



Machine learning assisted discovery of effective viscous material laws for shear-thinning fiber suspensions

Benedikt Sterr¹ · Andrew Hrymak² · Matti Schneider³ · Thomas Böhlke¹

Received: 1 February 2024 / Accepted: 17 April 2024
© The Author(s) 2024

Abstract

In this article, we combine a Fast Fourier Transform based computational approach and a supervised machine learning strategy to discover models for the anisotropic effective viscosity of shear-thinning fiber suspensions. Using the Fast Fourier Transform based computational approach, we study the effects of the fiber orientation state and the imposed macroscopic shear rate tensor on the effective viscosity for a broad range of shear rates of engineering process interest. We visualize the effective viscosity in three dimensions and find that the anisotropy of the effective viscosity and its shear rate dependence vary strongly with the fiber orientation state. Combining the results of this work with insights from literature, we formulate four requirements a model of the effective viscosity should satisfy for shear-thinning fiber suspensions with a Cross-type matrix fluid. Furthermore, we introduce four model candidates with differing numbers of parameters and different theoretical motivations, and use supervised machine learning techniques for non-convex optimization to identify parameter sets for the model candidates. By doing so, we leverage the flexibility of automatic differentiation and the robustness of gradient based, supervised machine learning. Finally, we identify the most suitable model by comparing the prediction accuracy of the model candidates on the fiber orientation triangle, and find that multiple models predict the anisotropic shear-thinning behavior to engineering accuracy over a broad range of shear rates.

Keywords Effective viscosity · Fiber-reinforced composites · Non-Newtonian suspension · Supervised machine learning · Cross-fluid

1 Introduction

1.1 State of the art

Viscosity models for fiber polymer suspensions are widely used in molding process simulations of composite parts [1], which play a major role in the lightweight design of engineering systems, i.e., in the automotive, aerospace, and energy sectors [2, 3]. Molding process simulations are an estab-

lished tool in composite part engineering, partly because of the industrial benefits of digital twins [4] and virtual process chains [2, 5, 6]. In molding simulations for fiber reinforced plastics, accurate modeling of the suspension viscosity is crucial to predict various parameters of engineering interest. The suspension viscosity influences manufacturing process parameters [7, 8], as well as fiber orientation and fiber volume distributions [9]. Consequently, flow fields [10] and final part properties [10, 11] are also affected by the suspension viscosity. However, finding analytical models for the suspension viscosity is a challenging task, partly because of the locally inhomogeneous flow field inside the suspension, as well as the hydrodynamic interactions [12] and mechanical contacts [13] between the fibers. Furthermore, the suspension viscosity also depends strongly on the local microstructure, i.e., the fiber geometry [14], the fiber volume fraction [15], and the fiber orientation state [16]. As far as external influences are concerned, the loading direction, the shear rate [17, 18], and the melt temperature [19] also affect the suspension viscosity. Especially for suspensions with non-Newtonian

✉ Benedikt Sterr
benedikt.sterr@kit.edu

✉ Thomas Böhlke
thomas.boehlke@kit.edu

¹ Karlsruhe Institute of Technology (KIT), Institute of Engineering Mechanics, Karlsruhe, Germany

² Department of Chemical and Biochemical Engineering, University of Western Ontario, London, Canada

³ Institute of Engineering Mathematics, University of Duisburg-Essen, Essen, Germany

solvents and fiber concentrations beyond the semi-dilute regime, capturing the large variety of effects in a single analytical model is a difficult task.

Based on the work by Batchelor [20, 21] on fiber suspensions with Newtonian matrix behavior, Goddard [22, 23] proposed self-consistent analytical models of dilute and semi-dilute fiber suspensions with power-law shear-thinning matrix behavior. While the model predictions agree well with experimental results qualitatively, as found by Goddard [24] and Mobuchon et al. [25], quantitative accuracy can still be improved. Later, Souloumiac and Vincent [26] incorporated fiber orientation distributions into a self-consistent modeling approach for dilute, semi-dilute and concentrated fiber suspensions. However, in a comparison with experimental results in a convergent channel flow, the prediction accuracy of the model varies strongly with the shear rate and the fiber volume fraction. More recently, Ferec et al. [27] proposed semi-analytical and numerical self-consistent models for fiber suspensions with Ellis and Carreau-type matrix behavior. Even though the semi-analytical model was able to accurately replicate steady state solutions of a simple shear flow simulation, the models have yet to be successfully applied in other flow scenarios. Focusing on the concentrated and the hyperconcentrated regime, Pipes et al. [28–31] developed models of collimated fiber suspensions, including uniformly distributed fiber misalignments with an orientation averaging [32, 33] approach. The model predictions agree well with experimental results by Binding [34], but the applicability of the model is restricted to collimated fiber arrays. To improve existing molding simulation solvers through small modifications, Favalaro et al. [35] combined orientation averaging with a deformation mode and microstructure dependent informed isotropic viscosity, and successfully predicted the shell-core effect common in fiber suspension molding. However, the error of the model depends strongly on the applied deformation mode and approximated anisotropic viscosity. Thus, including recent developments [36, 37] it proves difficult to achieve high model prediction accuracy over the wide variety of application requirements in engineering systems.

In addition to the challenging analytical treatment, the investigation of fiber suspensions via rheological experiments is also difficult. Because of the transient effects during fiber suspension molding, including fiber breakage and change of orientation state, it is difficult to determine the suspension viscosity for a particular microstructure and load case [34, 38]. Furthermore, interaction of fibers with measurement devices [39] can affect the measured quantities.

In light of the analytical and experimental difficulties involved when studying the viscosity of fiber suspensions with non-Newtonian solvents, computational approaches provide insights and observations that are otherwise hard to obtain. Švec et al. [40] combined the lattice Boltzmann

method for fluid flow, an immersed boundary procedure for the interaction between fluid and rigid particles, and a mass tracking algorithm for the free surface representation to simulate slump tests of a suspension of rigid spherical particles and fibers suspended in a Bingham-type fluid. In comparison with a slump test of pure matrix material, they observed a smaller spread and increased height in the slump test of the suspension, suggesting an increased effective yield stress. Using a Finite Element Method (FEM) based approach, Domurath et al. [41] investigated the rheological coefficients of the transversely isotropic fluid equation by Ericksen [32] in the context of rigid fibers suspended in a power-law fluid. In a simple shear flow, they found the model by Souloumiac and Vincent [26] to overpredict the orientation dependence of a rheological coefficient. Combining the RVE method [42] and Fast Fourier Transform (FFT) based computational techniques [43], Sterr et al. [18] extended work by Bertóti et al. [44] on Newtonian fiber suspensions to suspensions with non-Newtonian solvents. They visualized the anisotropic viscosity tensor in the case of a Cross-type matrix fluid and studied the effects of fiber volume fraction and shear rate on the suspension viscosity.

1.2 Contributions

In this article, we combine high fidelity, FFT-based computational methods and a supervised machine learning strategy to discover material models for the effective viscosity of shear-thinning fiber suspensions. To do so, we study the anisotropic and shear rate dependent effective viscosity of fiber suspensions with a Cross-type matrix behavior and a fiber volume fraction of 25%. In section 2 we outline the setup for the FFT-based computational investigations, where we compute the effective material response for 109 different fiber orientation states and a variety of loading states. In terms of computational scope, this constitutes a significant extension to previous FFT-based work on shear-thinning fiber suspensions, in which a single fiber orientation state was considered [18]. Based on the computational data, we visualize the suspension viscosity tensor in three dimensions and formulate an anisotropy criterion applicable in a non-linear setting using a loading direction and shear rate dependent scalar viscosity, see section 3. We find that the anisotropy of the suspension viscosity shows a significantly different shear rate dependence for different fiber orientation states. Also, for all investigated microstructures, we confirm that the anisotropy of the suspension viscosity depends on the degree of non-linearity of the matrix material, which was previously only studied for a transversely isotropic microstructure [18]. In section 4, based on the insights into the shear rate and orientation state dependence of the suspension viscosity, we formulate requirements a model of the suspension viscosity needs to fulfill. According to the formulated requirements we

propose four models with different numbers of parameters, as well as different phenomenological and theoretical motivations. Using the ADAM algorithm, we employ supervised machine learning techniques for non-convex optimization to learn the parameters of the model candidates. Overall, three models achieve similar maximum validation errors below 5.15 %, while one model based on superposition and orientation averaging performs unfavorably.

2 A computational study of the effective behavior of shear-thinning fiber suspensions

2.1 Description of fiber orientation states

Since the effective behavior of a fiber suspension depends on the fiber orientation statistics in the suspension, a mathematical description of the orientation state is essential to the study of fiber suspensions. In the following, we use a fiber orientation distribution function [45]

$$\rho : S^2 \rightarrow \mathbb{R}, \quad \mathbf{n} \mapsto \rho(\mathbf{n}), \tag{2.1}$$

to encode the probability that fibers are oriented in direction \mathbf{n} on the 2-sphere S^2 , and thus describe the orientation state of a fiber suspension. The temporal and spatial evolution of the orientation distribution function ρ is governed by a partial differential equation, the Fokker–Planck equation [46]. However, in component scale molding simulations, computing the evolution of the orientation distribution function ρ using the Fokker-Planck equation requires huge computational effort. Instead, it is common [1] to use the second order fiber orientation tensor [45, 47]

$$\mathbf{N} = \int_{S^2} \mathbf{n} \otimes \mathbf{n} \rho(\mathbf{n}) \, dS(\mathbf{n}), \tag{2.2}$$

as a measure for the fiber orientation state. The tensor \mathbf{N} is symmetric, positive semi-definite, and has unit trace, such that

$$N_{ii} = \lambda_1 + \lambda_2 + \lambda_3 = 1, \tag{2.3}$$

where λ_1, λ_2 , and λ_3 denote the eigenvalues of the tensor \mathbf{N} . By sorting the eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \lambda_3, \tag{2.4}$$

and considering equation (2.3), the bounds for the largest eigenvalue λ_1 are found as

$$1 \geq \lambda_1 \geq \frac{1}{3}. \tag{2.5}$$

Additionally, an upper bound for the eigenvalue λ_2 follows from positive semi-definiteness of the tensor \mathbf{N} , and a lower bound may be derived by eliminating λ_3 from equation (2.4) using equation (2.3), such that

$$\min(\lambda_1, 1 - \lambda_1) \geq \lambda_2 \geq \frac{1 - \lambda_1}{2}. \tag{2.6}$$

Thus, we may parametrize the tensor \mathbf{N} by its two eigenvalues λ_1 and λ_2 , and the rotation \mathbf{Q} in terms of an eigendecomposition

$$\mathbf{N} = \mathbf{Q} \, \text{diag}(\lambda_1, \lambda_2, \lambda_3) \, \mathbf{Q}^T, \tag{2.7}$$

where the **diag** operator constructs a second order tensor in the standard basis of \mathbb{R}^3 . By objectivity, the results of this article generalize to all rotations \mathbf{Q} , and for simplicity we choose

$$\mathbf{Q} = \mathbf{I}, \tag{2.8}$$

where \mathbf{I} is the second order unit tensor. Consequently, any second order fiber orientation tensor \mathbf{N} may be encoded by a vector $\underline{\lambda} = (\lambda_1, \lambda_2)^T$, and represents a point in the fiber orientation triangle S_T defined through equations (2.5) and (2.6), such that

$$S_T = \left\{ \underline{\lambda} = (\lambda_1, \lambda_2)^T \mid \begin{array}{l} 1 \geq \lambda_1 \geq \frac{1}{3} \text{ and} \\ \min(\lambda_1, 1 - \lambda_1) \geq \lambda_2 \geq \frac{1 - \lambda_1}{2} \end{array} \right\}. \tag{2.9}$$

Following Köbler et al. [48], we use a CMYK coloring scheme to visualize the fiber orientation state within the fiber orientation triangle S_T . The isotropic, unidirectional, and planar isotropic orientation states represent the corners of the fiber orientation triangle and are colored cyan, magenta, and yellow, respectively.

