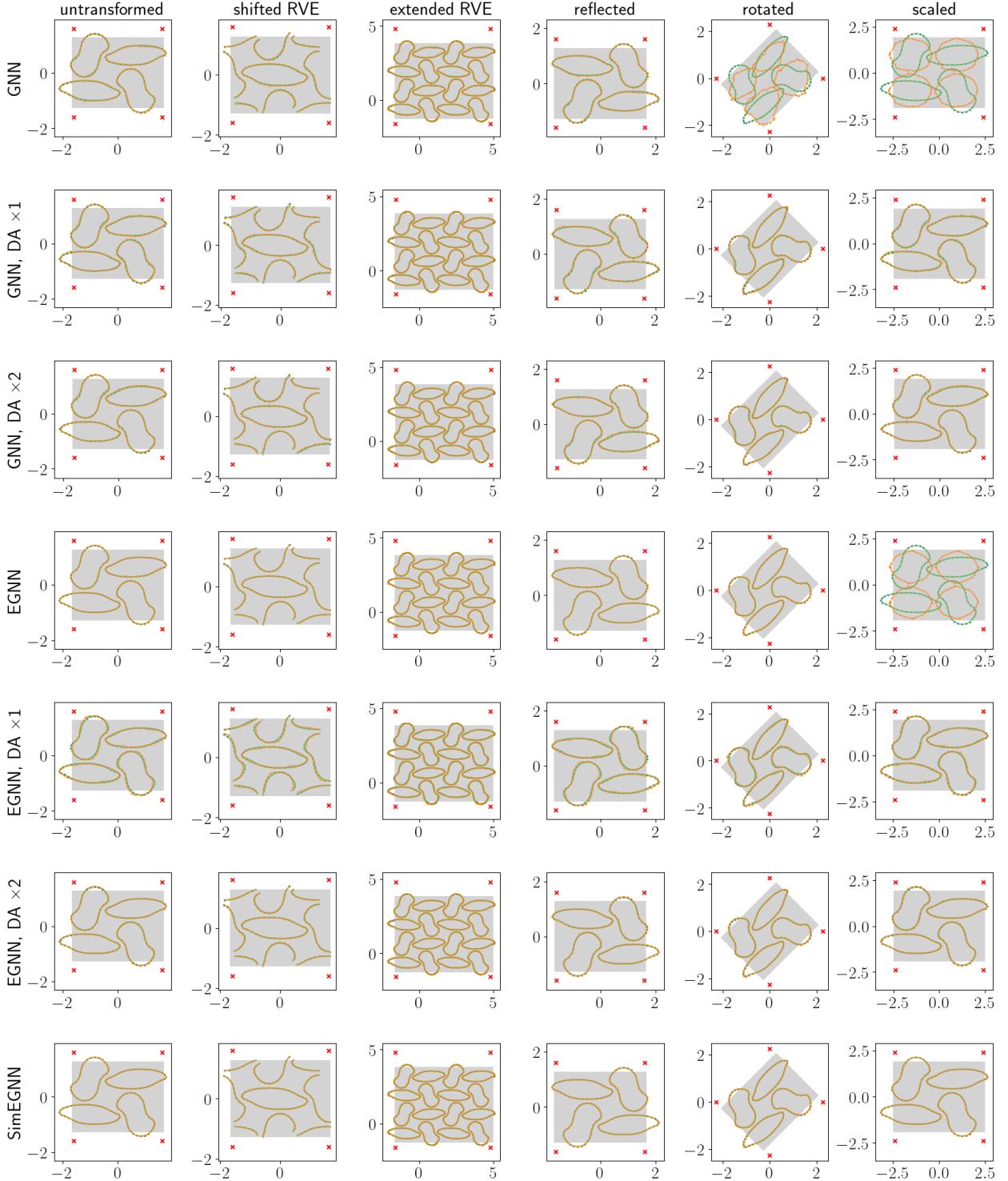


Figure B.18: Model-predicted deformation (orange) compared to the FEM ground truth (green), for $\mathbf{F} = \begin{bmatrix} 1.04 & 0 \\ 0 & 0.8 \end{bmatrix}$ (combination of tension and compression, resulting in left/right bifurcation as well as rotational bifurcation). The grey square indicates how \mathbf{F} deforms the square RVE through an affine transformation. The red crosses indicate the original positions of the corners.

