Machine-Learning-Based Multi-scale Modeling for Complex Fluids

Pei Ge

Michigan State University / Worcester Polytechnic Institute

June 30, 2025

Research

- Machine-learning-based multi-scale modeling
  - Non-Newtonian Flow
  - Molecular Kinetics
- Main challenge:
  - High dimensionality
  - Physical interpretation (e.g., structure preserving, micro to macro mapping)
  - Numerical stability

Outline

- Data-driven Learning of Generalized Langevin Equations with State-dependent Memory
- A variational-informed machine-learning model of non-Newtonian fluids

Generalized Langevin Equations (GLE)

- Full system: $\mathbf{z_q}, \mathbf{z_p} \in \mathbb{R}^{N_{\text{full}}}$, reduced system: $\mathbf{q}, \mathbf{p} \in \mathbb{R}^{N_{\text{reduced}}}$, $N_{\text{reduced}} \ll N_{\text{full}}$
- Following Zwanzig's projection formalism

$$\begin{cases} \dot{\mathbf{q}} = M^{-1}\mathbf{p} \\ \dot{\mathbf{p}} = -\nabla U(\mathbf{q}) - \int_0^t \mathbf{K}(\mathbf{q}(\tau), t-\tau)\dot{\mathbf{q}}(\tau)\mathrm{d}\tau + \mathcal{R}(t). \end{cases}$$

- Standard GLE: Homogeneous memory kernel. $\mathbf{K}(\mathbf{q}(\tau), t-\tau) \approx \mathbf{K}(t-\tau)$.
- Not enough for accurately predicting the collective behaviors.

Outline

- Data-driven Learning of Generalized Langevin Equations with State-dependent Memory
  - Background
  - **State-dependent Memory**
  - Data-driven Method
  - Numerical Result

GLE with State-dependent Memory

- The state-dependent GLE model

$$\begin{cases} \dot{\mathbf{q}} = \mathbf{M}^{-1}\mathbf{p} \\ \dot{\mathbf{p}} = -\nabla U(\mathbf{q}) - \int_0^t \phi(\mathbf{q}_t)\Theta(t-\tau)\phi(\mathbf{q}_\tau)\mathbf{v}_\tau \, \mathrm{d}\tau + \phi(\mathbf{q}_t)^T \mathbf{R}(t) \end{cases}$$

- **Main idea**: $\phi(\mathbf{q}) \in \mathbb{R}^{N_{\text{reduced}}\times 1} \longrightarrow \mathbb{R}^{n\times N_{\text{reduced}}}$ encodes n state-dependent features
- $\Theta(t) \in \mathbb{R} \to \mathbb{R}^{n\times n}$ encodes the non-Markovian effects
- $\mathbf{R}(t)$ is the thermal noise satisfied $\langle \mathbf{R}(t), \mathbf{R}(\tau)\rangle = k_B T\Theta(t-\tau)$
- $\Theta(t) = \mathrm{e}^{-\alpha t} \sum_{k=0}^{N_\omega} \tilde{\Theta}_k^S \cos(\omega_k t) - \tilde{\Theta}_k^A \sin(\omega_k t)$

Invariant Density Distribution

- **Proposition**: Invariant distribution: $\rho_{\mathrm{eq}}(\mathbf{q},\mathbf{p}) \propto \exp\left\{-\left[U(\mathbf{q}) + \mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}/2\right]/k_BT\right\}$

- **Proof**: Rewrite $\tilde{\Theta}_k^S = \Gamma_{k,1}^T\Gamma_{k,1} + \Gamma_{k,2}^T\Gamma_{k,2}$, $\tilde{\Theta}_k^A = \Gamma_{k,1}^T\Gamma_{k,2} - \Gamma_{k,2}^T\Gamma_{k,1}$

$$\Theta(t) = \mathrm{e}^{-\alpha t}\sum_{k=0}^{N_\omega}\begin{pmatrix}\Gamma_{k,1}\\ \Gamma_{k,2}\end{pmatrix}^T\begin{pmatrix}\cos(\omega_k t)\mathbf{I} & \sin(\omega_k t)\mathbf{I}\\ -\sin(\omega_k t)\mathbf{I} & \cos(\omega_k t)\mathbf{I}\end{pmatrix}\underbrace{\begin{pmatrix}\Gamma_{k,1}\\ \Gamma_{k,2}\end{pmatrix}}_{\triangleq\,\Gamma_k} = \sum_{k=0}^{N_\omega}\begin{pmatrix}\Gamma_{k,1}\\ \Gamma_{k,2}\end{pmatrix}^T\exp\left(\underbrace{\begin{pmatrix}-\alpha\mathbf{I} & \omega_k\mathbf{I}\\ -\omega_k\mathbf{I} & -\alpha\mathbf{I}\end{pmatrix}}_{\triangleq\,\mathbf{J}_k}t\right)\begin{pmatrix}\Gamma_{k,1}\\ \Gamma_{k,2}\end{pmatrix}$$

- Rewrite $\dot{\mathbf{p}} = -\nabla U(\mathbf{q}) - \int_0^t \phi(\mathbf{q}_t)^T\Theta(t-\tau)\phi(\mathbf{q}_\tau)\mathbf{v}_\tau\,\mathrm{d}\tau + \phi(\mathbf{q}_t)^T\mathbf{R}(t)$ into

$$\dot{\mathbf{p}} = -\nabla U(\mathbf{q}) + \sum_{k=0}^{N_\omega}\phi(\mathbf{q}_t)^T\Gamma_k^T\underbrace{\left(-\int_0^t \exp\left(\mathbf{J}_k(t-\tau)\right)\Gamma_k\phi(\mathbf{q}_\tau)\mathbf{v}_\tau\,\mathrm{d}\tau + \mathbf{R}_k(t)\right)}_{\triangleq\,\mathbf{z}_k} \tag{1}$$

where $\langle\mathbf{R}_k(t),\mathbf{R}_k(\tau)\rangle = k_BT\exp\left(\mathbf{J}_k(t-\tau)\right)$

