Received: 7 October 2022 | Accepted: 10 November 2022

DOI: 10.1002/pamm.202200143

# Material-informed training of viscoelastic deep material networks

**Sebastian Gajek**[1,*], **Matti Schneider**[1], and **Thomas Böhlke**[1]

[1] Karlsruhe Institute of Technology (KIT)

Deep material networks (DMN) are a data-driven homogenization approach that show great promise for accelerating concurrent two-scale simulations. As a salient feature, DMNs are solely identified by linear elastic precomputations on representative volume elements. After parameter identification, DMNs act as surrogates for full-field simulations of such volume elements with inelastic constituents.

In this work, we investigate how the training on linear elastic data, i.e., how the choice of the loss function and the sampling of the training data, affects the accuracy of DMNs for inelastic constituents. We investigate linear viscoelasticity and derive a material-informed sampling procedure for generating the training data and a loss function tailored to the problem at hand. These ideas improve the accuracy of an identified DMN and allow for significantly reducing the number of samples to be generated and labeled.



## 1 Introduction

To harness the full lightweight potential of composite materials, the rich microstructure information needs to be accounted for in mechanical simulations. Concurrent multiscale methods couple simulations on the microscopic and macroscopic scale to include microstructure information into component scale simulations. However, approaches such as the $FE^2$ [2] or the FE-FFT [3] method are typically too computational demanding for problems of industrial complexity.

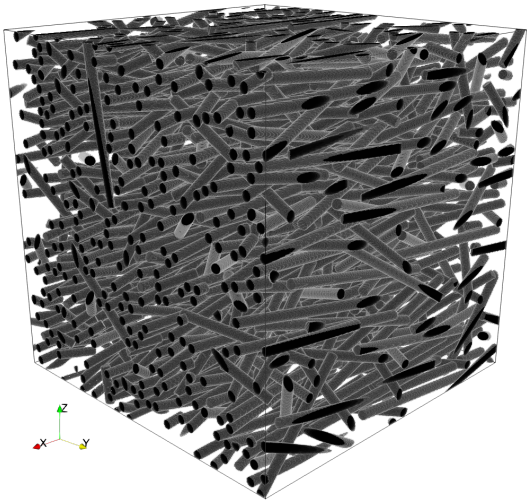


**Fig. 1:** Short-fiber reinforced polymer microstructure [1]

Deep material networks [4–6] are a data-driven and micromechanics-inspired homogenization approach. DMNs seek to approximate a fixed microstructure by means of hierarchical laminates. The volume fractions of the individual laminates and their lamination directions serve as fitting parameters of the approach. DMNs are inversely identified by approximating the effective linear elastic material behavior of a fixed microstructure, which is considered as a function of the input stiffnesses of the phases. Thus, the parameter identification is exclusively based on linear elastic precomputations and independent of the nonlinear and inelastic material behavior which the model is later evaluated for. After parameter identification, the DMN acts as a surrogate for the microscopic problem. Solving the (implicit) DMN in every Gauss point of a macroscopic simulation gives rise to the FE-DMN method [7, 8], which allows for incorporating microstructure information into a macroscopic simulation. Most notably, DMNs predict the effective properties for arbitrarily nonlinear constituents and complex three-dimensional microstructures with high accuracy and provide speed-ups of up to five orders of magnitude [8, 9], allowing for conducting DMN-accelerated large-scale component simulations on commodity hardware. Extensions to treat thermomechanically coupled problems [1] and problems with fluctuating microstructure characteristics [8, 10] are possible.

However, besides their apparent success in practice, there is still a lack of understanding the underlying mechanisms of DMNs. Indeed, Gajek et al. [6] showed that the effective stress of a microstructure with arbitrarily nonlinear constituents is determined to first order in the strain rate by linear elastic localization. However, an in-depth understanding of how training on linear elastic data influences the accuracy of the DMN for arbitrarily nonlinear constituents is still missing. For instance, open questions are: Might there be a way to tailor data sampling to the underlying material behavior of the phases in order to improve the inelastic result? As we know from supervised machine learning, there is a plethora of possible loss functions which typically excel in a specific task and fall behind for others. Is there a specific form the loss function should have in order to improve the inelastic results? Are sampling and loss possibly linked and should not be considered separately?

In this work, we investigate these questions, at least for simple material behavior. We restrict to linear viscoelasticity and derive a material-informed sampling procedure of the training data and a novel loss function which directly links the linear training to the DMN's accuracy for inelastic constituents.