2.2 Computational homogenization procedure

In this article, we consider incompressible, shear-thinning fiber suspensions, occupying a rectangular volume $Y \subseteq \mathbb{R}^3$. First, we give a short summary of the numerical procedure for the homogenization of shear-thinning fiber suspensions following Sterr et al. [18, §2]. We are interested in computing the effective viscous stress $\bar{\boldsymbol{\tau}} \in \text{Sym}_0(3)$ of the suspension in response to an applied, effective shear rate tensor $\bar{\mathbf{D}} \in \text{Sym}_0(3)$, where $\text{Sym}_0(3)$ denotes the vector space of symmetric and traceless second-order tensors. For a pressure field $p : Y \rightarrow \mathbb{R}$, a solenoidal velocity field $\mathbf{v} : Y \rightarrow \mathbb{R}^3$ and a local viscous stress field

$$\boldsymbol{\tau} = \mathcal{T}(\cdot, \bar{\mathbf{D}} + \nabla^s \mathbf{v}), \tag{2.10}$$

with associated stress operator

$$\mathcal{T} : Y \times \text{Sym}_0(3) \rightarrow \text{Sym}_0(3), \tag{2.11}$$

$$(\mathbf{x}, \mathbf{D}) \mapsto \frac{\partial \Psi}{\partial \mathbf{D}}(\mathbf{x}, \mathbf{D}), \tag{2.12}$$

and dissipation potential $\Psi : Y \times \text{Sym}_0(3) \rightarrow \mathbb{R}$, the local balance of linear momentum without inertial effects

$$\text{div } \boldsymbol{\tau} - \nabla p = \mathbf{0}, \tag{2.13}$$

needs to be fulfilled. In the context of suspensions with rigid inclusions, the local viscous stress $\boldsymbol{\tau}$ in equation (2.13) is, constitutively, not well defined inside the rigid inclusions. We address this issue by changing to a dual formulation as follows. With the complementary dissipation potential

$$\Phi(\boldsymbol{\tau}) := \sup \{ \boldsymbol{\tau} \cdot \mathbf{D} - \Psi(\mathbf{D}) \mid \mathbf{D} \in \text{Sym}_0(3) \}, \tag{2.14}$$

arising as the Legendre–Fenchel dual of the dissipation potential Ψ , we search for minimizers of the variational problem

$$\langle \Phi(\cdot, \boldsymbol{\tau}) - \boldsymbol{\tau} \cdot \bar{\mathbf{D}} \rangle_Y \longrightarrow \inf. \tag{2.15}$$

Here, $\langle \cdot \rangle_Y$ denotes the spatial average

$$\langle \cdot \rangle_Y \equiv \frac{1}{|Y|} \int_Y (\cdot) \, d\mathbf{x} \quad \text{with} \quad |Y| \equiv \int_Y d\mathbf{x}, \tag{2.16}$$

and the infimum is taken over all stress fields $\boldsymbol{\tau}$ satisfying the equilibrium equation (2.13). Minimizers of the variational problem (2.15) satisfy the Euler–Lagrange equation

$$\mathbb{P}_{\mathcal{E}} \left[\frac{\partial \Phi}{\partial \boldsymbol{\tau}}(\boldsymbol{\tau}) - \bar{\mathbf{D}} \right] = \mathbf{0}, \tag{2.17}$$

where $\mathbb{P}_{\mathcal{E}}$ refers to the L^2 -projector onto the shear stresses satisfying the equilibrium equation (2.13). We refer to Bertóti et al. [44, §2] for a closed form expression of the action of $\mathbb{P}_{\mathcal{E}}$ in Fourier space. In case of suspensions with rigid particles the shear rate tensor \mathbf{D} and hence the derivative $\partial \Phi / \partial \boldsymbol{\tau}$ vanishes inside the particles. Thus, the formulation of the optimization problem (2.15) in terms of the complementary dissipation potential Φ is advantageous for numerical schemes based on the Euler–Lagrange equation (2.17). If the optimization problem (2.15) is convex, a minimizer can be found with, e.g., gradient descent [18], the Barzilai–Borwein method [49], or Newton-CG [50, 51] approaches. Finally, the effective viscous stress $\bar{\boldsymbol{\tau}}$ can be computed by spatial averaging, such that

$$\bar{\boldsymbol{\tau}} = \langle \boldsymbol{\tau} \rangle_Y. \tag{2.18}$$

Table 1 Parameters of the Cross-type material law (2.19) for a commercially available polyamide 6 [53]

η_0	η_∞	k	m
288.9 Pa s	15.0 Pa s	$10.9 \cdot 10^{-4}$	1.1

2.3 Computational study setup and material parameters

To study the material response of shear-thinning fiber suspensions with a variety of microstructures, we generated fiber suspension microstructures for 109 points of the fiber orientation triangle S_T (2.9), see Figs. 1a and 2, using the sequential addition and migration method [52]. Building upon the investigations in Sterr et al. [18], a commercially available polyamide 6 [53] was chosen as the matrix material, and a Cross-type material law

$$\eta(\dot{\gamma}) = \eta_\infty + \frac{\eta_0 - \eta_\infty}{1 + (k\dot{\gamma})^m}, \tag{2.19}$$

was fitted to the available material data for shear rates $\dot{\gamma}$ in the interval $[1.7, 16300] \text{ s}^{-1}$ at a temperature of 250°C . The resulting model parameters are collected in Table 1.

The viscosities η_0 and η_∞ define the material behavior for shear rates $\dot{\gamma} = 0$ and $\dot{\gamma} \rightarrow \infty$, respectively, while the parameters k and m control the non-linear transition between the viscosities η_0 and η_∞ . The suspension microstructures were discretized on a staggered grid [54] using composite voxels [55] with a general dual mixing rule for the special case of rigid particles [18]. The resulting non-linear system of equations was solved with a Newton-CG approach. To limit the required computational effort, we restricted to microstructures with a fiber volume fraction $c_F = 25\%$, where all fibers have equal length ℓ and diameter d . More precisely, we prescribed an aspect ratio $r_a = \ell/d$ of 10. The resolutions and sizes of the microstructure volume elements were chosen according to the investigations in Sterr et al. [18], such that the number of voxels per fiber diameter is $v/d = 15$ and the edge length of the cubic volume elements is $L = 2.2\ell$.

For each macroscopic scalar shear rate $\dot{\gamma}$ in the set of studied shear rates $S_{\dot{\gamma}}$, such that

$$\dot{\gamma} \in S_{\dot{\gamma}} = \left\{ a \cdot 10^b \text{ s}^{-1} \mid a = 1, 2, 5; b = 1, 2, 3, 4 \right\} \cup \left\{ 10^5 \text{ s}^{-1} \right\}, \tag{2.20}$$

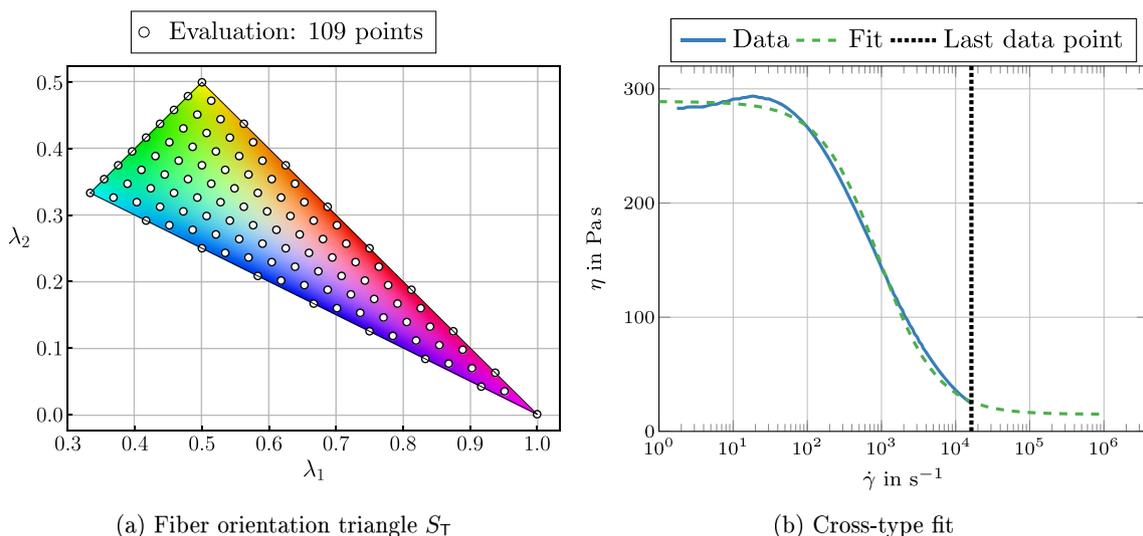


Fig. 1 Fiber orientation triangle S_T in CMYK coloring with 109 evaluation points (a), and material data with Cross-type fit for Ultramid@B3K (b)

we investigate the six load cases collected in the matrix $\underline{\underline{D}}$ in Mandel notation

$$\underline{\underline{D}} = \dot{\gamma} \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\frac{3}{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\frac{3}{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{2}} \end{bmatrix}. \tag{2.21}$$

Here, the set of studied shear rates $S_{\dot{\gamma}}$ is intended to cover a broad variety of engineering process shear rates, and is motivated by the typical shear rates in compression molding, standard injection molding, as well as thin-wall and micro molding [56, 57]. For shear rates $\dot{\gamma}$, where the matrix behavior is mostly Newtonian and the superposition of material responses is valid, we compute and collect the components of the effective viscosity tensor $\underline{\underline{V}}$ in a matrix $\underline{\underline{V}}$ with

$$\underline{\underline{\tau}} = \underline{\underline{V}} \underline{\underline{D}}, \quad \text{i.e.,} \quad \underline{\underline{V}} = \underline{\underline{\tau}} \underline{\underline{D}}^\dagger. \tag{2.22}$$

Here, $(\cdot)^\dagger$ stands for the Moore–Penrose pseudoinverse, and $\underline{\underline{\tau}}$ collects the computed effective stresses

$$\underline{\underline{\tau}} = \begin{bmatrix} \bar{\tau}_{11}^{(1)} & \bar{\tau}_{11}^{(2)} & \bar{\tau}_{11}^{(3)} & \bar{\tau}_{11}^{(4)} & \bar{\tau}_{11}^{(5)} & \bar{\tau}_{11}^{(6)} \\ \bar{\tau}_{22}^{(1)} & \bar{\tau}_{22}^{(2)} & \bar{\tau}_{22}^{(3)} & \bar{\tau}_{22}^{(4)} & \bar{\tau}_{22}^{(5)} & \bar{\tau}_{22}^{(6)} \\ \bar{\tau}_{33}^{(1)} & \bar{\tau}_{33}^{(2)} & \bar{\tau}_{33}^{(3)} & \bar{\tau}_{33}^{(4)} & \bar{\tau}_{33}^{(5)} & \bar{\tau}_{33}^{(6)} \\ \sqrt{2}\bar{\tau}_{23}^{(1)} & \sqrt{2}\bar{\tau}_{23}^{(2)} & \sqrt{2}\bar{\tau}_{23}^{(3)} & \sqrt{2}\bar{\tau}_{23}^{(4)} & \sqrt{2}\bar{\tau}_{23}^{(5)} & \sqrt{2}\bar{\tau}_{23}^{(6)} \\ \sqrt{2}\bar{\tau}_{13}^{(1)} & \sqrt{2}\bar{\tau}_{13}^{(2)} & \sqrt{2}\bar{\tau}_{13}^{(3)} & \sqrt{2}\bar{\tau}_{13}^{(4)} & \sqrt{2}\bar{\tau}_{13}^{(5)} & \sqrt{2}\bar{\tau}_{13}^{(6)} \\ \sqrt{2}\bar{\tau}_{12}^{(1)} & \sqrt{2}\bar{\tau}_{12}^{(2)} & \sqrt{2}\bar{\tau}_{12}^{(3)} & \sqrt{2}\bar{\tau}_{12}^{(4)} & \sqrt{2}\bar{\tau}_{12}^{(5)} & \sqrt{2}\bar{\tau}_{12}^{(6)} \end{bmatrix}. \tag{2.23}$$

3 Spatial representation and anisotropy of the effective suspension viscosity

With the goal of modeling the effective behavior of shear-thinning fiber suspensions in mind, it is essential to understand the anisotropic effective viscosity of such suspensions, as well as its dependence on the fiber orientation state $\underline{\lambda}$ and the shear rate $\dot{\gamma}$ first. Accordingly, we begin by visualizing the effective viscosity and studying its anisotropy for low shear rates $\dot{\gamma}$ in the following, before studying the effects of non-linear shear-thinning. To reduce the dimensional complexity when studying the effective viscosity, we use a scalar elongational viscosity η_{app} . Following Sterr et al. [18, §4], we define the viscosity η_{app} based on a modified approach by Böhlke and Brüggemann [58], such that

$$\eta_{app}(\dot{\gamma}, \underline{\lambda}, \mathbf{d}) = \underline{\underline{V}}(\dot{\gamma}, \underline{\lambda}) \cdot (\mathbf{d} \otimes \mathbf{d} \otimes \mathbf{d} \otimes \mathbf{d}), \tag{3.1}$$

where $\mathbf{d} \in S^2$ denotes the direction of elongation, $\underline{\lambda} \in S_T$ refers to the fiber orientation state, and S_T is the fiber orientation triangle defined in equation (2.9). We use equation (2.22) to approximate the effective viscosity $\underline{\underline{V}}$ and thus the elongational viscosity η_{app} with

$$\eta_{app}(\dot{\gamma}, \underline{\lambda}, \mathbf{d}) \approx \underline{\underline{a}}(\mathbf{d})^T \underline{\underline{V}}(\dot{\gamma}, \underline{\lambda}) \underline{\underline{a}}(\mathbf{d}), \tag{3.2}$$

where $\underline{\underline{a}}(\mathbf{d})$ denotes the components of $\mathbf{d} \otimes \mathbf{d}$ in Mandel notation. In addition to the elongational viscosity η_{app} , the bulk viscosity

$$\eta_b(\dot{\gamma}, \underline{\lambda}, \mathbf{d}) = \mathbf{I} \cdot \underline{\underline{V}}(\dot{\gamma}, \underline{\lambda})(\mathbf{d} \otimes \mathbf{d}) = (\mathbf{d} \otimes \mathbf{d}) \cdot \underline{\underline{V}}[\mathbf{I}], \tag{3.3}$$

is required to capture all information contained in the effective viscosity $\underline{\underline{V}}$ [59, Sec. 4.3] if the effective material is