Invariant Density Distribution

- The form with the extended variable is

$$\dot{\mathbf{q}} = \mathbf{M}^{-1}\mathbf{p} \qquad \dot{\mathbf{p}} = -\nabla U(\mathbf{q}) + \sum_{k=0}^{N_\omega} (\Gamma_k \phi(\mathbf{q}_t))^T \mathbf{z}_k$$
$$\mathbf{z}_k = -\int_0^t \exp\left(\mathbf{J}_k(t-\tau)\right) \Gamma_k \phi(\mathbf{q}_\tau)\mathbf{v}_\tau \, \mathrm{d}\tau + \mathbf{R}_k(t) \tag{2}$$
$$\frac{\mathrm{d}\mathbf{z}_k}{\mathrm{d}t} = \mathbf{J}_k\mathbf{z}_k - \Gamma_k\phi(\mathbf{q}_t)\mathbf{v}_t - \Lambda_k\dot{\mathbf{W}}_k \qquad \Lambda_k\Lambda_k^T = -k_B T(\mathbf{J}_k + \mathbf{J}_k^T)$$

which is a Langevin equation

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{q}, \mathbf{p}, \cdots, \mathbf{z}_k, \cdots) = \mathbf{J}\nabla F(\mathbf{q}, \mathbf{p}, \cdots, \mathbf{z}_k, \cdots) + \Lambda\dot{\mathbf{W}}_t \tag{3}$$

Invariant Density Distribution

- Langevin equation

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{q}, \mathbf{p}, \cdots, \mathbf{z}_k, \cdots) = \mathbf{J}\nabla F(\mathbf{q}, \mathbf{p}, \cdots, \mathbf{z}_k, \cdots) + \Lambda \dot{\mathbf{W}}_t$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \\ \cdots \\ \mathbf{z}_k \\ \cdots \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & \mathbf{I} & \cdots & 0 & \cdots \\ -\mathbf{I} & 0 & \cdots & (\Gamma_k \phi(\mathbf{q}))^T & \cdots \\ 0 & \cdots & \cdots & 0 & \cdots \\ 0 & -\Gamma_k \phi(\mathbf{q}) & 0 & \mathbf{J}_k & 0 \\ 0 & \cdots & \cdots & 0 & \cdots \end{pmatrix}}_{\mathbf{J}} \underbrace{\begin{pmatrix} \nabla U(\mathbf{q}) \\ \mathbf{v} \\ \cdots \\ \mathbf{z}_k \\ \cdots \end{pmatrix}}_{\nabla F} + \sqrt{2k_B T \alpha}\begin{pmatrix} 0 \\ 0 \\ \cdots \\ \dot{\mathbf{W}}_k \\ \cdots \end{pmatrix} \tag{4}$$

- The invariant density function $\rho_{\mathrm{eq}}(\mathbf{q}, \mathbf{p}, \mathbf{z}) = \exp\left[-F(\mathbf{q}, \mathbf{p}, \mathbf{z})/k_B T\right]$
- How to parameterize Γ_k and $\phi(\mathbf{q})$?

Outline

- Data-driven Learning of Generalized Langevin Equations with State-dependent Memory
  - Background
  - State-dependent Memory
  - **Data-driven Method**
  - Numerical Result

Data-driven Method

- Recall the state-dependent GLE

$$\dot{\mathbf{p}}_t + \nabla U(\mathbf{q}_t) = -\int_0^t \phi(\mathbf{q}_t)^T \Theta(t-\tau)\phi(\mathbf{q}_\tau)\mathbf{v}_\tau \,\mathrm{d}\tau + \phi(\mathbf{q}_t)^T \mathbf{R}(t)$$

- Conditional correlation function

$$\left\langle \dot{\mathbf{p}}_t + \nabla U(\mathbf{q}), \mathbf{v}_0^T \middle| \mathbf{q}_0 = \mathbf{q}^* \right\rangle = -\int_0^t \left\langle \phi(\mathbf{q}_t)^T \Theta(t-\tau)\phi(\mathbf{q}_\tau)\mathbf{v}_\tau, \mathbf{v}_0^T \middle| \mathbf{q}_0 = \mathbf{q}^* \right\rangle$$

- Represent $\phi(\mathbf{q})$ with a set of sparse bases $\psi(\mathbf{q})$, such that $\phi(\mathbf{q}) = \mathbf{H}\psi(\mathbf{q})$

$$\text{RHS} = -\int_0^t \mathrm{Tr}\left[\Theta(t-\tau)\mathbf{H}\left\langle \psi(\mathbf{q}_\tau)\mathbf{v}(\tau)\mathbf{v}_0^T \psi(\mathbf{q}_t)^T \middle| \mathbf{q}_0 = \mathbf{q}^* \right\rangle \mathbf{H}^T\right] \mathrm{d}\tau$$

Model Parameterization Steps

- Pre-compute

$$\begin{aligned} g(t;\mathbf{q}^*) &= \left\langle \dot{\mathbf{p}}_t + \nabla U(\mathbf{q}_t), \mathbf{q}_0^T \middle| \mathbf{q}_0 = \mathbf{q}^* \right\rangle \\ C_{\psi,\psi}(t,\tau;\mathbf{q}^*) &= \left\langle \psi(\mathbf{q}_\tau)\mathbf{v}(\tau)\mathbf{v}_0^T \psi(\mathbf{q}_t)^T \middle| \mathbf{q}_0 = \mathbf{q}^* \right\rangle \end{aligned} \tag{5}$$

- Step 1: compute $\Theta(t) = \mathrm{e}^{-\alpha t} \sum_{k=0}^{N_\omega} \tilde{\Theta}_k^S \cos(\omega_k t)$
- Step 2: compute $\tilde{g}(t;\mathbf{q}^*) = -\int_0^t \mathrm{Tr}\left[\Theta(t-\tau)\mathbf{H} C_{\psi,\psi}(t,\tau;\mathbf{q}^*)\mathbf{H}^T\right] \mathrm{d}\tau$
- Step 3: optimize $\text{loss}(\mathbf{q}^*, \mathbf{t}) = \left\| g(t;\mathbf{q}^*) - \tilde{g}(t;\mathbf{q}^*) \right\|_2^2$

Simulation

- Recall that $\langle \mathbf{R}(t), \mathbf{R}(\tau)\rangle = k_B T \Theta(t-\tau)$ and

$$\dot{\mathbf{p}} = -\nabla U(\mathbf{q}) - \phi(\mathbf{q}_t)^T \int_0^t \Theta(t-\tau)\phi(\mathbf{q}_\tau)\mathbf{v}(\tau)\,\mathrm{d}\tau + \phi(\mathbf{q}_t)^T \mathbf{R}(t)$$

- The noise $\mathbf{R}(t)$ term is pre-generated by

$$\mathbf{R}(t) = \frac{1}{\sqrt{k_B T}} \sum_{j=0}^{2N} \hat{\Theta}_j^{1/2} \left[\cos(\omega_j t)\xi_j + \sin(\omega_j t)\eta_j\right]$$

$$\hat{\Theta}_j = \sum_{k=0}^{N_\omega} \left(\frac{\alpha\tilde{\Theta}_k}{\alpha^2 + (\omega_k - \omega_j)^2} + \frac{\alpha\tilde{\Theta}_k}{\alpha^2 + (\omega_k + \omega_j)^2} \right)$$

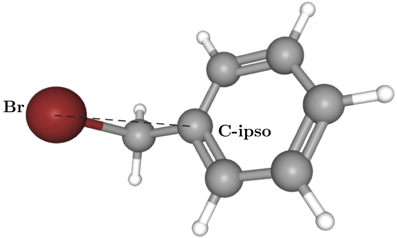
- The convolution term can be computed using the fast convolution algorithm.
- Only requires $O(N \log N)$ complexity.