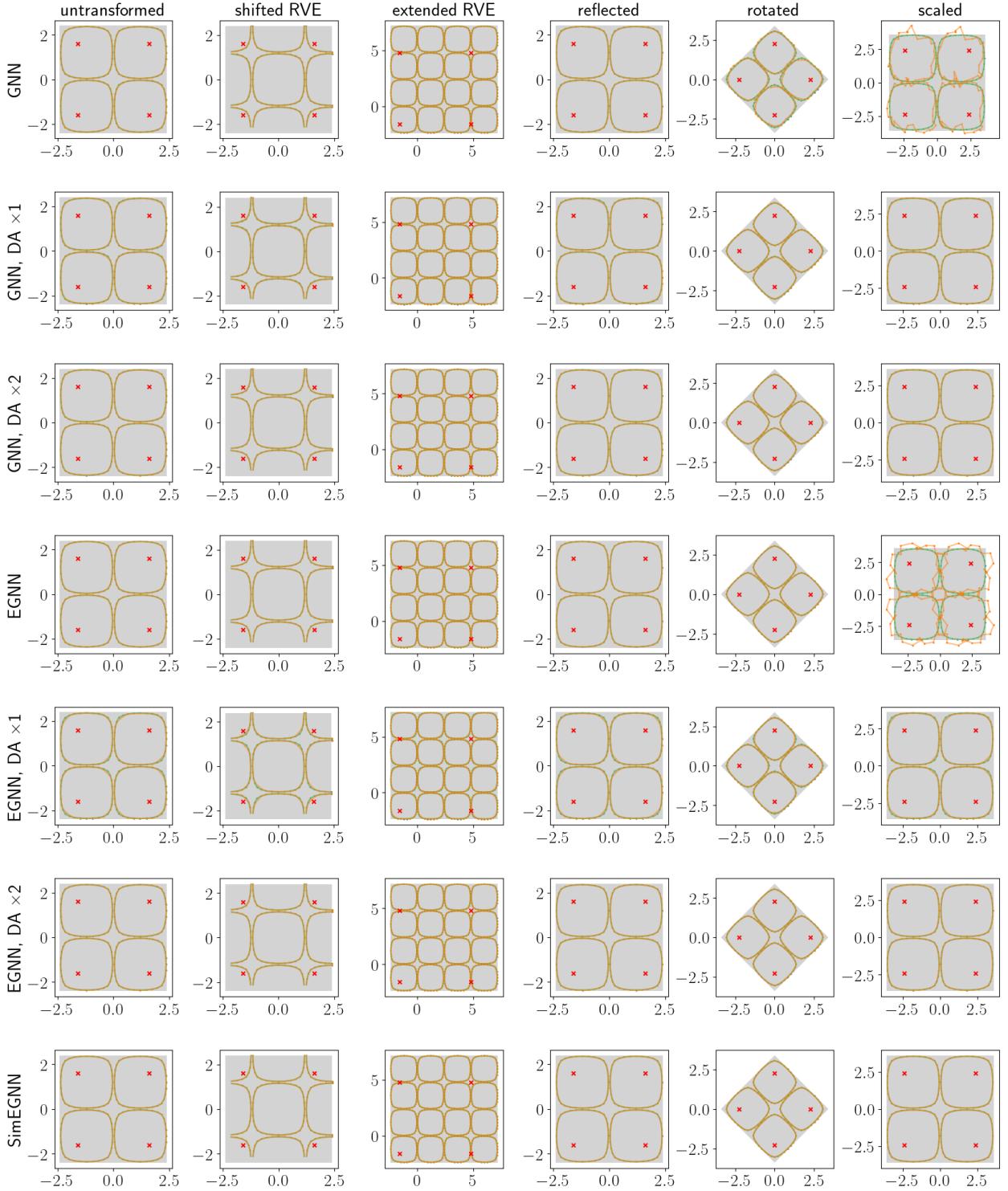


Figure B.19: Model-predicted deformation (orange) compared to the FEM ground truth (green), for $\mathbf{F} = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}$ (biaxial tension). The grey square indicates how \mathbf{F} deforms the square RVE through an affine transformation. The red crosses indicate the original positions of the corners.