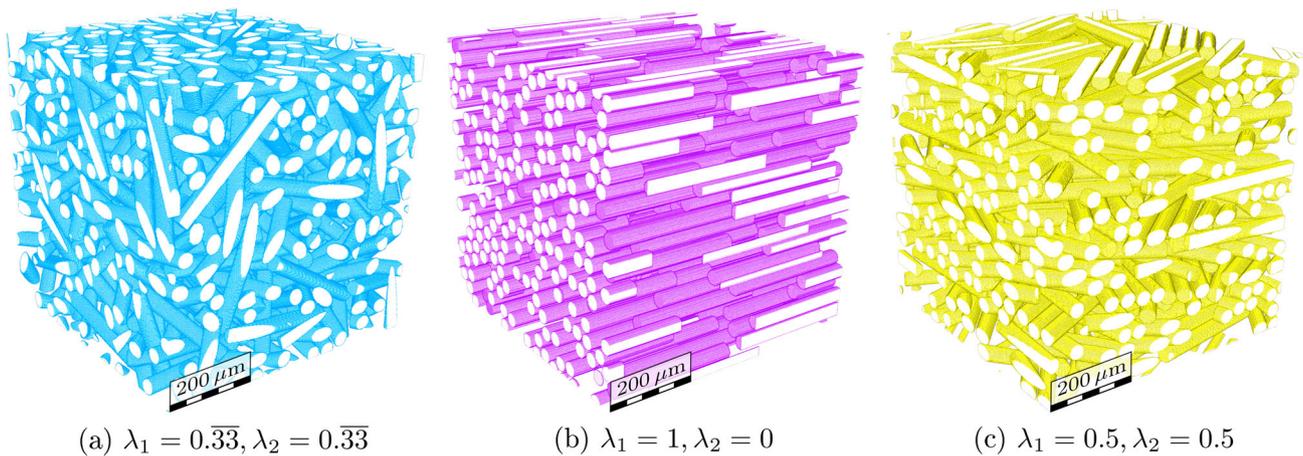


Fig. 2 Investigated microstructures for isotropic (a), unidirectional (b), and planar isotropic (c) fiber orientation states

compressible. However, as we consider incompressible material behavior, the bulk viscosity η_b vanishes, since

$$\bar{\nabla}[I] = 0. \quad (3.4)$$

Hence, the elongational viscosity η_{app} encodes all information contained in the matrix $\bar{\underline{V}}(\dot{\gamma}, \underline{\lambda})$, and allows us to study the anisotropic flow resistance of the suspension in a complete manner. For the elongational viscosity and related quantities we omit denoting the dependence on the shear rate $\dot{\gamma}$, the fiber orientation state $\underline{\lambda}$, and the direction of elongation \underline{d} explicitly. Because of the interpolating property of equation (2.22), we may visualize the direction-dependent material behavior for load cases not collected in $\bar{\underline{D}}$, and develop intuition for the influence of different fiber orientation states on the elongational viscosity η_{app} . For a discussion on the applicability of equation (2.22) in the case of non-linear material behavior, we refer to Sterr et al. [18, §2.3, §4]. For the visualization of the elongational viscosity η_{app} , we restrict to a shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$, where the material behavior is mostly linear and the superposition principle encoded in equation (2.22) holds. The elongational viscosity bodies for the isotropic, the planar isotropic, the transversely isotropic, and the unidirectional fiber orientation states are visualized in Fig. 3 for a shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$. To study the suspension anisotropy at the same shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$, for each orientation state and over all directions \underline{d} , we compute the range $\Delta\eta_{app}$ and the ratio r_{app}

$$\Delta\eta_{app} = \max(\eta_{app}) - \min(\eta_{app}), \quad r_{app} = \frac{\max(\eta_{app})}{\min(\eta_{app})}, \quad (3.5)$$

and collect the results in Table 2. For the isotropic state, the elongational viscosity η_{app} lies between 1513 Pa s to 1576 Pa s, resulting in a range $\Delta\eta_{app} = 63 \text{ Pa s}$ and a

ratio $r_{app} = 1.04$. Because there is no principal fiber orientation axis in the isotropic state, the maximum viscosity η_{app} , the range $\Delta\eta_{app}$, and the ratio r_{app} are lower than in the more strongly oriented states, see Fig. 3a. In contrast, the maxima of the elongational viscosity η_{app} occur in the principal fiber orientation axes clearly visible in Fig. 3b–d. For the planar isotropic state, the maximum of the elongational viscosity η_{app} occurs in the x - y plane with 1947 Pa s, and for the transversely isotropic and unidirectional cases, the maxima occur in the x direction with 2973 Pa s and 4845 Pa s, respectively. Also, with increasing degree of orientation, the range $\Delta\eta_{app}$ and ratio r_{app} grow from 977 Pa s and 2.01 in the planar isotropic state, 1810 Pa s and 2.56 in the transversely isotropic state, and up to 4130 Pa s and 6.78 in the unidirectional state. The strong dependence of the maximum elongational viscosity, the range $\Delta\eta_{app}$, and the ratio r_{app} on the orientation state highlights the influence of the fibers on the anisotropy and magnitude of the effective viscous material behavior. Furthermore, the location and magnitude of the minimum elongational viscosity η_{app} depend on the orientation state as well. In an isotropic volume element, the fibers increase the elongational viscosity η_{app} uniformly across orientation space. Consequently, with a magnitude of 1513 Pa s, the minimum of the effective viscosity η_{app} in the isotropic state is larger than in the other orientation states. In comparison, the minima of the effective viscosity η_{app} are 970 Pa s in the planar isotropic state, 1163 Pa s in the transversely isotropic state, and 715 Pa s in the unidirectional state. Interestingly, the minima of the elongational viscosity η_{app} occur in directions where there is the least flow along the main fiber orientation directions. Because of incompressibility, the direction with the least flow in fiber direction is not perpendicular to the principal fiber orientation axis. Rather, it occurs at a specific angle which depends on the orientation state.

So far, we investigated the effective viscosity of selected fiber suspensions for a relatively low shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$

Table 2 Range $\Delta\eta_{app}$ (3.5) and ratio r_{app} (3.5) of the elongational viscosity η_{app} at a shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$ for isotropic, planar isotropic, transversely isotropic, and unidirectional fiber orientation states

	Isotropic	Planar isotropic	Transversely isotropic	Unidirectional
$\Delta\eta_{app}$	63	977	1810	4130
r_{app}	1.04	2.01	2.56	6.78

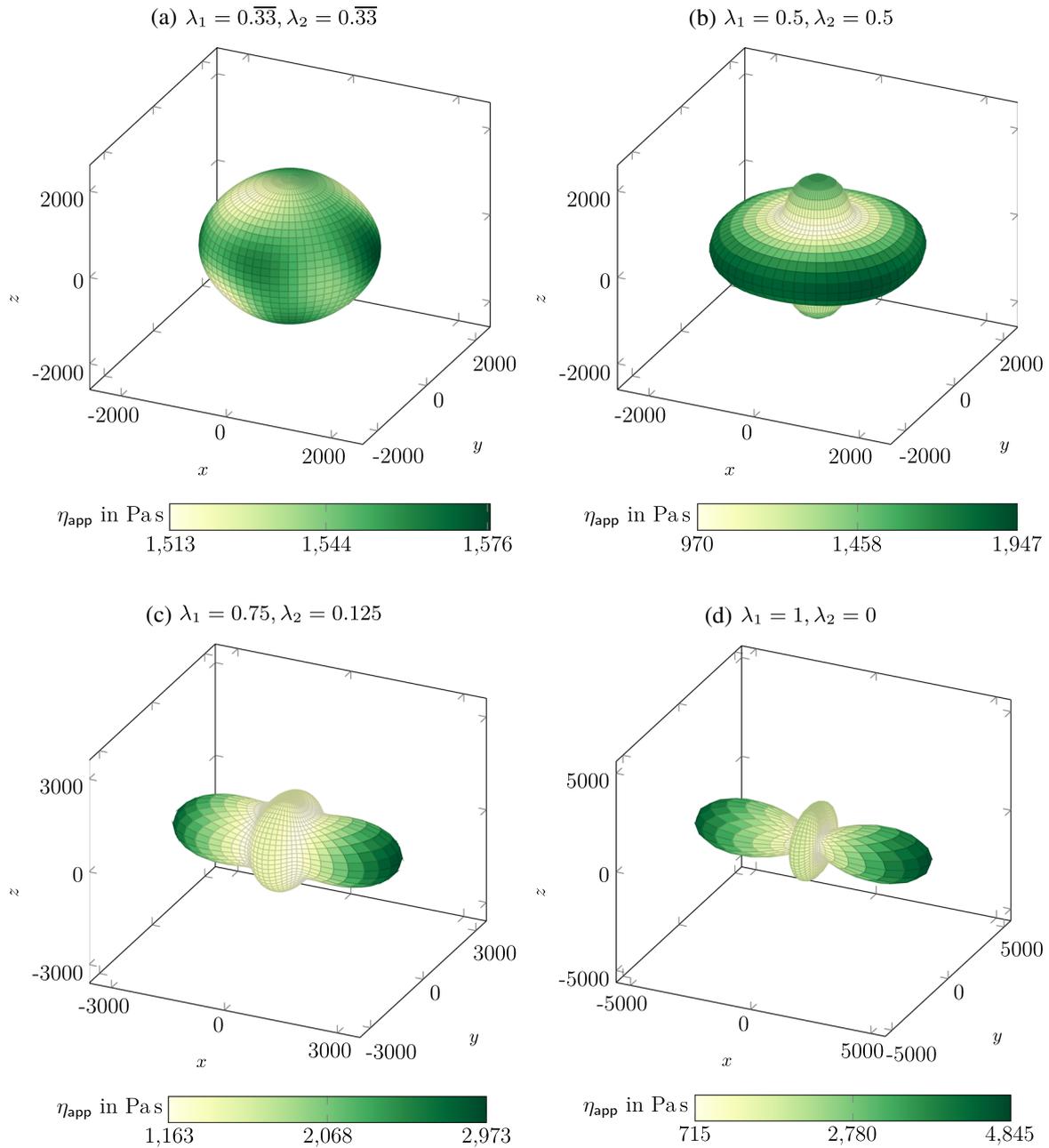


Fig. 3 Elongational viscosity η_{app} at a shear rate $\dot{\gamma} = 10 \text{ 1/s}$ for isotropic (a), planar isotropic (b), transversely isotropic (c), and unidirectional (d) fiber orientation states

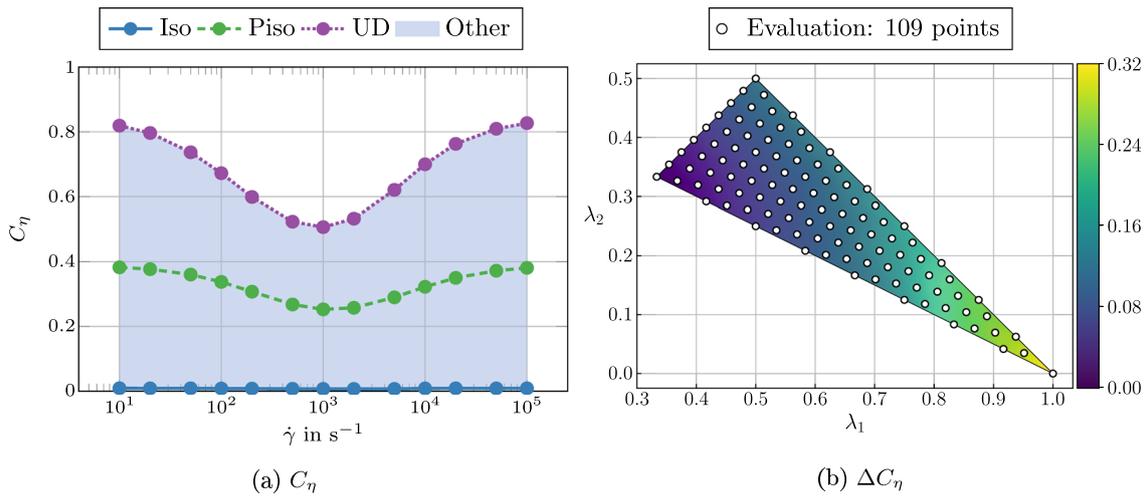


Fig. 4 Coefficient of variation C_η for shear rates $\dot{\gamma} \in [10, 10^5] \text{ s}^{-1}$ (a), and range of the coefficient of variation ΔC_η (b) over the fiber orientation triangle S_τ

using the elongational viscosity η_{app} , the range $\Delta\eta_{\text{app}}$, and the ratio r_{app} . However, the ratio r_{app} or popular anisotropy measures used in crystal elasticity [60–62] rely on the existence of load independent stiffness and compliance tensors. To model the effective suspension viscosity, we want to study its load dependence for all shear rates $\dot{\gamma} \in S_{\dot{\gamma}}$ in the set $S_{\dot{\gamma}}$, which includes shear rates where the effective material behavior is non-linear. Thus, a way of computing the suspension anisotropy without relying on the interpolative properties of equation (2.22) is necessary and is introduced in the following. A convenient expression for the elongational viscosity η_{app} in terms of the investigated load cases collected in the matrix $\bar{\underline{\underline{D}}}$ follows from equation (3.1) with $\bar{\underline{\underline{D}}} = \dot{\gamma}/\sqrt{2} \underline{\underline{d}} \otimes \underline{\underline{d}}$ and $\bar{\underline{\underline{\tau}}} = \bar{\underline{\underline{V}}}[\bar{\underline{\underline{D}}}]$, such that

$$\eta_{\text{app}}(\dot{\gamma}, \underline{\underline{\lambda}}, \bar{\underline{\underline{D}}}) = 2 \frac{\bar{\underline{\underline{D}}} \cdot \bar{\underline{\underline{\tau}}}}{\bar{\underline{\underline{D}}} \cdot \bar{\underline{\underline{D}}}}. \quad (3.6)$$

We use this expression (3.6) to define a vector $\underline{\underline{\eta}}$ for all studied shear rates $\dot{\gamma}$

$$\underline{\underline{\eta}}(\dot{\gamma}, \underline{\underline{\lambda}}) = (\eta_{\text{app}}(\dot{\gamma}, \underline{\underline{\lambda}}, \bar{\underline{\underline{D}}}_1), \eta_{\text{app}}(\dot{\gamma}, \underline{\underline{\lambda}}, \bar{\underline{\underline{D}}}_2), \dots, \eta_{\text{app}}(\dot{\gamma}, \underline{\underline{\lambda}}, \bar{\underline{\underline{D}}}_6))^T, \quad (3.7)$$

which contains the elongational viscosity η_{app} for each load case $\bar{\underline{\underline{D}}}_i$ collected in the columns of the matrix $\bar{\underline{\underline{D}}}$. To measure the average magnitude and the dispersion of the elongational viscosities η_{app} collected in the vector $\underline{\underline{\eta}}$, we compute the mean μ_η and the standard deviation s_η

$$\begin{aligned} \mu_\eta(\dot{\gamma}, \underline{\underline{\lambda}}) &= \frac{1}{6} \sum_{i=1}^6 \eta_{\text{app},i}(\dot{\gamma}, \underline{\underline{\lambda}}), \\ s_\eta(\dot{\gamma}, \underline{\underline{\lambda}}) &= \sqrt{\frac{1}{6} \sum_{i=1}^6 (\eta_{\text{app},i}(\dot{\gamma}, \underline{\underline{\lambda}}) - \mu_\eta(\dot{\gamma}, \underline{\underline{\lambda}}))^2}. \end{aligned} \quad (3.8)$$

The standard deviation s_η is not a dimensionless quantity and its magnitude is a nominal measure of dispersion. Thus, the standard deviation s_η is unsuitable to interpret the dispersion of the values collected in the vector $\underline{\underline{\eta}}$ in relation to their magnitude. Instead, we use the coefficient of variation C_η defined by the equation

$$C_\eta(\dot{\gamma}, \underline{\underline{\lambda}}) = \frac{s_\eta(\dot{\gamma}, \underline{\underline{\lambda}})}{\mu_\eta(\dot{\gamma}, \underline{\underline{\lambda}})}, \quad (3.9)$$

to relate the standard deviation s_η to the mean μ_η and thus compute a dimensionless anisotropy measure of the suspension viscosity for a given fiber orientation state.