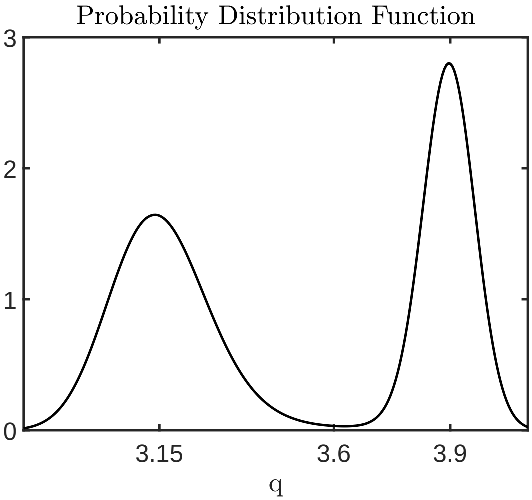
Outline

- Data-driven Learning of Generalized Langevin Equations with State-dependent Memory
    - Background
    - State-dependent Memory
    - Data-driven Method
    - **Numerical Result**

Numerical Result: Full Model

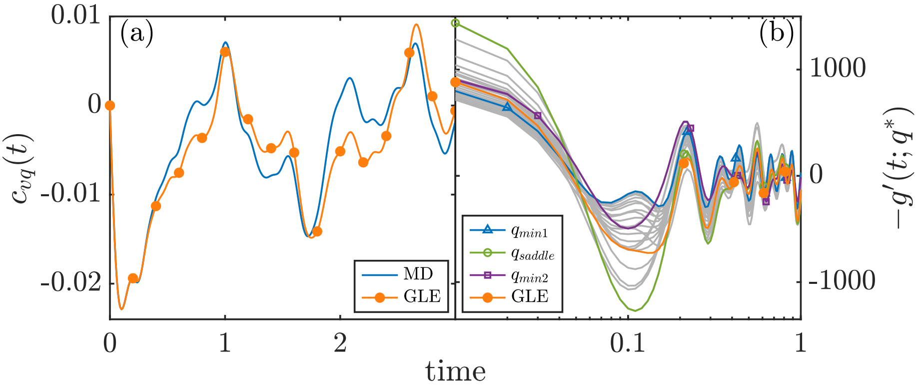
- A Benzyl bromide molecule in water.
- **q** is the distance between the bromine atom and the ipso-carbon atom.
- $U(\mathbf{q})$ is evaluated from the PDF.





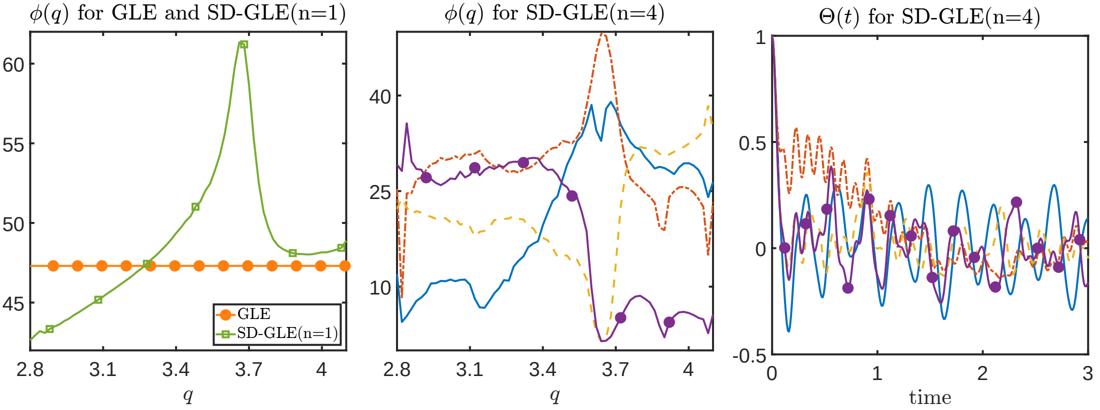
Numerical Result: Limitation of Standard GLE

$$g(t;\mathbf{q}^*) = \left\langle \dot{\mathbf{p}}_t + \nabla U(\mathbf{q}_t), \mathbf{q}_0^T \middle| \mathbf{q}_0 = \mathbf{q}^* \right\rangle, \quad g'(0;q^*) = -k_B T \mathbf{K}(q^*,0)/m$$

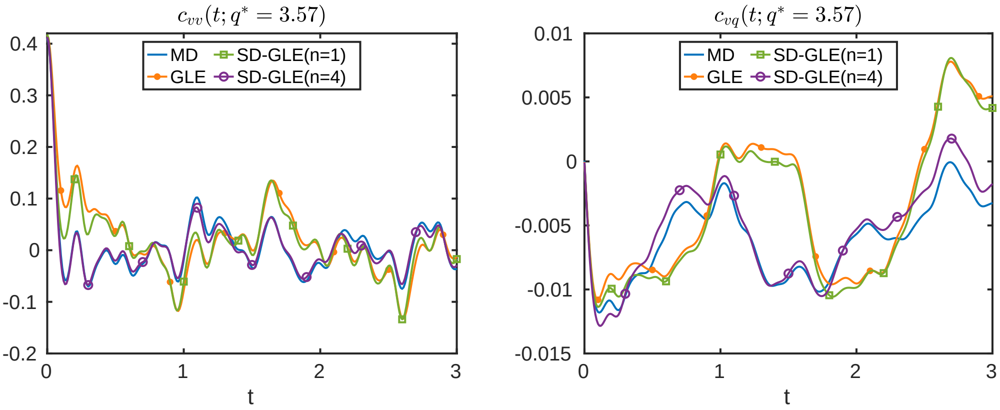


Numerical Result: $\Theta(t)$ and $\phi(\mathbf{q})$

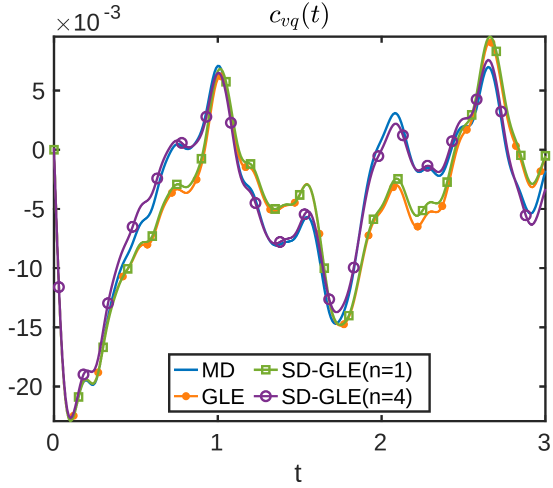
$$\dot{\mathbf{p}} = -\nabla U(\mathbf{q}) - \int_0^t \phi(\mathbf{q}_t)^T \Theta(t-\tau)\phi(\mathbf{q}_\tau)\mathbf{v}(\tau)\,\mathrm{d}\tau + \phi(\mathbf{q}_t)^T \mathbf{R}(t)$$



