Reduction and closure of dynamical systems using deep learning

Qianxiao Li

Department of Mathematics

Institute for Functional Intelligent Materials

blog.nus.edu.sg/qianxiaoli

Hammer & Nails 2023, Swiss Edition
Congressi Stefano Franscini, Ascona, Switzerland
31 Oct 2023

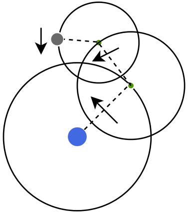
NUS
National University of Singapore

Introduction

Constructing dynamics from data is an old problem

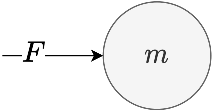
Constructing dynamical models from observations is a fundamental problem

Ptolemy's epicycle models



Newton's laws

$$m\ddot{x} = F$$



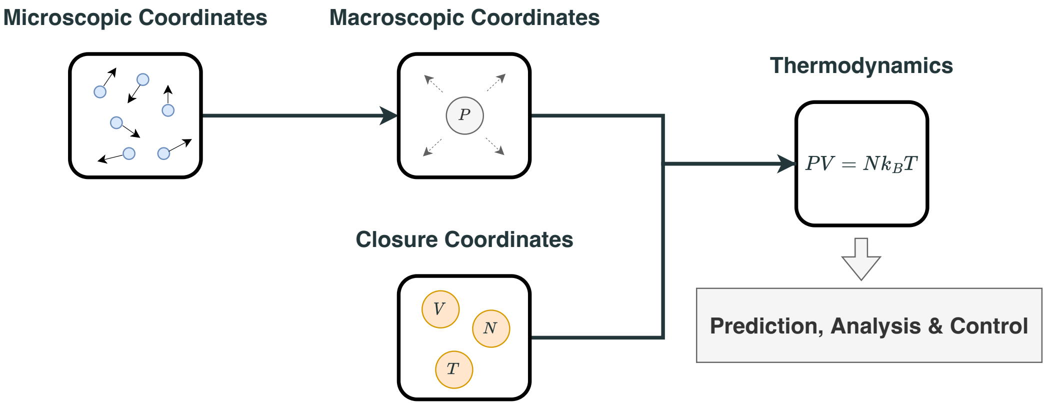
Schrödinger equation

$$i\hbar\dot{\psi} = H\psi$$

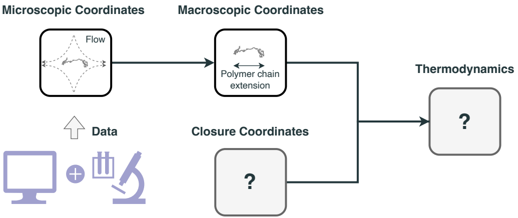
“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.”

– Paul Dirac

From microscopic to macroscopic



Constructing macroscopic dynamics



For dynamical processes, finding closure coordinates and constructing macroscopic dynamics relies on deep theoretical insight + trial and error

Can machine learning automate this process of discovery?

K. Hippalgaonkar, Q. Li, X. Wang, J. W. Fisher, J. Kirkpatrick, T. Buonassisi, *Nature Reviews Materials*, 1–20 (Jan. 24, 2023)

Mathematical formulation

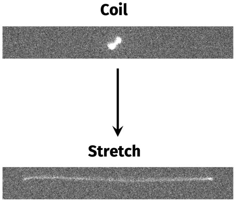
We are given

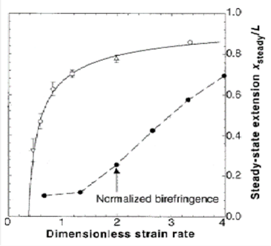
- Microscopic degrees of freedom: $X(t) \in \mathbb{R}^D$
- Macroscopic state of interest: $Z^*(t) = \varphi^*(X(t)) \in \mathbb{R}^{d_1}$ $(d_1 \ll D)$
- Goal – model the evolution of $Z^*(t)$

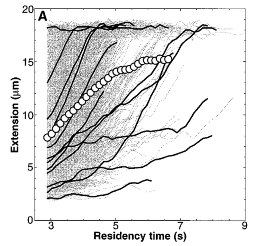
We aim to find

- Closure variables $\hat{Z}(t) = \hat{\varphi}(X(t)) \in \mathbb{R}^{d_2}$ $(d_2 \ll D)$
- A closed equation for $Z(t) = (Z^*(t), \hat{Z}(t)) \in \mathbb{R}^d$ $(d = d_1 + d_2)$