* Corresponding author: e-mail sebastian.gajek@kit.edu, phone +49 721 608-46904

## 2 Deep material networks

Let $Y \subseteq \mathbb{R}^3$ denote a two-phase periodic rectangular microstructure in three spatial dimensions. The microstructure at hand [1], which is depicted in Fig. 1, comprises $16\,\mathrm{vol}\%$ of straight fibers with an aspect ratio of 20 and features a volume element to fiber length ratio of around two and a second-order fiber orientation tensor of

$$\boldsymbol{A}_2 = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix} \boldsymbol{e}_i \otimes \boldsymbol{e}_j. \tag{1}$$

Let $\mathcal{GSM}$ denote the set of all generalized standard materials (GSMs) [11]. It is well known that GSMs are closed under homogenization. More precisely, periodic homogenization gives rise to the generally nonlinear homogenization function [6]

$$\mathcal{M} : \mathcal{GSM} \times \mathcal{GSM} \to \mathcal{GSM}, \quad (\mathcal{G}_\mathrm{F}, \mathcal{G}_\mathrm{M}) \mapsto \bar{\mathcal{G}},$$

which maps the material behavior of the fibers $\mathcal{G}_\mathrm{F} \in \mathcal{GSM}$ and the material behavior of the matrix $\mathcal{G}_\mathrm{M} \in \mathcal{GSM}$ to the effective material behavior $\bar{\mathcal{G}} \in \mathcal{GSM}$ of the composite. The purpose of homogenization theories is to find $\mathcal{M}$ for arbitrarily nonlinear phases $\mathcal{G}_\mathrm{F}$ and $\mathcal{G}_\mathrm{M}$. However, for practical reasons, only an approximation of $\mathcal{M}$ might be provided whose error needs to be balanced with the effort of obtaining the approximation.

Deep material networks seek to approximate the nonlinear homogenization function $\mathcal{M}$, associated to microstructure $Y$, by the homogenization function

$$\mathcal{M}^{\mathrm{DMN}}_{\boldsymbol{p}} : \mathcal{GSM} \times \mathcal{GSM} \to \mathcal{GSM}, \quad (\mathcal{G}_\mathrm{F}, \mathcal{G}_\mathrm{M}) \mapsto \bar{\mathcal{G}}^{\mathrm{DMN}},$$

where the underlying microstructure is chosen as a hierarchical laminate [4–6]. More precicely, $\mathcal{M}^{\mathrm{DMN}}_{\boldsymbol{p}}$ is given by a binary tree of two-phase rank-one laminates of predefined depth $K$, see Gajek et al. [6] for a schematic illustration. The fitting parameters of the approach are given by the laminates' volume fractions and directions of lamination and are collected in the parameter vector $\boldsymbol{p} \in \mathcal{P}$.

Using that two-phase linear elastic homogenization determines nonlinear homogenization to first order in the strain rate [6], DMNs seek to approximate $\mathcal{M}'$ by $\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}$, where

$$\mathcal{M}' : \mathcal{C} \times \mathcal{C} \to \mathcal{C}, \quad (\mathbb{C}_\mathrm{F}, \mathbb{C}_\mathrm{M}) \mapsto \bar{\mathbb{C}} \quad \text{and} \quad \mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}} : \mathcal{C} \times \mathcal{C} \to \mathcal{C}, \quad (\mathbb{C}_\mathrm{F}, \mathbb{C}_\mathrm{M}) \mapsto \bar{\mathbb{C}}^{\mathrm{DMN}}$$

denote the linear elastic homogenization functions associated to $\mathcal{M}$ and $\mathcal{M}^{\mathrm{DMN}}_{\boldsymbol{p}}$, respectively. The former and latter map the two stiffness tensors of fibers and matrix $\mathbb{C}_\mathrm{F}, \mathbb{C}_\mathrm{M} \in \mathcal{C}$ to the effective stiffness $\bar{\mathbb{C}} \in \mathcal{C}$. For parameter identification, $\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}$ is determined by computing an approximate minimizer of the regression problem [4]

$$J(\boldsymbol{p}) = \sum_{s=1}^{N_\mathrm{s}} J_s(\boldsymbol{p}) + \psi(\boldsymbol{p}) \longrightarrow \min_{\boldsymbol{p} \in \mathcal{P}}, \tag{2}$$

where

$$J_s(\boldsymbol{p}) = \frac{1}{N_s} \frac{\|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}^s_\mathrm{F}, \mathbb{C}^s_\mathrm{M}) - \mathcal{M}'(\mathbb{C}^s_\mathrm{F}, \mathbb{C}^s_\mathrm{M})\|}{\|\mathcal{M}'(\mathbb{C}^s_\mathrm{F}, \mathbb{C}^s_\mathrm{M})\|} \tag{3}$$