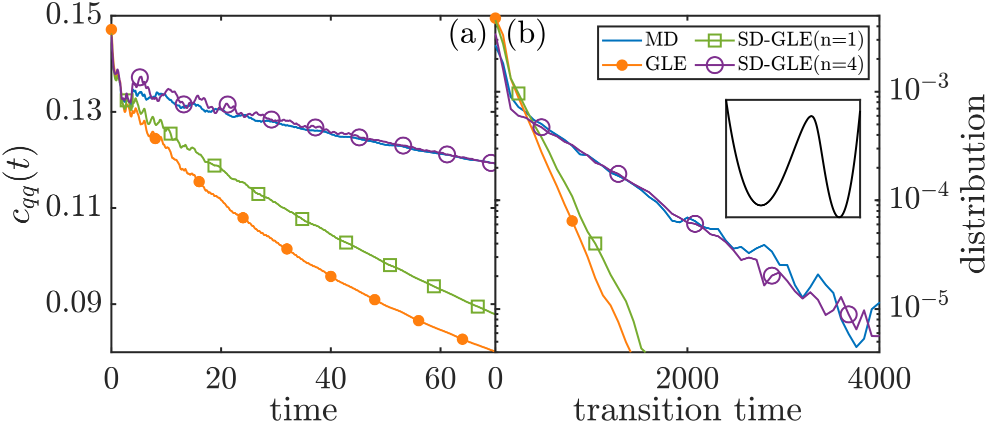
Numerical Result on Saddle Point



Numerical Result: Standard Correlation Function



Numerical Result: Transition Time



Conclusion

- State-dependent memory kernel is crucial on the reduced dynamics.
- Our model has consistent density distribution.
- Efficient training is achieved by pre-computed three-point correlation function.

Outline

- A variational-informed machine-learning model of non-Newtonian fluids
  - **Background**
  - Micro-scale molecular fidelity and interpretation
  - Macro-model variational structure and energy stability
  - Frame-indifference and physical constraints
  - Numerical Result

Hydrodynamics of non-Newtonian fluids

- Continuum hydrodynamic model of incompressible non-Newtonian fluids

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 \\ \rho \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} &= -\nabla p + \nabla \cdot (\tau_s + \tau_p) + \mathbf{f}_{\text{ext}} \end{aligned} \tag{6}$$

- Newtonian model of solvent stress τ_s

$$\tau_s = \eta_s(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$$

- Polymer stress τ_p is generally unknown

Why Difficult?

- Polymer stress τ_p determined by the micro-scale interactions

$$\tau_p = \langle \mathbf{r} \otimes \nabla V(\mathbf{r}) \rangle \qquad \mathbf{r} \text{ – bond vector} \qquad V \text{ – potential function}$$

- Conformation tensor $\mathbf{c}$

$$\tau_p = \mathbf{G}(\mathbf{c}), \quad \frac{\mathcal{D}\mathbf{c}}{\mathcal{D}t} = \mathbf{H}(\mathbf{c}), \quad \frac{\mathcal{D}\mathbf{c}}{\mathcal{D}t} \text{ – objective tensor derivative} \tag{7}$$

- Frame-indifference: $\widetilde{\dfrac{\mathcal{D}\mathbf{c}}{\mathcal{D}t}} = \mathcal{U}\dfrac{\mathcal{D}\mathbf{c}}{\mathcal{D}t}\mathcal{U}^T, \quad \mathcal{U}\mathcal{U}^T = \mathbf{I}$

- Multiple choices of the objective tensor derivative

  Upper convected: $\overset{\triangledown}{\mathbf{c}} = \dfrac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} - \nabla\mathbf{u}^T\mathbf{c} - \mathbf{c}\nabla\mathbf{u}$

  Lower convected: $\overset{\vartriangle}{\mathbf{c}} = \dfrac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} + \nabla\mathbf{u}\mathbf{c} + \mathbf{c}\nabla\mathbf{u}^T$

  $\ldots$

Hookean Model

- **Hookean Model**: $\mathbf{c} = \left\langle \mathbf{r}\mathbf{r}^T \right\rangle$ and harmonic bond potential $V(\mathbf{r}) \propto (\mathbf{r} - \mathbf{r}_0)^2$

$$\underbrace{\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} - \nabla\mathbf{u}^T\mathbf{c} - \mathbf{c}\nabla\mathbf{u}}_{\mathcal{D}\mathbf{c}/\mathcal{D}t} = \underbrace{\frac{1}{\lambda}(\mathbf{I} - \mathbf{c})}_{\mathbf{H}(\mathbf{c})} \qquad \tau_p = \langle \mathbf{r} \otimes \nabla V(\mathbf{r}) \rangle \propto \mathbf{c}$$

- What if $\nabla V(\mathbf{r})$ is non-linear to $\mathbf{r}$?
  - τ_p becomes non linear, so multiple $\mathbf{c}$ is needed
  - Each $\mathbf{c}$ needs a unique $\mathcal{D}\mathbf{c}/\mathcal{D}t$
  - $\mathbf{c}$ may rotate even without flow

Why difficult?

