

# DeePN<sup>2</sup>: A Deep Learning-Based Non-Newtonian Hydrodynamic Model\*

Lidong Fang<sup>1</sup>, Pei Ge<sup>1</sup>, Lei Zhang<sup>2</sup>, Weinan E<sup>3,4,5</sup>, and Huan Lei<sup>† 1,6</sup>

<sup>1</sup>Department of Computational Mathematics, Science and Engineering, Michigan State University, MI 48824, USA.

<sup>2</sup>School of Mathematical Sciences, Institute of Natural Sciences, and MOE-LSC, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China.

<sup>3</sup>Center for Machine Learning Research and School of Mathematical Sciences, Peking University, Beijing 100871, China.

<sup>4</sup>AI for Science Institute, Beijing 100080, China.

<sup>5</sup>Department of Mathematics and Program in Applied and Computational Mathematics, Princeton University, NJ 08544, USA.

<sup>6</sup>Department of Statistics and Probability, Michigan State University, MI 48824, USA.

**Abstract.** A long standing problem in the modeling of non-Newtonian hydrodynamics of polymeric flows is the availability of reliable and interpretable hydrodynamic models that faithfully encode the underlying micro-scale polymer dynamics. The main complication arises from the long polymer relaxation time, the complex molecular structure and heterogeneous interaction. DeePN<sup>2</sup>, a deep learning-based non-Newtonian hydrodynamic model, has been proposed and has shown some success in systematically passing the micro-scale structural mechanics information to the macro-scale hydrodynamics for suspensions with simple polymer conformation and bond potential. The model retains a multi-scaled nature by mapping the polymer configurations into a set of symmetry-preserving macro-scale features. The extended constitutive laws for these macro-scale features can be directly learned from the kinetics of their micro-scale counterparts. In this paper, we develop DeePN<sup>2</sup> using more complex micro-structural models. We show that DeePN<sup>2</sup> can faithfully capture the broadly overlooked viscoelastic differences arising from the specific molecular structural mechanics without human intervention.

## Keywords:

Non-Newtonian fluids,  
Machine learning,  
Multi-scale modeling,  
Fluid mechanics.

## Article Info.:

Volume: 1  
Number: 1  
Pages: 114- 140  
Date: March/2022  
doi.org/10.4208/jml.220115

## Article History:

Received: 4/1/2022  
Accepted: 19/3/2022

## Communicated by:

Zhi-Qin John Xu

## 1 Introduction

Accurate modeling of non-Newtonian hydrodynamics plays a central role in the modeling of the transport, diffusion, and synthesis processes in many scientific and engineering applications. Unlike simple fluids, non-Newtonian fluids may exhibit enormously complex flow behavior as a result of the micro-scale polymer dynamics. In particular, the polymer relaxation time often becomes comparable to the hydrodynamic time scale. As a result, the macro-scale fluid evolution can not be uniquely determined by the instantaneous flow

\*The first two authors contributed equally.

†Corresponding author. leihuan@msu.edu.

field and the memory effect is generally important. To close the hydrodynamic equations, existing models are primarily based on the following two approaches. The first approach relies on empirical constitutive models [1, 2]. Notable examples include the Hookean model [3, 4], the FENE-P model [5, 6], the Giesekus model [7], and the Phan-Thien and Tanner models [8]. Despite their popularity, the accuracy of these models is almost always in doubt. The second approach resorts to various sophisticated micro-macro coupling algorithms, e.g., by directly solving the Fokker-Planck equation using Lattice Boltzmann method [9], Galerkin method [10–13], and particle method [14–18], or sampling the polymer configuration via micro-scale simulations [19–21]. While the effects of the polymer interaction can be carried over to the macro-scale model, the computational cost can be exceedingly large due to the retaining of the micro-scale description. Methods based on asymptotic analysis [22, 23] or the direct fitting of the strain-stress relationship [24] are limited to simple flows such as the steady flow. Several semi-analytical approaches have been proposed [25–31] using moment closure to approximate the micro-scale polymer configuration probability density function (PDF) and to derive the constitutive equations for the FENE dumbbell solution [29–31]. However, these approaches are all based on restricted ansatz for the PDF and therefore are not reliable for more general flow regimes.

To construct truly reliable and interpretable hydrodynamic models with molecular-level fidelity, it is essential to be able to efficiently code the information from the micro-scale interaction into the macro-scale transport equations. Ideally, the construction should meet the following requirements:

- be interpretable;
- be reliable – it should be accurate for all kinds of practical situations that one might encounter;
- respect physical constraints, including symmetries and conservation laws;
- be numerically robust and efficient.

As a first step towards constructing models that meet these requirements, we developed a machine learning-based approach [32], “deep learning-based non-Newtonian hydrodynamic model” or DeePN<sup>2</sup>, that learns the non-Newtonian hydrodynamic model from the underlying micro-scale description of the dumbbell solution. Rather than approximating the closure with standard moments, DeePN<sup>2</sup> finds a set of encoders, i.e., a set of macro-scale features that best represent the micro-scale dumbbell structure. It also finds accurate closed-form equation for these macro-scale features. The constructed model retains a clear physical interpretation and accurately captures the nonlinear viscoelastic responses, where the conventional Hookean and FENE-P models show limitations.

Beyond dumbbell suspensions, one major challenge towards constructing truly reliable hydrodynamic models arises from the heterogeneous polymer micro-structural mechanics. In this work, we aim to fill the gap by developing the generalized DeePN<sup>2</sup> model for multi-bead polymer molecules with arbitrary structure and interaction. Firstly, with the proper design of the generalized micro-macro encoders and the machine learning-based symmetry-preserving constitutive dynamics, we demonstrate that the heterogeneous molecular structural-induced interaction can be systematically encoded into the

macro-scale hydrodynamics. Unlike moment closure approximations, the encoders are not designed to recover the high-dimensional configuration PDF. Instead, they take an *interpretable* form and are learned to probe the optimal approximation of the polymer stress and constitutive dynamics. This essential difference enables DeePN<sup>2</sup> to circumvent the high-dimensionality of the polymer configuration PDF. Secondly, the explicit form of the micro-macro encoders enables us to reliably learn the dynamics of the macro-scale features directly from the kinetic equations of their micro-scale analog. In this sense, this learning framework retains a multi-scaled nature where micro-scale interaction and physical constraints can be seamlessly inherited. Moreover, the learning only requires instantaneous micro-scale samples. This unique property differs from the common sophisticated data-driven approaches [33–40], where time-derivative samples are often needed to learn the governing dynamics. This is particularly suited for multi-scale fluid models where accurate time-derivative samples may not be readily accessible. We demonstrate the power of the DeePN<sup>2</sup> model for polymer molecules of three distinct shapes with training samples collected from one-dimensional (1D) homogeneous shear flow. Numerical results show that the broadly overlooked heterogeneous molecular structural mechanics plays an important role in the rheology of non-Newtonian fluids, which, fortunately, can be faithfully encoded into DeePN<sup>2</sup>. The constructed model successfully captures the hydrodynamics with different viscoelastic responses for a variety of 1D and 2D flows when compared with the micro-scale simulation results. The present work also paves the way towards constructing truly reliable non-Newtonian hydrodynamic models for general 3D flows.

## 2 Methods

### 2.1 Micro-scale and continuum hydrodynamic models

Let us start with the micro-scale description of the semi-dilute polymer suspension. We assume each molecule consists of  $N$  particles with the position vector  $\mathbf{q} = [\mathbf{q}_1; \mathbf{q}_2; \dots; \mathbf{q}_N]$ , where  $\mathbf{q}_i \in \mathbb{R}^3$  is the position of the  $i$ -th particle. The intramolecular potential energy  $V(\mathbf{q})$  takes the form

$$V(\mathbf{q}) = \sum_{j=1}^{N_b} V_b(|\mathbf{q}_{j_1} - \mathbf{q}_{j_2}|), \quad V_b(l) = -\frac{k_s l_0^2}{2} \log \left[ 1 - \frac{l^2}{l_0^2} \right], \quad (2.1)$$

where  $N_b$  is the bond number and  $(j_1, j_2)$  represents the indices of beads associated with the  $j$ -th bond. Without loss of generality, the individual bond interaction  $V_b$  takes the form of the FENE potential [41], where  $k_s$  is the spring constant and  $l_0$  is the maximum of the extension length. It is worth mentioning that the polymer molecule is not restricted to the dumbbell shape. Instead, it generally consists of multiple particles with arbitrary structure and bond connection. Fig. 2.1 shows a sketch of the polymer molecules with three different structures. As we will show, given the same form of the individual bond interaction  $V_b$ , the different polymer micro-structural mechanics leads to distinct non-Newtonian hydrodynamics.

In principle, the viscoelastic response of the system is determined by the full micro-scale interaction. However, direct simulation for the full micro-scale interaction is often

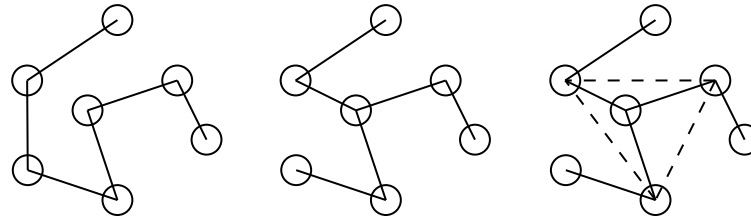


Figure 2.1: A sketch of 7-bead polymer molecule with chain-, star- and net-shaped structures (from left to right). The solid lines represent the FENE bond potential with the *same* interaction parameters. The dashed lines of the net-shaped molecule represent the three additional side chains connecting the polymer arms. While both the chain- and the star-shaped molecules are connected with six bonds; the suspensions exhibit different hydrodynamics due to the different micro-structural mechanics as shown below.

limited by the prohibited computational cost. Continuum hydrodynamics models based on various empirical constitutive models are often used, with the general form

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0, \\ \rho \frac{d\mathbf{u}}{dt} &= -\nabla p + \nabla \cdot (\boldsymbol{\tau}_s + \boldsymbol{\tau}_p) + \mathbf{f}_{\text{ext}}, \end{aligned} \tag{2.2}$$

where  $\rho$ ,  $\mathbf{u}$  and  $p$  represent the fluid density, velocity and pressure field, respectively.  $\mathbf{f}_{\text{ext}}$  is the external body force and  $\boldsymbol{\tau}_s = \eta_s(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$  is the solvent stress tensor with shear viscosity  $\eta_s$ .  $\boldsymbol{\tau}_p$  is the polymer stress tensor whose detailed form is generally unknown. To construct  $\boldsymbol{\tau}_p$ , the DeePN<sup>2</sup> model seeks the approximation in terms of a set of macro-scale features  $\mathbf{c}_1, \dots, \mathbf{c}_n$ , and simultaneously, the constitutive dynamics of these features, i.e.,

$$\boldsymbol{\tau}_p = \mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n), \tag{2.3a}$$

$$\frac{D\mathbf{c}_i}{Dt} = \mathbf{H}_i(\mathbf{c}_1, \dots, \mathbf{c}_n), \quad i = 1, \dots, n, \tag{2.3b}$$

where  $\mathbf{G}$  and  $\mathbf{H}_i$  represent the stress and constitutive models, respectively.  $\frac{D}{Dt}$  denotes the objective tensor derivative.

Eqs. (2.2) and (2.3) take the form similar to the conventional hydrodynamics. Instead of using empirical approximation to close the equation, we aim to construct a model directly from the micro-scale description (2.1) with the help of machine learning, such that the constructed model can naturally encode the molecular-specific interaction beyond empirical approximations with clear physical interpretation.

## 2.2 DeePN<sup>2</sup> for arbitrary molecular structural mechanics

To learn Eq. (2.2) from the full model (2.1), one essential problem lies in how to seamlessly pass the micro-scale interaction to the continuum model. To bridge the scales, we learn a set of micro-to-macro encoders, denoted by  $\{\mathbf{b}_i(\mathbf{q})\}_{i=1}^n$ , such that the continuum modeling terms (e.g., the polymer stress  $\boldsymbol{\tau}_p$ ) can be well approximated in terms of the corresponding macro-scale features  $\{\mathbf{c}_i(\mathbf{q})\}_{i=1}^n$  via Eq. (2.3a), where  $\boldsymbol{\tau}_p := n_p \sum_j \langle \mathbf{q}_j \otimes \nabla_{\mathbf{q}_j} V(\mathbf{q}) \rangle$ ,

$\mathbf{c}_i = \langle \mathbf{b}_i(\mathbf{q}) \rangle$ ,  $n_p$  is the polymer number density and  $\langle \cdot \rangle$  denotes the average with respect to the configuration PDF. In particular, the features  $\mathbf{c}_i$  need to satisfy the proper invariant and symmetry conditions inherited from the encoders  $\mathbf{b}_i(\cdot)$  such that the constructed continuum model can strictly preserve frame-indifference condition:

$$\widetilde{\boldsymbol{\tau}}_p = \mathbf{Q}\boldsymbol{\tau}_p\mathbf{Q}^T, \quad \mathbf{G}(\widetilde{\mathbf{c}}_1, \dots, \widetilde{\mathbf{c}}_n) = \mathbf{Q}\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n)\mathbf{Q}^T, \quad (2.4)$$

where the superscript  $\widetilde{\cdot}$  denotes the corresponding values under an arbitrary orthogonal transformation by  $\mathbf{Q} \in \text{SO}(3)$ .