measures the proximity of $\mathcal{M}'$ and $\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}$ for specific stiffness samples $(\mathbb{C}^s_\mathrm{F}, \mathbb{C}^s_\mathrm{M})$, which might be sampled as axis-aligned orthotropic stiffness tensors [4–6]. The second contribution $\psi(\boldsymbol{p})$ in Eq. (3) denotes a regularization term, which enforces that the DMN's volume fractions are consistent w.r.t. the underlying microstructure $Y$. After parameter identification, the DMN's nonlinear homogenization function $\mathcal{M}^{\mathrm{DMN}}_{\boldsymbol{p}}$ may then be used as an approximation of the nonlinear homogenization function $\mathcal{M}$ for arbitrarily nonlinear constituents $\mathcal{G}_\mathrm{F}$ and $\mathcal{G}_\mathrm{M}$, see Gajek et al. [6]. The specific structure of the DMN allows for computing the effective properties (effective stress and algorithmic tangent) with excellent efficiency.

## 3 Considering linear viscoelastic composites

We seek to investigate if we might improve upon the aforementioned sampling of orthotropic stiffnesses and the specified form of the loss function. Therefore, we consider a setup where the relationship between linear and nonlinear homogenization is completely understood, namely linear viscoelasticity. We assume the fibers to be isotropic and linear elastic

$$\boldsymbol{\sigma}_\mathrm{F}(t) = \mathbb{C}_\mathrm{F} : \boldsymbol{\varepsilon}(t) \quad \text{with} \quad \mathbb{C}_\mathrm{F} = 3K_\mathrm{F}\,\mathbb{P}_1 + 2G_\mathrm{F}\,\mathbb{P}_2, \tag{4}$$

where $\mathbb{P}_1 : \mathrm{Sym}_2(3) \to \mathrm{Sym}_2(3)$ and $\mathbb{P}_2 : \mathrm{Sym}_2(3) \to \mathrm{Sym}_2(3)$ project onto the spherical and deviator subspaces of $\mathrm{Sym}_2(3)$, the set of symmetric second order tensors in three dimensions. For an initially stress-free material, i.e., $\boldsymbol{\sigma}_{\mathrm{M}}(0) = 0$, the stress-strain rate relationship of a viscoelastic matrix material reads

$$\boldsymbol{\sigma}_{\mathrm{M}}(t) = \int_0^\infty \mathbb{C}_{\mathrm{M}}(\tau) : \dot{\boldsymbol{\varepsilon}}(t-\tau)\,\mathrm{d}\tau. \tag{5}$$

Thus, the effective stress of microstructure $Y$ homogenizes to

$$\bar{\boldsymbol{\sigma}}(t) = \int_0^\infty \mathcal{M}'(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) : \dot{\bar{\boldsymbol{\varepsilon}}}(t-\tau)\,\mathrm{d}\tau, \tag{6}$$

whereas the effective stress of the DMN computes to

$$\bar{\boldsymbol{\sigma}}^{\mathrm{DMN}}(t) = \int_0^\infty \mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) : \dot{\bar{\boldsymbol{\varepsilon}}}(t-\tau)\,\mathrm{d}\tau. \tag{7}$$

In particular, we obtain the *exact* representation

$$\bar{\boldsymbol{\sigma}}^{\mathrm{DMN}}(t) - \bar{\boldsymbol{\sigma}}(t) = \int_0^\infty \left[\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) - \mathcal{M}'(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau))\right] : \dot{\bar{\boldsymbol{\varepsilon}}}(t-\tau)\,\mathrm{d}\tau \tag{8}$$

for the error in the effective stress. By Hölder's inequality [12], we deduce

$$\sup_t \|\bar{\boldsymbol{\sigma}}^{\mathrm{DMN}}(t) - \bar{\boldsymbol{\sigma}}(t)\| \le \int_0^\infty \|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) - \mathcal{M}'(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau))\|\,\mathrm{d}\tau \sup_t \|\dot{\bar{\boldsymbol{\varepsilon}}}(t)\|. \tag{9}$$