- **Motivation**: Learn the macro-scale dynamics directly from the micro-scale model

- **Caution & criticism**:
  - Time-series $\mathbf{c}_{t_i}$ may not be feasible to learn $\frac{\mathcal{D}\mathbf{c}}{\mathcal{D}t} = \mathbf{H}(\mathbf{c})$

$$\mathbf{c}_{t+\mathrm{d}t} - \mathbf{c}_t = \widetilde{\mathbf{H}}(\nabla\mathbf{u}, \mathbf{c})\mathrm{d}t$$

  - Retain physical interpretation and frame-indifference constraints

$$\widetilde{\frac{\mathcal{D}\mathbf{c}}{\mathcal{D}t}} = \mathcal{U}\frac{\mathcal{D}\mathbf{c}}{\mathcal{D}t}\mathcal{U}^T, \quad \mathcal{U}\mathcal{U}^T = \mathbf{I}$$

  - Numerical stability and generalization ability

Outline

- A variational-informed machine-learning model of non-Newtonian fluids
    - Background
    - **Micro-scale molecular fidelity and interpretation**
    - Macro-model variational structure and energy stability
    - Frame-indifference and physical constraints
    - Numerical Result

Micro-macro Correspondence

- Seek a set of explicit micro-macro correspondences

$$\mathbf{c}_i = \langle \mathbf{b}_i \rangle \quad \mathbf{b}_i = g_i^2(|\mathbf{r}^*|)(\omega_i \mathbf{r})^T(\omega_i \mathbf{r}) \in \mathbb{R}^{3\times 3}$$

$\langle\cdot\rangle$ - discrete samples collected from micro-scale molecular dynamics simulations

- $\mathbf{b}_i$ preserves frame-indifference

- By the meso-scale model of $\mathbf{r}$ and Fokker–Planck equation:

$$\underbrace{\frac{\mathrm{d}\mathbf{c}_i}{\mathrm{d}t} - \nabla\mathbf{u}^T : \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle}_{\text{interpretable objective tensor derivative}} = \underbrace{\frac{k_B T}{\gamma} \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{r_k} V(\mathbf{r}) \cdot \nabla_{r_j} \mathbf{b}_i \right\rangle - \frac{1}{\gamma} \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{r_j} \cdot \nabla_{r_k} \mathbf{b}_i \right\rangle}_{\text{dynamics directly from instantaneous discrete time samples}} \tag{8}$$

- A_{jk} and $V(\mathbf{r})$ encode the micro-scale structure

Micro-macro Correspondence

- **Question**: How should we build the macro-model?

$$\frac{\mathrm{d}\mathbf{c}_i}{\mathrm{d}t} - \nabla\mathbf{u}^T : \underbrace{\left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle}_{\mathcal{E}_i(\mathbf{c}_1,\cdots,\mathbf{c}_n)} = \frac{k_B T}{\gamma} \underbrace{\left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{r_k} V(\mathbf{r}) \cdot \nabla_{r_j} \mathbf{b}_i \right\rangle}_{\mathbf{H}_{1,i}(\mathbf{c}_1,\cdots,\mathbf{c}_n)} - \frac{1}{\gamma} \underbrace{\left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{r_j} \cdot \nabla_{r_k} \mathbf{b}_i \right\rangle}_{\mathbf{H}_{2,i}(\mathbf{c}_1,\cdots,\mathbf{c}_n)}$$

- Ensure physical constraints
  - Energy stability
  - Positive definite of conformation tensor

Outline

- A variational-informed machine-learning model of non-Newtonian fluids
  - Background
  - Micro-scale molecular fidelity and interpretation
  - **Macro-model variational structure and energy stability**
  - Frame-indifference and physical constraints
  - Numerical Result

Variational structure of the macro-scale hydrodynamic model

- Macro-scale hydrodynamics governed by the coupling of a reversible and an irreversible process

$$\frac{\mathrm{d}\mathcal{X}}{\mathrm{d}t} = \mathcal{L}\frac{\delta E}{\delta \mathcal{X}} + \mathcal{M}\frac{\delta S}{\delta \mathcal{X}} \tag{9}$$

$$\mathcal{L}\frac{\delta S}{\delta \mathcal{X}} = \mathcal{M}\frac{\delta E}{\delta \mathcal{X}} \equiv \mathbf{0} \tag{10}$$

$\mathcal{X} : \Omega \to \mathbb{R}^d$ - field variables $\quad E[\mathcal{X}] = \int_\Omega \hat{E}(\mathcal{X})\mathrm{d}\Omega$ - energy $\quad S[\mathcal{X}] = \int_\Omega \hat{S}(\mathcal{X})\mathrm{d}\Omega$ - entropy

$\mathcal{L} = -\mathcal{L}^T$ - poisson matrix $\quad \mathcal{M} \succ 0$ - friction matrix

- The degeneracy condition (4) ensures the energy conservation $\mathrm{d}E/\mathrm{d}t \equiv 0$ and the entropy production $\mathrm{d}S/\mathrm{d}t \geq 0$ and therefore the free energy stability

Onsager, Physical review, 1931; Morrison., Physica D, 1986; Grmela-Öttinger, Phys. Rev. E, 1997; Lin-Liu-Zhang, Commun. Pure Appl. Math., 2005; Yu-Tian-E-Li, Phys Rev. Fluids, 2021

Variational structure of the Hookean model

- **Example**: the PDE form of the Hookean non-Newtonian model

$$\rho_t = -\nabla \cdot (\rho \mathbf{u}) \qquad \frac{\mathrm{d}\rho \mathbf{u}}{\mathrm{d}t} = -\nabla p + \nabla \cdot (\tau_s + \tau_p)$$

$$\overset{\triangledown}{\mathbf{c}} = \frac{1}{\lambda}(\mathbf{I} - \mathbf{c}) \qquad \tau_s = \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^T), \quad \tau_p = \mathbf{c}$$

is governed by the variational structure

$$\mathcal{X} = (\rho, \mathbf{u}, \epsilon, \mathbf{c}) \qquad E = E_{\text{Newton}} = \int_\Omega \left(\frac{1}{2}\rho|\mathbf{u}|^2 + \epsilon\right) \mathrm{d}\Omega \qquad \mathcal{L} = \begin{pmatrix} \mathcal{L}_{\text{Newton}} & -\mathcal{L}_{\mathbf{c}}^T \\ \mathcal{L}_{\mathbf{c}} & \end{pmatrix}$$

$$S = \frac{1}{2}\int_\Omega S_{\text{Newton}} + \mathrm{Tr}(\mathbf{I} - \mathbf{c}) + \ln(\det(\mathbf{c}))\mathrm{d}\Omega \qquad \mathcal{M} = \begin{pmatrix} \mathcal{M}_{\text{Newton}} & \\ & \frac{1}{\lambda}\{\delta_{ik}\mathbf{c}_{jl} + \mathbf{c}_{ik}\delta_{jl}\} \end{pmatrix}$$

$\mathcal{L}_{\text{Newton}}, \mathcal{M}_{\text{Newton}}$ - poisson and friction operators of Newtonian fluid

$\mathcal{L}_{\mathbf{c}}$ - poisson operator of the upper-convected derivative

Grmela-Öttinger, Phys. Rev. E, 1997; Lin-Liu-Zhang, Commun. Pure Appl. Math., 2005