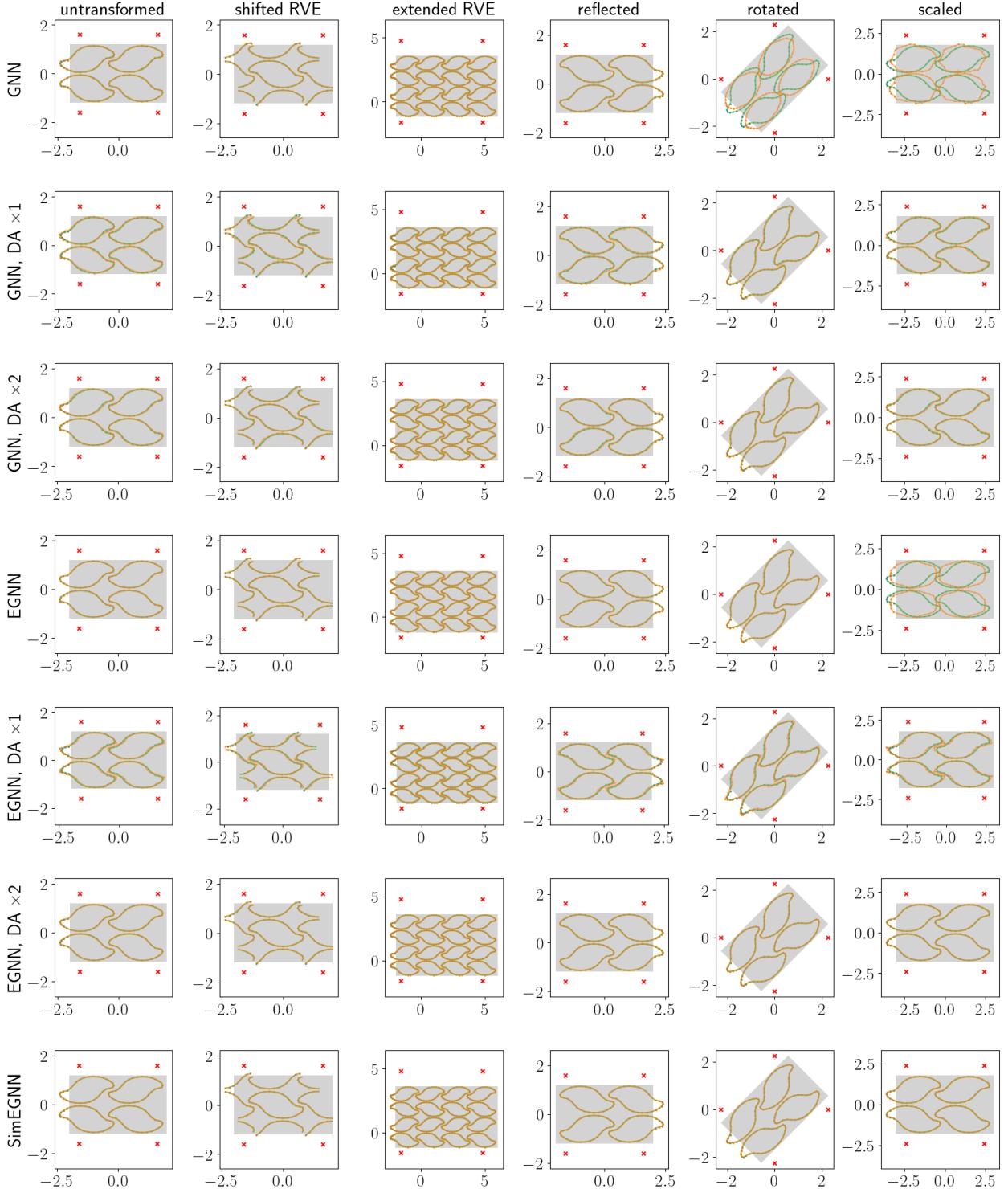


Figure B.20: Model-predicted deformation (orange) compared to the FEM ground truth (green), for $\mathbf{F} = \begin{bmatrix} 1.25 & 0 \\ 0 & 0.75 \end{bmatrix}$ (combination of tension and compression, resulting in left/right bifurcation). The grey square indicates how \mathbf{F} deforms the square RVE through an affine transformation. The red crosses indicate the original positions of the corners.

References

- [1] K. Bertoldi, V. Vitelli, J. Christensen, M. Van Hecke, Flexible mechanical metamaterials (oct 2017). doi:10.1038/natrevmats.2017.66.
URL <https://www.nature.com/articles/natrevmats201766>
- [2] X. Yu, J. Zhou, H. Liang, Z. Jiang, L. Wu, Mechanical metamaterials associated with stiffness, rigidity and compressibility: A brief review, *Progress in Materials Science* 94 (2018) 114–173. doi:10.1016/j.pmatsci.2017.12.003.
URL <https://doi.org/10.1016/j.pmatsci.2017.12.003>
- [3] S. Babaee, J. Shim, J. C. Weaver, E. R. Chen, N. Patel, K. Bertoldi, 3D soft metamaterials with negative poisson's ratio, *Advanced Materials* 25 (36) (2013) 5044–5049. doi:10.1002/adma.201301986.
URL <https://onlinelibrary.wiley.com/doi/10.1002/adma.201301986>
- [4] Z. G. Nicolaou, A. E. Motter, Mechanical metamaterials with negative compressibility transitions, *Nature Materials* 11 (7) (2012) 608–613. arXiv:1207.2185, doi:10.1038/nmat3331.
URL <https://www.nature.com/articles/nmat3331>
- [5] K. Bertoldi, M. C. Boyce, S. Deschanel, S. M. Prange, T. Mullin, Mechanics of deformation-triggered pattern transformations and superelastic behavior in periodic elastomeric structures, *Journal of the Mechanics and Physics of Solids* 56 (8) (2008) 2642–2668. doi:10.1016/j.jmps.2008.03.006.
URL <http://arxiv.org/abs/2301.11355>
- [6] J. T. Overvelde, K. Bertoldi, Relating pore shape to the non-linear response of periodic elastomeric structures, *Journal of the Mechanics and Physics of Solids* 64 (1) (2014) 351–366. doi:10.1016/j.jmps.2013.11.014.
URL <http://dx.doi.org/10.1016/j.jmps.2013.11.014>
- [7] D. Yang, B. Mosadegh, A. Ainla, B. Lee, F. Khashai, Z. Suo, K. Bertoldi, G. M. Whitesides, Buckling of Elastomeric Beams Enables Actuation of Soft Machines, *Advanced Materials* 27 (41) (2015) 6323–6327. doi:10.1002/adma.201503188.
URL <https://onlinelibrary.wiley.com/doi/full/10.1002/adma.201503188>
<https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201503188>
<https://onlinelibrary.wiley.com/doi/10.1002/adma.201503188>
- [8] Y. Kim, H. Yuk, R. Zhao, S. A. Chester, X. Zhao, Printing ferromagnetic domains for untethered fast-transforming soft materials, *Nature* 558 (7709) (2018) 274–279. doi:10.1038/s41586-018-0185-0.
- [9] A. Guell Izard, L. Valdevit, Magnetoelastic Metamaterials for Energy Dissipation and Wave Filtering, *Advanced Engineering Materials* 22 (2) (2020) 1901019. doi:10.1002/adem.201901019.