Also, the coefficient of variation C_η is a useful quantity to compare the anisotropy of fiber orientation states with elongational viscosities η_{app} of varying magnitudes. To study the maximum load dependent change of the anisotropy for a given fiber orientation state, we define the range over all shear rates ΔC_η

$$\Delta C_\eta(\underline{\underline{\lambda}}) = \max_{\dot{\gamma}} (C_\eta(\dot{\gamma}, \underline{\underline{\lambda}})) - \min_{\dot{\gamma}} (C_\eta(\dot{\gamma}, \underline{\underline{\lambda}})). \quad (3.10)$$

The development of the coefficient of variation C_η for selected orientation states over the shear rate $\dot{\gamma}$, and the range ΔC_η for all fiber orientation states are shown in Fig. 4. For all investigated shear rates and orientation states, the coefficient of variation C_η is largest in the unidirectional state, and smallest in the isotropic state, see Fig. 4a. The coefficient of variation C_η for all other investigated orientation states lies between these bounds. Qualitatively, the coefficient of variation C_η , and hence, the degree of anisotropy at a given shear rate, decreases up to a shear rate $\dot{\gamma} \approx 10^3 \text{ s}^{-1}$ for all orientation states. This decrease in the coefficient C_η suggests increased local velocity gradients, and thus stronger shear-thinning in the matrix, during flow along the fiber ori-

entation directions. Quantitatively the range ΔC_η depends on the orientation state, and is largest in the unidirectional state, see Fig. 4(b).

In the isotropic state, the coefficient of variation C_η and its range ΔC_η are close to zero for all investigated shear rates, showing that non-linear shear thinning effects occur at equivalent strengths, independent of the load direction. With increasing degree of orientation, the coefficient of variation C_η and its range ΔC_η grow. In the planar isotropic state, the coefficient of variation C_η varies with a range $\Delta C_\eta = 0.13$ between $C_\eta = 0.38$ at a shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$ and $C_\eta = 0.25$ at a shear rate $\dot{\gamma} = 10^3 \text{ s}^{-1}$. The range ΔC_η is largest in the unidirectional state, where the coefficient of variation C_η varies strongly with a range $\Delta C_\eta = 0.32$ between $C_\eta = 0.82$ at a shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$ and $C_\eta = 0.50$ at a shear rate $\dot{\gamma} = 10^3 \text{ s}^{-1}$. The distinct differences in the coefficient of variation C_η and the range ΔC_η between the studied fiber orientation states highlight the strong influence of the fiber orientation on the magnitude and anisotropy of the effective material behavior. As a consequence of the Cross-type material law (2.19) and increasing velocity gradients in the polymer matrix, the coefficient of variation C_η increases for shear rates $\dot{\gamma}$ larger than 10^3 s^{-1} up to similar values as observed for a shear rate $\dot{\gamma} = 10 \text{ s}^{-1}$. The results discussed in this section thus underline the need to account for the shear rate and load direction dependent shear thinning behavior of the suspension when modeling the suspension viscosity. This is in line with findings in the literature [29, 31, 35, 63], where the effect of the suspended fibers on the matrix shear rate is estimated using fiber orientation statistics. In the next section we will use a different approach, and discuss how the anisotropic shear thinning effect can be characterized with supervised machine learning based on the computational results of the FFT-based homogenization and knowledge of the local Cross-type material law.

4 Modeling the effective suspension viscosity

4.1 Model requirements

With the results presented in the previous section 3 at hand, we wish to model the effective viscosity of shear-thinning fiber suspensions with analytical means and identify the model parameters using supervised machine learning. To do so, we first summarize the key criteria a model of the effective viscous behavior should fulfill. The model should

1. be tensorial, i.e., capture shear rate and load direction dependence objectively.
2. replicate the local Cross-type material behavior on the macro-scale, i.e., the model should show Newtonian

behavior in the shear rate limits $\dot{\gamma} \rightarrow 0 \text{ s}^{-1}$ and $\dot{\gamma} \rightarrow \infty \text{ s}^{-1}$, and capture the shear rate and load direction dependent, non-linear transition between the two Newtonian limits. This requirement is based on the investigations in Sterr et al. [18, §4], the results of the anisotropy investigation in the previous section 3, and considerations in literature [29, 31, 35, 63].

3. yield an incompressible and orthotropic effective viscosity $\bar{\mathbb{V}}$. Because the fourth order fiber orientation tensors of the generated microstructures are orthotropic, see Schneider [52] and Montgomery-Smith et al. [64], the effective viscosity $\bar{\mathbb{V}}$ of the suspensions is orthotropic as well.
4. be applicable on the whole fiber orientation triangle S_T , as defined in equation (2.9).

In accordance with requirements 1 and 2, we restrict our investigations to tensorial models of the type

$$\bar{\mathbb{V}}(\mathbf{D}, \underline{a}) = \bar{\mathbb{V}}_\infty(\underline{a}) + \mathbb{T}_8(\mathbf{D}, \underline{a}) [\bar{\mathbb{V}}_0(\underline{a}) - \bar{\mathbb{V}}_\infty(\underline{a})], \quad (4.1)$$

where $\underline{a} \in \mathbb{R}^m$, $m \in \mathbb{N}$, is the vector of model parameters, $\mathbb{T}_8 : \text{Sym}_0(3) \times \mathbb{R}^m \rightarrow (\mathbb{R}^3)^{\otimes 8}$ stands for an eighth order tensor function, and $\bar{\mathbb{V}}_0 : \mathbb{R}^m \rightarrow \text{Sym}_0(3)$ as well as $\bar{\mathbb{V}}_\infty : \mathbb{R}^m \rightarrow \text{Sym}_0(3)$ denote fourth order tensor functions. The functions $\bar{\mathbb{V}}_0$ and $\bar{\mathbb{V}}_\infty$ are used to construct the Newtonian viscosity tensors $\bar{\mathbb{V}}_0(\underline{a})$ and $\bar{\mathbb{V}}_\infty(\underline{a})$ in the shear rate limits $\dot{\gamma} \rightarrow 0 \text{ s}^{-1}$ and $\dot{\gamma} \rightarrow \infty \text{ s}^{-1}$, respectively. Equation (4.1) is a tensor-valued generalization of the Cross-type material law (2.19), where the scalar Newtonian viscosities are replaced by the viscosity tensors $\bar{\mathbb{V}}_0(\underline{a})$ and $\bar{\mathbb{V}}_\infty(\underline{a})$, and the function \mathbb{T}_8 is introduced to model the non-linear transition between the two Newtonian viscosity tensors $\bar{\mathbb{V}}_0(\underline{a})$ and $\bar{\mathbb{V}}_\infty(\underline{a})$. The function \mathbb{T}_8 varies between the individual models, and encodes the direction dependent non-linearity of the models as stated in requirement 2. In summary, three objects in the Ansatz (4.1) need to be modeled: the Newtonian viscosity tensors $\bar{\mathbb{V}}_0(\underline{a})$ and $\bar{\mathbb{V}}_\infty(\underline{a})$, as well as the eighth order tensor function \mathbb{T}_8 controlling the anisotropic and non-linear shear rate dependence. In the following sections, we discuss the respective modeling approaches for these three objects.

4.2 Modeling the Newtonian limits of the effective suspension viscosity

To integrate orthotropic symmetry and incompressibility according to requirement 3 into the models, we first consider

the vector space V_0 of fourth order incompressible tensors with minor and major symmetries

$$V_0 = \{\mathbb{X} \mid \mathbb{X} \in (\mathbb{R}^3)^{\otimes 4}, \text{ and } \mathbb{X}[\mathbf{I}] = 0 \text{ and } \mathbb{X}^{\text{T}_H} = \mathbb{X}^{\text{T}_L} = \mathbb{X}^{\text{T}_R} = \mathbb{X}\}, \tag{4.2}$$

where $(\cdot)^{\text{T}_H}$, $(\cdot)^{\text{T}_L}$, and $(\cdot)^{\text{T}_R}$ stand for major transposition, left transposition, and right transposition, respectively. Second, we then define a basis \mathcal{B} of the space V_0

$$\underline{\underline{\mathcal{B}}} = \begin{bmatrix} 0 & \frac{-2}{\sqrt{6}} & 0 & 0 & 0 \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.3}$$

such that the columns of the matrix $\underline{\underline{\mathcal{B}}}$ represent the basis vectors of \mathcal{B} in Mandel notation. Finally, using the basis \mathcal{B} , we define a function

$$\mathbb{A} : \mathbb{R}^6 \rightarrow V_0^+(3), \tag{4.4}$$

$$\underline{a} \mapsto \mathcal{P}_{\mathcal{B}}^+ \mathbb{M}(\underline{a}) = \mathcal{P}_{\mathcal{B}}^+ \begin{bmatrix} a_1 & a_2 & 0 & 0 & 0 \\ a_2 & a_3 & 0 & 0 & 0 \\ 0 & 0 & a_4 & 0 & 0 \\ 0 & 0 & 0 & a_5 & 0 \\ 0 & 0 & 0 & 0 & a_6 \end{bmatrix}_{\mathcal{B}}, \tag{4.5}$$

which constructs an orthotropic fourth order tensor using the components a_i of the model parameter vector \underline{a} , and the non-unique projector $\mathcal{P}_{\mathcal{B}}^+$ into the space V_0^+ of fourth order tensors that are positive definite on the space spanned by the basis \mathcal{B} . Employing eigendecomposition, we define the action of the projector $\mathcal{P}_{\mathcal{B}}^+$ as

$$\mathcal{P}_{\mathcal{B}}^+ [\mathbb{X}] = \mathcal{P}_{\mathcal{B}}^+ \left[\sum_{i=0}^5 \lambda_i \mathbf{p}_i \otimes \mathbf{p}_i \right] = \sum_{i=0}^5 \max(\lambda_i, \beta) \mathbf{p}_i \otimes \mathbf{p}_i, \tag{4.6}$$

where λ_i and \mathbf{p}_i denote eigenvalue eigentensor pairs of a fourth order tensor \mathbb{X} on the space spanned by the basis \mathcal{B} , and $\beta \in \mathbb{R}_{>0}$ is a small positive and real constant. The definition of the projection operator $\mathcal{P}_{\mathcal{B}}^+$ involves a tunable constant β , which is required for numerical purposes. In fact, formally setting β to zero leads to the projector onto positive semi-definite tensors, i.e., those which may be degenerate. However, positive definiteness on the incompressible subspace \mathcal{B} is preferred over positive semi-definiteness, because of the following physical and numerical reasons. We wish to use the tensor function \mathbb{A} to build fourth order tensors $\mathbb{X} \in V_0^+$ from six model parameters, i.e., the viscosity tensors $\bar{V}_0(\underline{a})$

and $\bar{V}_{\infty}(\underline{a})$. Therefore, encoding a vanishing stress response through vanishing eigenvalues would not adhere to the physical model considered in the context of this article.