A generalized extendable variational structure

- **Main idea**: seek the energy variational form of our model

$$\mathcal{X} = (\rho, \mathbf{u}, \epsilon, \mathbf{c}_1, \cdots, \mathbf{c}_n) \qquad E = \int_\Omega \left(\frac{1}{2}\rho|\mathbf{u}|^2 + \epsilon\right)\mathrm{d}\Omega \qquad S = \int_\Omega S_{\text{Newton}} + \hat{S}(\mathbf{c}_1, \cdots, \mathbf{c}_n)\mathrm{d}\Omega$$

$$\mathcal{L} = \begin{pmatrix} \mathcal{L}_{\text{Newton}} & -\mathcal{L}_{\mathbf{c}_1}^T & \cdots \\ \mathcal{L}_{\mathbf{c}_1} & & \\ \vdots & & \end{pmatrix} \qquad \mathcal{M} = \begin{pmatrix} \mathcal{M}_{\text{Newton}} & \\ & \underbrace{\begin{matrix} \mathcal{M}_{44} & \cdots \\ \vdots & \end{matrix}}_{\mathcal{M}^S \in \mathbb{R}^{3n\times 3n}} \end{pmatrix} + \begin{pmatrix} & & -(M^A)^T \\ \underbrace{\begin{matrix} \mathcal{M}_{42} & \mathcal{M}_{43} \\ \vdots & \vdots \end{matrix}}_{\mathcal{M}^A \in [\mathbb{R}^{3n\times 3\times 3} \quad \mathbb{R}^{3n\times 3}]} & & \end{pmatrix}$$

$$\mathcal{M}^S \text{ - dynamics of } \mathbf{c} \qquad \mathcal{M}^A \text{ - objective tensor derivative } \frac{\mathcal{D}\mathbf{c}}{\mathcal{D}t}$$

- To strictly preserve the degeneracy condition with $\mathcal{M}\dfrac{\delta E}{\delta \mathcal{X}} = 0$

$$\mathcal{M}_{42} \cdot \mathbf{u} + \mathcal{M}_{43} = 0, \qquad \mathcal{M}_{42} = \frac{\partial}{\partial \mathbf{x}} \cdot \widetilde{\mathcal{E}}_i \qquad \widetilde{\mathcal{E}}_i \in \mathbb{R}^{3\times3\times3\times3}$$

A generalized extendable variational structure

- The final constitutive dynamics takes the PDE form:

$$\underbrace{\frac{\mathrm{d}\mathbf{c}_i}{\mathrm{d}t} - \nabla\mathbf{u}^T\mathbf{c}_i - \mathbf{c}_i\nabla\mathbf{u} - \nabla\mathbf{u}^T : \widetilde{\mathcal{E}}_i}_{\frac{\mathcal{D}\mathbf{c}_i}{\mathcal{D}t}} = \underbrace{\mathcal{M}_i^S : \frac{\partial\hat{S}}{\partial\mathbf{c}}}_{\mathbf{H}_i(\cdot)} \qquad \tau_p = \underbrace{\sum_i \left(2\mathbf{c}_i\frac{\partial\hat{S}}{\partial\mathbf{c}_i} + \widetilde{\mathcal{E}}_i : \frac{\partial\hat{S}}{\partial\mathbf{c}_i}\right)}_{\mathbf{G}(\cdot)} \tag{11}$$

- For simplicity, let $\mathcal{E}_i = \widetilde{\mathcal{E}}_i + \mathcal{L}_{\mathbf{c}_i}$

$$\frac{\mathrm{d}\mathbf{c}_i}{\mathrm{d}t} - \nabla\mathbf{u}^T : \mathcal{E}_i = \mathcal{M}_i^S : \frac{\partial\hat{S}}{\partial\mathbf{c}} \qquad \tau_p = \mathcal{E} : \frac{\partial\hat{S}}{\partial\mathbf{c}}$$

DeePN^2

- DeePN^2: A **variational-informed** machine-learning model of non-Newtonian fluids

$$\begin{aligned} \mathcal{X} &= (\rho, \mathbf{u}, \epsilon, \mathbf{c}_1, \cdots, \mathbf{c}_n) \\ \frac{\mathrm{d}\mathcal{X}}{\mathrm{d}t} &= \mathcal{L}\frac{\delta E}{\delta \mathcal{X}} + \mathcal{M}\frac{\delta S}{\delta \mathcal{X}} \\ \mathcal{L}\frac{\delta S}{\delta \mathcal{X}} &= \mathcal{M}\frac{\delta E}{\delta \mathcal{X}} \equiv \mathbf{0} \end{aligned} \quad \xrightarrow[\tau_p = \mathcal{E} : \frac{\partial \hat{S}}{\partial \mathbf{c}}]{\mathbf{H}_i = \mathcal{M}_i^S : \frac{\partial \hat{S}}{\partial \mathbf{c}}} \quad \begin{aligned} \rho_t &= -\nabla \cdot (\rho \mathbf{u}) \\ \frac{\mathrm{d}\rho_t \mathbf{u}}{\mathrm{d}t} &= -\nabla p + \nabla \cdot (\tau_s + \tau_p) \\ \frac{\mathcal{D}\mathbf{c}_i}{\mathcal{D}t} &:= \frac{\mathrm{d}\mathbf{c}_i}{\mathrm{d}t} - \nabla \mathbf{u}^T : \mathcal{E}_i = \mathbf{H}_i(\mathbf{c}) \end{aligned} \tag{12}$$

- $\mathcal{M}^S$, $\mathcal{E}$ and $\hat{S}$ represented by neural networks preserving physical constraints

$$\mathcal{M}_i^S : \frac{\partial \hat{S}}{\partial \mathbf{c}_i} = \frac{k_B T}{\gamma}\left\langle \sum_j^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{r_k} V(\mathbf{r}) \cdot \nabla_{r_j} \mathbf{b}_i \right\rangle - \frac{1}{\gamma}\left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{r_j} \cdot \nabla_{r_k} \mathbf{b}_i \right\rangle$$

$$\mathcal{E}_i = \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle \qquad \tau_p = \mathcal{E} : \frac{\partial \hat{S}}{\partial \mathbf{c}} = \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} V(\mathbf{r}) \right\rangle \tag{13}$$

Outline

- A variational-informed machine-learning model of non-Newtonian fluids
  - Background
  - Micro-scale molecular fidelity and interpretation
  - Macro-model variational structure and energy stability
  - **Frame-indifference and physical constraints**
  - Numerical Result