Let us define the auxilary function $f : \mathcal{GSM} \times \mathcal{GSM} \to \mathbb{R}$,

$$(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) \mapsto \|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) - \mathcal{M}'(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau))\|, \tag{10}$$

and let us consider a Maxwell model

$$\mathbb{C}_{\mathrm{M}}(\tau) = \mathbb{C}_{\mathrm{M},0} + \sum_{i=1}^{N_{\mathrm{M}}} \mathbb{C}_{\mathrm{M},i}\,\mathrm{e}^{-\frac{\tau}{\lambda_i}} \quad \text{with} \quad \mathbb{C}_{\mathrm{M},0} = 3K_{\mathrm{K}}\,\mathbb{P}_1 + 2G_{\mathrm{M}}\,\mathbb{P}_2 \quad \text{and} \quad \mathbb{C}_{\mathrm{M},i} = 2G_{\mathrm{M},i}\,\mathbb{P}_2, \tag{11}$$

comprising $N_{\mathrm{M}}$ dashpots with relaxation times $\lambda_i \in \mathbb{R}_{>0}$. Using the substitution $y = 1 - \mathrm{e}^{-\frac{\tau}{\lambda}}$ in terms of the positive constant $\lambda \in \mathbb{R}_{>0}$, the integral

$$\int_0^\infty \|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) - \mathcal{M}'(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau))\|\,\mathrm{d}\tau \tag{12}$$

transforms into the expression

$$\int_0^\infty f\left(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M},0} + \sum_{i=1}^{N_{\mathrm{M}}} \mathbb{C}_{\mathrm{M},i}\,\mathrm{e}^{-\frac{\tau}{\lambda_i}}\right)\mathrm{d}\tau = \int_0^1 \frac{\lambda}{1-y} f\left(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M},0} + \sum_{i=1}^{N_{\mathrm{M}}} (1-y)^{\frac{\lambda}{\lambda_i}} \mathbb{C}_{\mathrm{M},i}\right)\mathrm{d}y. \tag{13}$$

Thus, the accuracy of the DMN for this specific setup is determined by a non-finite distribution on a line segment connecting $\mathbb{C}_{\mathrm{M},0}$ and $\mathbb{C}_{\mathrm{M},0} + \sum_{i=1}^{N_{\mathrm{M}}} \mathbb{C}_{\mathrm{M},i}\, e^{-\frac{\tau}{\lambda_i}}$ in stiffness space.

Applying a Monte Carlo integration, the integral (12) transforms via Eq. (13) into

$$\int_0^\infty \|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau)) - \mathcal{M}'(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}}(\tau))\|\,\mathrm{d}\tau \approx \frac{1}{N_s}\sum_{s=1}^{N_s} \frac{\lambda}{1-y_s} f\left(\mathbb{C}_{\mathrm{F}}, \mathbb{C}_{\mathrm{M}} + \sum_{i=1}^{N_{\mathrm{M}}} (1-y_s)^{\frac{\lambda}{\lambda_i}} \mathbb{C}_i\right) \tag{14}$$

which suggests that we should rather consider a loss function of the form

$$J(\boldsymbol{p}) = \sum_{s=1}^{N_s} J_s(\boldsymbol{p}) + \psi(\boldsymbol{p}) \quad \text{with} \quad J_s(\boldsymbol{p}) = \frac{\lambda}{N_s} \frac{\|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}^s_{\mathrm{F}}, \mathbb{C}^s_{\mathrm{M}}) - \mathcal{M}'(\mathbb{C}^s_{\mathrm{F}}, \mathbb{C}^s_{\mathrm{M}})\|}{1-y^s} \tag{15}$$

with stiffness samples

$$\mathbb{C}^s_{\mathrm{F}} = 3K_{\mathrm{F}}\,\mathbb{P}_1 + 2G_{\mathrm{F}}\,\mathbb{P}_2 \quad \text{and} \quad \mathbb{C}^s_{\mathrm{M}} = 3K_{\mathrm{M}}\,\mathbb{P}_1 + 2\left(G_{\mathrm{M}} + \sum_{i=1}^{N_{\mathrm{M}}} G_{\mathrm{M},i}\,(1-y_s)^{\frac{\lambda}{\lambda_i}}\right)\mathbb{P}_2 \tag{16}$$

where we draw $y_s \in [0, 1)$.