To construct the encoder  $\mathbf{b}(\cdot)$ , we note that the micro-scale potential  $V(\mathbf{q})$  is translational and rotational invariant. Accordingly, let  $\mathbf{r}^*(\mathbf{q}) \in \mathbb{R}^{3N-6}$  (we consider the general case  $N \geq 3$  here) denote the translational-rotational-invariant configuration vector and  $\mathbf{r}(\mathbf{q}) \in \mathbb{R}^{3N-3}$  denote the translational-invariant configuration vector consisting of  $N - 1$  linearly independent position vectors. Since  $N_b \geq N - 1$  for all molecules, one straightforward choice is the first  $N - 1$  bond connection vectors, i.e.,

$$\begin{aligned} \mathbf{r} &= [\mathbf{r}_1; \mathbf{r}_2; \dots; \mathbf{r}_{N-1}], \quad \mathbf{r}_j = \mathbf{q}_{j1} - \mathbf{q}_{j2}, \quad 1 \leq j \leq N - 1, \\ \mathbf{r}^* &= \left[ |\mathbf{r}_1|, |\mathbf{r}_2|, |\mathbf{r}_{12}|, |\mathbf{r}_3|, |\mathbf{r}_{13}|, |\mathbf{r}_{23}|, |\mathbf{r}_4|, |\mathbf{r}_{24}|, |\mathbf{r}_{34}|, \dots, |\mathbf{r}_{N-1}|, |\mathbf{r}_{(N-2)(N-1)}| \right], \end{aligned} \quad (2.5)$$

where  $\mathbf{r}_{jk} := \mathbf{r}_j - \mathbf{r}_k$ . We note that this form applies to general molecular structures;  $\mathbf{r}$  determines the molecular structure up to translations. Specifically,  $\mathbf{r}^*$  represents the  $3N - 6$  degrees of freedom after eliminating translational and rotational degrees of freedom, and  $\mathbf{r}$  suffices to fully determine the translational invariant polymer configuration and strictly retains the rotational symmetry in accordance with  $\mathbf{q}$ , i.e.,

$$\mathbf{r}_j(\mathbf{Q}\mathbf{q}) = \mathbf{Q}\mathbf{r}_j(\mathbf{q}), \quad \mathbf{r}^*(\mathbf{Q}\mathbf{q}) = \mathbf{r}^*(\mathbf{q}).$$

To preserve rotational symmetry, one straightforward approach is to represent  $\mathbf{b}(\cdot)$  in the linear space spanned by  $\{\mathbf{r}_j\}_{j=1}^{N-1}$ . However, this choice yields the trivial macro-scale feature, i.e.,  $\langle \mathbf{r}_j \rangle \equiv 0$ , due to the rotational symmetry. Alternatively, we construct the following second-order tensor

$$\begin{aligned} \mathbf{c}_i &= \langle \mathbf{b}_i(\mathbf{r}) \rangle, \quad \mathbf{b}_i = \mathbf{f}_i \mathbf{f}_i^T, \quad 1 \leq i \leq n, \\ \mathbf{f}_i &= g_i(\mathbf{r}^*) \sum_{j=1}^{N-1} w_{ij} \mathbf{r}_j, \end{aligned} \quad (2.6)$$

where  $[w_{ij}]_{1 \leq i \leq n, 1 \leq j \leq N-1}$  are the weights and  $\{g_i(\cdot)\}_{i=1}^n$  is a set of scalar functions that encodes the polymer intramolecular interaction. Both terms will be learned from the micro-scale description and represented by deep neural networks (DNNs). Rotational symmetries can be naturally inherited, i.e.,  $\widetilde{\mathbf{c}} = \langle \mathbf{b}(\widetilde{\mathbf{r}}) \rangle \equiv \mathbf{Q}\mathbf{c}\mathbf{Q}^T$ . Compared with the special form for dumbbell molecules in Ref. [32], Eq. (2.6) provides a general form of  $\mathbf{c}$  applicable to multi-bead molecules of arbitrary structure since  $\mathbf{r}$  and  $\mathbf{r}^*$  fully determine the  $3N - 3$  translational invariant polymer configuration. In the remaining of the paper, we will abuse the notation and denote  $\mathbf{b}(\mathbf{q})$  as  $\mathbf{b}(\mathbf{r})$ .

Besides the polymer stress model (2.3a), the remaining task to close Eq. (2.2) is the construction of the constitutive dynamics (2.3b) of the macro-scale features  $\{\mathbf{c}_i\}_{i=1}^n$ . There are two issues to deal with: the proper form of the objective time derivative of  $\mathbf{c}_i$  and the accurate estimation of their time evolution. In the literature, the objective tensor derivative, denoted by  $\frac{D\mathbf{c}_i}{Dt}$ , is often chosen to take some heuristic forms (e.g. the convected [3] and corotational [42] forms). Moreover, the time-series samples collected from the micro-scale simulations are generally super-imposed with pronounced sampling error; direct estimation of the time derivative as was done in [33,35,38] will end with noisy data. Fortunately, both challenges are addressed in DeePN<sup>2</sup> using an explicit micro-macro correspondence. The dynamics of  $\mathbf{c}_i$  can be derived from the its micro-scale correspondence  $\mathbf{b}_i(\mathbf{r})$  in the form of the micro-scale configuration  $\mathbf{r}$ , i.e.,

$$\begin{aligned} \frac{d}{dt}\mathbf{c}_i - \boldsymbol{\kappa} : \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle &= \frac{k_B T}{\gamma} \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b}_i \right\rangle \\ &\quad - \frac{1}{\gamma} \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V(\mathbf{r}_1, \dots, \mathbf{r}_{N_b}) \cdot \nabla_{\mathbf{r}_j} \mathbf{b}_i \right\rangle, \end{aligned} \quad (2.7)$$

where  $\boldsymbol{\kappa} := \nabla \mathbf{u}^T$ ,  $\gamma$  is the friction coefficient and  $\mathbf{r}_j$  is the connection vector as defined in Eq. (2.5) for  $j > N - 1$ . We abuse the notation and denote  $V(\mathbf{q})$  as  $V(\mathbf{r}_1, \dots, \mathbf{r}_{N_b}) = \sum_{j=1}^{N_b} V_b(r_j)$ . The molecular structure and interaction are specified via  $\mathbf{A} \in \mathbb{R}^{N_b \times N_b}$ , which is defined by

$$\mathbf{A} = \mathbf{S}\mathbf{S}^T, \quad S_{jk} = \begin{cases} +1, & k = j_1, \\ -1, & k = j_2, \\ 0, & \text{else,} \end{cases} \quad 1 \leq j \leq N_b, \quad 1 \leq k \leq N, \quad (2.8)$$

where  $j_1$  and  $j_2$  are the same notations as those in Eq. (2.1). We note that Eq. (2.7) only requires the first  $(N - 1)$  rows of  $\mathbf{A}$  since the polymer configuration can be fully determined by  $\mathbf{r}_1, \dots, \mathbf{r}_{N-1}$ . As a special case, if the molecule takes the chain shape,  $\mathbf{A}$  recovers the standard Rouse matrix [43,44].

Eq. (2.7) defines the dynamics for the features  $\{\mathbf{c}_i\}_{i=1}^n$ , derived from their micro-scale correspondences. In particular, given the proposed form of the encoder functions (2.6), we can show that the two combined terms of the left-hand-side of Eq. (2.7) strictly preserve rotational symmetry (see Appendix A). This leads to an important observation that the two combined terms provide the generalized form for the macro-scale objective tensor derivative  $\frac{D\mathbf{c}_i}{Dt}$ . Unlike the heuristic choices in empirical models, the new form retains a clear micro-scale physical interpretation. Furthermore, all the modeling terms in the form of  $\langle \cdot \rangle$  can be directly evaluated using samples collected from the micro-scale simulations under the corresponding flow condition. This enables us to avoid estimating the time derivative values from the noise-prone time-series samples. Accordingly, the macro-scale constitutive dynamics takes the form

$$\frac{d\mathbf{c}_i}{dt} - \boldsymbol{\kappa} : \mathcal{E}_i = \frac{k_B T}{\gamma} \mathbf{H}_{1,i}(\mathbf{c}_1, \dots, \mathbf{c}_n) - \frac{1}{\gamma} \mathbf{H}_{2,i}(\mathbf{c}_1, \dots, \mathbf{c}_n), \quad (2.9)$$

where the individual terms will be represented by proper neural networks and parameterized by matching their micro-scale correspondences, i.e.,

$$\begin{aligned}
 \mathcal{E}_i(\mathbf{c}_1, \dots, \mathbf{c}_n) &= \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle, \\
 \mathbf{H}_{1,i}(\mathbf{c}_1, \dots, \mathbf{c}_n) &= \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b}_i \right\rangle, \\
 \mathbf{H}_{2,i}(\mathbf{c}_1, \dots, \mathbf{c}_n) &= \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V(\mathbf{r}_1, \dots, \mathbf{r}_{N-1}) \cdot \nabla_{\mathbf{r}_j} \mathbf{b}_i \right\rangle.
 \end{aligned} \tag{2.10}$$

### 2.3 Symmetry-preserving DNN models

To complete the DeePN<sup>2</sup> model, we need to specify the DNN models. These DNN models should also strictly preserve rotational symmetry. Different from the rotational-invariant scalar stress model considered in Ref. [45], the second-order tensors  $\mathbf{G}$ ,  $\mathbf{H}_{1,i}$ ,  $\mathbf{H}_{2,i}$  need to satisfy the symmetry condition (2.4) and the fourth-order tensors  $\mathcal{E}_i$  need to retain the objectivity of  $\frac{\mathcal{D}\mathbf{c}_i}{\mathcal{D}t}$ . However, there does not exist such a reference frame in which these symmetry constraints can be satisfied by the macro-scale modeling terms.

To handle this problem, we consider the eigen-space of the feature  $\mathbf{c}_1$  with a fixed form of the encoder  $\mathbf{b}_1(\cdot)$ , e.g., by setting  $g_1(\cdot) = w_{1,\cdot} \equiv 1$  and let other  $\mathbf{b}_i(\cdot)$  involved in the training. Let us consider the eigen-decomposition  $\mathbf{c}_1 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ , assuming that it has distinct eigenvalues. We introduce the following matrices

$$\begin{aligned}
 \mathcal{S}^{(1)} &= \begin{pmatrix} +1 & & \\ & +1 & \\ & & +1 \end{pmatrix}, & \mathcal{S}^{(2)} &= \begin{pmatrix} +1 & & \\ & -1 & \\ & & +1 \end{pmatrix}, \\
 \mathcal{S}^{(3)} &= \begin{pmatrix} +1 & & \\ & +1 & \\ & & -1 \end{pmatrix}, & \mathcal{S}^{(4)} &= \begin{pmatrix} +1 & & \\ & -1 & \\ & & -1 \end{pmatrix}.
 \end{aligned}$$

We denote  $\mathbf{U}^{(j)} = \mathbf{U}\mathcal{S}^{(j)}$  and  $\hat{\mathbf{c}}_i^{(j)} = \mathbf{U}^{(j)T} \mathbf{c}_i \mathbf{U}^{(j)}$ . We can show that the formulation of the stress model  $\mathbf{G} = \frac{1}{4} \sum_{j=1}^4 \mathbf{U}^{(j)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j)}, \dots, \hat{\mathbf{c}}_n^{(j)}) \mathbf{U}^{(j)T}$  satisfies Eq. (2.4) (see Appendix C).

During simulation, the eigenvalues of  $\mathbf{c}_1$  may cross each other. To account for this, we consider all the 6 permutations of the three eigenvalues, i.e.,

$$\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \frac{1}{24} \sum_{k=0}^5 \sum_{j=1}^4 \mathbf{U}^{(j,k)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j,k)}, \dots, \hat{\mathbf{c}}_n^{(j,k)}) \mathbf{U}^{(j,k)T}, \tag{2.11}$$

where  $k$  represents the rank of permutation (e.g., in lexicographical order) and  $\mathbf{U}^{(j,k)}$  is a variation of  $\mathbf{U}^{(j)}$  with corresponding column permutation. Furthermore, to avoid the

eigenvector degeneracy, we set a threshold value  $\epsilon$  for the eigenvalues. When two eigenvalues approach each other, e.g.,  $|\lambda_2 - \lambda_3| < \epsilon$ , we freeze all the eigenvectors until  $|\lambda_2 - \lambda_3| \geq \epsilon$ . In this work, we take  $\epsilon = 10^{-3}$ , and we refer to Appendix D for detailed numerical studies.