4.3 Modeling the anisotropic and non-linear shear rate dependence of the effective suspension viscosity

To capture the anisotropic and non-linear viscous behavior in accordance with requirement 2, we incorporate a generalized distance g from zero load

$$g : \text{Sym}_0(3) \times \mathbb{R}^7 \rightarrow \mathbb{R}, \tag{4.7}$$

$$(\mathbf{D}, \underline{a}) \mapsto (\mathbf{D} \cdot \mathbb{A}(\underline{a})[\mathbf{D}])^{a_7}, \tag{4.8}$$

into the models, which is similar to the Mahalanobis distance [65] popular in statistics. The generalized distance g depends on the components a_i collected in the model parameter vector \underline{a} , as well as the magnitude and direction of the shear rate tensor \mathbf{D} . Because of these properties, we use the generalized distance g as a model building block to encode the fiber orientation state specific, anisotropic shear rate dependence of the effective suspension viscosity. However, for an exponent $a_7 < 1$ the gradient of this generalized distance g is singular whenever $\mathbf{D} \cdot \mathbb{A}(\underline{a})[\mathbf{D}]$ vanishes, posing a problem during gradient based learning of the model parameters \underline{a} . This singularity is circumvented by the enforced positive definiteness of the tensors constructed by the function \mathbb{A} and non-vanishing load \mathbf{D} . Here, the constant β acts as a lower bound on the eigenvalues of the tensor $\mathbb{A}(\underline{a})$, as can be seen from equations (4.4) and (4.6), and restricts the space of possible tensors $\mathbb{A}(\underline{a})$ that can be constructed from the model parameters \underline{a} . Improper choice of the constant β , such that β is greater than the smallest eigenvalue of the optimal tensor $\mathbb{A}(\underline{a})$, could therefore influence the quality of fit. However, because we wish to learn the model parameters \underline{a} from homogenization data, the optimal tensor $\mathbb{A}(\underline{a})$ and its smallest eigenvalue are not known a priori. Therefore, the constant β should be chosen to be a small number, and we nominally select $\beta = 10^{-6}$ in the unit of the associated eigenvalue λ_i . To model the anisotropic Cross-type non-linearity using the generalized distance g , we define two non-linear eighth order tensor functions $\mathbb{T}_8^{(1)}$ and $\mathbb{T}_8^{(2)}$ through their actions on an orthotropic tensor $\mathbb{X} \in \text{Sym}_0$. We define the function $\mathbb{T}_8^{(1)}$ through

$$\mathbb{T}_8^{(1)}(\mathbf{D}, \underline{a}) [\mathbb{X}] = \frac{\mathbb{X}}{1 + g(\mathbf{D}, \underline{a})}, \quad \underline{a} \in \mathbb{R}^7, \tag{4.9}$$

such that the scalar shear rate dependence of the Cross-type model (2.19) is replaced by the generalized distance g (4.7). Thus, the function $\mathbb{T}_8^{(1)}$ scales all the components of the orthotropic tensor \mathbb{X} equally, depending on the load \mathbf{D} and

the model parameter vector \underline{a} . Because all components of the tensor \mathbb{X} are scaled equally, the anisotropy of the tensor \mathbb{X} remains unchanged under the function $\mathbb{T}_8^{(1)}$. However, the non-linear shear rate dependence encoded in the function $\mathbb{T}_8^{(1)}$ may be anisotropic since the magnitude of the scaling depends on the function g , and hence on a possibly anisotropic tensor $\mathbb{A}(\underline{a})$ and the load \mathbf{D} . Of the seven model parameters collected in the model parameter vector \underline{a} , six model parameters control the influence of the direction of the load \mathbf{D} on the scaling, and one parameter controls the rate of transition between the Newtonian limits for shear rates $\dot{\gamma} \rightarrow 0$ and $\dot{\gamma} \rightarrow \infty$.

To allow for greater flexibility in the scaling of the orthotropic tensor \mathbb{X} , we also define the tensor function $\mathbb{T}_8^{(2)}$ through

$$\mathbb{T}_8^{(2)}(\mathbf{D}, \underline{a})[\mathbb{X}] = \begin{bmatrix} h_1(\mathbf{D}, \underline{a})X_{11} & h_2(\mathbf{D}, \underline{a})X_{12} & 0 & 0 & 0 & 0 \\ h_2(\mathbf{D}, \underline{a})X_{12} & h_3(\mathbf{D}, \underline{a})X_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & h_4(\mathbf{D}, \underline{a})X_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & h_5(\mathbf{D}, \underline{a})X_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & h_6(\mathbf{D}, \underline{a})X_{55} & 0 \end{bmatrix}_{\mathcal{B}}, \quad \underline{a} \in \mathbb{R}^{42}, \quad (4.10)$$

such that the six components of the tensor \mathbb{X} are scaled individually by a scalar $h_i(\mathbf{D}, \underline{a})$, $i \in \{1, 2, 3, 4, 5, 6\}$, defined as

$$h_i(\mathbf{D}, \underline{a}) = \frac{1}{1 + g(\mathbf{D}, \underline{f}_i)}, \quad \underline{f}_i = (a_{7(i-1)+1}, \dots, a_{7i}) \in \mathbb{R}^7. \quad (4.11)$$

Like the scaling factor in the function $\mathbb{T}_8^{(1)}$, the scalars $h_i(\mathbf{D}, \underline{a})$ are each computed from seven model parameters, the shear rate tensor \mathbf{D} , and the function g . Therefore, the tensor \mathbb{X} is not only scaled under the function $\mathbb{T}_8^{(2)}$, but the anisotropy of the tensor \mathbb{X} may also change under the function $\mathbb{T}_8^{(2)}$. In this sense, the function $\mathbb{T}_8^{(1)}$ is a special case of the function $\mathbb{T}_8^{(2)}$, where all scalars $h_i(\mathbf{D}, \underline{a})$ are equal. The increased flexibility in the scaling of the tensor \mathbb{X} with function $\mathbb{T}_8^{(2)}$ requires 36 parameters to control the influence of the direction of the load \mathbf{D} , and six parameters to control the rate of transition between the Newtonian limits for shear rates $\dot{\gamma} \rightarrow 0$ and $\dot{\gamma} \rightarrow \infty$. This results in a total of 42 parameters for the function $\mathbb{T}_8^{(2)}$.

4.4 Definitions of effective suspension viscosity models

In the following, we combine the definitions of the previous sections 4.2 and 4.3 with the Ansatz (4.1) to build four models for the effective suspension viscosity. With the symbol $\#$

denoting concatenation of vectors the four models presented in Table 3 are considered and described in the following.

Model 1

In Model 1, we use the function \mathbb{A} (4.4) to construct the Newtonian viscosity tensors $\bar{\mathbb{V}}_0(\underline{b})$ and $\bar{\mathbb{V}}_\infty(\underline{c})$ from twelve model parameters collected in the vectors $\underline{b} \in \mathbb{R}^6$ and $\underline{c} \in \mathbb{R}^6$. Also, we use the function $\mathbb{T}_8^{(1)}$ (4.9) with seven parameters collected in the vector $\underline{d} \in \mathbb{R}^7$ to model the anisotropic and non-linear shear rate dependence of the effective suspension viscosity. Overall, Model 1 has 19 parameters and is the most general considered model that uses the function $\mathbb{T}_8^{(1)}$.

Model 2

Like in Model 1, we use the function \mathbb{A} to construct the Newtonian viscosity tensor $\bar{\mathbb{V}}_\infty(\underline{b})$ in the shear rate limit $\dot{\gamma} \rightarrow \infty \text{ s}^{-1}$ from six model parameters, and the non-linear function $\mathbb{T}_8^{(1)}$ to model the non-linear transition between the Newtonian limits with seven model parameters. These 13 parameters are collected in the vectors $\underline{c} \in \mathbb{R}^6$ and $\underline{d} \in \mathbb{R}^7$. However, in Model 1, the viscosity tensor $\bar{\mathbb{V}}_0(\underline{b})$ in the shear rate limit $\dot{\gamma} \rightarrow 0 \text{ s}^{-1}$ is also constructed with the function \mathbb{A} (4.4), using six model parameters. In Model 2, we exploit the fact that the viscosity tensors $\bar{\mathbb{V}}_0(\underline{b})$ and $\bar{\mathbb{V}}_\infty(\underline{c})$ share the same anisotropy in the shear rate limits $\dot{\gamma} \rightarrow 0 \text{ s}^{-1}$ and $\dot{\gamma} \rightarrow \infty \text{ s}^{-1}$, see Fig. 4 and Sterr et al. [18], and express the viscosity tensor $\bar{\mathbb{V}}_0(\underline{b})$ as a scalar multiple of the tensor $\bar{\mathbb{V}}_\infty(\underline{c})$, such that the condition

$$\bar{\mathbb{V}}_0(\underline{b}) = b_1 \bar{\mathbb{V}}_\infty(\underline{c}), \quad (4.12)$$

is satisfied. This relation reduces the number of parameters from 19 in Model 1 to 14 in Model 2. We found that this adaptation introduces only small errors, as we discuss in section 4.6 in more detail, and use equation (4.12) in Model 3 and Model 4 as well.

Model 3

In Model 3, like in Model 1 and Model 2, we use the function $\mathbb{T}_8^{(1)}$ to model the non-linear transition between the Newtonian limits with seven model parameters collected

Table 3 Number of parameters m , choices for the functions $\bar{\mathbb{V}}_0$, $\bar{\mathbb{V}}_\infty$, and \mathbb{T}_8 , as well as the complete expression for the effective viscosity tensor $\bar{\mathbb{V}}(\mathbf{D}, \underline{a})$ for the Models 1, 2, 3, and 4

Model 1		Model 2	
parameters: $m = 19$,	$\underline{b}, \underline{c} \in \mathbb{R}^6, \underline{d} \in \mathbb{R}^7$	parameters: $m = 14$,	$\underline{b} \in \mathbb{R}^1, \underline{c} \in \mathbb{R}^6, \underline{d} \in \mathbb{R}^7$
\underline{a}	$= \underline{b} + \underline{c} + \underline{d}$	\underline{a}	$= \underline{b} + \underline{c} + \underline{d}$
$\bar{\mathbb{V}}_0(\underline{b})$	$= \mathbb{A}(\underline{b})$	$\bar{\mathbb{V}}_0(\underline{b})$	$= b_1 \bar{\mathbb{V}}_\infty(\underline{c})$
$\bar{\mathbb{V}}_\infty(\underline{c})$	$= \mathbb{A}(\underline{c})$	$\bar{\mathbb{V}}_\infty(\underline{c})$	$= \mathbb{A}(\underline{c})$
$\mathbb{T}_8(\mathbf{D}, \underline{d})$	$= \mathbb{T}_8^{(1)}(\mathbf{D}, \underline{d})$	$\mathbb{T}_8(\mathbf{D}, \underline{d})$	$= \mathbb{T}_8^{(1)}(\mathbf{D}, \underline{d})$
$\bar{\mathbb{V}}(\mathbf{D}, \underline{a}) = \mathbb{A}(\underline{c}) + \frac{\mathbb{A}(\underline{b}) - \mathbb{A}(\underline{c})}{1 + (\mathbf{D} \cdot \mathbb{A}(\underline{d})[\mathbf{D}])^{d_7}}$		$\bar{\mathbb{V}}(\mathbf{D}, \underline{a}) = \mathbb{A}(\underline{c}) \left(1 + \frac{b_1 - 1}{1 + (\mathbf{D} \cdot \mathbb{A}(\underline{d})[\mathbf{D}])^{d_7}} \right)$	
Model 3		Model 4	
parameters: $m = 11$,	$\underline{b} \in \mathbb{R}^1, \underline{c} \in \mathbb{R}^3, \underline{d} \in \mathbb{R}^7$	parameters: $m = 49$,	$\underline{b} \in \mathbb{R}^1, \underline{c} \in \mathbb{R}^6, \underline{d} \in \mathbb{R}^{42}$
\underline{a}	$= \underline{b} + \underline{c} + \underline{d}$	\underline{a}	$= \underline{b} + \underline{c} + \underline{d}$
$\bar{\mathbb{V}}_0(\underline{b})$	$= b_1 \bar{\mathbb{V}}_\infty(\underline{c})$	$\bar{\mathbb{V}}_0(\underline{b})$	$= b_1 \bar{\mathbb{V}}_\infty(\underline{c})$
$\bar{\mathbb{V}}_\infty(\underline{c})$	$= \mathbb{P}_2(c_1 \mathbb{P}_2 + c_2 \mathbf{I} \square_s \mathbf{N} + c_3 \mathbb{N}) \mathbb{P}_2$	$\bar{\mathbb{V}}_\infty(\underline{c})$	$= \mathbb{A}(\underline{c})$
$\mathbb{T}_8(\mathbf{D}, \underline{d})$	$= \mathbb{T}_8^{(1)}(\mathbf{D}, \underline{d})$	$\mathbb{T}_8(\mathbf{D}, \underline{d})$	$= \mathbb{T}_8^{(2)}(\mathbf{D}, \underline{d})$
$\bar{\mathbb{V}}(\mathbf{D}, \underline{a}) = \bar{\mathbb{V}}_\infty(\underline{c}) \left(1 + \frac{b_1 - 1}{1 + (\mathbf{D} \cdot \mathbb{A}(\underline{d})[\mathbf{D}])^{d_7}} \right)$		$\bar{\mathbb{V}}(\mathbf{D}, \underline{a}) = \mathbb{A}(\underline{c}) + \mathbb{T}_8^{(2)}(\mathbf{D}, \underline{d}) [(b_1 - 1)\mathbb{A}(\underline{c})]$	

in the vector $\underline{d} \in \mathbb{R}^7$. Also, as in Model 2, we use equation (4.12) to relate the viscosity tensors $\bar{\mathbb{V}}_0(\underline{b})$ and $\bar{\mathbb{V}}_\infty(\underline{c})$ via a scalar coefficient collected in the vector $\underline{b} \in \mathbb{R}^1$. To further reduce the number of model parameters through physical considerations motivated by superposition and orientation averaging [32, 33, 44], we introduce the equation

$$\bar{\mathbb{V}}_\infty(\underline{c}) = \mathbb{P}_2(c_1 \mathbb{P}_2 + c_2 \mathbf{I} \square_s \mathbf{N} + c_3 \mathbb{N}) \mathbb{P}_2, \quad (4.13)$$

for the viscosity tensor $\bar{\mathbb{V}}_\infty(\underline{c})$, where \mathbb{P}_2 denotes the identity on the space $\text{Sym}_0(3)$. The operator \square_s stands for the symmetrized box product, which, for the second order tensors $\mathbf{A}, \mathbf{B}, \mathbf{C} \in (\mathbb{R}^3)^{\otimes 2}$, is defined as

$$\mathbf{A} \square_s \mathbf{B} = \frac{1}{2}(\mathbf{A} \square \mathbf{B} + \mathbf{B} \square \mathbf{A}), \quad \text{where } (\mathbf{A} \square \mathbf{B})[\mathbf{C}] = \mathbf{A} \mathbf{C} \mathbf{B}. \quad (4.14)$$

Overall, this reduces the number of parameters from 14 in Model 2 to 11 in Model 3.