Example: stretching of polymers in elongational flow







- $X(t)$: coordinates of each molecule in the polymer at time t
- $Z^*(t)$: length/extension of the polymer at time t

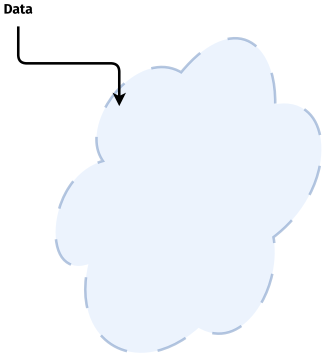
Strong heterogeneity in extension dynamics[1,2]

[1] T. T. Perkins, D. E. Smith, S. Chu, *Science* **276**, 2016–2021 (1997).

[2] D. E. Smith, S. Chu, *Science* **281**, 1335–1340 (Aug. 28, 1998).

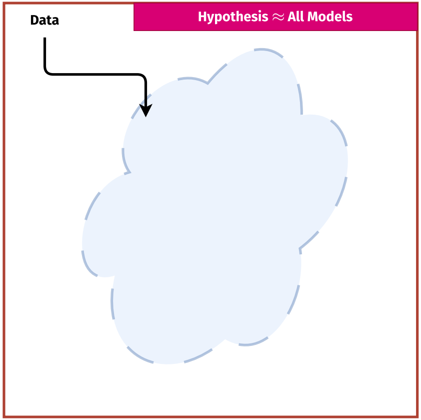
Two viewpoints for learning dynamics from data

Data-based approach



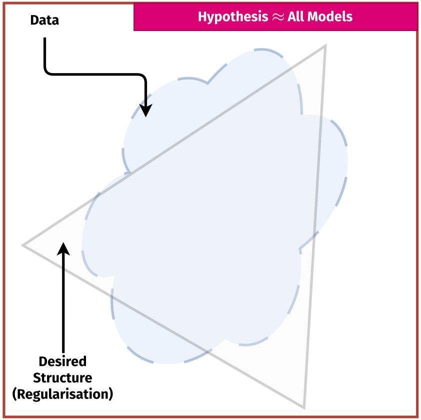
Two viewpoints for learning dynamics from data