Eq. (2.11) provides the rotation-symmetric form for the second-order stress tensor  $\mathbf{G}$ , where  $\hat{\mathbf{G}}$  is represented by DNNs. The constitutive model terms  $\mathbf{H}_{1,i}$  and  $\mathbf{H}_{2,i}$  can be constructed in a similar manner. Finally, we can show the fourth-order tensors  $\{\mathcal{E}_i\}_{i=1}^n$  associated with the encoders (2.6) can be constructed in the form

$$\boldsymbol{\kappa} : \mathcal{E}_i = \boldsymbol{\kappa} \mathbf{c}_i + \mathbf{c}_i \boldsymbol{\kappa}^T + \boldsymbol{\kappa} : \left( \sum_{j=1}^9 \mathbf{E}_{1,i}^{(j)} \otimes \mathbf{E}_{2,i}^{(j)} \right), \quad (2.12)$$

where  $\mathbf{E}_{1,i}^{(j)}$  and  $\mathbf{E}_{2,i}^{(j)}$  are second-order tensors which respect the symmetry condition (2.4) and can be constructed in the form of Eq. (2.11) (see Appendix C). The constructed DeePN<sup>2</sup> model takes the form similar to the general hydrodynamic equations (2.2) and (2.3), where some of the model terms are represented by DNNs in the form of Eqs. (2.11) and (2.12).

## 2.4 Algorithm

We summarize the DeePN<sup>2</sup> model in Algorithm 1.

---

**Algorithm 1** DeePN<sup>2</sup> for polymer suspensions retaining micro-structural fidelity.

---

- 1: Conduct the micro-scale simulations (see Appendix E) and collect time-discrete training samples (see Appendix F).
  - 2: Pre-process the training samples by pre-computing the first conformation tensor  $\mathbf{c}_1 = \langle \mathbf{b}_1(\mathbf{r}) \rangle$ , its eigen-decomposition, and the polymer stress based on the micro-scale polymer configurations for each training sample set.
  - 3: End-to-end training: Establish the joint learning of the symmetry-preserving encoders  $(\mathbf{b}_2(\cdot), \dots, \mathbf{b}_n(\cdot))$  (see Appendix C), the macro-scale DNN functions (the stress  $\mathbf{G}(\cdot)$  and the other constitutive modeling terms  $\{\mathcal{E}_i(\cdot), \mathbf{H}_{1,i}(\cdot), \mathbf{H}_{2,i}(\cdot)\}_{i=1}^n$ ) by matching the micro-scale counterparts derived from the structure-specific kinetic equations (see Appendix G).
  - 4: Solve the macro-scale hydrodynamic Eqs. (2.2) and (2.9).
- 

## 3 Numerical results

The present DeePN<sup>2</sup> model is trained using micro-scale samples collected from the homogeneous shear flow. We demonstrate the model accuracy and generalization ability by considering various flows in comparison with the results of the micro-scale simulations for the suspensions with three different polymer structural models as shown in Fig. 2.1. As we will see, the micro-scale structure does play an important role in the viscoelastic response. We will use this to examine the DeePN<sup>2</sup> model fidelity.



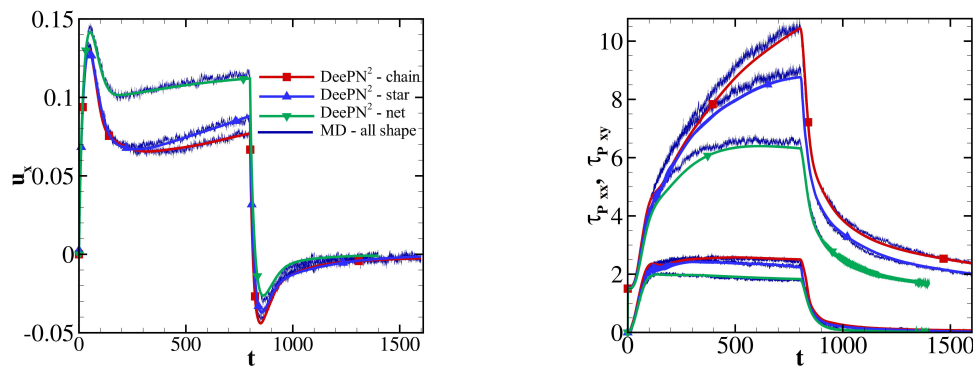


Figure 3.1: The velocity  $u_x$  (left) and polymer stress  $\tau_p$  (right) of the reverse Poiseuille flow ( $y = 6$ ) of the polymer suspensions of three different molecule structures shown in Fig. 2.1.  $\tau_p$  is normalized by polymer number density  $n_p$ , i.e., it is the stress energy per polymer (the same for the remaining figures). With the same FENE bond, the polymer suspensions exhibit different flow responses due to the different molecule structural mechanics. The dark blue lines with rough oscillations denote the micro-scale simulation results; the solid lines with symbols denote the DeePN<sup>2</sup> predictions.

First, we consider the reverse Poiseuille flow in a  $60 \times 100 \times 60$  domain (in reduced unit) with the opposite body force  $\mathbf{f}_{\text{ext}} = (0.016, 0, 0)$  applied to each half of the domain divided by the plane  $y = 50$  starting from  $t = 0$ . At  $t = 800$ , the external force is removed. The relaxation process of the flow field is recorded until the total simulation time  $t = 1600$ . For all the three systems, the predictions from DeePN<sup>2</sup> agree well with the micro-scale simulation results, as shown in Fig. 3.1. In particular, the flow velocity fields of the three systems are nearly identical at the initial stage  $t \in [0, 200]$ , as the development of the flow field is dominated by the solvent and the near-equilibrium responses of the polymer molecules in this regime. Starting from  $t = 250$ , the velocity fields of the three systems exhibit distinct evolution processes. The velocity of the chain-shaped molecule suspension exhibits the largest oscillation and the longest development stage during  $t \in [250, 800]$ . In contrast, the velocity of the star-shaped molecule suspension exhibits moderate oscillation and shows an apparent increase during  $t \in [400, 800]$ , indicating that the polymer elastic energy reaches a plateau earlier than the chain-shaped system. Moreover, the velocity of the net-shaped molecule suspension exhibits the smallest oscillation, indicating that the three additional side-chains further affect the rheological properties of the polymer suspension.

Such differences can also be studied by examining the polymer stress development. As shown in Fig. 3.1, the value of  $\tau_{p_{xx}}$  for the chain-shaped molecule suspension keeps increasing through the development stage  $t \in [0, 800]$  while for the star-shaped molecule,  $\tau_{p_{xx}}$  shows only a moderate increase. In contrast, the net-shaped molecule suspension reaches steady state at about  $t = 400$ . Moreover, the steady value of the shear stress  $\tau_{p_{xy}}$  of the chain-shaped molecule is also larger than the star-shaped and the net-shaped molecules, indicating the largest restored elastic energy. This result is also consistent with

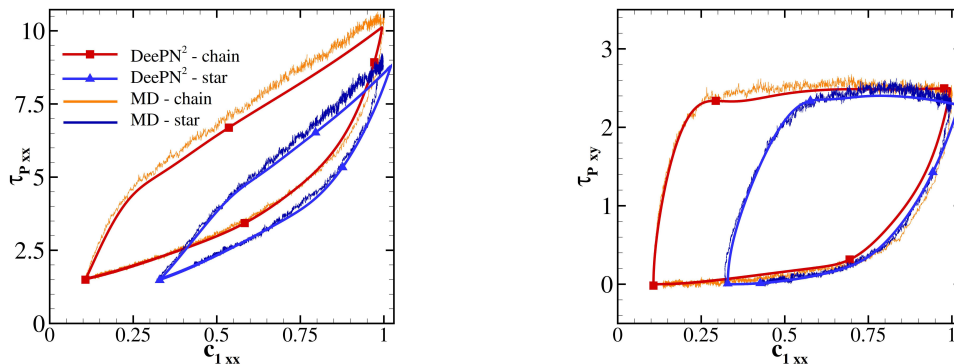


Figure 3.2: The evolution of the polymer stress  $\tau_p$  and conformation tensor  $c_1$  obtained from the reverse Poiseuille flow ( $y = 6$ ) of the polymer suspensions. The clockwise loops represent the development and relaxation processes. For the visualization, the conformation tensor  $c_{1,xx}$  is rescaled by the maximum value obtained from the micro-scale simulation.

the larger velocity oscillation from the minimal values to 0 during the relaxation process with  $t \in [800, 1000]$ .

The different rheological properties of the three polymer suspensions can be understood as follows. Although both the chain-shaped and star-shaped molecules have 6 *identical* FENE bonds, the chain-shaped molecule is less symmetric than the star-shaped molecule. Accordingly, it shows larger dispersion in the  $\mathbb{R}^{18}$  configuration space, and hence, is more flexible than the star-shaped molecule. The elastic response time of the chain-shaped molecule suspension is longer than that of the star-shaped molecule suspension; larger elastic energy can be restored during the relaxation stage. On the other hand, the net-shaped molecule is more rigid than the star-shaped molecule due to the additional bond interaction.

Another important feature of non-Newtonian fluids is the hysteresis effect. Classical models such as Hookean and FENE-P cannot capture such effects [46, 47]. Fig. 3.2 shows the evolution of the polymer stress and conformation tensor for the chain- and star-shaped molecule suspensions. The clockwise loops show the hysteresis effects during the development and relaxation processes; the non-unique stress values indicate that linear and mean field approximations are insufficient in describing the viscoelastic response of the system. In contrast, these effects are accurately captured with the DeePN<sup>2</sup> model. Similar to Fig. 3.1, the chain-shaped molecule suspension shows more pronounced hysteresis effect due to the larger dispersion in the configuration space, reflected as the larger “loop area” than the results for star-shaped molecule suspension.

Next, we investigate the Womersley flow [48] by applying the opposite oscillating body force  $\mathbf{f}_{\text{ext}} = (\pm f_0 \cos(2\pi\omega t), 0, 0)$  to each half of the domain along the  $z$ -direction, where we set  $f_0 = 0.012$  and  $\omega = 1/3000$ . Fig. 3.3 shows the velocity development of the star- and net-shaped molecule suspensions. Similar to the reverse Poiseuille flow, the net-shaped molecule suspension shows less pronounced viscoelastic responses, reflected as the slower

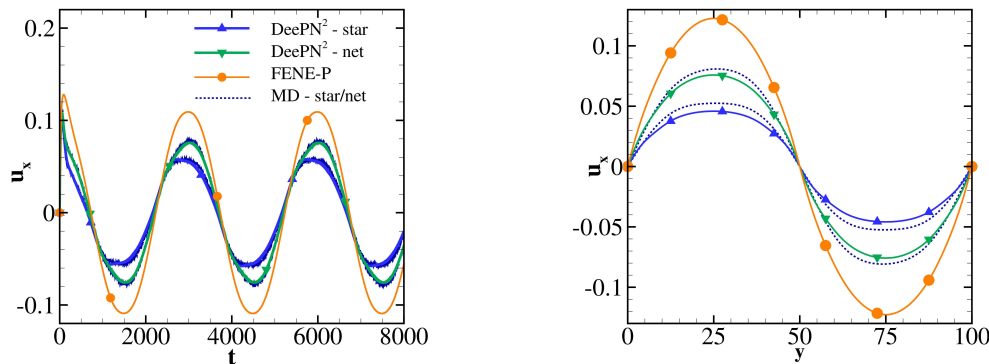


Figure 3.3: The oscillating Womersley flow of the star- and net-shaped molecule suspensions predicted from the micro-scale simulation, DeePN<sup>2</sup> and the FENE-P model. The FENE-P model parameters are chosen to match the dynamics of the orientation tensor (the vector between two free-end particles) near equilibrium. Left: the velocity evolution  $u_x(y, t)$  at  $y = 6$ . Right: the velocity profile  $u_x(y, t)$  at  $t = 6450$ .

decay near  $t \in [200, 400]$  and the larger oscillation due to the less elastic energy storage. For comparison, we also show the prediction from the conventional FENE-P model. The parameters are chosen to match the dynamics of the orientation tensor (the vector between two free-end particles) near equilibrium. As expected, the FENE-P model shows limitations for predicting the flow responses of the two suspensions.