Model 4

In Model 4, in contrast to the Models 1, 2, and 3, we use the function $\mathbb{T}_8^{(2)}$ instead of the function $\mathbb{T}_8^{(1)}$ to model the non-linear transition between the Newtonian limits. As discussed in the previous section 4.3, this allows for greater flexibility in exchange for more model parameters, making Model 4 the most general of the considered models. Like in Model 2 and Model 3, we use seven parameters collected in the vectors $\underline{b} \in \mathbb{R}^1$ and $\underline{c} \in \mathbb{R}^6$ to model the viscosity tensors $\bar{\mathbb{V}}_0(\underline{b})$ and $\bar{\mathbb{V}}_\infty(\underline{c})$. In combination with the 42 input parameters of the function $\mathbb{T}_8^{(2)}$, which we collect in the vector \underline{d} , this leads to a total of 49 model parameters for Model 4. Overall, each model parameter controls a distinct effect or quantity of the model. In other words, Model 4 could not

be expressed with fewer parameters while maintaining the same type of non-linearity and the same degree of modeling flexibility.

4.5 Supervised learning of model parameters

In the previous sections 4.1 and 4.4, we described the requirements for models of the effective suspension viscosity and presented four model candidates satisfying these requirements. In this section, we discuss the supervised machine learning strategy we use to identify the model parameters. To learn the model parameters from the data obtained with FFT-based computational homogenization, we first construct a loss function \mathcal{L} measuring the difference between the model predictions and the FFT-based computational results, and then minimize the loss \mathcal{L} using the ADAM [66] optimization algorithm. We consider all fiber orientation states $\underline{\lambda} \in S_T^{109}$ in the set $S_T^{109} \subset S_T$ of 109 points of the discretized fiber orientation triangle shown in Figs. 1a and 4b. For every load case $\bar{\mathbf{D}}$, fiber orientation state $\underline{\lambda}$, and parameter vector \underline{a} , we define the loss \mathcal{L}

$$\mathcal{L}(\bar{\mathbf{D}}, \underline{\lambda}, \underline{a}) = \frac{\|\bar{\boldsymbol{\tau}}^{\text{FFT}}(\bar{\mathbf{D}}, \underline{\lambda}) - \bar{\boldsymbol{\tau}}^{\text{Model}}(\bar{\mathbf{D}}, \underline{\lambda}, \underline{a})\|_2}{\|\bar{\boldsymbol{\tau}}^{\text{FFT}}(\bar{\mathbf{D}}, \underline{\lambda})\|_2}, \quad (4.15)$$

where $\|\cdot\|_2$ stands for the Euclidean norm, $\bar{\boldsymbol{\tau}}^{\text{model}} : \text{Sym}_0(3) \times \mathbb{R}^2 \times \mathbb{R}^m \rightarrow \text{Sym}_0(3)$ refers to the effective stress predicted by a particular model, and $\bar{\boldsymbol{\tau}}^{\text{FFT}} : \text{Sym}_0(3) \times \mathbb{R}^2 \rightarrow \text{Sym}_0(3)$ refers to the effective stress obtained with FFT-based computational homogenization. Similar to section 3, we omit the dependence of the loss \mathcal{L} and the derived quantities on various variables for improved readability and compactness. For each model and fiber orientation state $\underline{\lambda}$, we wish to identify the model parameters \underline{a} minimizing the worst case loss \mathcal{L} for the set of investigated load cases $\bar{\mathbf{D}}_{\dot{\gamma}}$, such that

$$\underline{a}(\underline{\lambda}) = \underset{\underline{a} \in \mathbb{R}^m}{\text{argmin}} \max_{\bar{\mathbf{D}} \in \bar{\mathbf{D}}_{\dot{\gamma}}} \mathcal{L}(\bar{\mathbf{D}}, \underline{\lambda}, \underline{a}). \quad (4.16)$$

Here, the set of investigated load cases $\bar{\mathbf{D}}_{\dot{\gamma}}$ is defined via equations (2.20) and (2.21). However, since the models described in the previous section are not convex in the model parameters \underline{a} and have multiple local minima, the exact minimizer \underline{a} is difficult to determine. We wish to start our investigations with a word of warning: identifying the model parameters from computational multiscale simulations does not qualify as a mathematically well-posed problem [67], for a number of reasons. For a start, there might not be a unique solution, and the outcome of the learning process, i.e., the model parameters, might change drastically for small variations in the data. Also, depending on the optimization algorithm and the optimization hyperparameters, the learned

model parameters can vary as well. Accordingly, the model parameters obtained with the supervised learning procedure we present in this article are possibly not the global minimizers of the optimization problem (4.16). However, as we show in the next section 4.6, the presented supervised learning procedure can be employed to successfully identify model parameters which result in model predictions with engineering accuracy. To capitalize on the advances in the field of non-convex optimization and machine learning, we use the Python programming language and the machine learning framework PyTorch [68] with version 1.12 to identify the model parameters \underline{a} . When this article was written, more recent PyTorch versions were available that feature Just In Time (JIT) compilation, which could possibly increase the performance of custom code modules. However, for compatibility with existing code, we use version 1.12. For each model and fiber orientation state $\underline{\lambda}$, we randomly initialized 5000 realizations of the parameter vector \underline{a} and then used the field-tested ADAM [66] algorithm to improve the initial guesses, thus generating the set S_{min} of minimizing parameter vectors \underline{a} . We selected the ADAM algorithm from the optimization algorithms available in the PyTorch framework, since the ADAM algorithm combines an adaptive learning rate with a classical momentum [69] based approach. This provides advantages over the standard gradient descent method [70, 71], and algorithms relying on adaptive learning rates alone, such as AdaGrad [72], AdaDelta [73] and RMSprop [74]. For a detailed discussion on the convergence properties of the ADAM algorithm and its variants, we refer the reader to an article by Chen et al. [75]. Also, for an overview on a variety of other optimization algorithms in machine learning, we refer the reader to a review by Sun et al. [76]. Conveniently, the ADAM algorithm’s learning rate hyperparameters are approximate bounds of the optimization step size. Exploiting this feature, we use PyTorch’s option to define parameter groups with individually varying learning options for parameters with different magnitudes, such that the step sizes for nominally large parameters are not bounded by the step sizes for nominally small ones. As a viable alternative to first-order optimization techniques, PyTorch also offers the powerful second-order optimization algorithm L-BFGS [77]. However, when this article was written, the PyTorch L-BFGS algorithm did not support multiple learning rates and parameter groups, which we have found to be valuable in accelerating convergence. Therefore, we opted to use the first-order optimization algorithm ADAM. Overall, we use four parameter groups, with one group each for the parameters of the functions $\bar{\mathbb{V}}_0$ and $\bar{\mathbb{V}}_\infty$, as well as one group each for the parameters associated with the anisotropy and the rate of non-linear transition encoded in the tensor function \mathbb{T}_8 , see equations (4.9), (4.10) and (4.11). We distinguish the two parameter groups associated with the tensor function \mathbb{T}_8 with the symbols G_8^{aniso} , which contains six

Table 4 Learning rate hyperparameter by parameter group for use with the ADAM algorithm

	\tilde{V}_0	\tilde{V}_∞	G_8^{aniso}	G_8^{expo}
Learning rate	$5 \cdot 10^{-3}$	$5 \cdot 10^{-1}$	$5 \cdot 10^{-7}$	$5 \cdot 10^{-2}$

parameters in case of the function $\mathbb{T}_8^{(1)}$ and 36 parameters in the case of the function $\mathbb{T}_8^{(2)}$, as well as the group G_8^{expo} containing one and six parameters for the functions $\mathbb{T}_8^{(1)}$ and $\mathbb{T}_8^{(2)}$, respectively. Except for Model 1, the learning rates for the respective parameter groups were chosen as shown in Table 4. For Model 1 in particular, the learning rate for the parameters associated with the function \tilde{V}_0 was chosen as 0.5, because Model 1 does not use the scalar relationship (4.12). We determined the numerical values of the learning rates for each parameter group by trial and error. To improve convergence towards minima as the learning process proceeds [78, 79], we used PyTorch’s ReduceLROnPlateau learning rate scheduler to half the learning rates if the loss \mathcal{L} has not improved for 200 epochs. This rule was applied with a cooldown of 400 epochs and all other options were kept standard. For the ADAM algorithm, other than the learning rates, standard hyperparameters were used such that the two momentum coefficients β_1 and β_2 are set to $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and the stabilization constant ε is set to $\varepsilon = 10^{-8}$. Finally, we consider the learning process to be finished when the loss \mathcal{L} has not improved for 2000 epochs.

4.6 Model accuracy

To compare the accuracy of the presented models, we define error measures and describe a validation procedure in the following. Since finding a global minimum for the problem (4.16) is hard, we restrict to the set of minimizing parameters S_{\min} , and identify the parameters $\underline{a}^{\min} \in S_{\min}$ minimizing the worst case loss \mathcal{L} over the set of investigated load cases $\tilde{D}_{\dot{\gamma}}$, such that

$$\underline{a}^{\min}(\underline{\lambda}) = \operatorname{argmin}_{\underline{a} \in S_{\min}} \max_{\tilde{D} \in \tilde{D}_{\dot{\gamma}}} \mathcal{L}(\tilde{D}, \underline{\lambda}, \underline{a}). \tag{4.17}$$

The minimizing parameters \underline{a}^{\min} are only available on the given discretization points $\underline{\lambda}$ where homogenization data is available as well. Yet, in accordance with requirement 4, we are interested in generalizing a given model for fiber orientation states $\underline{\lambda}$ where no minimizing parameters \underline{a}^{\min} are available. For this purpose, we follow Köbler et al. [48] and use a convex linear combination to interpolate stresses as follows. Suppose the nodes $\underline{\lambda}_1, \underline{\lambda}_2,$ and $\underline{\lambda}_3$ form a triangle and the minimizing parameters \underline{a}^{\min} are known in those nodes.

Then we compute the effective stresses $\bar{\tau}^C$ at some point $\underline{\lambda}$

$$\bar{\tau}^C(\tilde{D}, \underline{\lambda}) = \sum_{i=1}^3 s_i \bar{\tau}^{\text{model}}(\tilde{D}, \underline{\lambda}_i, \underline{a}^{\min}(\underline{\lambda}_i)), \tag{4.18}$$

$$\underline{\lambda} = s_1 \underline{\lambda}_1 + s_2 \underline{\lambda}_2 + s_3 \underline{\lambda}_3, \quad s_i \geq 0, \quad i \in 1, 2, 3,$$

$$\sum_{i=1}^3 s_i = 1, \tag{4.19}$$

through interpolation of the model stresses $\bar{\tau}^{\text{model}}$ at the points $\underline{\lambda}_1, \underline{\lambda}_2,$ and $\underline{\lambda}_3$. In addition to generalizing the models over the entire fiber orientation triangle S_T , we use equation (4.18) to validate the models against the data obtained with FFT-based computational homogenization. For this purpose, we define the validation error e_V similarly to the loss \mathcal{L} , i.e., we define

$$e_V(\tilde{D}, \underline{\lambda}) = \frac{\|\bar{\tau}^{\text{FFT}}(\tilde{D}, \underline{\lambda}) - \bar{\tau}^C(\tilde{D}, \underline{\lambda})\|_2}{\|\bar{\tau}^{\text{FFT}}(\tilde{D}, \underline{\lambda})\|_2}, \tag{4.20}$$

for each considered model. Of the 109 triangulation points $\underline{\lambda}$ contained in the set S_T^{109} we designate 45 points as the set $S_T^F \subset S_T^{109}$ of fitting points $\underline{\lambda}^F$, see Fig. 5. For validation purposes, we designate the centroids of the triangles formed by the fitting points $\underline{\lambda}^F$ as validation points, defining the set $S_T^V \subset S_T^{109}$ of 64 validation points $\underline{\lambda}^V$. By definition of the interpolation (4.18), the loss \mathcal{L} coincides with the validation error e_V at the fitting points $\underline{\lambda}^F$. To study the quality of fit and the quality of the stress interpolation, we define the largest loss \mathcal{L}^{\max} and the largest validation error e_V^{\max}

$$\begin{aligned} \mathcal{L}^{\max} &= \max_{\underline{\lambda} \in S_T^{109}} \max_{\tilde{D} \in \tilde{D}_{\dot{\gamma}}} \mathcal{L}(\tilde{D}, \underline{\lambda}, \underline{a}^{\min}(\underline{\lambda})), \\ e_V^{\max} &= \max_{\underline{\lambda} \in S_T^{109}} \max_{\tilde{D} \in \tilde{D}_{\dot{\gamma}}} e_V(\tilde{D}, \underline{\lambda}). \end{aligned} \tag{4.21}$$