Neural Networks with Physical Constraints

- Recall the formulation

$$\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} - \nabla\mathbf{u}^T : \mathcal{E} = \mathcal{M}^S : \frac{\partial \hat{S}}{\partial \mathbf{c}}$$

- **Question**: What constraints should $\hat{S}$, $\mathcal{M}^S$, $\mathcal{E}$ follow?
  - Frame-indifference
  - $\mathcal{M}^S$ - positive definite
  - $\hat{S}$ - concave
  - Keep **c** positive definite

Positive definite and frame-indifferent friction matrix

- **Proposition**: Represent $\mathcal{M}^S$ and $\mathcal{E}$ in the polynomial function space of $\{\mathbf{c}_i\}$
- **Key observation**: Let $\mathbf{A} \in \mathbb{R}^{3\times 3}$, the Cayley–Hamilton theorem yields

$$\mathbf{A}^{p+1} = \mathrm{Tr}(\mathbf{A})\mathbf{A}^p - \frac{1}{2}(\mathrm{Tr}(\mathbf{A})^2 - \mathrm{Tr}(\mathbf{A}^2))\mathbf{A}^{p-1} + \det(\mathbf{A})\mathbf{A}^{p-2}$$

Therefore, higher-order polynomials can be represented by

$$\mathbf{c}^p = \xi_1(\mathbf{x_c})\mathbf{I} + \xi_2(\mathbf{x_c})\mathbf{c} + \xi_3(\mathbf{x_c})\mathbf{c}^2 \tag{14}$$

where $\mathbf{x_c}$ are the rotation invariants.

Positive definite and frame-indifferent friction matrix

- Representation of polynomial basis functions $\mathbf{W}(\mathbf{c})$ and frame-indifference variables $\mathbf{x_c}$

$$
\begin{aligned}
\mathbf{B}(\mathbf{c}) &= \left\{\mathbf{c}_i, \mathbf{I} + (\mathbf{c}_i - \mathbf{c}_{\mathrm{j}})^2\right\} \in \mathbb{R}^{m\times 3\times 3} \\
\mathbf{W}(\mathbf{c}) &= \left\{\mathbf{I}, \mathbf{B}_i, \mathbf{B}_i^2\right\}_{\mathbf{B}_i \in \mathbf{B}(\mathbf{c})} \in \mathbb{R}^{(2m+1)\times 3\times 3} \\
\mathbf{x_c} &= \left\{\mathrm{Tr}(\mathbf{B}_i), \mathrm{Tr}(\mathbf{B}_i^2), \mathrm{Tr}(\mathbf{B}_i^3)\right\}_{\mathbf{B}_i \in \mathbf{B}(\mathbf{c})} \in \mathbb{R}^{3m\times 1}
\end{aligned}
$$

- Neural network representation of $\mathcal{M}^S = \mathrm{diag}(\mathcal{M}_1^S, \cdots, \mathcal{M}_n^S)$ and $\mathcal{E}_i$

$$
\begin{aligned}
\mathcal{M}_i^S &= \mathbf{W}(\mathbf{c})^T \cdot \left(\Gamma_i^{\mathcal{M}}(\mathbf{x_c})\Gamma_i^{\mathcal{M}}(\mathbf{x_c})^T\right) \cdot \mathbf{W}(\mathbf{c}) \in \mathbb{R}^{3\times 3\times 3\times 3} \\
\mathcal{E}_i &= \mathbf{W}(\mathbf{c})^T \cdot \Gamma_i^{\mathcal{E}}(\mathbf{x_c}) \cdot \mathbf{W}(\mathbf{c}) \in \mathbb{R}^{3\times 3\times 3\times 3}
\end{aligned}
\tag{15}
$$

where $\Gamma \in \mathbb{R}^{3m} \to \mathbb{R}^{(2m+1)\times(2m+1)}$ is represented by the neural networks.

Concave and frame-indifferent entropy $\hat{S}$

- **Lemma**: If f and g are convex and g is non-decreasing, then $h(x) = g(f(x))$ is convex.
- Frame-indifference and convex inputs:

$$\mathbf{x_c} = \{\mathrm{Tr}(\mathbf{c}_i), \mathrm{Tr}(\mathbf{c}_i^2), \mathrm{Tr}(\mathbf{c}_i^3)\} \in \mathbb{R}^{3n\times 1} \tag{16}$$

- Convex and non-decreasing neural network

$$\begin{aligned}
\mathbf{z}_0 &= g_0(\mathbf{W}_0^x \mathbf{x_c} + \mathbf{b}_0) \\
\mathbf{z}_{l+1} &= g_{l+1}(\mathbf{W}_{l+1}^z \mathbf{z}_l + \mathbf{W}_{l+1}^x \mathbf{x_c} + \mathbf{b}_{l+1}) \\
\tilde{S}(\mathbf{x_c}) &= \mathbf{W}_{L+1}^z \mathbf{z}_L + \mathbf{W}_{L+1}^x \mathbf{x_c} + \mathbf{b}_{L+1} \qquad l = 0, \cdots, L-1
\end{aligned} \tag{17}$$

where $\mathbf{W}_l^z \geq 0$, $\mathbf{W}_l^x \geq 0$, g_l is convex and non-decreasing activation functions.

Amos-Xu-Kolter, International Conference on Machine Learning, 2017

Positive Definite $\mathbf{c}$

- **Assumption**: $\hat{S} = \hat{S}_{\text{bound}}(\text{Tr}(\mathbf{c}), \text{Tr}(\mathbf{c}^2)) + \log(\det(\mathbf{c}_i))$ and $\mathcal{E}_i$, $\mathbf{M}_i$, $\dfrac{\mathrm{d}\hat{S}_{\text{bound}}}{\mathrm{d}\mathbf{c}_i}$ bounded, $\mathbf{c}$ has upper bound
- **Proposition**: Given $\mathbf{c}(0)$ positive definite , $\mathbf{c}(t)$ will always be positive definite.
- **Sketch of Proof**: Recall the formulation $\dfrac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} - \nabla\mathbf{u}^T : \mathcal{E} = \mathcal{M}^S : \dfrac{\partial \hat{S}}{\partial \mathbf{c}}$
- Consider the dynamic of $\det(\mathbf{c})$

$$\frac{\mathrm{d}\det(\mathbf{c}_i)}{\mathrm{d}t} = \det(\mathbf{c}_i)\mathbf{c}_i^{-1} : \frac{\mathrm{d}\mathbf{c}_i}{\mathrm{d}t} = \det(\mathbf{c}_i)\left(\mathbf{c}_i^{-1} : (\nabla\mathbf{u} : \mathcal{E}_i) + \mathbf{c}_i^{-1} : (\mathbf{M}_i : \frac{\mathrm{d}\hat{S}}{\mathrm{d}\mathbf{c}})\right)$$