## 4 Numerical results

For the numerical investigations, we use the material parameters for the linear elastic fibers [13] and the viscoelastic matrix [14] with $N_\mathrm{M} = 9$ Maxwell elements as reported in Table 1.

| | | | | | | |
|---|---|---|---|---|---|---|
| Fibers [13] | $K_\mathrm{F} = 50.0\,\mathrm{GPa}$ | $G_\mathrm{F} = 28.6\,\mathrm{GPa}$ | | | | |
| Matrix [14] | $K_{\mathrm{M},0} = 4.93\,\mathrm{GPa}$ | $G_{\mathrm{M},0} = 415.6\,\mathrm{MPa}$ | | | | |
| | $G_{\mathrm{M},1} = 154.8\,\mathrm{MPa}$ | $\tau_1 = 10^{-4}\,\mathrm{s}$ | $G_{\mathrm{M},4} = 72.3\,\mathrm{MPa}$ | $\tau_4 = 10^{-1}\,\mathrm{s}$ | $G_{\mathrm{M},7} = 36.7\,\mathrm{MPa}$ | $\tau_7 = 10^{2}\,\mathrm{s}$ |
| | $G_{\mathrm{M},2} = 127.0\,\mathrm{MPa}$ | $\tau_2 = 10^{-3}\,\mathrm{s}$ | $G_{\mathrm{M},5} = 50.9\,\mathrm{MPa}$ | $\tau_5 = 10^{0}\,\mathrm{s}$ | $G_{\mathrm{M},8} = 31.5\,\mathrm{MPa}$ | $\tau_8 = 10^{3}\,\mathrm{s}$ |
| | $G_{\mathrm{M},3} = 97.6\,\mathrm{MPa}$ | $\tau_3 = 10^{-2}\,\mathrm{s}$ | $G_{\mathrm{M},6} = 38.4\,\mathrm{MPa}$ | $\tau_6 = 10^{1}\,\mathrm{s}$ | $G_{\mathrm{M},9} = 30.6\,\mathrm{MPa}$ | $\tau_9 = 10^{4}\,\mathrm{s}$ |

**Table 1:** Material parameters for the linear elastic fibers and the viscoelastic polymer matrix

We start with the identification of a DMN surrogate model using the *conventional* loss

$$J_\mathrm{C}(\boldsymbol{p}) = \frac{1}{N_s}\sum_{s=1}^{N_\mathrm{s}} \frac{\|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M}) - \mathcal{M}'(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M})\|_1}{\|\mathcal{M}'(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M})\|_1} + \psi(\boldsymbol{p}). \tag{17}$$

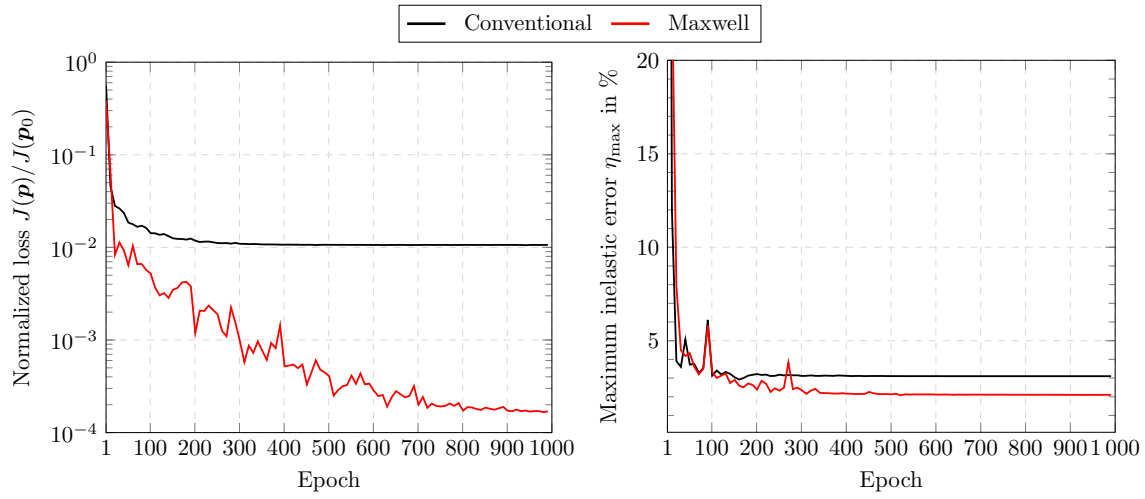
To generate the training data, we sample tuples of orthotropic stiffnesses $\{(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M})\}_{s=1}^{N_\mathrm{s}}$, see Gajek et al. [6] for details. Furthermore, we identify a DMNs using the *Maxwell-informed* loss

$$J_\mathrm{M}(\boldsymbol{p}) = \frac{\lambda}{N_s}\sum_{s=1}^{N_\mathrm{s}} \frac{\|\mathcal{M}'^{\mathrm{DMN}}_{\boldsymbol{p}}(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M}) - \mathcal{M}'(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M})\|_1}{1-y^s} + \psi(\boldsymbol{p}), \tag{18}$$