- URL <https://onlinelibrary.wiley.com/doi/full/10.1002/adem.201901019>
<https://onlinelibrary.wiley.com/doi/abs/10.1002/adem.201901019>
- [10] S. Ning, D. Chu, H. Jiang, F. Yang, Z. Liu, Z. Zhuang, The role of material and geometric nonlinearities and damping effects in designing mechanically tunable acoustic metamaterials, International Journal of Mechanical Sciences 197 (January) (2021) 106299. doi:10.1016/j.ijmecsci.2021.106299.
URL <https://doi.org/10.1016/j.ijmecsci.2021.106299>
- [11] S. M. Montgomery, S. Wu, X. Kuang, C. D. Armstrong, C. Zemelka, Q. Ze, R. Zhang, R. Zhao, H. J. Qi, Magneto-Mechanical Metamaterials with Widely Tunable Mechanical Properties and Acoustic Bandgaps, Advanced Functional Materials 31 (3) (2021) 2005319. doi:10.1002/adfm.202005319.
URL <https://onlinelibrary.wiley.com/doi/full/10.1002/adfm.202005319>
<https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.202005319>
- [12] L. Wu, Y. Wang, K. Chuang, F. Wu, Q. Wang, W. Lin, H. Jiang, A brief review of dynamic mechanical metamaterials for mechanical energy manipulation, Materials Today 44 (April) (2021) 168–193. doi:10.1016/j.mattod.2020.10.006.
URL <https://doi.org/10.1016/j.mattod.2020.10.006>
- [13] S. Terryn, J. Brancart, D. Lefebvre, G. Van Assche, B. Vanderborght, Self-healing soft pneumatic robots, Science Robotics 2 (9) (2017) 16. doi:10.1126/scirobotics.aan4268.
URL <https://www.science.org/doi/10.1126/scirobotics.aan4268>
- [14] Y. Kim, G. A. Parada, S. Liu, X. Zhao, Ferromagnetic soft continuum robots, Science Robotics 4 (33) (2019) 7329. doi:10.1126/SCIROBOTICS.AAX7329.
URL <https://www.science.org/doi/10.1126/scirobotics.aax7329>
- [15] U. Veerabagu, H. Palza, F. Quero, Review: Auxetic Polymer-Based Mechanical Metamaterials for Biomedical Applications (jul 2022). doi:10.1021/acsbiomaterials.2c00109.
URL <https://pubs.acs.org/doi/full/10.1021/acsbiomaterials.2c00109>
- [16] Y. Jiang, Q. Wang, Highly-stretchable 3D-architected Mechanical Metamaterials, Scientific Reports 6 (1) (2016) 1–11. doi:10.1038/srep34147.
URL <https://www.nature.com/articles/srep34147>
- [17] L. D'Alessandro, E. Belloni, R. Arditò, A. Corigliano, F. Braghin, Modeling and experimental verification of an ultra-wide bandgap in 3D phononic crystal, Applied Physics Letters 109 (22) (2016) 221907. doi:10.1063/1.4971290.
URL <http://dx.doi.org/10.1063/1.4971290><http://aip.scitation.org/toc/apl/109/22>

- [18] Q. Shi, K. Yu, X. Kuang, X. Mu, C. K. Dunn, M. L. Dunn, T. Wang, H. Jerry Qi, Recyclable 3D printing of vitrimer epoxy, *Materials Horizons* 4 (4) (2017) 598–607. doi:10.1039/c7mh00043j.
 URL <https://pubs.rsc.org/en/content/articlehtml/2017/mh/c7mh00043j>
<https://pubs.rsc.org/en/content/articlelanding/2017/mh/c7mh00043j>
- [19] X. Ren, J. Shen, P. Tran, T. D. Ngo, Y. M. Xie, Design and characterisation of a tuneable 3D buckling-induced auxetic metamaterial, *Materials and Design* 139 (2018) 336–342. doi:10.1016/j.matdes.2017.11.025.
- [20] S. M. Montgomery, X. Kuang, C. D. Armstrong, H. J. Qi, Recent advances in additive manufacturing of active mechanical metamaterials, *Current Opinion in Solid State and Materials Science* 24 (5) (2020) 100869. doi:10.1016/j.cossms.2020.100869.
- [21] M. Bodaghi, A. Serjouei, A. Zolfagharian, M. Fotouhi, H. Rahman, D. Durand, Reversible energy absorbing meta-sandwiches by FDM 4D printing, *International Journal of Mechanical Sciences* 173 (2020) 105451. doi:10.1016/j.ijmecsci.2020.105451.
- [22] Z. P. Wang, L. H. Poh, J. Dirrenberger, Y. Zhu, S. Forest, Isogeometric shape optimization of smoothed petal auxetic structures via computational periodic homogenization, *Computer Methods in Applied Mechanics and Engineering* 323 (2017) 250–271. doi:10.1016/j.cma.2017.05.013.
 URL www.sciencedirect.com/www.elsevier.com/locate/cma
- [23] E. Medina, C. H. Rycroft, K. Bertoldi, Nonlinear shape optimization of flexible mechanical metamaterials, *Extreme Mechanics Letters* 61 (2023) 102015. doi:10.1016/j.eml.2023.102015.
 URL <https://doi.org/10.1016/j.eml.2023.102015>
- [24] M. P. Bendsøe, N. Kikuchi, Generating optimal topologies in structural design using a homogenization method, *Computer Methods in Applied Mechanics and Engineering* 71 (2) (1988) 197–224. doi:10.1016/0045-7825(88)90086-2.
- [25] M. P. Bendsøe, O. Sigmund, Topology optimization: Theory, methods and applications, 2004.
 URL <https://link.springer.com/book/10.1007/978-3-662-05086-6#toc>
- [26] O. Sigmund, Materials with prescribed constitutive parameters: An inverse homogenization problem, *International Journal of Solids and Structures* 31 (17) (1994) 2313–2329. doi:10.1016/0020-7683(94)90154-6.
- [27] E. Andreassen, B. S. Lazarov, O. Sigmund, Design of manufacturable 3D extremal elastic microstructure, *Mechanics of Materials* 69 (1) (2014) 1–10. doi:10.1016/j.mechmat.2013.09.018.