The distinct viscoelastic responses of the different suspensions can be further elucidated by examining the elongation flow. We impose the traceless flow gradient  $\nabla \mathbf{u} = \text{diag}(\dot{\epsilon}, -\dot{\epsilon}, 0)$  where the strain rate  $\dot{\epsilon}$  is set to be  $4 \times 10^{-4}$ . Fig. 3.4 shows the stress development of the chain- and star-shaped molecule suspensions. The micro-scale simulations are imposed by the generalized uniaxial extension flow boundary conditions [49,50]. Compared with the shear flow, the elongation flow yields larger extension and longer processes, as was shown in experimental studies [51]; the steady state is achieved at about  $t = 2.5 \times 10^3$  and  $t = 10^4$  for the star- and chain-shaped molecule suspensions, respectively. Moreover, the steady stress value  $\tau_{p,xx}$  of the chain-shaped molecule suspension is much larger than the value of the star-shaped molecule suspension. Such differences are also due to the larger flexibility of the chain-shaped molecule, which produces a stronger extension under external flow. DeePN<sup>2</sup> successfully captures the different responses and shows good agreement with the micro-scale simulations for both cases.

Finally, we consider the Taylor-Green vortex flow [52,53] in a  $100 \times 100 \times 160$  domain (in reduced unit) of the micro-scale simulation. The external force  $\mathbf{f}_{\text{ext}} = (f_x, f_y, 0)$  is applied to the domain following

$$f_x(x, y) = -2f_0 \sin\left(\frac{2\pi x}{L}\right) \cos\left(\frac{2\pi y}{L}\right), \quad f_y(x, y) = 2f_0 \cos\left(\frac{2\pi x}{L}\right) \sin\left(\frac{2\pi y}{L}\right),$$

where  $L = 100$  and  $f_0 = 6 \times 10^{-3}$ . Periodic boundary conditions are imposed along all of the three directions. The force field imposes an elongation to the flow field along the

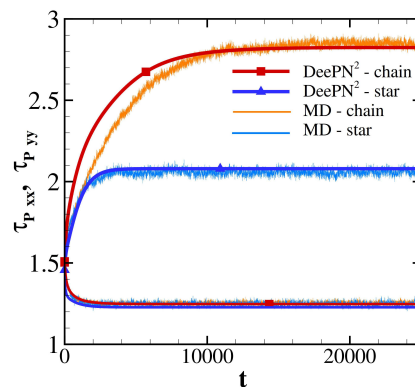


Figure 3.4: The elongation flow of the chain- and star-shaped molecule suspensions predicted from the micro-scale simulation and DeePN<sup>2</sup>. With the same bond potential and strain rate, the chain-shaped molecule suspension yields larger elongation stress. The lines with rough oscillations denote the micro-scale simulation results; the solid lines with symbols denote the DeePN<sup>2</sup> predictions.

$x$ -direction and a compression along the  $y$ -direction. The flow near the center  $(L/2, L/2)$  resembles the planar elongation flow. Four vortices appear at  $(L/2 \pm L/4, L/2 \pm L/4)$ . Fig. 3.5(a-b) shows the steady-state velocity field. Compared with the star-shaped molecule suspension, the velocity field of the chain-shaped molecule suspension shows larger deviation from the symmetric structure of the Newtonian flow (i.e.,  $\propto [-\sin(2\pi x/L) \cos(2\pi y/L), \cos(2\pi x/L) \sin(2\pi y/L)]$ ) due to the larger polymer stress across the flow regime. Furthermore, the two suspensions yield different velocity magnitude, as shown in Fig. 3.5(c). Fig. 3.5(d) shows the velocity development at  $(75, 49)$ . The velocities of both suspensions achieve a similar maximum value near  $t = 30$  and decay along with the polymer stress development. However, the star-shaped molecule suspension reaches the steady state much earlier with a larger velocity than the chain-shaped molecule suspension.

Fig. 3.6 (a-b) shows the steady-state stress field for the two suspensions. We see that the chain-shaped molecule suspension exhibits larger polymer stress variation along the elongation and contraction directions, reflected in the larger loop area in Fig. 3.6(b). Such difference is also consistent with the more pronounced asymmetric velocity field shown in Fig. 3.5(a-b). In addition, we also examine the transient states where the flow undergoes intricate and heterogeneous process. Fig. 3.6(c) shows the stress development at point  $(49, 35)$ , where  $\tau_{P,xx}$  and  $\tau_{P,yy}$  cross over during the evolution. During the initial stage,  $\tau_{P,yy}$  increases along with the flow development towards to the stagnation point. At  $t > 150$ ,  $\tau_{P,yy}$  decreases due to the compression along the  $y$ -direction. Meanwhile,  $\tau_{P,xx}$  increases and achieves a steady state slightly larger than  $\tau_{P,yy}$  for the star-shaped solution. On the other hand, the chain-shaped solution ends up with a significantly larger value of  $\tau_{P,xx}$  due to the larger molecule flexibility and further extension along the  $x$ -direction. The different viscoelastic responses are also reflected in the stress development at point  $(49, 49)$ . As shown in Fig. 3.6(d), the chain-shaped solution exhibits longer evolution of  $\tau_{P,xx}$  and larger

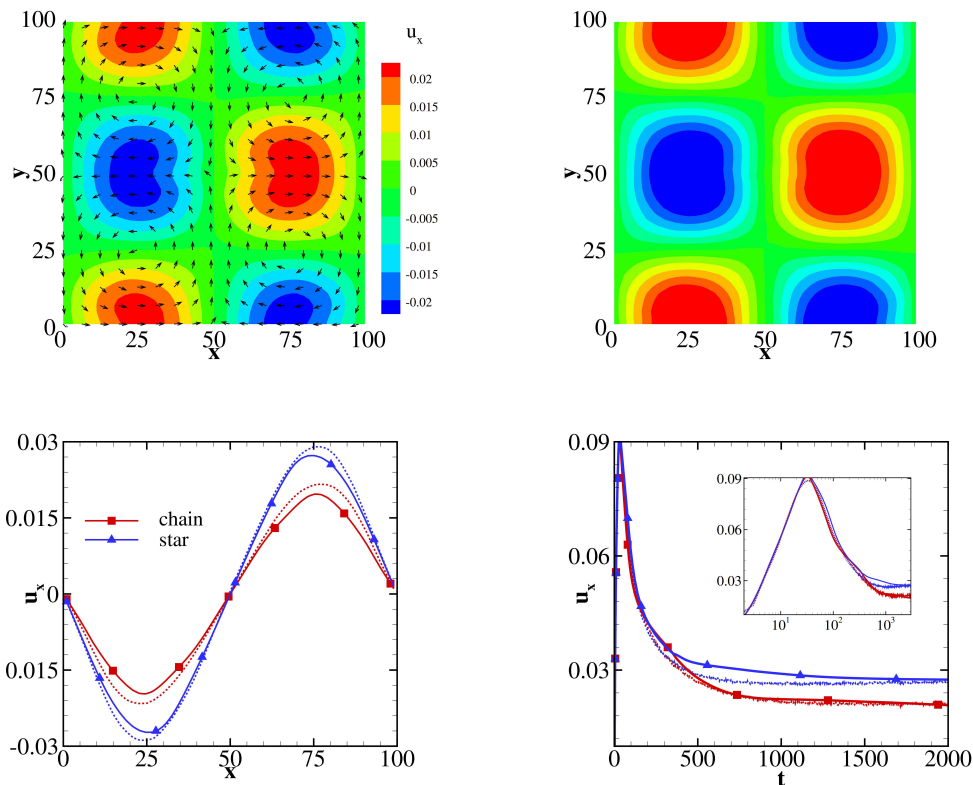


Figure 3.5: The velocity field of the Taylor-Green vortex flow of the chain- and star-shaped molecule suspensions predicted from the micro-scale simulations and DeePN<sup>2</sup>. (a-b) The 2D steady-state velocity field of the chain- and star-shaped molecule suspensions from the micro-scale simulations. The velocity field of the chain-shaped system yields more pronounced deviations from the symmetric Newtonian flow due to the more pronounced polymer stress across the flow regime. (c) The steady-state 1D velocity profile  $u_x(x, y = 49)$ . The solid and dashed lines represent the predictions from the micro-scale simulations and the DeePN<sup>2</sup> model, respectively. (d) The time history of  $u_x(x = 75, y = 49)$ .

steady value than the star-shaped solution. DeePN<sup>2</sup> successfully captures such micro-structure-induced rheological differences and shows good agreement with the micro-scale simulation results.

## 4 Discussion

We have developed a general machine-learning based model, DeePN<sup>2</sup>, for describing the non-Newtonian hydrodynamics for polymer solutions with arbitrary molecular structure and interaction. The constructed model retains a clear physical interpretation and faithfully encodes the micro-scale structural information into the macro-scale hydrodynamics, where conventional models based on empirical closures generally show limitations. In par-

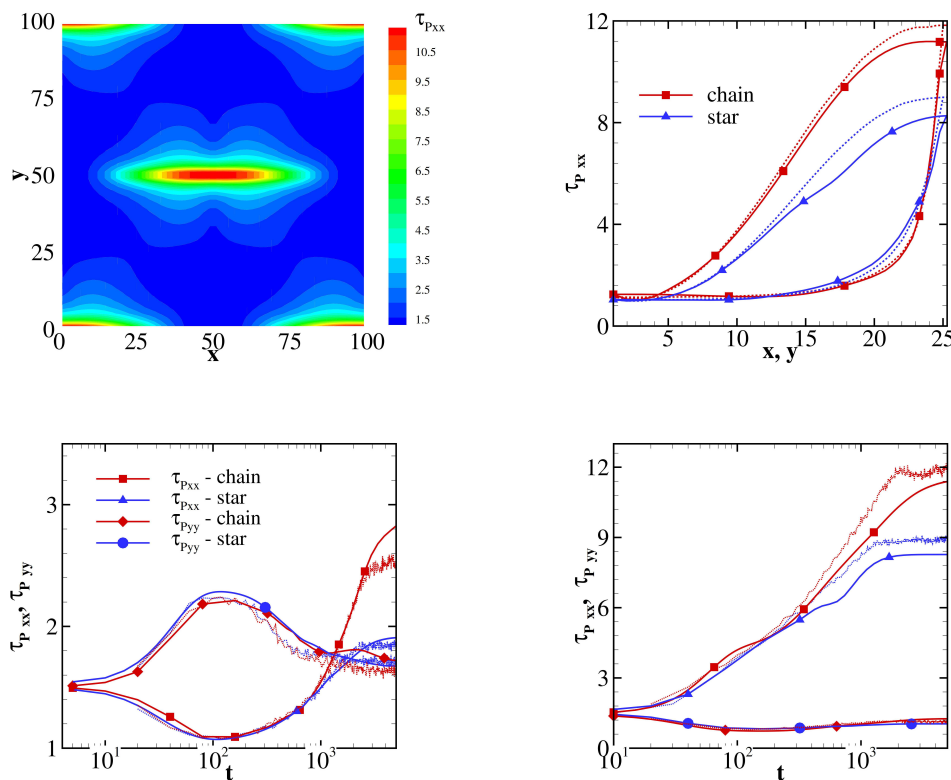


Figure 3.6: The stress field of the Taylor-Green vortex flow of the chain- and star-shaped molecule suspensions predicted from the micro-scale simulations and DeePN<sup>2</sup>. (a) The 2D steady-state stress field of the chain-shaped molecule suspension from the micro-scale simulations. (b) The 1D steady-state stress profiles  $\tau_{P_{xx}}(x, y = 49)$  and  $\tau_{P_{xx}}(x = 49, y)$ . The chain-shaped molecule suspension yields larger stress variations (i.e., the "loop area") along the flow domain. (c-d) The stress evolution of  $\tau_{P_{xx}}(t)$  and  $\tau_{P_{yy}}(t)$  at the points  $(49, 35)$  and  $(49, 49)$ , respectively. The dashed and the solid lines denote the micro-scale simulations and the DeePN<sup>2</sup> predictions, respectively.

ticular, for the chain- and star-shaped molecule suspensions with the same bead number and bond interaction, DeePN<sup>2</sup> successfully captures the different viscoelastic responses arising from the different molecular structural symmetry (i.e., the effective rigidity) in the configuration space without additional human intervention. Unlike the direct evaluation or moment-closure representations of the configurational PDF, the present DeePN<sup>2</sup> model directly learns a set of micro-to-macro mappings to probe the optimal approximations of the constitutive dynamics in terms of the macro-scale features, and thereby circumventing the numerical challenges due to the high-dimensionality of the polymer configuration space. This multi-scaled nature enables us to learn the constitutive dynamics of the macro-scale features directly from the kinetic equations of their micro-scale counterparts using only discrete rather than the time-derivative samples commonly used in the ma-

chine learning-based models of complex dynamic problems.