sample $\{y^s\}_{i=1}^{N_\mathrm{s}}$ using the Sobol sequence [15] and assemble the corresponding stiffness samples $\{(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{M})\}_{s=1}^{N_\mathrm{s}}$ by evaluating relations (16). The positive constant $\lambda$ is set to $\lambda = \min_i(\lambda_i)$. For both procedures, we sample $N_\mathrm{s} = 1024$ stiffness tuples. For efficiently labeling the training data, we precompute the effective stiffnesses $\bar{\mathbb{C}}^s = \mathcal{M}'(\mathbb{C}^s_\mathrm{F},\mathbb{C}^s_\mathrm{s})$ by means of an FFT-based micromechanics code [16] using the conjugate gradient solver [17].

For both aforementioned losses, sampled and labeled training data, we train a DMN with $K = 8$ layers for 1000 epochs using minibatches of size 8 and using the Adam optimizer [18]. The initial fitting parameters $\boldsymbol{p}_0 \in \mathcal{P}$, i.e., volume fractions and lamination directions, are initialized as explained in Gajek et al. [6]. The learning rate, i.e., step size of stochastic gradient descent, is set to 0.015. The learning rate is halved every 100 epochs to improve convergence.

Fig. 2a illustrates the training progression. For better comparability, both losses $J_\mathrm{C}(\boldsymbol{p})$ and $J_\mathrm{M}(\boldsymbol{p})$ are normalized with their initial values $J_\mathrm{C}(\boldsymbol{p}_0)$ and $J_\mathrm{M}(\boldsymbol{p}_0)$, respectively. We observe that the Maxwell-informed loss drops by up to four orders of magnitude, whereas the conventional approach shows an improvement of two orders of magnitude.




**(a)** Normalized loss vs. epochs

**(b)** Maximum inelastic error vs. epochs

**Fig. 2:** Loss and maximum inelastic error during model identification

To quantify the performance of the two surrogate models for the viscoelastic case, we define the following inelastic error

$$\eta(t) = \frac{\|\bar{\boldsymbol{\sigma}}^{\mathrm{DMN}}(t) - \bar{\boldsymbol{\sigma}}(t)\|_1}{\max_\tau \|\bar{\boldsymbol{\sigma}}(\tau)\|_1}. \tag{19}$$

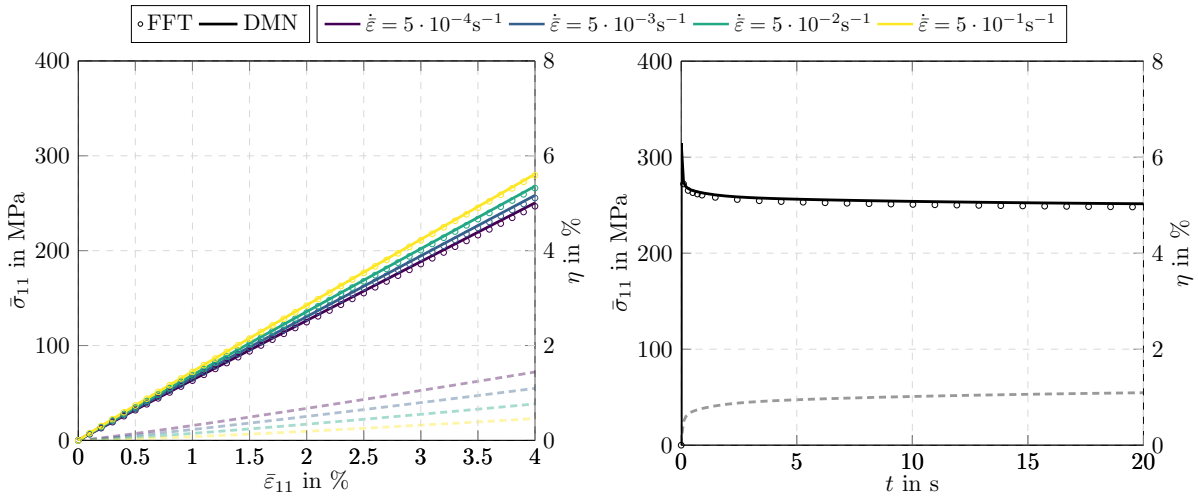
The full-field solutions $\bar{\boldsymbol{\sigma}}(t)$, which serve as reference, are computed by an FFT-based micromechanics code [16] using an inexact Newton-CG solver [19]. At every tenth epoch during training, we compute 24 tensile tests and six relaxation tests. For tensile and relaxation tests, we use mixed boundary conditions [20] and apply the load in the following six loading directions separately

$$\bar{\boldsymbol{\varepsilon}} = \frac{\bar{\varepsilon}}{2}(\boldsymbol{e}_i \otimes \boldsymbol{e}_j + \boldsymbol{e}_j \otimes \boldsymbol{e}_i) \quad \text{with} \quad (i,j) \in \{(1,1),(2,2),(3,3),(1,2),(1,3),(2,3)\}\,. \tag{20}$$