Data-based approach

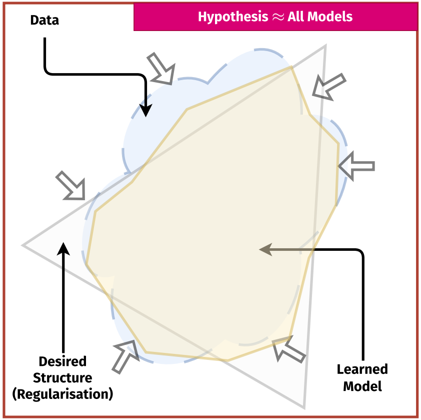


Two viewpoints for learning dynamics from data

Data-based approach

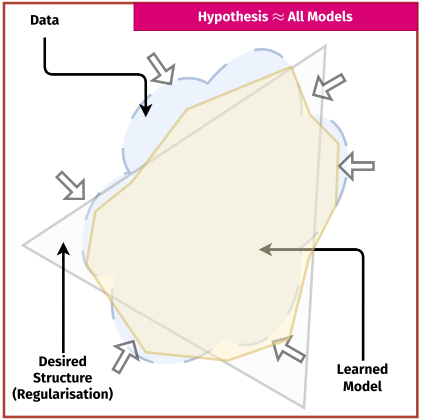


Data-based approach

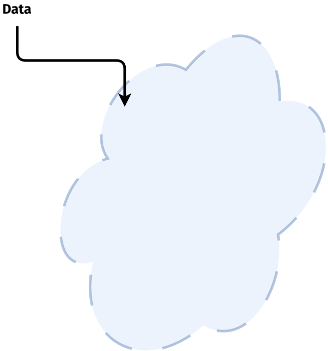


Two viewpoints for learning dynamics from data

Data-based approach

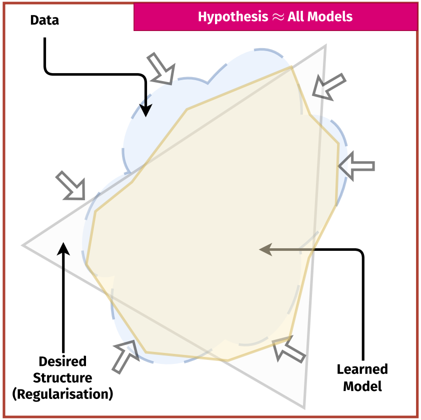


Model-based approach

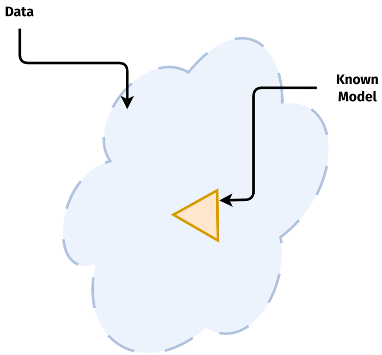


Two viewpoints for learning dynamics from data

Data-based approach

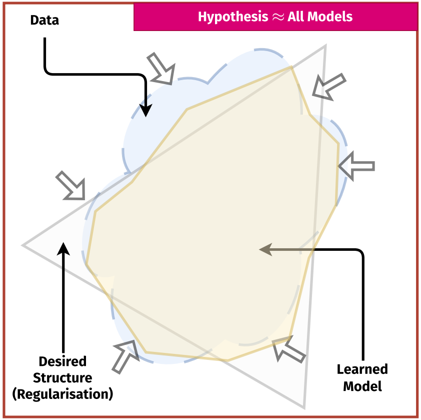


Model-based approach

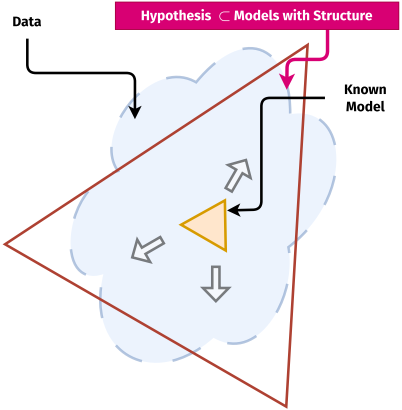


Two viewpoints for learning dynamics from data

Data-based approach

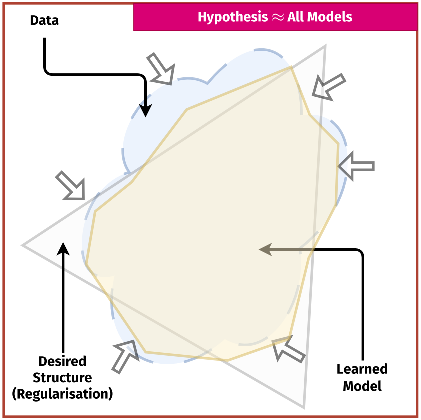


Model-based approach

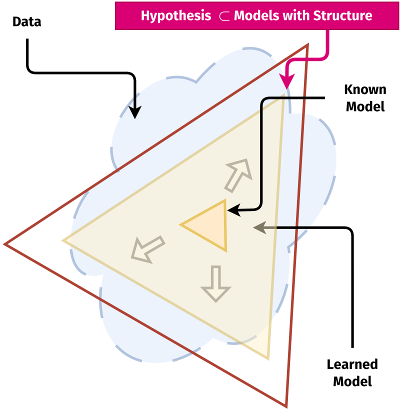


Two viewpoints for learning dynamics from data

Data-based approach



Model-based approach



Known model: the Onsager principle

The Onsager principle is a general description for near-equilibrium dynamics

$$\dot{Z}(t) = -M\nabla V(Z(t))$$

- Z is a generalised coordinate
- V is a generalised potential (or free energy)
- M models dissipation
- M is symmetric (reciprocal relations[3]) and positive definite (stability)

[3] L. Onsager, *Physical review* **38**, 2265 (1931).

Generalised model: generalised stochastic Onsager principle (GSOP)

We generalise the Onsager principle[4,5]

$$\boxed{dZ(t) = -\left[M(Z(t)) + W(Z(t))\right]\nabla V(Z(t))dt + \sigma(Z(t))dB(t)}$$

- M is symmetric positive semi-definite (dissipative, reciprocal relations)
- W is anti-symmetric (conservative)
- V is lower-bounded, sufficient growth (potential, free energy, -entropy)
- σ is matrix-valued (thermal fluctuations)

[4] H. Yu, X. Tian, W. E, Q. Li, *Physical Review Fluids* **6**, 114402 (Nov. 23, 2021).

[5] X. Chen, B. W. Soh, Z.-E. Ooi, E. Vissol-Gaudin, H. Yu, K. S. Novoselov, K. Hippalgaonkar, Q. Li, *Constructing Custom Thermodynamics Using Deep Learning*, (http://arxiv.org/abs/2308.04119), preprint.