Furthermore, we define the mean validation error e_V^{mean} and the maximum validation error $e_V^{\tilde{D}}$ over the set of investigated load cases $\tilde{D}_{\dot{\gamma}}$

$$\begin{aligned} e_V^{\text{mean}} &= \frac{1}{109} \sum_{\underline{\lambda} \in S_T^{109}} \max_{\tilde{D} \in \tilde{D}_{\dot{\gamma}}} e_V(\tilde{D}, \underline{\lambda}), \\ e_V^{\tilde{D}}(\underline{\lambda}) &= \max_{\tilde{D} \in \tilde{D}_{\dot{\gamma}}} e_V(\tilde{D}, \underline{\lambda}), \end{aligned} \tag{4.22}$$

that occurred for a given model on the discretized fiber orientation triangle. Before discussing the prediction accuracy of the presented models in detail, we briefly summarize the supervised learning procedure presented in the previous section 4.5, and the validation approach for the stress interpolation (4.18) discussed above. For each model, and in each of the 109 investigated triangulation points $\underline{\lambda} \in S_T^{109}$,

Table 5 Largest loss \mathcal{L}^{\max} , as well as validation errors e_V^{\max} , and e_V^{mean} for the Models 1 to 4

	Model 1	Model 2	Model 3	Model 4
\mathcal{L}^{\max} in %	5.00	5.00	15.36	5.00
e_V^{\max} in %	5.05	5.15	15.36	5.00
e_V^{mean} in %	2.65	2.73	7.76	2.28

we conducted 5000 optimization runs for the minimization problem (4.16) and thus obtained the set S_{\min} of minimizing parameter vectors \underline{a} . From this set S_{\min} , we identified the single best parameter vector $\underline{a}^{\min} \in S_{\min}$ (4.17), which is associated with the smallest worst case loss of a model in a fixed triangulation point $\underline{\lambda}$. To study whether stress interpolation (4.18) can be used to generalize a model beyond points where the parameters \underline{a}^{\min} are available, we designated the 45 points $\underline{\lambda}^F \in S_{\Gamma}^F$ as fitting points and the 64 points $\underline{\lambda}^V \in S_{\Gamma}^V$ as validation points. On these points $\underline{\lambda}^F$ and $\underline{\lambda}^V$, we compute the validation error e_V (4.20) using the stress interpolation procedure defined through the equations (4.18) and (4.19), such that only the parameter vectors \underline{a}^{\min} in the fitting points are used. Finally, to investigate whether a model captures the underlying material behavior appropriately, how stress interpolation affects the prediction quality, and how the prediction quality varies over the fiber orientation triangle S_{Γ} , we are interested in the error measures \mathcal{L}^{\max} , e_V^{\max} , e_V^{mean} , and $e_V^{\bar{D}}$.

Thus, we list the largest Loss \mathcal{L}^{\max} , as well as the validation errors e_V^{\max} , and e_V^{mean} in Table 5, visualize the validation error $e_V^{\bar{D}}$ over the fiber orientation triangle S_{Γ} in Fig. 5, and discuss the errors in the following.

Model 1 and Model 2 differ only in their modeling approach for the viscosity tensor $\bar{\mathbb{V}}_0(\underline{b})$, where the parameter based construction of an orthogonal tensor $\bar{\mathbb{V}}_0(\underline{b}) = \mathbb{A}(\underline{b})$ in Model 1 is replaced by the scalar relationship (4.12) in Model 2. This reduction in dimensionality does not increase the loss \mathcal{L} , as the largest loss \mathcal{L}^{\max} for both Model 1 and Model 2 is 5.00%. Also, the largest validation error $e_V^{\max} = 5.15\%$ and mean validation error $e_V^{\text{mean}} = 2.73\%$ of Model 2 are only 0.10% and 0.08% larger than the errors $e_V^{\max} = 5.05\%$ and $e_V^{\text{mean}} = 2.65\%$ of Model 1. Hence, the scalar relationship defined by equation (4.12) appears to be a valid assumption in the context of the investigated physics and microstructures. The largest loss \mathcal{L}^{\max} occurred for the orientation state $\underline{\lambda} = (0.83, 0.08)^T$ for both Model 1 and Model 2. In contrast, the largest validation errors e_V^{\max} occurred at orientation states $\underline{\lambda} = (0.90, 0.07)^T$ for Model 1 and $\underline{\lambda} = (0.59, 0.38)^T$ for Model 2. For the chosen discretization of the fiber orientation triangle S_{Γ} , a small additional error is introduced by stress interpolation, as shown by the slight differences between the largest losses \mathcal{L}^{\max} and the largest validation errors e_V^{\max} for Model 1 and Model 2. Among the investigated models, Model 3 shows the largest loss and the largest valida-

tion error with $\mathcal{L}^{\max} = 15.36\%$ and $e_V^{\max} = 15.36\%$, occurring at the orientation state $\underline{\lambda} = (0.69, 0.31)^T$. For orientation states towards the lower and left edges of the fiber orientation triangle S_{Γ} , the magnitude of the validation error $e_V^{\bar{D}}$ for Model 3 is comparable to that of the other models, see Fig. 5. However, a mean validation error $e_V^{\text{mean}} = 7.76\%$ in combination with the relatively large validation errors $e_V^{\bar{D}}$ occurring at orientation states towards the upper edge of the fiber orientation triangle S_{Γ} indicate that equation (4.13) is not a sufficiently accurate approximation over the whole fiber orientation triangle S_{Γ} . For the most general model, Model 4, the largest loss \mathcal{L}^{\max} and the largest validation error e_V^{\max} are both 5.00% for the orientation state $\underline{\lambda} = (0.83, 0.08)^T$. Hence, the increased anisotropic capability of Model 4 improves the quality of fit and prediction only slightly compared to Models 1 and 2, albeit Model 4 uses more parameters. The mean validation error $e_V^{\text{mean}} = 2.28\%$ of Model 4 is also only 0.37% and 0.45% lower than for Models 1 and 2. In summary, Model 3 performed the worst among the investigated models and Model 4 performed the best. The orientation averaging incorporated in Model 3 reduced the prediction accuracy, while the anisotropic function $\mathbb{T}^{(2)}$ improved the prediction accuracy of Model 4. However, Models 1 and 2 use fewer parameters than Model 4 and show rather similar prediction accuracy. Consequently, the degree of anisotropic non-linearity encoded in Model 1 and Model 2 seems to be sufficient to capture the effective viscous behavior to engineering accuracy in the investigated load cases. In terms of practical implementation, computational efficiency and balanced prediction accuracy we consider Model 2 the best of the investigated models, since it uses a moderate amount of parameters and yields accuracy comparable to Model 4.

5 Conclusions

In this work, we used supervised machine learning and FFT-based computational techniques to discover models for the effective suspension viscosity of fiber suspensions with shear-thinning matrix behavior. We first extended the computational investigations of previous work to a broad variety of fiber orientation states. For all considered orientation states, we studied the anisotropy and shear rate dependence of the suspension viscosity over a wide range of shear rates of engineering interest. Confirming previous observations in the case of a transversely isotropic orientation state, we found that the anisotropy of the suspension viscosity for a particular orientation state varies substantially depending on the load direction and shear rate. Furthermore, the degree of non-linearity of the matrix material in the studied volume elements influences the anisotropy of the suspension viscosity strongly. Based on the observed material behavior,

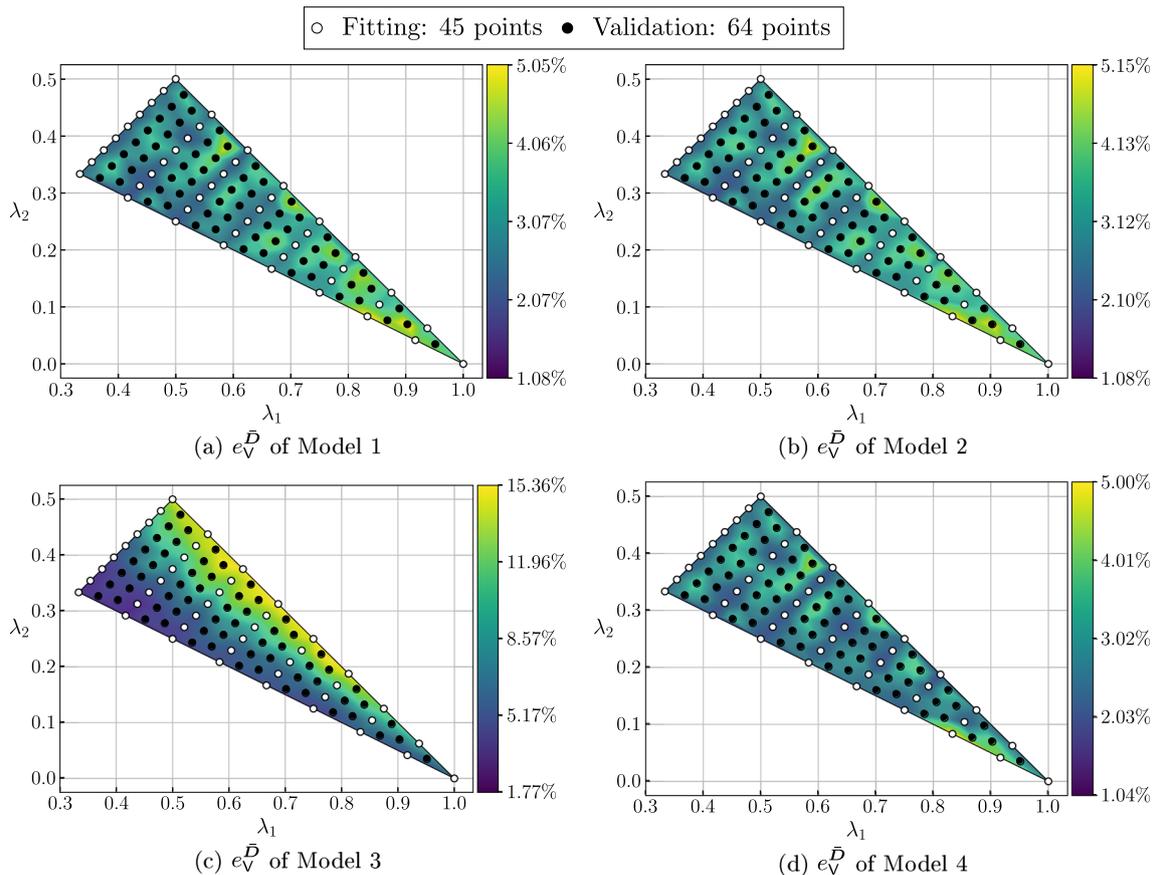


Fig. 5 Validation error e_V^D over the fiber orientation triangle S_T for Models 1 to 4

we introduced four requirements a model of the suspension viscosity should satisfy, and formulated four model candidates according to these requirements. Using supervised machine learning techniques for non-convex optimization, we identified the model parameters based on the high-fidelity FFT-based computational results, and found that three of the four presented models achieve validation errors below 5.15%. One model containing an approximation of the suspension viscosity based on superposition and orientation averaging did not perform favorably when compared with the other presented models.

In future work, the models presented in this article could be employed to enhance the prediction accuracy of engineering process simulations, such as compression and injection molding simulations. In component scale simulations, the presented models could provide substantial reductions in computational cost when compared with multiscale computational approaches, such as FE² [80–82], or combinations of the finite element method with FFT-based methods [83]. Prediction capabilities of the presented approach could be further extended by considering additional physical effects, such as temperature dependence and polymer crystallization in the presented models, and in the FFT-based computa-

tional approach. Also, with the procedure presented in this article, models for microstructures with curved fibers, fiber bundles or fibers with higher aspect ratios could be developed to facilitate the development of engineering systems in the context of long fiber reinforced systems. Furthermore, the supervised learning procedure presented in this article could be modified to potentially identify more suitable model parameters, and thus increase the prediction accuracy of the presented models further. For example, the effect of using different optimization algorithms, such as AdamW [84] and NAdam [85], on the model prediction accuracy could be explored. Also, different approaches to determine the optimization hyperparameters, such as gradient-based techniques, Bayesian optimization, and metaheuristic algorithms could be used [86], and their effect on the model prediction accuracy could be studied as well. Additionally, relationships between the model parameters and the model prediction quality could be investigated using visualization techniques for high-dimensional data, such as t-Distributed Stochastic Neighbor Embedding [87] or uniform manifold approximation and projection [88].

Acknowledgements The research documented in this manuscript was 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. We thank the anonymous reviewers for their helpful comments.