For every loading direction, we load to a macroscopic strain of $\bar{\varepsilon} = 4\,\%$ with four different strain rates, i.e., 24 tensile loadings in total, which are logarithmically spaced from $\dot{\bar{\varepsilon}} = 5 \cdot 10^{-4}\,\mathrm{s}^{-1}$ to $\dot{\bar{\varepsilon}} = 5 \cdot 10^{-1}\,\mathrm{s}^{-1}$. For the relaxation tests, we ramp the macroscopic strain to $\bar{\varepsilon} = 4\,\%$ in one millisecond and hold the load for twenty more seconds.

Fig. 2b shows the progression of the maximum inelastic error $\eta_{\max}$, i.e., the maximum over all times steps, loadings and loading directions vs. the epochs. We observe that for both approaches, the maximum inelastic error decreases rapidly during training and no further improvement is observed above 400 epochs. At the end of training, the conventional approach leads to a maximum inelastic error of around $3.5\,\%$, whereas the Maxwell-informed loss provides a model that exhibits an inelastic error of around $2\,\%$.

To get a better impression of how a maximum inelastic error of $2\,\%$ translates into actual stress-strain curves, Fig. 3a summarizes four uniaxial extensions in the 11-direction. We focus on the effective stress in principal fiber direction since this case is typically rather challenging, i.e., strong morphological anisotropy and strain and stress peaks at the fiber ends, even for such simple material models as viscoelasticity. Indeed, the rate dependence of the material is clearly seen and the agreement between full-field solutions and DMN predictions is excellent with inelastic errors well below $2\,\%$. In Fig. 3b, the result for the relaxation test in principal fiber directions is shown. We observe that the surrogate model also reproduces the relaxation behavior with high accuracy.




**(a)** Uniaxial extension in principal fiber direction

**(b)** Relaxation test in principal fiber direction

**Fig. 3:** Comparing the effective stresses computed by an FFT-based micromechanics solver and predicted by the DMN, trained with the Maxwell-informed loss, for a uniaxial extension in principal fiber direction and a relaxation test in principal fiber direction

Having demonstrated that the Maxwell-informed loss leads to lower inelastic errors, we next investigate the influence of the number of samples on the inelastic error. For this purpose, we train DMNs with a different number of samples. We train in each case eight DMNs with 16, 32, 64, 128, 256, 512 and 1024 samples to assess mean and standard deviation of the maximum inelastic error $\eta_{\max}$. Note that the batch size is set to eight and the learning rate is scaled inversely with the number of samples to obtain comparable results.

The results are shown in Fig. 4. For the conventional approach, the inelastic error depends on the number of samples. If DMNs are trained with only 16 samples, inelastic errors of around $7\,\%$ will be attained. With an increasing number of samples, the mean and standard deviation of the maximum inelastic error decrease. Only for 512 samples and above, no significant improvement can be observed. Considering the Maxwell-informed approach, we observe a completely different behavior. Already for 16 samples, a maximum inelastic error of $2\,\%$ may be attained. Increasing the number of samples does not improve the results any further. Indeed, a closer look at the models trained with 512 and 1024 samples reveals a slight increase of the mean and standard deviation of the inelastic error.

## 5 Conclusion

In the article at hand, we were concerned with further shedding light on the linear elastic training of deep material networks and its influence on the inelastic performance of the surrogate model.

To derive exact expressions for the inelastic model error, we restricted to linear viscoelasticity. We showed that data sampling and loss function should not be chosen independently, as has been the case so far for training DMNs. We derived a novel material-informed sampling procedure and loss, which allowed for reducing the inelastic error by a factor of almost two compared to a conventionally trained DMN. Furthermore, we demonstrated that the novel sampling allows for further reducing the computational costs for data labeling, i.e., computing the effective stiffnesses. The inelastic results proved to be almost independent of the number of samples. For the special case of viscoelasticity, it was sufficient to train with only 16 samples, representing a speed-up factor of 64. Especially in combination with the superconvergence properties [21] of modern computational micromechanics solvers, the effort of labeling the training date may be drastically reduced.