- [28] Q. Chen, X. Zhang, B. Zhu, Design of buckling-induced mechanical metamaterials for energy absorption using topology optimization, *Structural and Multidisciplinary Optimization* 58 (4) (2018) 1395–1410. doi:10.1007/s00158-018-1970-y.
 URL <https://link.springer.com/article/10.1007/s00158-018-1970-y>
- [29] C. R. Thomsen, F. Wang, O. Sigmund, Buckling strength topology optimization of 2D periodic materials based on linearized bifurcation analysis, *Computer Methods in Applied Mechanics and Engineering* 339 (2018) 115–136. doi:10.1016/j.cma.2018.04.031.
 URL www.sciencedirect.com/www.elsevier.com/locate/cma
- [30] F. Wang, O. Sigmund, 3D architected isotropic materials with tunable stiffness and buckling strength, *Journal of the Mechanics and Physics of Solids* 152 (2021) 104415. arXiv:2012.01359, doi:10.1016/j.jmps.2021.104415.
- [31] T. Xue, S. Mao, Mapped shape optimization method for the rational design of cellular mechanical metamaterials under large deformation, *International Journal for Numerical Methods in Engineering* 123 (10) (2022) 2357–2380. doi:10.1002/nme.6941.
 URL <https://onlinelibrary.wiley.com/doi/full/10.1002/nme.6941>
<https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6941>
<https://onlinelibrary.wiley.com/doi/10.1002/nme.6941>
- [32] R. Hill, Elastic properties of reinforced solids: Some theoretical principles, *Journal of the Mechanics and Physics of Solids* 11 (5) (1963) 357–372. doi:10.1016/0022-5096(63)90036-X.
- [33] L. J. Walpole, The determination of the elastic field of an ellipsoidal inclusion in an anisotropic medium, *Mathematical Proceedings of the Cambridge Philosophical Society* 81 (2) (1977) 283–289. doi:10.1017/S0305004100053366.
 URL <https://www.jstor.org/stable/100095>
- [34] R. Hill, A self-consistent mechanics of composite materials, *Journal of the Mechanics and Physics of Solids* 13 (4) (1965) 213–222. doi:10.1016/0022-5096(65)90010-4.
- [35] T. Mori, K. Tanaka, Average stress in matrix and average elastic energy of materials with misfitting inclusions, *Acta Metallurgica* 21 (5) (1973) 571–574. doi:10.1016/0001-6160(73)90064-3.
- [36] D. J. Lucia, P. S. Beran, W. A. Silva, Reduced-order modeling: New approaches for computational physics, *Progress in Aerospace Sciences* 40 (1-2) (2004) 51–117. doi:10.1016/j.paerosci.2003.12.001.
- [37] G. Kerschen, J.-C. Golinval, A. F. Vakakis, L. A. Bergman, The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview, Tech. rep. (2005).

- [38] F. El Halabi, D. González, A. Chico, M. Doblaré, FE2 multiscale in linear elasticity based on parametrized microscale models using proper generalized decomposition, Computer Methods in Applied Mechanics and Engineering 257 (2013) 183–202. doi:10.1016/j.cma.2013.01.011.
 URL <http://dx.doi.org/10.1016/j.cma.2013.01.011>
- [39] J. Yvonnet, Q. C. He, The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains, Journal of Computational Physics 223 (1) (2007) 341–368. doi:10.1016/j.jcp.2006.09.019.
- [40] J. A. Hernández, J. Oliver, A. E. Huespe, M. A. Caicedo, J. C. Cante, High-performance model reduction techniques in computational multiscale homogenization, Computer Methods in Applied Mechanics and Engineering 276 (2014) 149–189. doi:10.1016/j.cma.2014.03.011.
 URL <http://dx.doi.org/10.1016/j.cma.2014.03.011>www.sciencedirect.com/www.elsevier.com/locate/cma
- [41] C. Gogu, Improving the efficiency of large scale topology optimization through on-the-fly reduced order model construction, International Journal for Numerical Methods in Engineering 101 (4) (2014) 281–304. doi:10.1002/nme.4797.
 URL <https://onlinelibrary.wiley.com/doi/full/10.1002/nme.4797><https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.4797><https://onlinelibrary.wiley.com/doi/10.1002/nme.4797>
- [42] M. Guo, J. S. Hesthaven, Reduced order modeling for nonlinear structural analysis using Gaussian process regression, Computer Methods in Applied Mechanics and Engineering 341 (2018) 807–826. doi:10.1016/j.cma.2018.07.017.
- [43] T. Guo, O. Rokoš, K. Veroy, Learning constitutive models from microstructural simulations via a non-intrusive reduced basis method, Computer Methods in Applied Mechanics and Engineering 384 (2021) 113924. doi:10.1016/j.cma.2021.113924.
 URL <https://linkinghub.elsevier.com/retrieve/pii/S0045782521002619>
- [44] T. Guo, F. A. Silva, O. Rokoš, K. Veroy, Learning constitutive models from microstructural simulations via a non-intrusive reduced basis method: Extension to geometrical parameterizations, Computer Methods in Applied Mechanics and Engineering 401 (2022) 115636. doi:10.1016/j.cma.2022.115636.
- [45] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, Computer Methods in Applied Mechanics and Engineering 304 (2016) 81–101. arXiv:1510.04232, doi:10.1016/j.cma.2016.02.001.
- [46] Z. Liu, M. A. Bessa, W. K. Liu, Self-consistent clustering analysis: An efficient multiscale scheme for inelastic heterogeneous materials, Computer Methods in Applied Mechanics and Engineering 306 (2016) 319–341. doi:10.1016/j.cma.2016.04.004.
 URL <http://dx.doi.org/10.1016/j.cma.2016.04.004>