Properties of GSOP I - generality

The GSOP form includes many well-known physical dynamical systems

- Langevin equation for molecular dynamics[6]

$$m\ddot{x} = -\nabla U(x) - \gamma \dot{x} + \sqrt{2\gamma k_B T}\dot{B}(t)$$

- Stochastic Poisson systems[7]

$$\dot{F} = \{F, H\} - [F, H] + \sigma(F)\dot{B}(t)$$

[6] M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, (Oxford university press, 2010).

[7] N. S. Goel, S. C. Maitra, E. W. Montroll, *Reviews of modern physics* **43**, 231 (1971).

Properties of GSOP II - invariance

GSOP is approximately invariant under coordinate transformation

- Let X follow GSOP with some $M(\cdot), W(\cdot), V(\cdot), \sigma(\cdot)$
- Let $\phi : \mathbb{R}^D \to \mathbb{R}^d$ is some approximately invertible transformation

Then, $Z = \phi(X)$ follows approximately another GSOP with some $\tilde{M}(\cdot), \tilde{W}(\cdot), \tilde{V}(\cdot), \tilde{\sigma}(\cdot)$

Properties of GSOP III - stability

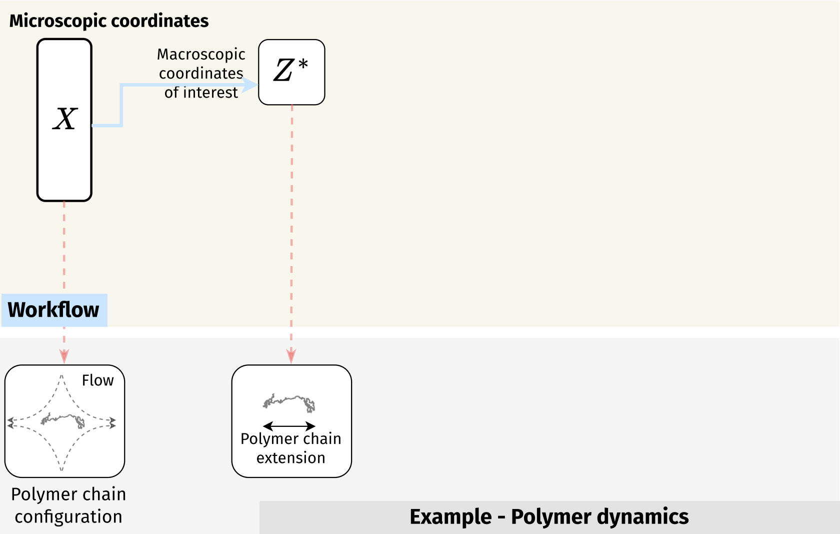
The GSOP gives rise to stable dynamics

- Let Z follows GSOP with some $M(\cdot), W(\cdot), V(\cdot), \sigma(\cdot)$

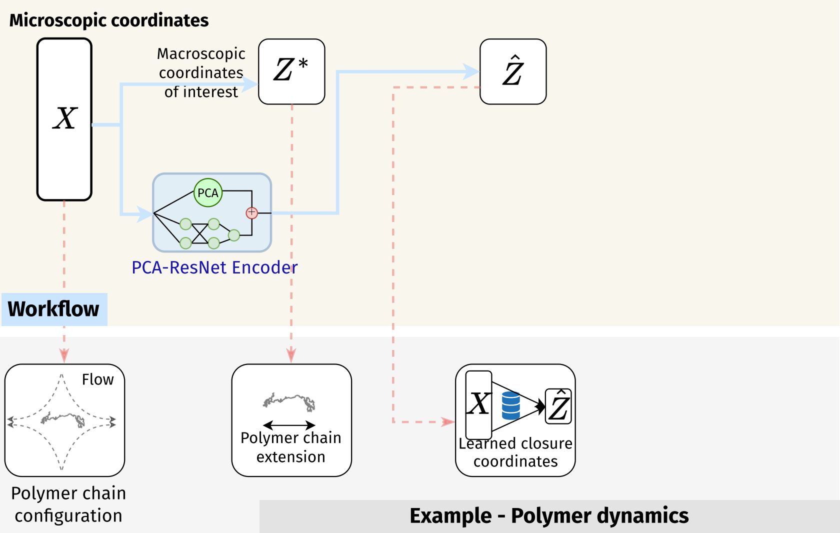
Then,

- $\mathbb{E}V(Z(t))$ decreases monotonically to $\mathcal{O}(\|\sigma\|^2)$
- The rate of dissipation is controlled by $M(\cdot)$
- If sub-level-sets of V are bounded and V has sufficient growth then the dynamics is stable