One thing we have not investigated systematically is the generation of training samples. For DeePN<sup>2</sup> to be truly reliable, the training samples should be representative enough for all the practical situations that one might encounter. However, due to the cost associated with generating such training samples, we would also like the training set to be as small as possible. This calls for an adaptive procedure for generating the training sample, such as the concurrent learning procedure discussed in [54]. The present DeePN<sup>2</sup> models are trained with samples collected from homogeneous shear flow. Even though the numerical predictions show good agreement with micro-scale simulations for a variety of flows, one should not expect this to be generally the case. Further work on sampling is needed to make sure that one can produce truly reliable DeePN<sup>2</sup> models. Furthermore, instead of the general form (2.6), a specific design of the encoders  $\mathbf{b}(\cdot)$  accounting for the molecule symmetry and rigidity may facilitate the extraction of the macro-scale features  $\mathbf{c}$ . In addition, more accurate micro-scale kinetic models accounting for the heterogeneous hydrodynamic interactions [55] and non-Markovianity [56, 57] can be used to construct the macro-scale constitutive dynamics. Finally, the adaptive choice of the number of features and the enhanced sampling of the discrete micro-scale configurations may further improve the performance of the DeePN<sup>2</sup> model. We leave these issues for future work.

## Acknowledgments

We thank Lei Wu and Liyao Lyu for helpful discussions. The work of Fang, Ge, and Lei is supported in part by the Extreme Science and Engineering Discovery Environment (XSEDE) Bridges at the Pittsburgh Supercomputing Center through allocation MTH210005 and the High Performance Computing Center at Michigan State University. Fang acknowledges the support from Shanghai Jiao Tong University during 08/2020–07/2021. The work of E is supported in part by a gift from iFlytek to Princeton University.

## Appendices

### A Rotational frame-indifference of the constitutive dynamics for the multi-bead encoder function

We consider a polymer molecule consisting of  $N$  particles. Let  $\mathbf{r} = [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_{N-1}]$  denote the polymer configuration, so that there exists an invertible linear transformation between  $[\mathbf{r}; \sum_{i=1}^N \mathbf{q}_i/N]$  and  $[\mathbf{q}_1; \mathbf{q}_2; \cdots; \mathbf{q}_N]$ , where  $\mathbf{q}_i$  is the position of the  $i$ -th particle. In fact, there are multiple choices for  $\mathbf{r}$ , including the one we have applied in Eq. (2.5), where  $\mathbf{r}$  consists of  $(N - 1)$  edges of a spanning tree in the bead-bond structure.

We consider a second-order tensor taking the general form

$$\mathbf{b} = \mathbf{f}^{(1)}(\mathbf{r})\mathbf{f}^{(2)}(\mathbf{r})^T, \quad \mathbf{f}^{(1)}(\mathbf{r}) = \sum_{j=1}^{N-1} g_j^{(1)}(\mathbf{r}^*)\mathbf{r}_j, \quad \mathbf{f}^{(2)}(\mathbf{r}) = \sum_{j=1}^{N-1} g_j^{(2)}(\mathbf{r}^*)\mathbf{r}_j, \quad (\text{A.1})$$

where  $\mathbf{r}^*$  is a translational-rotational-invariant vector and  $g^{(1)}$  and  $g^{(2)}$  are two scalar functions. We note that the encoder in the form of Eq. (A.1) is more general than Eq. (2.6).

In this appendix and the next, we consider two frames: frame 1 is static inertial, and frame 2 is rotating with respect to frame 1 with an time dependent orthogonal transformation  $\mathbf{Q}(t)$ . Let  $\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \tilde{\mathbf{b}}$  and  $\mathbf{x}, \mathbf{v}, \mathbf{b}$  denote the positions, velocities, and second-order tensors in frame 1 and 2 respectively. They have the following relations:

$$\tilde{\mathbf{x}} = \mathbf{Q}\mathbf{x}, \quad \tilde{\mathbf{v}} = \mathbf{Q}\mathbf{v} + \dot{\mathbf{Q}}\mathbf{x}, \quad \tilde{\mathbf{b}} = \mathbf{Q}\mathbf{b}\mathbf{Q}^T. \tag{A.2}$$

The material derivatives in both frames are

$$\left. \frac{d}{dt} \right|_{\text{frame 1}} := \frac{\partial}{\partial t} + \tilde{\mathbf{v}} \cdot \nabla_{\tilde{\mathbf{x}}}, \quad \left. \frac{d}{dt} \right|_{\text{frame 2}} := \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}}. \tag{A.3}$$

**Proposition A.1.** *With  $\mathbf{b}$  defined by Eq. (A.1), we have*

$$\begin{aligned} \left. \frac{d}{dt} \mathbf{c} - \boldsymbol{\kappa} : \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b} \right\rangle \right. &= \frac{k_B T}{\gamma} \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b} \right\rangle \\ &\quad - \frac{1}{\gamma} \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V_p(\mathbf{r}) \cdot \nabla_{\mathbf{r}_j} \mathbf{b} \right\rangle, \end{aligned} \tag{A.4}$$

*obeys rotational symmetry.*

*Proof.* Let us choose the vector  $\mathbf{r}^* = [|\mathbf{r}_1|, |\mathbf{r}_2|, |\mathbf{r}_{12}|, |\mathbf{r}_3|, |\mathbf{r}_{13}|, |\mathbf{r}_{23}|, \dots, |\mathbf{r}_{N-2, N-1}|]$ . Denote by  $r_i^*$  the  $i$ -th element of  $\mathbf{r}^*$  and  $\mathbf{r}_i^*$  the corresponding the 3-dimensional vector, i.e.,  $r_6^* = |\mathbf{r}_{23}|$  and  $\mathbf{r}_6^* = \mathbf{r}_{23}$ . Following Eq. (A.1),  $\mathbf{b}$  consists of

$$\mathbf{b} = \sum_{j,k=1}^{N-1} \mathbf{b}_{jk}, \quad \mathbf{b}_{jk} = g(\mathbf{r}^*) \mathbf{r}_j \mathbf{r}_k^T, \tag{A.5}$$

where  $g(\mathbf{r}^*)$  denotes  $g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*)$  for simplicity. With this general form, we have

$$\left. \frac{d}{dt} \langle \tilde{\mathbf{b}}_{jk} \rangle \right|_{\text{frame 1}} = \dot{\mathbf{Q}} \langle \mathbf{b}_{jk} \rangle \mathbf{Q}^T + \mathbf{Q} \langle \mathbf{b}_{jk} \rangle \dot{\mathbf{Q}}^T + \mathbf{Q} \left. \frac{d}{dt} \langle \mathbf{b}_{jk} \rangle \right|_{\text{frame 2}} \mathbf{Q}^T. \tag{A.6}$$

Moreover, we note that

$$\begin{aligned} \tilde{\boldsymbol{\kappa}} : \left( \sum_{i=1}^{N-1} \tilde{\mathbf{r}}_i \otimes \nabla_{\tilde{\mathbf{r}}_i} \otimes \tilde{\mathbf{b}}_{jk} \right) &= \sum_{i=1}^{N-1} \left[ \left( \mathbf{Q} \boldsymbol{\kappa} \mathbf{Q}^T + \dot{\mathbf{Q}} \mathbf{Q}^T \right) \cdot \mathbf{Q} \mathbf{r}_j \right] \cdot \mathbf{Q} \otimes \nabla_{\mathbf{r}_i} \otimes \left( \mathbf{Q} \mathbf{b}_{jk} \mathbf{Q}^T \right) \\ &= \sum_{i=1}^{N-1} \left( \boldsymbol{\kappa} \cdot \mathbf{r}_i \right) \cdot \nabla_{\mathbf{r}_i} \left( \mathbf{Q} \mathbf{b}_{jk} \mathbf{Q}^T \right) + \left( \mathbf{Q}^T \dot{\mathbf{Q}} \mathbf{r}_i \right) \cdot \nabla_{\mathbf{r}_i} \left( \mathbf{Q} \mathbf{b}_{jk} \mathbf{Q}^T \right) \end{aligned}$$



$$\begin{aligned}
 &= \sum_{i=1}^{N-1} \mathbf{Q}(\boldsymbol{\kappa} \cdot \mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} \mathbf{b}_{jk} \mathbf{Q}^T + \mathbf{Q} \left( \mathbf{Q}^T \dot{\mathbf{Q}} \mathbf{b}_{jk} + \mathbf{b}_{jk} \dot{\mathbf{Q}}^T \mathbf{Q} \right) \mathbf{Q}^T \\
 &\quad + \mathbf{Q} \left( \sum_{i=1}^{N-1} \mathbf{r}_i^T (\dot{\mathbf{Q}}^T \mathbf{Q}) \nabla_{\mathbf{r}_i} g(\mathbf{r}^*) \right) \mathbf{r}_j^T \mathbf{r}_k^T \mathbf{Q}^T \\
 &= \sum_{i=1}^{N-1} \mathbf{Q}(\boldsymbol{\kappa} \cdot \mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} \mathbf{b}_{jk} \mathbf{Q}^T + \dot{\mathbf{Q}} \mathbf{b}_{jk} \mathbf{Q}^T + \mathbf{Q} \mathbf{b}_{jk} \dot{\mathbf{Q}}^T, \tag{A.7}
 \end{aligned}$$

where we have used  $\mathbf{r}_i^T (\dot{\mathbf{Q}}^T \mathbf{Q}) \mathbf{r}_i \equiv 0$  since  $\mathbf{Q}^T \dot{\mathbf{Q}}$  is anti-symmetric. Eq. (A.6) and Eq. (A.7) shows that the combination of the two terms on the left-hand-side of Eq. (A.4) rigorously preserve the rotational symmetry, i.e.,

$$\begin{aligned}
 &\left( \frac{d}{dt} \langle \tilde{\mathbf{b}} \rangle - \tilde{\boldsymbol{\kappa}} : \sum_{i=1}^{N-1} \langle \tilde{\mathbf{r}}_i \otimes \nabla_{\tilde{\mathbf{r}}_i} \otimes \tilde{\mathbf{b}} \rangle \right) \Big|_{\text{frame 1}} \\
 &\equiv \mathbf{Q} \left( \frac{d}{dt} \langle \mathbf{b} \rangle - \boldsymbol{\kappa} : \sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle \right) \Big|_{\text{frame 2}} \mathbf{Q}^T. \tag{A.8}
 \end{aligned}$$

It is straightforward to prove rotational symmetry for the other terms in Eq. (A.4). □

## B Symmetry-preserving neural network representation of the objective tensor derivatives

**Proposition B.1.** *The following ansatz of  $\langle \sum_{i=1}^{N-1} \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle$  ensures that the dynamic of evolution of  $\mathbf{c}$  retains rotational invariance.*

$$\begin{aligned}
 \sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle &= \sum_{j,k=1}^{N-1} \langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) (\mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} + \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k}) \otimes \mathbf{r}_j \mathbf{r}_k^T \rangle \\
 &\quad + \sum_{k=1}^9 \mathbf{E}_1^{(k)}(\mathbf{c}) \otimes \mathbf{E}_2^{(k)}(\mathbf{c}), \tag{B.1}
 \end{aligned}$$

where  $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ ,  $\tilde{\mathbf{c}} = (\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_n)$ , and  $\mathbf{E}_1$  and  $\mathbf{E}_2$  satisfy

$$\tilde{\mathbf{E}}_1 := \mathbf{E}_1(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{E}_1(\mathbf{c}) \mathbf{Q}^T, \quad \tilde{\mathbf{E}}_2 := \mathbf{E}_2(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{E}_2(\mathbf{c}) \mathbf{Q}^T. \tag{B.2}$$

*Proof.* Without loss of generality, we represent the fourth order tensor by the following two bases

$$\begin{aligned}
 &\mathbf{F}_1(\mathbf{c}) \otimes \mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_3(\mathbf{c}) + \mathbf{F}_3(\mathbf{c}) \otimes (\mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_1(\mathbf{c}))^{T_{\{2,3\}}}, \quad \mathbf{F}_1(\mathbf{c}), \mathbf{F}_3(\mathbf{c}) \in \mathbb{R}^3, \mathbf{F}_2(\mathbf{c}) \in \mathbb{R}^{3 \times 3}, \\
 &\mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c}), \quad \mathbf{E}_1(\mathbf{c}), \mathbf{E}_2(\mathbf{c}) \in \mathbb{R}^{3 \times 3}, \tag{B.3}
 \end{aligned}$$

where the super-script  $T_{\{2,3\}}$  represents the transpose between the 2nd and 3rd indices; also  $\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3, \mathbf{E}_1$  and  $\mathbf{E}_2$  satisfy the symmetry conditions

$$\begin{aligned} \mathbf{F}_1(\tilde{\mathbf{c}}) &= \mathbf{Q}\mathbf{F}_1(\mathbf{c}), \quad \mathbf{F}_3(\tilde{\mathbf{c}}) = \mathbf{Q}\mathbf{F}_3(\mathbf{c}), \\ \mathbf{E}_1(\tilde{\mathbf{c}}) &= \mathbf{Q}\mathbf{E}_1(\mathbf{c})\mathbf{Q}^T, \quad \mathbf{E}_2(\tilde{\mathbf{c}}) = \mathbf{Q}\mathbf{E}_2(\mathbf{c})\mathbf{Q}^T, \quad \mathbf{F}_2(\tilde{\mathbf{c}}) = \mathbf{Q}\mathbf{F}_2(\mathbf{c})\mathbf{Q}^T. \end{aligned} \tag{B.4}$$