- [47] G. Liang, K. Chandrashekara, Neural network based constitutive model for elastomeric foams, *Engineering Structures* 30 (7) (2008) 2002–2011. doi:10.1016/j.engstruct.2007.12.021.
- [48] B. A. Le, J. Yvonnet, Q. He, Computational homogenization of nonlinear elastic materials using neural networks, *International Journal for Numerical Methods in Engineering* 104 (12) (2015) 1061–1084. doi:10.1002/nme.4953.
URL <http://onlinelibrary.wiley.com/doi/10.1002/nme.3279/fullhttps://onlinelibrary.wiley.com/doi/10.1002/nme.4953>
- [49] K. Linka, M. Hillgärtner, K. P. Abdolazizi, R. C. Aydin, M. Itskov, C. J. Cyron, Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning, *Journal of Computational Physics* 429 (2021) 110010. doi:10.1016/j.jcp.2020.110010.
URL <https://doi.org/10.1016/j.jcp.2020.110010https://linkinghub.elsevier.com/retrieve/pii/S0021999120307841>
- [50] T. Xue, S. Adriaenssens, S. Mao, Learning the nonlinear dynamics of soft mechanical metamaterials with graph networks (2022). arXiv:2202.13775.
URL <https://www.researchgate.net/publication/358885333http://arxiv.org/abs/2202.13775>
- [51] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals, and Systems* 2 (4) (1989) 303–314. doi:10.1007/BF02551274.
URL <https://link.springer.com/article/10.1007/BF02551274>
- [52] K. Hornik, Approximation capabilities of multilayer feedforward networks, *Neural Networks* 4 (2) (1991) 251–257. doi:10.1016/0893-6080(91)90009-T.
- [53] M. Leshno, V. Y. Lin, A. Pinkus, S. Schocken, Multilayer feedforward networks with a nonpolynomial activation function can approximate any function, *Neural Networks* 6 (6) (1993) 861–867. doi:10.1016/S0893-6080(05)80131-5.
URL <http://archive.nyu.edu/handle/2451/14329>
- [54] G. C. Peng, M. Alber, A. Buganza Tepole, W. R. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W. W. Lytton, P. Perdikaris, L. Petzold, E. Kuhl, Multiscale Modeling Meets Machine Learning: What Can We Learn?, *Archives of Computational Methods in Engineering* 28 (3) (2021) 1017–1037. arXiv:1911.11958, doi:10.1007/s11831-020-09405-5.
URL <https://link.springer.com/article/10.1007/s11831-020-09405-5>
- [55] E. B. Tadmor, R. E. Miller, R. S. Elliott, *Continuum Mechanics and Thermodynamics*, 2012.
- [56] A. L. Frankel, R. E. Jones, C. Alleman, J. A. Templeton, Predicting the mechanical response of oligocrystals with deep learning, *Computational Materials Science* 169 (2019).

arXiv:1901.10669, doi:10.1016/j.commatsci.2019.109099.
URL www.elsevier.com/locate/commatsci

- [57] J. K. Wilt, C. Yang, G. X. Gu, Accelerating Auxetic Metamaterial Design with Deep Learning, *Advanced Engineering Materials* 22 (5) (2020) 1901266. doi:10.1002/adem.201901266.
URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/adem.201901266>
- [58] A. Pandey, R. Pokharel, Machine learning based surrogate modeling approach for mapping crystal deformation in three dimensions, *Scripta Materialia* 193 (2021) 1–5. doi:10.1016/j.scriptamat.2020.10.028.
- [59] Z. Yang, C. H. Yu, M. J. Buehler, Deep learning model to predict complex stress and strain fields in hierarchical composites, *Science Advances* 7 (15) (apr 2021). doi:10.1126/SCIAADV.ABD7416.
URL <https://www.science.org/doi/10.1126/sciadv.abd7416>
- [60] J. R. Mianroodi, N. H. Siboni, D. Raabe, Teaching solid mechanics to artificial intelligence—a fast solver for heterogeneous materials, *npj Computational Materials* 7 (1) (2021). arXiv:2103.09147, doi:10.1038/s41524-021-00571-z.
URL <https://doi.org/10.1038/s41524-021-00571-z>
- [61] J. R. Mianroodi, S. Rezaei, N. H. Siboni, B. X. Xu, D. Raabe, Lossless multi-scale constitutive elastic relations with artificial intelligence, *npj Computational Materials* 8 (1) (2022) 1–12. arXiv:2108.02837, doi:10.1038/s41524-022-00753-3.
- [62] M. S. Khorrami, J. R. Mianroodi, N. H. Siboni, P. Goyal, B. Svendsen, P. Benner, D. Raabe, An artificial neural network for surrogate modeling of stress fields in viscoplastic polycrystalline materials, *npj Computational Materials* 9 (1) (2023) 1–10. arXiv:2208.13490, doi:10.1038/s41524-023-00991-z.
URL <https://www.nature.com/articles/s41524-023-00991-z>
- [63] T. S. Cohen, M. Welling, Group equivariant convolutional networks, in: 33rd International Conference on Machine Learning, ICML 2016, Vol. 6, 2016, pp. 4375–4386. arXiv:1602.07576.
- [64] N. N. Vlassis, R. Ma, W. C. Sun, Geometric deep learning for computational mechanics Part I: anisotropic hyperelasticity, *Computer Methods in Applied Mechanics and Engineering* 371 (2020) 113299. arXiv:2001.04292, doi:10.1016/j.cma.2020.113299.
URL <https://doi.org/10.1016/j.cma.2020.113299>
- [65] A. Thomas, A. Durmaz, M. Alam, P. Gumbsch, H. Sack, C. Eberl, Materials fatigue prediction using graph neural networks on microstructure representations, *Scientific Reports* 2023 13:1 13 (1) (2023) 1–16. doi:10.1038/s41598-023-39400-2.
URL <https://www.nature.com/articles/s41598-023-39400-2>