- By the Cayley-Hamilton theorem, let $\mathbf{P}(\mathbf{c}_i) := \det(\mathbf{c}_i)\mathbf{c}_i^{-1} = \mathbf{c}_i^2 - \text{Tr}(\mathbf{c}_i)\mathbf{c}_i + \frac{1}{2}(\text{Tr}(\mathbf{c}_i)^2 - \text{Tr}(\mathbf{c}_i^2))\mathbf{I}$

$$\frac{\mathrm{d}\det(\mathbf{c}_i)}{\mathrm{d}t} = \mathbf{P}(\mathbf{c}_i) : (\nabla\mathbf{u} : \mathcal{E}_i) + \mathbf{P}(\mathbf{c}_i) : (\mathbf{M}_i : \frac{\mathrm{d}\hat{S}}{\mathrm{d}\mathbf{c}_i})$$

Positive Definite $\mathbf{c}$

- Since $\hat{S} = \hat{S}_{\text{bound}}(\mathrm{Tr}(\mathbf{c}), \mathrm{Tr}(\mathbf{c}^2)) + \log(\det(\mathbf{c}_i))$, then we have

$$\frac{\mathrm{d}\hat{S}}{\mathrm{d}\mathbf{c}_i} = \frac{\mathrm{d}\hat{S}_{\text{bound}}}{\mathrm{d}\mathbf{c}_i} + \mathbf{c}_i^{-1}$$

$$\frac{\mathrm{d}\det(\mathbf{c}_i)}{\mathrm{d}t} = \mathbf{P}(\mathbf{c}_i) : \left(\nabla\mathbf{u} : \mathcal{E}_i + \mathbf{M}_i : \frac{\mathrm{d}\hat{S}_{\text{bound}}}{\mathrm{d}\mathbf{c}_i}\right) + \frac{1}{\det(\mathbf{c}_i)} \left(\mathbf{P}(\mathbf{c}_i) : \mathbf{M}_i : \mathbf{P}(\mathbf{c}_i)\right)$$

- $\mathbf{M}_i$ is positive definite so $\mathbf{P}(\mathbf{c}_i) : \mathbf{M}_i : \mathbf{P}(\mathbf{c}_i) > 0$
- So $\det(\mathbf{c})$ has lower bound that larger than 0
- $\mathbf{c}$ keeps positive definite

Outline

- A variational-informed machine-learning model of non-Newtonian fluids
  - Background
  - Micro-scale molecular fidelity and interpretation
  - Macro-model variational structure and energy stability
  - Frame-indifference and physical constraints
  - **Numerical Result**

Dumbbell solution: reverse-Poiseuille flow

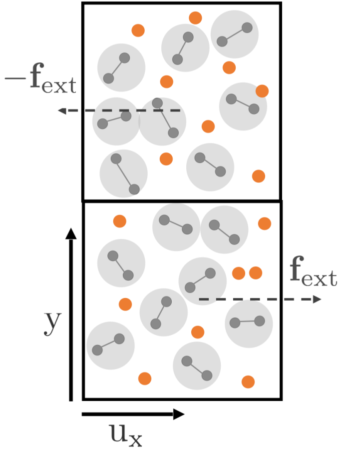
- Micro-scale polymer solution model
  - Elastic bond (FENE)

$$V_b(\mathbf{r}) = -\frac{k}{2} r_0^2 \log(1 - \frac{|\mathbf{r}|^2}{r_0^2})$$

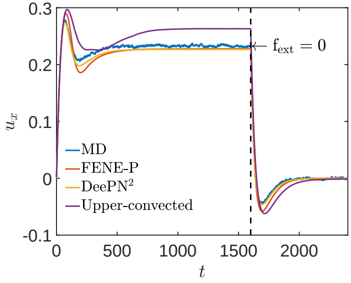
k_s - spring constant

r_0 - maximum bond extension

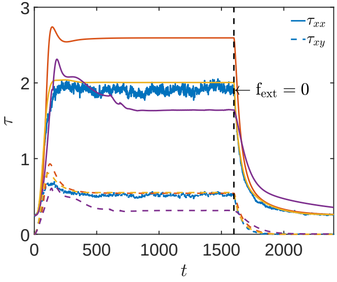
  - Pairwise interactions imposed between solvent-solvent, solvent-polymer
- Empirical continuum non-Newtonian fluid models
  - FENE-P model: $\tau_p \propto \dfrac{\mathbf{c}}{1 - \mathrm{Tr}(\mathbf{c})/r_0^2}$, $\overset{\triangledown}{\mathbf{c}} \propto \mathbf{I} - \lambda \dfrac{\mathbf{c}}{1 - \mathrm{Tr}(\mathbf{c})/r_0^2}$



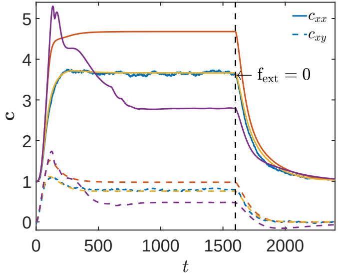
Numerical Results



velocity evolution



stress evolution



conformation tensor evolution

Conclusion

- **Retain micro-model fidelity**: systematically pass the micro-scale analytical form to macro-scale
    - time-series samples are not needed
    - interpretable objective tensor derivative

- **Respect frame-indifference and physical constraints**
    - micro-macro correspondence
    - neural networks (rotational symmetry, concave, positive definite)

- **Be reliable**: strictly preserve the energy structure and ensure numerical stability
    - existing energy stable schemes (e.g., scalar auxiliary variable, the convex splitting) can be naturally inherited
    - existing schemes for positive definite conformation tensors (e.g., matrix-logarithm of the conformation tensor)

Published Work

1. **Pei Ge**, Zhongqiang Zhang, and Huan Lei. Data-Driven Learning of the Generalized Langevin Equation with State-Dependent Memory. Physical Review Letters, 2024
2. **Pei Ge**, Linfeng Zhang, and Huan Lei. Machine learning assisted coarse-grained molecular dynamics modeling of meso-scale interfacial fluids. Journal of Chemical Physics, 2023
3. Zhiyuan She, **Pei Ge**, and Huan Lei. Data-driven construction of stochastic reduced dynamics encoded with non-Markovian features. Journal of Chemical Physics, 2023
4. Lidong Fang, **Pei Ge**, Lei Zhang, Weinan E, and Huan Lei. DeePN^2: A Deep Learning-Based Non-Newtonian Hydrodynamic Model. Journal of Machine Learning, 2022

Thank You!

MICHIGAN STATE
UNIVERSITY