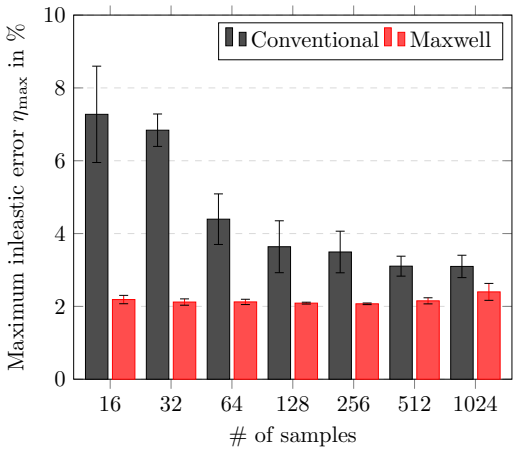



**Fig. 4:** Mean and standard deviation of the maximum inelastic error for DMNs trained with a varying number of samples

Summing up, our results indicate that it is essential to take into account the inelastic and nonlinear material behavior of the phases for training data sampling and formulating a suitable loss. In a next step, we seek to extend our approach to other material classes, such as plasticity or damage [22].

**Acknowledgements** The research documented in this manuscript has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 255730231, within the International Research Training Group "Integrated engineering of continuous-discontinuous long fiber reinforced polymer structures" (GRK 2078/2). The support by the German Research Foundation (DFG) is gratefully acknowledged. Open access funding enabled and organized by Projekt DEAL.

## References

[1] S. Gajek, M. Schneider, and T. Böhlke, Computational Mechanics **69**(5), 1087–1113 (2022).
[2] J. Renard and M. F. Marmonier, Aerospace Science and Technology **9**, 37–51 (1987).
[3] J. Spahn, H. Andrä, M. Kabel, and R. Müller, Computer Methods in Applied Mechanics and Engineering **268**, 871–883 (2014).
[4] Z. Liu, C. T. Wu, and M. Koishi, Computer Methods in Applied Mechanics and Engineering **345**, 1138–1168 (2019).
[5] Z. Liu and C. T. Wu, Journal of the Mechanics and Physics of Solids **127**, 20–46 (2019).
[6] S. Gajek, M. Schneider, and T. Böhlke, Journal of the Mechanics and Physics of Solids **142**, 103984 (2020).
[7] Z. Liu, H. Wei, T. Huang, and C. T. Wu, Intelligent multiscale simulation based on process-guided composite database, in: 16th International LS-DYNA Users Conference, (2020).
[8] S. Gajek, M. Schneider, and T. Böhlke, Computer Methods in Applied Mechanics and Engineering **384**, 113952 (2021).
[9] S. Gajek, M. Schneider, and T. Böhlke, PAMM **21**(1), e202100069 (2021).
[10] N. Meyer, S. Gajek, J. Görthofer, A. Hrymak, L. Kärger, F. Henning, M. Schneider, and T. Böhlke, Composites Part B: Engineering p. 110380 (2022).
[11] N. Halphen and Q. Nguyen, Journal de Mécanique **14**, 508–520 (1975).
[12] G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities (Cambridge University Press, New York, 1934).
[13] E. Tikarrouchine, G. Chatzigeorgiou, Y. Chemisky, and F. Meraghni, International Journal of Solids and Structures **164**, 120–140 (2019).
[14] L. Kehrer, D. Wicht, J. T. Wood, and T. Böhlke, GAMM-Mitteilungen **41**(1), e201800007 (2018).
[15] I. M. Sobol, USSR Computational Mathematics and Mathematical Physics **7**, 86–112 (1967).
[16] H. Moulinec and P. Suquet, Comptes Rendus de l'Académie des Sciences. Série II **318**(11), 1417–1423 (1994).
[17] J. Zeman, J. Vondřejc, J. Novak, and I. Marek, Journal of Computational Physics **229**(21), 8065–8071 (2010).
[18] D. P. Kingma and J. Ba, 3rd International Conference on Learning Representations, ICLR pp. 1–15 (2015).
[19] M. Kabel, T. Böhlke, and M. Schneider, Computational Mechanics **54**(6), 1497–1514 (2014).
[20] M. Kabel, S. Fliegener, and M. Schneider, Computational Mechanics **57**(2), 193–210 (2016).
[21] M. Schneider, International Journal for Numerical Methods in Engineering **123**(17), 4119–4135 (2022).
[22] J. Görthofer, M. Schneider, A. Hrymak, and T. Böhlke, International Journal of Damage Mechanics **31**(1), 43–86 (2022).