- [66] K. Karapiperis, D. M. Kochmann, Prediction and control of fracture paths in disordered architected materials using graph neural networks, Communications Engineering 2 (1) (2023). doi:10.1038/s44172-023-00085-0.
URL www.ae108.ethz.ch
- [67] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, P. W. Battaglia, Learning Mesh-Based Simulation with Graph Networks (oct 2020). arXiv:2010.03409.
URL <http://arxiv.org/abs/2010.03409>
- [68] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, A. Tkatchenko, Quantum-chemical insights from deep tensor neural networks, Nature Communications 8 (2017). arXiv:1609.08259, doi:10.1038/ncomms13890.
URL www.nature.com/naturecommunications
- [69] K. T. Schütt, P. J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, K. R. Müller, SchNet: A continuous-filter convolutional neural network for modeling quantum interactions, in: Advances in Neural Information Processing Systems, Vol. 2017-Decem, 2017, pp. 992–1002. arXiv:1706.08566.
URL www.quantum-machine.org.
- [70] K. T. Schütt, H. E. Sauceda, P. J. Kindermans, A. Tkatchenko, K. R. Müller, SchNet - A deep learning architecture for molecules and materials, Journal of Chemical Physics 148 (24) (2018). arXiv:1712.06113, doi:10.1063/1.5019779.
- [71] N. Lubbers, J. S. Smith, K. Barros, Hierarchical modeling of molecular energies using a deep neural network, Journal of Chemical Physics 148 (24) (2018) 241715. arXiv:1710.00017, doi:10.1063/1.5011181.
URL <https://doi.org/10.1063/1.5011181>
- [72] O. T. Unke, M. Meuwly, PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges, Journal of Chemical Theory and Computation 15 (6) (2019) 3678–3693. arXiv:1902.08408, doi:10.1021/acs.jctc.9b00181.
URL <https://pubs.acs.org/sharingguidelines>
- [73] J. Gasteiger, J. Groß, S. Günnemann, Directional Message Passing for Molecular Graphs (2020). arXiv:2003.03123.
URL <https://www.daml.in.tum.de/dimenet> <http://arxiv.org/abs/2003.03123>
- [74] T. Xie, J. C. Grossman, Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties, Physical Review Letters 120 (14) (2018). arXiv:1710.10324, doi:10.1103/PhysRevLett.120.145301.
- [75] N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, P. Riley, Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds (feb 2018). arXiv:1802.08219.
URL <http://arxiv.org/abs/1802.08219>