For the term  $\mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c})$ , we have

$$\boldsymbol{\kappa} : \mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c}) = \text{Tr}(\boldsymbol{\kappa}\mathbf{E}_1(\mathbf{c}))\mathbf{E}_2(\mathbf{c}) \tag{B.5}$$

and

$$\begin{aligned} \tilde{\boldsymbol{\kappa}} : \tilde{\mathbf{E}}_1 \otimes \tilde{\mathbf{E}}_2 \Big|_{\text{frame 1}} &= \left( \mathbf{Q}\boldsymbol{\kappa}\mathbf{Q}^T + \dot{\mathbf{Q}}\mathbf{Q}^T \right) : \left( \mathbf{Q}\mathbf{E}_1(\mathbf{c})\mathbf{Q}^T \otimes \tilde{\mathbf{E}}_2 \right) \\ &= \text{Tr}(\boldsymbol{\kappa}\mathbf{E}_1(\mathbf{c}))\tilde{\mathbf{E}}_2 + \text{Tr}(\dot{\mathbf{Q}}\mathbf{Q}^T\mathbf{Q}\mathbf{E}_1(\mathbf{c})\mathbf{Q}^T)\tilde{\mathbf{E}}_2 \\ &= \text{Tr}(\boldsymbol{\kappa}\mathbf{E}_1(\mathbf{c}))\tilde{\mathbf{E}}_2 \\ &\equiv \mathbf{Q} \left( \boldsymbol{\kappa} : \mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c}) \Big|_{\text{frame 2}} \right) \mathbf{Q}^T, \end{aligned} \tag{B.6}$$

where we have used  $\text{Tr}(\dot{\mathbf{Q}}\mathbf{Q}^T) \equiv 0$ .

For the term  $\mathbf{F}_1(\mathbf{c}) \otimes \mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_3(\mathbf{c}) + \mathbf{F}_3(\mathbf{c}) \otimes (\mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_1(\mathbf{c}))^{T_{\{2,3\}}}$ , we have

$$\boldsymbol{\kappa} : \mathbf{F}_1(\mathbf{c}) \otimes \mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_3(\mathbf{c}) = \mathbf{F}_2(\mathbf{c})^T \boldsymbol{\kappa} \mathbf{F}_1(\mathbf{c}) \mathbf{F}_3(\mathbf{c})^T \tag{B.7}$$

and

$$\tilde{\boldsymbol{\kappa}} : \tilde{\mathbf{F}}_1 \otimes \tilde{\mathbf{F}}_2 \otimes \tilde{\mathbf{F}}_3 = \mathbf{Q}\mathbf{F}_2(\mathbf{c})^T \boldsymbol{\kappa} \mathbf{F}_1(\mathbf{c}) \mathbf{F}_3(\mathbf{c})^T \mathbf{Q}^T + \mathbf{Q}\mathbf{F}_2(\mathbf{c})^T \mathbf{Q}^T \dot{\mathbf{Q}} \mathbf{F}_1(\mathbf{c}) \mathbf{F}_3(\mathbf{c})^T \mathbf{Q}^T. \tag{B.8}$$

On the other hand, we note that

$$\frac{d\tilde{\mathbf{b}}}{dt} \Big|_{\text{frame 1}} = \dot{\mathbf{Q}}\mathbf{b}\mathbf{Q}^T + \mathbf{Q}\mathbf{b}\dot{\mathbf{Q}}^T + \mathbf{Q} \frac{d\mathbf{b}}{dt} \Big|_{\text{frame 2}} \mathbf{Q}^T. \tag{B.9}$$

To ensure the rotational symmetry of  $\frac{D\mathbf{b}}{Dt}$ , we have

$$\mathbf{F}_2 \equiv \mathbf{I}, \quad \sum_i \mathbf{F}_1^{(i)} \otimes \mathbf{I} \otimes \mathbf{F}_3^{(i)} = \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) \mathbf{r}_j \otimes \mathbf{I} \otimes \mathbf{r}_k \right\rangle. \tag{B.10}$$

Hence, we have

$$\begin{aligned} \frac{d}{dt} \tilde{\mathbf{c}} - \tilde{\boldsymbol{\kappa}} : \left( \sum_i \tilde{\mathbf{F}}_1^{(i)} \otimes \tilde{\mathbf{F}}_2^{(i)} \otimes \tilde{\mathbf{F}}_3^{(i)} + \tilde{\mathbf{F}}_3^{(i)} \otimes \left( \tilde{\mathbf{F}}_2^{(i)} \otimes \tilde{\mathbf{F}}_1^{(i)} \right)^{T_{\{2,3\}}} \right) \Big|_{\text{frame 1}} \\ \equiv \mathbf{Q} \left( \frac{d}{dt} \mathbf{c} - \boldsymbol{\kappa} : \left( \sum_i \mathbf{F}_1^{(i)} \otimes \mathbf{F}_2^{(i)} \otimes \mathbf{F}_3^{(i)} + \mathbf{F}_3^{(i)} \otimes \left( \mathbf{F}_2^{(i)} \otimes \mathbf{F}_1^{(i)} \right)^{T_{\{2,3\}}} \right) \right) \Big|_{\text{frame 2}} \mathbf{Q}^T. \end{aligned} \tag{B.11}$$

Furthermore, using Eq. (B.10), we obtain

$$\begin{aligned} & \sum_i \mathbf{F}_1^{(i)} \otimes \mathbf{F}_2^{(i)} \otimes \mathbf{F}_3^{(i)} + \mathbf{F}_3^{(i)} \otimes (\mathbf{F}_2^{(i)} \otimes \mathbf{F}_1^{(i)})^{T_{\{2,3\}}} \\ &= \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) (\mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} + \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k}) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle. \end{aligned} \tag{B.12}$$

Accordingly, the remaining part of  $\sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle$  is expanded by

$$\left\langle \sum_{i=1}^{N-1} \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \sum_{j,k=1}^{N-1} g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle = \sum_{i=1}^9 \mathbf{E}_1^{(i)}(\mathbf{c}) \otimes \mathbf{E}_2^{(i)}(\mathbf{c}). \tag{B.13}$$

Combining Eqs. (B.11), (B.12) and (B.13), we conclude that the decomposition

$$\begin{aligned} \sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle &= \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) (\mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} + \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k}) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle \\ &\quad + \sum_{k=1}^9 \mathbf{E}_1^{(k)}(\mathbf{c}) \otimes \mathbf{E}_2^{(k)}(\mathbf{c}) \end{aligned} \tag{B.14}$$

ensures the objectivity of the time-derivative of  $\mathbf{c}$ . □

## C Symmetry-preserving neural network representation of the second-order tensor

In the DeePN<sup>2</sup> model, we construct the NN representations of the second-order tensors for the stress  $\mathbf{G}$ , constitutive terms  $\mathbf{H}_1, \mathbf{H}_2$ , and objective tensor derivative terms  $\mathbf{E}_1$  and  $\mathbf{E}_2$  that satisfy the rotational symmetry conditions, i.e.,

$$\mathbf{G}(\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_n) = \mathbf{Q} \mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) \mathbf{Q}^T, \tag{C.1}$$

where  $\tilde{\mathbf{c}}_i = \mathbf{Q} \mathbf{c}_i \mathbf{Q}^T$  and  $\mathbf{Q}$  is an orthogonal matrix.

To preserve the rotational symmetry condition (C.1), we fix the form of encoder  $\mathbf{b}_1$  and transfer the learning to the eigen-space of  $\mathbf{c}_1$ . Let us assume that the eigen-decomposition  $\mathbf{c}_1 = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$  has distinct eigenvalues, where  $\mathbf{U}$  is the matrix whose columns are the eigenvectors of  $\mathbf{c}_1$ .  $\mathbf{U}$  is not unique due to the non-uniqueness of the eigenvectors. Without loss of generality, we further assume that the first element of  $\mathbf{u}_1$  to be positive. With the following lemma, we show that the general form of  $\mathbf{U}$  can be always written as  $\mathbf{U}^{(j)} := \mathbf{U} \mathbf{S}^{(j)}$  with  $j = 1, \dots, 4$ , where  $\mathbf{S}^{(j)}$  is given by

$$\mathbf{S}^{(1)} = \begin{pmatrix} +1 & & \\ & +1 & \\ & & +1 \end{pmatrix}, \quad \mathbf{S}^{(2)} = \begin{pmatrix} +1 & & \\ & -1 & \\ & & +1 \end{pmatrix},$$

$$S^{(3)} = \begin{pmatrix} +1 & & \\ & +1 & \\ & & -1 \end{pmatrix}, \quad S^{(4)} = \begin{pmatrix} +1 & & \\ & -1 & \\ & & -1 \end{pmatrix}.$$

**Lemma C.1.** For a symmetry matrix  $M \in \mathbb{R}^{3 \times 3}$ , let  $S_M$  denote the set of matrices with the transformation of  $S^{(j)}$ , i.e.,  $S_M := \{S^{(1)}MS^{(1)}, \dots, S^{(4)}MS^{(4)}\}$ . For any  $M^{(j)} := S^{(j)}MS^{(j)} \in S_M$ ,  $S^{(k)}M^{(j)}S^{(k)} \in S_M$ ,  $1 \leq j, k \leq 4$ . Furthermore,  $S_M$  can be constructed by  $M^{(j)}$ , i.e.,  $S_M \equiv \{S^{(1)}M^{(j)}S^{(1)}, \dots, S^{(4)}M^{(j)}S^{(4)}\}$ .

*Proof.* By applying  $S^{(j)}$  to  $M$ , it is easy to see that the diagonal part of  $M^{(j)}$  remains the same. Since  $M^{(j)}$  is also symmetric, we only need to check the upper-triangular part, taking the four possible operations

$$\begin{pmatrix} * & + & + \\ & * & + \\ & & * \end{pmatrix} \quad \begin{pmatrix} * & - & + \\ & * & - \\ & & * \end{pmatrix} \quad \begin{pmatrix} * & + & - \\ & * & - \\ & & * \end{pmatrix} \quad \begin{pmatrix} * & - & + \\ & * & - \\ & & * \end{pmatrix},$$

where “+” represents that the element remains the same and “-” represents a sign change. We see that number of “-” operations is either 0 or 2. Starting from any of the above choice for  $M^{(j)}$ , all of the four operators yields either 0 or 2 “-” operations. Therefore,  $S^{(k)}M^{(j)}S^{(k)} \in S_M$ . Furthermore, if the upper triangular part of  $M$  has distinct absolute values, then  $\forall M^{(j)}$ ,  $S^k M^j S^k \neq S^{k'} M^j S^{k'}$  with  $k \neq k'$ , hence  $S_M$  can be constructed by  $M^j$ . Otherwise, if some upper triangular entries of  $M$  share the same absolute value, we can draw the same conclusion accordingly.  $\square$

Now we consider the matrix whose columns are the eigenvectors of  $\tilde{\mathbf{c}}_1 = \mathbf{Q}\mathbf{c}_1\mathbf{Q}^T$ , denoted by  $\tilde{\mathbf{U}}$ . We can write  $\tilde{\mathbf{U}} = \mathbf{Q}\mathbf{U}S^{(j)}$ , where  $j \in \{1, 2, 3, 4\}$ . Accordingly, the DNN input of  $\mathbf{c}_i$  takes the form

$$\tilde{\mathbf{U}}^T \tilde{\mathbf{c}}_i \tilde{\mathbf{U}} = \left(\mathbf{Q}\mathbf{U}S^{(j)}\right)^T \mathbf{Q}\mathbf{c}_i\mathbf{Q}^T \left(\mathbf{Q}\mathbf{U}S^{(j)}\right) = S^{(j)}\mathbf{U}^T \mathbf{c}_i \mathbf{U}S^{(j)}.$$

Let  $M = \mathbf{U}^T \mathbf{c}_i \mathbf{U}$ , by using Lemma C.1, it is easy to see that  $S_{\mathbf{U}^T \mathbf{c}_i \mathbf{U}}$  can be constructed by taking  $j = 1, \dots, 4$ .