Author Contributions The present study was conceptualized by B. Sterr, M. Schneider, and T. Böhlke. The presented models were derived by B. Sterr, M. Schneider, and T. Böhlke. B. Sterr and M. Schneider implemented and validated the software. B. Sterr performed the simulations, analyzed and visualized the data, designed the machine learning strategy, and drafted the manuscript. The original manuscript draft was extensively reviewed and edited by B. Sterr, A. Hrymak, M. Schneider, and T. Böhlke. Resources were provided by M. Schneider and T. Böhlke. The research project was supervised by A. Hrymak, M. Schneider and T. Böhlke.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kennedy P, Zheng R (2013) Flow analysis of injection molds. Carl Hanser Verlag GmbH Co KG, Munich, Germany
- Henning F, Kärger L, Dörr D, Schirmaier FJ, Seuffert J, Bernath A (2019) Fast processing and continuous simulation of automotive structural composite components. *Compos Sci Technol* 171:261–279
- Qureshi J (2022) A review of fibre reinforced polymer structures. *Fibers* 10(3):27
- Botín-Sanabria DM, Mihaita A-S, Peimbert-García RE, Ramírez-Moreno MA, Ramírez-Mendoza RA, Lozoya-Santos JdJ (2022) Digital twin technology challenges and applications: a comprehensive review. *Remote Sensing* 14(6):1335
- Görthofer J, Meyer N, Pallicity TD, Schöttl L, Trauth A, Schemmann M, Hohberg M, Pinter P, Elsner P, Henning F et al (2019) Virtual process chain of sheet molding compound: development, validation and perspectives. *Compos B Eng* 169:133–147
- Meyer N, Gajek S, Görthofer J, Hrymak A, Kärger L, Henning F, Schneider M, Böhlke T (2023) A probabilistic virtual process chain to quantify process-induced uncertainties in sheet molding compounds. *Compos B Eng* 249:110380
- Castro J, Tomlinson G (1990) Predicting molding forces in SMC compression molding. *Poly Eng Sci* 30(24):1568–1573
- Goodship V (2017) ARBURG practical guide to injection moulding. Smithers Rapra, Shawbury, United Kingdom
- Tseng H-C, Chang R-Y, Hsu C-H (2018) Predictions of fiber concentration in injection molding simulation of fiber-reinforced composites. *J Thermoplast Compos Mater* 31(11):1529–1544
- Karl T, Gatti D, Böhlke T, Frohnepfel B (2021) Coupled simulation of flow-induced viscous and elastic anisotropy of short-fiber reinforced composites. *Acta Mech* 232(6):2249–2268
- Böhlke T, Henning F, Hrymak AN, Kärger L, Weidenmann K, Wood JT (2019) Continuous-discontinuous fiber-reinforced polymers: an integrated engineering approach. Carl Hanser Verlag GmbH Co KG, Munich
- Rahnama M, Koch DL, Shaqfeh ES (1995) The effect of hydrodynamic interactions on the orientation distribution in a fiber suspension subject to simple shear flow. *Phys Fluids* 7(3):487–506
- Sundararajakumar R, Koch DL (1997) Structure and properties of sheared fiber suspensions with mechanical contacts. *J Nonnewton Fluid Mech* 73(3):205–239
- Karl T, Böhlke T (2022) Unified mean-field modeling of viscous short-fiber suspensions and solid short-fiber reinforced composites. *Arch Appl Mech* 92(12):3695–3727
- Sepehr M, Carreau PJ, Moan M, Ausias G (2004) Rheological properties of short fiber model suspensions. *J Rheol* 48(5):1023–1048
- Dinh SM, Armstrong RC (1984) A rheological equation of state for semiconcentrated fiber suspensions. *J Rheol* 28(3):207–227
- Cross MM (1965) Rheology of Non-Newtonian fluids: a new flow equation for pseudoplastic systems. *J Colloid Sci* 20(5):417–437
- Sterr B, Wicht D, Hrymak A, Schneider M, Böhlke T (2023) Homogenizing the viscosity of shear-thinning fiber suspensions with an FFT-based computational method. *J Nonnewton Fluid Mech* 321:105101
- Williams ML, Landel RF, Ferry JD (1955) The temperature dependence of relaxation mechanisms in amorphous polymers and other glass-forming liquids. *J Am Chem Soc* 77(14):3701–3707
- Batchelor G (1970) The stress system in a suspension of force-free particles. *J Fluid Mech* 41(3):545–570
- Batchelor G (1971) The stress generated in a non-dilute suspension of elongated particles by pure straining motion. *J Fluid Mech* 46(4):813–829
- Goddard JD (1976) Tensile stress contribution of flow-oriented slender particles in Non-Newtonian fluids. *J Nonnewton Fluid Mech* 1(1):1–17
- Goddard J (1976) The stress field of slender particles oriented by a Non-Newtonian extensional flow. *J Fluid Mech* 78(1):177–206
- Goddard J (1978) Tensile Behavior of Power-Law Fluids Containing Oriented Slender Fibers. *J Rheol* 22(6):615–622
- Mobuchon C, Carreau PJ, Heuzey M-C, Sepehr M, Ausias G (2005) Shear and extensional properties of short glass fiber reinforced polypropylene. *Polym Compos* 26(3):247–264
- Souloumiac B, Vincent M (1998) Steady shear viscosity of short fiber suspensions in thermoplastics. *Rheol Acta* 37(3):289–298
- Férec J, Bertevas E, Khoo BC, Ausias G, Phan-Thien N (2016) The effect of shear-thinning behaviour on rod orientation in filled fluids. *J Fluid Mech* 798:350–370
- Pipes RB, Hearle J, Beaussart A, Sastry A, Okine R (1991) A constitutive relation for the viscous flow of an oriented fiber assembly. *J Compos Mater* 25(9):1204–1217
- Pipes RB (1992) Anisotropic viscosities of an oriented fiber composite with a power-law matrix. *J Compos Mater* 26(10):1536–1552

30. Coffin D, Pipes RB (1991) Constitutive relationships for aligned discontinuous fibre composites. *Compos Manuf* 2(3–4):141–146
31. Pipes RB, Coffin DW, Shuler SF, Šimáček P (1994) Non-Newtonian constitutive relationships for hyperconcentrated fiber suspensions. *J Compos Mater* 28(4):343–351
32. Ericksen J (1960) Transversely isotropic fluids. *Kolloid-Zeitschrift* 173:117–122
33. Tucker CL III (1991) Flow regimes for fiber suspensions in narrow gaps. *J Nonnewton Fluid Mech* 39(3):239–268
34. Binding D (1991) Capillary and contraction flow of long-(glass) fibre filled polypropylene. *Compos Manuf* 2(3–4):243–252
35. Favaloro AJ, Tseng H-C, Pipes RB (2018) A new anisotropic viscous constitutive model for composites molding simulation. *Compos A Appl Sci Manuf* 115:112–122
36. Tseng H-C (2021) A constitutive equation for fiber suspensions in viscoelastic media. *Phys Fluids* 33(7):071702
37. Khan M, More RV, Ardekani AM (2023) A constitutive model for sheared dense fiber suspensions. *Physics of Fluids*, DOI 10.1063/5:0134728
38. Schelleis C, Hrymak A, Henning F (2023) “Optimizing processing parameters for glass fiber reinforced polycarbonate lft-d composites,” Society for the Advancement of Material and Process Engineering Europe Conference
39. Eschbach AR (1993) “Dynamic shear rheometer and method,” US Patent 5,271,266
40. Švec O, Skoček J, Stang H, Geiker MR, Roussel N (2012) Free surface flow of a suspension of rigid particles in a non-Newtonian fluid: A lattice Boltzmann approach. *J Nonnewton Fluid Mech* 179:32–42
41. Domurath J, Ausias G, Férec J, Saphiannikova M (2020) Numerical investigation of dilute suspensions of rigid rods in power-law fluids. *J Nonnewton Fluid Mech* 280:104280
42. Kanit T, Forest S, Galliet I, Mounoury V, Jeulin D (2003) Determination of the size of the representative volume element for random composites: statistical and numerical approach. *Int J Solids Struct* 40(13–14):3647–3679
43. Schneider M (2021) A review of nonlinear FFT-based computational homogenization methods. *Acta Mech* 232(6):2051–2100
44. Bertóti R, Wicht D, Hrymak A, Schneider M, Böhlke T (2021) A computational investigation of the effective viscosity of short-fiber reinforced thermoplastics by an FFT-based method. *European J Mech-B/Fluids* 90:99–113
45. Kanatani K (1984) Distribution of directional data and fabric tensors. *Int J Eng Sci* 22(2):149–164
46. Fokker AD (1914) Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld. *Ann Phys* 348(5):810–820
47. Advani SG, Tucker CL III (1987) The use of tensors to describe and predict fiber orientation in short fiber composites. *J Rheol* 31(8):751–784
48. Köbler J, Schneider M, Ospald F, Andrá H, Müller R (2018) Fiber orientation interpolation for the multiscale analysis of short fiber reinforced composite parts. *Comput Mech* 61:729–750
49. Barzilai J, Borwein JM (1988) Two-point step size gradient methods. *IMA J Numer Anal* 8(1):141–148
50. Dembo RS, Eisenstat SC, Steihaug T (1982) Inexact Newton methods. *SIAM J Numer Anal* 19(2):400–408
51. Schneider M (2020) A dynamical view of nonlinear conjugate gradient methods with applications to FFT-based computational micromechanics. *Comput Mech* 66(1):239–257
52. Schneider M (2017) The sequential addition and migration method to generate representative volume elements for the homogenization of short fiber reinforced plastics. *Comput Mech* 59(2):247–263
53. “Ultramid®B3K Polyamide 6 Material data.” <https://www.campusplastics.com/campus/de/datasheet/Ultramid%C2%AE+B3K/BASF/20/3a22f000>. Accessed: 2020-09-26
54. Harlow FH, Welch JE (1965) Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys Fluids* 8(12):2182–2189
55. Kabel M, Fink A, Schneider M (2017) The composite voxel technique for inelastic problems. *Comput Methods Appl Mech Eng* 322:396–418
56. Valero JRL (2020) *Plastics injection molding: scientific molding, recommendations, and best practices*. Carl Hanser Verlag GmbH Co KG, Munich
57. Friesenbichler W, Duretek I, Rajganes J, Kumar SR (2011) Measuring the pressure dependent viscosity at high shear rates using a new rheological injection mould. *Polimery* 56(1):58–62
58. Böhlke T, Brüggemann C (2001) Graphical representation of the generalized Hooke’s law. *Tech Mech* 21(2):145–158
59. He Q-C, Curnier A (1995) A more fundamental approach to damaged elastic stress-strain relations. *Int J Solids Struct* 32(10):1433–1457
60. Kube CM (2016) Elastic anisotropy of crystals. *AIP advances*, DOI 10.1063/1:4962996
61. Ranganathan SI, Ostoja-Starzewski M (2008) Universal elastic anisotropy index. *Phys Rev Lett* 101(5):055504
62. Zener CM, Siegel S (1949) Elasticity and Anelasticity of Metals. *J Phys Chem* 53(9):1468–1468
63. Koch DL (1995) A model for orientational diffusion in fiber suspensions. *Phys Fluids* 7(8):2086–2088
64. Montgomery-Smith S, He W, Jack DA, Smith DE (2011) Exact tensor closures for the three-dimensional Jeffery’s equation. *J Fluid Mech* 680:321–335
65. Chandra MP et al (1936) On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India* 2:49–55
66. Kingma DP, Ba J (2014) “Adam: A method for stochastic optimization,” arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)
67. Kabanikhin SI (2008) Definitions and examples of inverse and ill-posed problems. *J Inverse Illposed Prob* 16:317–357
68. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S (2019) “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Advances in Neural Information Processing Systems* 32, pp. 8024–8035, Curran Associates, Inc
69. Polyak BT (1964) Some methods of speeding up the convergence of iteration methods. *USSR Comput Math Math Phys* 4(5):1–17
70. Kabel M, Böhlke T, Schneider M (2014) Efficient fixed point and Newton-Krylov solvers for FFT-based homogenization of elasticity at large deformations. *Comput Mech* 54(6):1497–1514
71. Ruder S (2016) “An overview of gradient descent optimization algorithms,” arXiv preprint [arXiv:1609.04747](https://arxiv.org/abs/1609.04747)
72. Duchi J, Hazan E, Singer Y (2011) Adaptive subgradient methods for online learning and stochastic optimization. *J Mach Learn Res* 12(7):2121–2159
73. Zeiler MD (2012) “Adadelta: an adaptive learning rate method,” arXiv preprint [arXiv:1212.5701](https://arxiv.org/abs/1212.5701)
74. Tieleman T, Hinton G et al (2012) Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. COURSE: *Neural networks for machine learning* 4(2):26–31
75. Chen X, Liu S, Sun R, Hong M (2018) “On the convergence of a class of ADAM-type algorithms for non-convex optimization,” arXiv preprint [arXiv:1808.02941](https://arxiv.org/abs/1808.02941)
76. Sun S, Cao Z, Zhu H, Zhao J (2019) A survey of optimization methods from a machine learning perspective. *IEEE transactions on cybernetics* 50(8):3668–3681
77. Liu DC, Nocedal J (1989) On the limited memory BFGS method for large scale optimization. *Math Program* 45(1):503–528
78. Darken C, Moody J (1990) Note on learning rate schedules for stochastic optimization. *Adv Neural inf Process Syst* 3:832–838

79. Darken C, Chang J, Moody J et al (1992) Learning rate schedules for faster stochastic gradient search. *Neural networks for signal processing* 2:3–12 (**Citeseer**)
80. Feyel F (1999) Multiscale FE2 elastoviscoplastic analysis of composite structures. *Comput Mater Sci* 16(1–4):344–354
81. Miehe C (2002) Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation. *Int J Numer Meth Eng* 55(11):1285–1322
82. Fritzen F, Hodapp M (2016) The finite element square reduced (FE2R) method with GPU acceleration: towards three-dimensional two-scale simulations. *Int J Numer Meth Eng* 107(10):853–881
83. Spahn J, Andrä H, Kabel M, Müller R (2014) A multiscale approach for modeling progressive damage of composite materials using fast Fourier transforms. *Comput Methods Appl Mech Eng* 268:871–883
84. Loshchilov I, Hutter F (2017) “Decoupled weight decay regularization,” arXiv preprint [arXiv:1711.05101](https://arxiv.org/abs/1711.05101)
85. Dozat T (2016) “Incorporating Nesterov momentum into ADAM,” *International Conference on Learning Representations Workshop.*, 1–4
86. Yang L, Shami A (2020) On hyperparameter optimization of machine learning algorithms: theory and practice. *Neurocomputing* 415:295–316
87. Van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9(11):2579–2605
88. McInnes L, Healy J, Melville J (2018) “Umap: Uniform manifold approximation and projection for dimension reduction,” arXiv preprint [arXiv:1802.03426](https://arxiv.org/abs/1802.03426)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.