- [76] I. Batatia, D. P. Kovács, G. N. Simm, C. Ortner, G. Csányi, MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, in: Advances in Neural Information Processing Systems, Vol. 35, 2022, pp. 11423–11436. [arXiv:2206.07697](https://arxiv.org/abs/2206.07697).
- [77] B. Jing, S. Eismann, P. Suriana, R. J. Townshend, R. O. Dror, LEARNING FROM PROTEIN STRUCTURE WITH GEOMETRIC VECTOR PERCEPTRONS, in: ICLR 2021 - 9th International Conference on Learning Representations, International Conference on Learning Representations, ICLR, 2021. [arXiv:2009.01411v3](https://arxiv.org/abs/2009.01411v3)
- [78] K. T. Schütt, O. T. Unke, M. Gastegger, S. Schütt, O. T. Unke, M. Gastegger, Equivariant Message Passing for the Prediction of Tensorial Properties and Molecular Spectra, in: Proceedings of Machine Learning Research, Vol. 139, PMLR, 2021, pp. 9377–9388. [arXiv:2102.03150](https://proceedings.mlr.press/v139/schutt21a.html).
URL <https://proceedings.mlr.press/v139/schutt21a.html>
- [79] J. Gasteiger, F. Becker, S. Günnemann, GemNet: Universal Directional Graph Neural Networks for Molecules, in: Advances in Neural Information Processing Systems, Vol. 9, 2021, pp. 6790–6802. [arXiv:2106.08903](https://www.daml.in.tum.de/gemnet).
URL <https://www.daml.in.tum.de/gemnet>
- [80] V. G. Satorras, E. Hoogeboom, M. Welling, E(n) Equivariant Graph Neural Networks, in: Proceedings of Machine Learning Research, Vol. 139, 2021, pp. 9323–9332. [arXiv:2102.09844](https://arxiv.org/abs/2102.09844).
- [81] M. G. D. Geers, V. G. Kouznetsova, W. A. M. Brekelmans, Multi-scale computational homogenization: Trends and challenges, in: Journal of Computational and Applied Mathematics, Vol. 234, 2010, pp. 2175–2182. doi:[10.1016/j.cam.2009.08.077](https://doi.org/10.1016/j.cam.2009.08.077).
- [82] V. G. Kouznetsova, M. G. D. Geers, W. A. M. Brekelmans, Multi-scale second-order computational homogenization of multi-phase materials: A nested finite element solution strategy, Computer Methods in Applied Mechanics and Engineering 193 (48-51) (2004) 5525–5550. doi:[10.1016/j.cma.2003.12.073](https://doi.org/10.1016/j.cma.2003.12.073).
- [83] L. Kaczmarczyk, C. J. Pearce, N. Bićanić, Scale transition and enforcement of RVE boundary conditions in second-order computational homogenization, International Journal for Numerical Methods in Engineering 74 (3) (2008) 506–522. doi:[10.1002/nme.2188](https://doi.org/10.1002/nme.2188).
URL <https://onlinelibrary.wiley.com/doi/full/10.1002/nme.2188>
<https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.2188>
<https://onlinelibrary.wiley.com/doi/10.1002/nme.2188>
- [84] O. Rokoš, M. M. Ameen, R. H. J. Peerlings, M. G. D. Geers, Micromorphic computational homogenization for mechanical metamaterials with patterning fluctuation

- fields, Journal of the Mechanics and Physics of Solids 123 (2019) 119–137. doi: 10.1016/j.jmps.2018.08.019.
- [85] S. E. H. M. van Bree, O. Rokoš, R. H. J. Peerlings, M. Doškář, M. G. D. Geers, A Newton solver for micromorphic computational homogenization enabling multiscale buckling analysis of pattern-transforming metamaterials, Computer Methods in Applied Mechanics and Engineering 372 (dec 2020). arXiv:2008.12850, doi:10.1016/j.cma.2020.113333.
- [86] O. Rokoš, M. M. Ameen, R. H. J. Peerlings, M. G. D. Geers, Extended micromorphic computational homogenization for mechanical metamaterials exhibiting multiple geometric pattern transformations, Extreme Mechanics Letters 37 (may 2020). arXiv:2004.05226, doi:10.1016/j.eml.2020.100708.
- [87] K. Bertoldi, P. M. Reis, S. Willshaw, T. Mullin, Negative poisson’s ratio behavior induced by an elastic instability, Advanced Materials 22 (3) (2010) 361–366. doi: 10.1002/adma.200901956.
- [88] V. G. Kouznetsova, Computational homogenization for the multi-scale analysis of multi-phase materials, Ph.D. thesis (2002). doi:10.6100/IR560009.
URL <https://doi.org/10.6100/IR560009><http://www.narcis.nl/publication/RecordID/oai:library.tue.nl:560009>
- [89] C. Miehe, Computational micro-to-macro transitions for discretized micro-structures of heterogeneous materials at finite strains based on the minimization of averaged incremental energy, Computer Methods in Applied Mechanics and Engineering 192 (5-6) (2003) 559–591. doi:10.1016/S0045-7825(02)00564-9.
- [90] C. Miehe, Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation, International Journal for Numerical Methods in Engineering 55 (11) (2002) 1285–1322. doi:10.1002/nme.515.
URL <https://onlinelibrary.wiley.com/doi/full/10.1002/nme.515><https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.515><https://onlinelibrary.wiley.com/doi/10.1002/nme.515>
- [91] C. Geuzaine, J. F. Remacle, Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities, International Journal for Numerical Methods in Engineering 79 (11) (2009) 1309–1331. doi:10.1002/nme.2579.
- [92] M. M. Bronstein, J. Bruna, T. Cohen, P. Veličković, Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges (apr 2021). arXiv:2104.13478.
URL <http://arxiv.org/abs/2104.13478>
- [93] D. Chen, Y. Lin, W. Li, P. Li, J. Zhou, X. Sun, Measuring and relieving the over-smoothing problem for graph neural networks from the topological view, AAAI 2020 -

34th AAAI Conference on Artificial Intelligence (2020) 3438–3445arXiv:1909.03211,
doi:10.1609/aaai.v34i04.5747.

- [94] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, G. E. Dahl, Neural message passing for quantum chemistry, in: 34th International Conference on Machine Learning, ICML 2017, Vol. 3, 2017, pp. 2053–2070. arXiv:1704.01212.
- [95] W. M. Czarnecki, S. Osindero, M. Jaderberg, G. Swirszcz, R. Pascanu, Sobolev training for neural networks, in: Advances in Neural Information Processing Systems, Vol. 2017-Decem, 2017, pp. 4279–4288. arXiv:1706.04859.
- [96] C. K. Joshi, C. Bodnar, S. V. Mathis, T. Cohen, P. Liò, On the Expressive Power of Geometric Graph Neural Networks (2023). arXiv:2301.09308.
URL <https://github.com/chaitjo/geometric-gnn-dojo> <http://arxiv.org/abs/2301.09308>