**Proposition C.1.** Let  $\mathbf{U}$  be the matrix whose columns are the eigenvectors of  $\mathbf{c}_1$ . Let the DNN input be  $\hat{\mathbf{c}}_i^{(j)} = S^{(j)}\mathbf{U}^T \mathbf{c}_i \mathbf{U}S^{(j)}$ . The following form of  $\tau_p$

$$\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \frac{1}{4} \sum_{j=1}^4 \mathbf{U}^{(j)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j)}, \dots, \hat{\mathbf{c}}_n^{(j)}) \mathbf{U}^{(j)T}, \quad \mathbf{U}^{(j)} = \mathbf{U}S^{(j)}. \tag{C.2}$$

satisfies the rotational symmetry constraint (C.1).

Finally, to account for the swap of the eigenvectors when the eigenvalues cross over, we consider the 6 permutations of the three eigenvalues of  $\mathbf{c}_1$ , i.e.,

$$\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \frac{1}{24} \sum_{k=0}^5 \sum_{j=1}^4 \mathbf{U}^{(j,k)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j,k)}, \dots, \hat{\mathbf{c}}_n^{(j,k)}) \mathbf{U}^{(j,k)T}, \tag{C.3}$$

where  $k$  represents the rank of permutation (e.g., in lexicographical order) and  $\mathbf{U}^{(j,k)}$  is a variation of  $\mathbf{U}^{(j)}$  with corresponding column permutation.

## D Validation of the rotational-symmetry preserving NN representation

To validate the performance of the proposed DNN representation, we check the accuracy of the modeling terms given a set of conformation tensors  $\mathbf{c}_1, \dots, \mathbf{c}_n$  under different unitary transformations. Fig. D.1 shows the relative error under each transformation. The DNN representation (C.2) yields the same results under all the transformation. In contrast, the DNN without accounting for the four transformations yields significant error due to the non-uniqueness of the eigenvectors of  $\mathbf{c}_1$ .

In addition, we examine the 2D Taylor-Green vortex flow where the evolution of  $\mathbf{c}_1$  becomes degenerate at certain points. Fig. D.2 shows the stress evolution at (45, 37). At  $t = 1080$ , the eigenvalues  $\lambda_2$  and  $\lambda_3$  cross over. Concurrently, the prediction of the polymer stress  $\tau_p$  from the model without considering the swap of  $\mathbf{u}_2$  and  $\mathbf{u}_3$  shows apparent deviations near the regime as shown in Fig. D.2. In contrast, the prediction from the model retaining the eigenvalue permutation trained by Eq. (2.11) shows good agreement with the MD results.

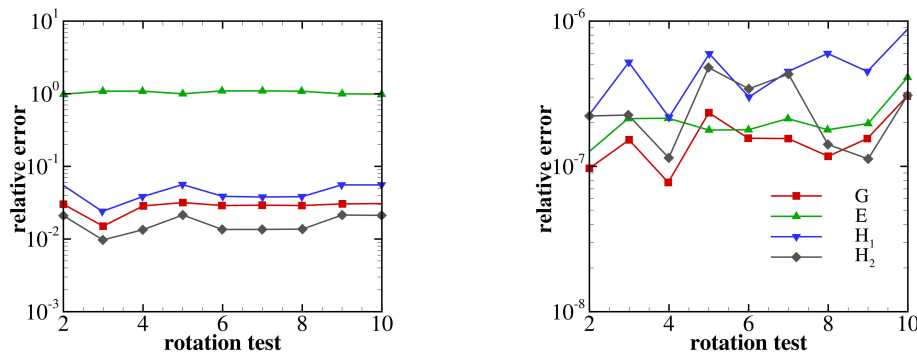


Figure D.1: The relative  $l_\infty$  error of the model prediction under randomly chosen orthogonal transformations without (left) and with (right) accounting for the four eigen-space transformations in Eq. (C.2).

## E Micro-scale model of the polymer solutions

In the present study, we consider suspensions with three different polymer structures as shown in Fig. 2.1. Each polymer molecule consists of  $N = 7$  beads connected with  $N_b$

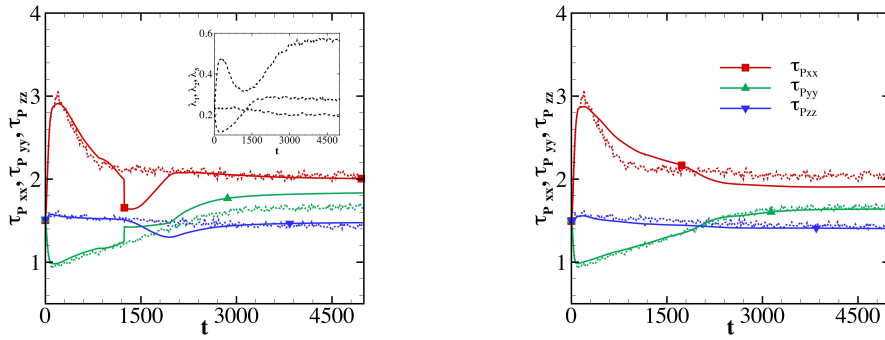


Figure D.2: Stress evolution of the Taylor-Green vortex flow at position (45,37) of the chain-shaped molecule suspension. Left: prediction without considering the swap of eigenvectors when the two eigenvalues approaches near  $t = 1255$  as shown in the inset plot. Right: predictions from the model retaining the eigenvalue permutation trained by Eq. (2.11). The dashed and the solid lines denote the micro-scale simulations and the DeePN<sup>2</sup> predictions, respectively.

FENE bonds, i.e.,

$$V(\mathbf{q}) = \sum_{j=1}^{N_b} V_b(|\mathbf{q}_{j_1} - \mathbf{q}_{j_2}|), \quad V_b(l) = -\frac{k_s}{2} l_0^2 \log \left[ 1 - \frac{l^2}{l_0^2} \right], \quad (\text{E.1})$$

where  $k_s$  represents the spring constant and  $l_0$  is the maximum of the extension length. The chain- and star-shaped molecules have  $N_b = 6$  bonds with the same bond parameters  $k_s = 0.1$  and  $l_0 = 2.3$  (in reduced unit). The net-shaped molecule is similar to the star-shaped molecule with the same parameters for the first 6 bonds; 3 additional bonds connect the side chain particles with  $k_s = 0.1$  and  $l_0 = 3.7$ . The polymer number density of the three suspensions is  $n_p = 0.3$ . The solvent is modeled by the dissipative particle dynamics (DPD) [58, 59] with number density  $n_s = 4.0$ . The pairwise interaction between particle  $i$  and  $j$  takes the standard form

$$\begin{aligned} \mathbf{F}_{ij} &= \mathbf{F}_{ij}^C + \mathbf{F}_{ij}^D + \mathbf{F}_{ij}^R, & \mathbf{F}_{ij}^C &= \begin{cases} a(1.0 - r_{ij}/r_c)\mathbf{e}_{ij}, & r_{ij} < r_c, \\ 0, & r_{ij} > r_c, \end{cases} \\ \mathbf{F}_{ij}^D &= \begin{cases} -\gamma\omega^D(r_{ij})(\mathbf{v}_{ij} \cdot \mathbf{e}_{ij})\mathbf{e}_{ij}, & r_{ij} < r_c, \\ 0, & r_{ij} > r_c, \end{cases} & \mathbf{F}_{ij}^R &= \begin{cases} \sigma\omega^R(r_{ij})\xi_{ij}\mathbf{e}_{ij}, & r_{ij} < r_c, \\ 0, & r_{ij} > r_c, \end{cases} \end{aligned}$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ ,  $r_{ij} = |\mathbf{r}_{ij}|$ ,  $\mathbf{e}_{ij} = \mathbf{r}_{ij}/r_{ij}$ , and  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ ,  $\xi_{ij}$  are independent identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance.  $\gamma$  and  $\sigma$  are related with the system temperature by the second fluctuation-dissipation theorem [60] as  $\sigma^2 = 2\gamma k_B T$ , where  $k_B T$  is set to 0.25. The detailed parameters are given in Table E.1.

Table E.1: Parameters (in reduced unit) of the micro-scale model of the polymer solution (S-solvent, P-polymer).

	$a$	$\gamma$	$\sigma$	$k$	$r_c$
S-S	4.0	5.0	1.58	0.25	1.0
S-P	0.0	40.0	4.47	0.0	1.0
P-P	4.0	0.01	0.071	1.0	0.7

## F Collecting training samples

Collecting training samples is one of the most important steps in the construction of DeePN<sup>2</sup>. To obtain reliable models, we need to ensure that the training sample set is representative enough of all the practical situations that the model is intended for. In the present study, we collect the training samples in shear flow with shear rate  $\dot{\gamma} \in [0, 0.09]$ . Since the training of the DeePN<sup>2</sup> model only requires discrete polymer configurations rather than time-series samples, one convenient approach is to consecutively increase the shear rate and collect the discrete configurations during the shear extension and relaxation process, where the inclusion of the relaxation process can facilitate the sampling of polymer configuration phase space due to the viscoelastic hysteresis effect. 32000 samples are collected where each sample consists of 5000 polymer configurations, which will be employed to evaluate the constitutive dynamics terms  $\langle \cdot \rangle$ . Due to the permutation symmetry of the the particle label, the effective number of configurations per sample is  $1 \times 10^4$  for the chain-shaped molecule and  $3 \times 10^4$  for the star- and net-shaped molecules.

## G Training procedure

The DeePN<sup>2</sup> model is constructed via the training of the NN representations of the encoder mappings  $\{g_j(\mathbf{r}^*)\}_{j=1}^n$ , stress model  $\mathbf{G}$ , evolution dynamics  $\{\mathbf{H}_{1,j}\}_{j=1}^n$ ,  $\{\mathbf{H}_{2,j}\}_{j=1}^n$  and the 4th order tensors  $\{\mathcal{E}_j\}_{j=1}^n$  of the objective tensor derivatives. In this study, we choose  $n = 3$  encoders and fix  $g_1(\mathbf{r}^*) \equiv 1$ . For the chain-shaped molecule, we set  $w_{1,i} = 1 - i/N$ ,  $1 \leq i \leq N - 1$  and  $\sum_i w_{1,i} \mathbf{r}_i$  represents the orientation between the free-end particle and the center of mass. For the star- and net-shaped molecules, we set  $w_{1,1} = 1$  and  $w_{1,i} = 0$  for  $i \geq 2$ . All terms are represented by the fully connected NN. The number of hidden layers are set to be (120, 120, 120), (300, 300, 300), (400, 400, 400), (450, 450, 450), (560, 560, 560), respectively. The activation function is taken to be the hyperbolic tangent. We emphasize that the mappings  $\{g_j(\mathbf{r}^*)\}_{j=1}^n$  and weights  $w \in \mathbb{R}^{n \times (N-1)}$  involve in the training process for the joint learning of the encoders  $\{\mathbf{b}_j(\mathbf{r})\}_{j=1}^n$  defined in Eq. (2.6) and the macro-scale features  $\{\mathbf{c}_j\}_{j=1}^n$ , although they do not appear explicitly in the macro-scale hydrodynamic equations.

The DNNs are trained by the Adam stochastic gradient descent method [61] for 20 epochs, using 5 samples per batch size. The initial learning rate is  $2.8 \times 10^{-4}$  and decay rate is 0.75 per 20000 steps.

Similar to Ref. [32], the loss function is defined by

$$L = \lambda_G L_G + \lambda_{H_1} L_{H_1} + \lambda_{H_2} L_{H_2} + \lambda_{\mathcal{E}} L_{\mathcal{E}},$$

where  $\lambda_G = 0.2$ ,  $\lambda_{H_1} = 0.1$ ,  $\lambda_{H_2} = 0.6$  and  $\lambda_{\mathcal{E}} = 0.1$  are hyperparameters. For each training batch of  $m$  training samples,  $L_G, L_{H_1}, L_{H_2}, L_{\mathcal{E}}$  of the system are given by

$$\begin{aligned} L_G &= \sum_{l=1}^m \sum_{i=1}^n \left\| \mathbf{G}_i(\mathbf{c}^{(l)}) - \left\langle \sum_{k=1}^{N_b} \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k} V \right\rangle^{(l)} \right\|^2, \\ L_{H_1} &= \sum_{l=1}^m \sum_{i=1}^n \left\| \mathbf{H}_{1,i}(\mathbf{c}^{(l)}) - \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b}_i \right\rangle^{(l)} \right\|^2, \\ L_{H_2} &= \sum_{l=1}^m \sum_{i=1}^n \left\| \mathbf{H}_{2,i}(\mathbf{c}^{(l)}) - \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V \cdot \nabla_{\mathbf{r}_j} \mathbf{b}_i \right\rangle^{(l)} \right\|^2, \\ L_{\mathcal{E}} &= \sum_{l=1}^m \sum_{i=1}^n \left\| \sum_{s=1}^9 \mathbf{E}_{1,i}^{(s)}(\mathbf{c}^{(l)}) \otimes \mathbf{E}_{2,i}^{(s)}(\mathbf{c}^{(l)}) - \left\langle \sum_{k=1}^{N-1} \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k} g_i^2 \otimes \sum_{j,j'=1}^{N-1} w_{ij} w_{ij'} \mathbf{r}_j \mathbf{r}_{j'}^T \right\rangle^{(l)} \right\|^2, \end{aligned} \tag{G.1}$$

where  $\| \cdot \|^2$  denotes the total sum of squares of the entries in the tensor, and  $\mathbf{c}^{(l)} = (\mathbf{c}_1^{(l)}, \dots, \mathbf{c}_n^{(l)})$ .

## References

- [1] Ronald G. Larson. *Constitutive Equations for Polymer Melts and Solutions*. Butterworth-Heinemann Press, 1988.
- [2] Robert G Owens and Timothy N Phillips. *Computational Rheology*. Imperial College Press, 2002.
- [3] James G Oldroyd and Alan Herries Wilson. On the formulation of rheological equations of state. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 200(1063):523–541, 1950.
- [4] Fang-Hua Lin, Chun Liu, and Ping Zhang. On hydrodynamics of viscoelastic fluids. *Communications on Pure and Applied Mathematics*, 58(11):1437–1471, 2005.
- [5] Anton Peterlin. Hydrodynamics of macromolecules in a velocity field with longitudinal gradient. *Journal of Polymer Science Part B: Polymer Letters*, 4(4):287–291, 1966.
- [6] Robert Byron Bird, Paul J Dotson, and N.L. Johnson. Polymer solution rheology based on a finitely extensible bead–spring chain model. *Journal of Non-Newtonian Fluid Mechanics*, 7(2):213–235, 1980.
- [7] Hanswalter Giesekus. A simple constitutive equation for polymer fluids based on the concept of deformation-dependent tensorial mobility. *Journal of Non-Newtonian Fluid Mechanics*, 11(1):69–109, 1982.
- [8] Nhan Phan Thien and Roger I. Tanner. A new constitutive equation derived from network theory. *Journal of Non-Newtonian Fluid Mechanics*, 2(4):353–365, 1977.
- [9] Amine Ammar. Lattice Boltzmann method for polymer kinetic theory. *Journal of non-Newtonian fluid mechanics*, 165(19-20):1082–1092, 2010.



- [10] Xijun Fan. Molecular models and flow calculations: II. Simulation of steady planar flow. *Acta Mechanica Sinica*, 1:49, 1989.
- [11] Alexei Lozinski and Cédric Chauvière. A fast solver for Fokker-Planck equation applied to viscoelastic flows calculations: 2D FENE model. *Journal of Computational Physics*, 189(2):607–625, 2003.
- [12] Cédric Chauvière and Alexei Lozinski. Simulation of dilute polymer solutions using a Fokker-Planck equation. *Computers & Fluids*, 33(5-6):687–696, 2004.
- [13] Jie Shen and Haijun Yu. On the approximation of the Fokker-Planck equation of the finitely extensible nonlinear elastic dumbbell model I: A new weighted formulation and an optimal spectral-Galerkin algorithm in two dimensions. *SIAM Journal on Numerical Analysis*, 50(3):1136–1161, 2012.
- [14] José Antonio Carrillo, Katy Craig, and Francesco S Patacchini. A blob method for diffusion. *Calculus of Variations and Partial Differential Equations*, 58(2):1–53, 2019.
- [15] Pierre Degond and Francisco-José Mustieles. A deterministic approximation of diffusion equations using particles. *SIAM Journal on Scientific and Statistical Computing*, 11(2):293–310, 1990.
- [16] G Lacombe and S Mas-Gallic. Presentation and analysis of a diffusion-velocity method. In *ESAIM: Proceedings*, volume 7, pages 225–233. EDP Sciences, 1999.
- [17] Yiwei Wang, Jiuhai Chen, Chun Liu, and Lulu Kang. Particle-based energetic variational inference. *Statistics and Computing*, 31(3):1–17, 2021.
- [18] Xuelian Bao, Chun Liu, and Yiwei Wang. On a deterministic particle-FEM discretization to micro-macro models of dilute polymeric fluids. *arXiv preprint arXiv:2112.10970*, 2021.
- [19] Manuel Laso and Hans Christian Öttinger. Calculation of viscoelastic flow using molecular models: The connffessit approach. *Journal of Non-Newtonian Fluid Mechanics*, 47:1–20, 1993.
- [20] M.A Hulsen, A.P.G van Heel, and B.H.A.A van den Brule. Simulation of viscoelastic flows using brownian configuration fields. *Journal of Non-Newtonian Fluid Mechanics*, 70(1):79–101, 1997.
- [21] Weiqing Ren and Weinan E. Heterogeneous multiscale method for the modeling of complex fluids and micro-fluidics. *Journal of Computational Physics*, 204(1):1–26, 2005.
- [22] Harold R. Warner. Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. *Industrial & Engineering Chemistry Fundamentals*, 11(3):379–387, 1972.
- [23] Harold R. Warner. PhD thesis, University of Wisconsin, 1971.
- [24] Lifei Zhao, Zhen Li, Bruce Caswell, Jie Ouyang, and George Em Karniadakis. Active learning of constitutive relation from mesoscopic dynamics for macroscopic modeling of non-Newtonian flows. *Journal of Computational Physics*, 363:116–127, 2018.
- [25] Massimiliano Grosso, PL Maffettone, P Halin, Roland Keunings, and Vincent Legat. Flow of nematic polymers in eccentric cylinder geometry: Influence of closure approximations. *Journal of Non-Newtonian Fluid Mechanics*, 94(2):119–134, 2000.
- [26] J. Feng, C. V. Chaubal, and L. G. Leal. Closure approximations for the doi theory: Which to use in simulating complex flows of liquid-crystalline polymers? *Journal of Rheology*, 42(5):1095–1119, 1998.
- [27] Qi Wang. Comparative studies on closure approximations in flows of liquid crystal polymers: I. Elongational flows. *Journal of Non-Newtonian Fluid Mechanics*, 72(2):141–162, 1997.
- [28] Gregory M. Forest, Ruhai Zhou, and Qi Wang. Full-tensor alignment criteria for sheared nematic polymers. *Journal of Rheology*, 47(1):105–127, 2003.
- [29] Gregory Lielens, Roland Keunings, and Vincent Legat. The FENE-L and FENE-LS closure approximations to the kinetic theory of finitely extensible dumbbells. *Journal of Non-Newtonian Fluid Mechanics*, 87(2):179–196, 1999.
- [30] Peng Yu, Qiang Du, and Chun Liu. From micro to macro dynamics via a new closure approximation to the FENE model of polymeric fluids. *Multiscale Modeling & Simulation*, 3(4):895–917, 2005.

- [31] YunKyong Hyon, Qiang Du, and Chun Liu. An enhanced macroscopic closure approximation to the micro-macro fene model for polymeric materials. *Multiscale Modeling & Simulation*, 7(2):978–1002, 2008.
- [32] Huan Lei, Lei Wu, and Weinan E. Machine learning based non-Newtonian fluid model with molecular fidelity. *Physics Review E*, 102:043309, 2020.
- [33] Samuel H. Rudy, Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Data-driven discovery of partial differential equations. *Science Advances*, 3(4), 2017.
- [34] Hayden Schaeffer, Giang Tran, and Rachel Ward. Extracting sparse high-dimensional dynamics from limited data. *SIAM Journal on Applied Mathematics*, 78(6):3279–3295, 2018.
- [35] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- [36] Tong Qin, Kailiang Wu, and Dongbin Xiu. Data driven governing equations approximation using deep neural networks. *Journal of Computational Physics*, 395:620–635, 2019.
- [37] Jiequn Han, Chao Ma, Zheng Ma, and Weinan E. Uniformly accurate machine learning-based hydrodynamic models for kinetic equations. *Proceedings of the National Academy of Sciences*, 116(44):21983–21991, 2019.
- [38] Naoki Seryo, Takeshi Sato, John J. Molina, and Takashi Taniguchi. Learning the constitutive relation of polymeric flows with memory. *Phys. Rev. Research*, 2:033107, Jul 2020.
- [39] Haijun Yu, Xinyuan Tian, Weinan E, and Qianxiao Li. OnsagerNet: Learning stable and interpretable dynamics using a generalized onsager principle. 2020.
- [40] Juntao Huang, Zhiting Ma, Yizhou Zhou, and Wen-An Yong. Learning thermodynamically stable and galilean invariant partial differential equations for non-equilibrium flows. *Journal of Non-Equilibrium Thermodynamics*, 46(4):355–370, 2021.
- [41] Harold R. Warner. Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. *Industrial & Engineering Chemistry Fundamentals*, 11(3):379–387, 1972.
- [42] Stanislas Zaremba. Sur une forme perfectionnee de la theorie de la relaxation. *Bull. Int. Acad. Sci. Cracovie*, pages 594–614, 1903.
- [43] Robert Byron Bird, Charles F Curtiss, Robert C Armstrong, and Ole Hassager. *Dynamics of Polymeric Liquids, Volume 2: Kinetic Theory, 2nd Edition*. Wiley, 2nd edition, 1987.
- [44] Prince E. Rouse. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *The Journal of Chemical Physics*, 21(7):1272–1280, 1953.
- [45] Xu-Hui Zhou, Jiequn Han, and Heng Xiao. Frame-independent vector-cloud neural network for nonlocal constitutive modeling on arbitrary grids. 2021.
- [46] Patrick S. Doyle, Eric S.G. Shaqfeh, Gareth H. McKinley, and Stephen H. Spiegelberg. Relaxation of dilute polymer solutions following extensional flow. *Journal of Non-Newtonian Fluid Mechanics*, 76(1):79–110, 1998.
- [47] Gregory Lielens, Pierre Halin, Ingrid Jaumain, Roland Keunings, and Vincent Legat. New closure approximations for the kinetic theory of finitely extensible dumbbells. *Journal of Non-Newtonian Fluid Mechanics*, 76(1):249–279, 1998.
- [48] John. R. Womersley. Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. *The Journal of Physiology*, 127(3):553–563, 1955.
- [49] David A. Nicholson and Gregory C. Rutledge. Molecular simulation of flow-enhanced nucleation in n-eicosane melts under steady shear and uniaxial extension. *The Journal of Chemical Physics*, 145(24):244903, 2016.
- [50] Takahiro Murashima, Katsumi Hagita, and Toshihiro Kawakatsu. Elongational viscosity of weakly entangled polymer melt via coarse-grained molecular dynamics simulation. *Nihon Reoroji Gakkaishi*, 46(5):207–220, 2018.

- [51] Douglas E. Smith, Hazen P. Babcock, and Steven Chu. Single-polymer dynamics in steady shear flow. *Science*, 283(5408):1724–1727, 1999.
- [52] Geoffrey Ingram Taylor. The formation of emulsions in definable fields of flow. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 146(858):501–523, 1934.
- [53] Becca Thomases and Michael Shelley. Emergence of singular structures in oldroyd-b fluids. *Physics of Fluids*, 19(10):103103, 2007.
- [54] Weinan E, Jiequn Han, and Linfeng Zhang. Machine-learning-assisted modeling. *Physics Today*, 74(7):36–41, 2021.
- [55] Bruno H. Zimm. Dynamics of polymer molecules in dilute solution: Viscoelasticity, flow birefringence and dielectric loss. *The Journal of Chemical Physics*, 24(2):269–278, 1956.
- [56] Huan Lei, Nathan A. Baker, and Xiantao Li. Data-driven parameterization of the generalized Langevin equation. *Proc. Natl. Acad. Sci.*, 113(50):14183–14188, 2016.
- [57] Huan Lei and Xiantao Li. Petrov-Galerkin methods for the construction of non-Markovian dynamics preserving nonlocal statistics. *J. Chem. Phys.*, 154:184108, 2021.
- [58] P. J. Hoogerbrugge and J. M. V. A. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.*, 19(3):155–160, 1992.
- [59] Robert D Groot and Patrick B Warren. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *Journal of Chemical Physics*, 107(11):4423–4435, 1997.
- [60] Pep Español and Patrick B. Warren. Statistical mechanics of dissipative particle dynamics. *Europhysics Letters*, 30(4):191–196, 1995.
- [61] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, 12 2015.