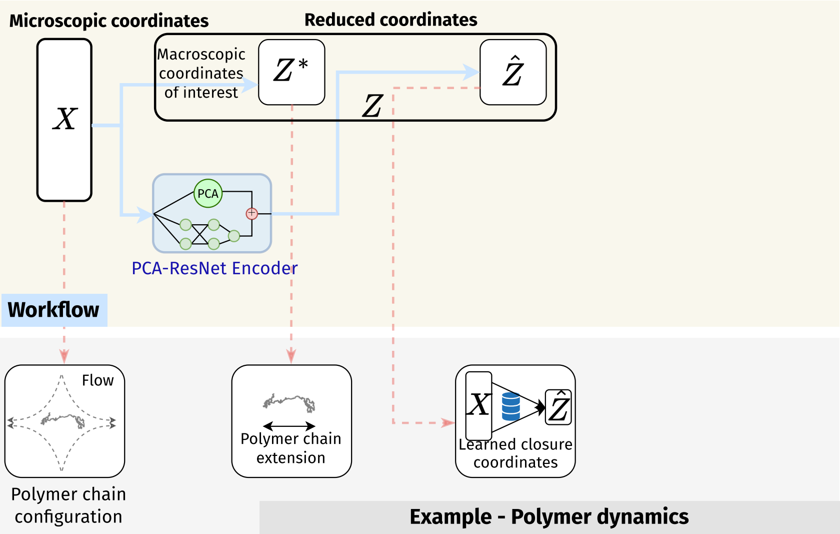
The stochastic OnsagerNet workflow



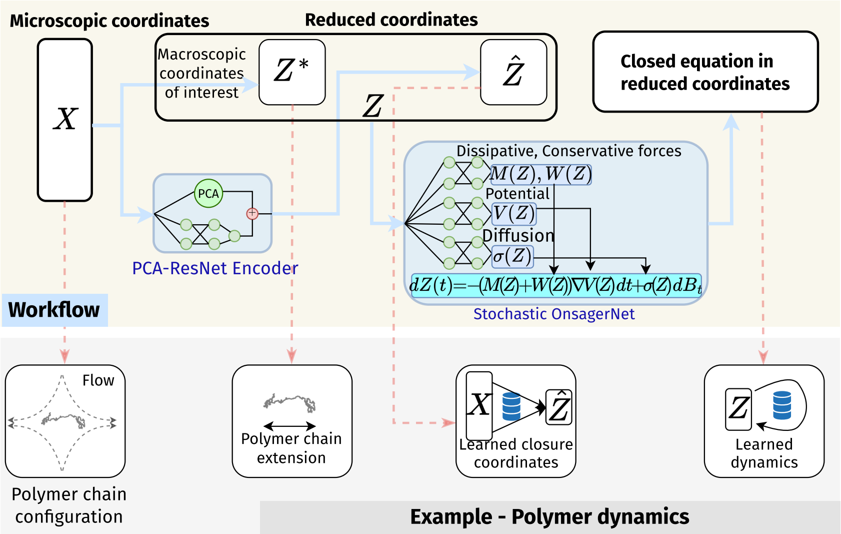
The stochastic OnsagerNet workflow



The stochastic OnsagerNet workflow

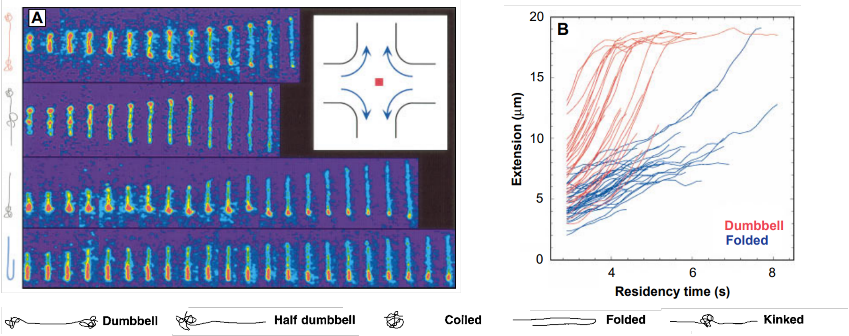


The stochastic OnsagerNet workflow



Case study:
polymer stretching dynamics

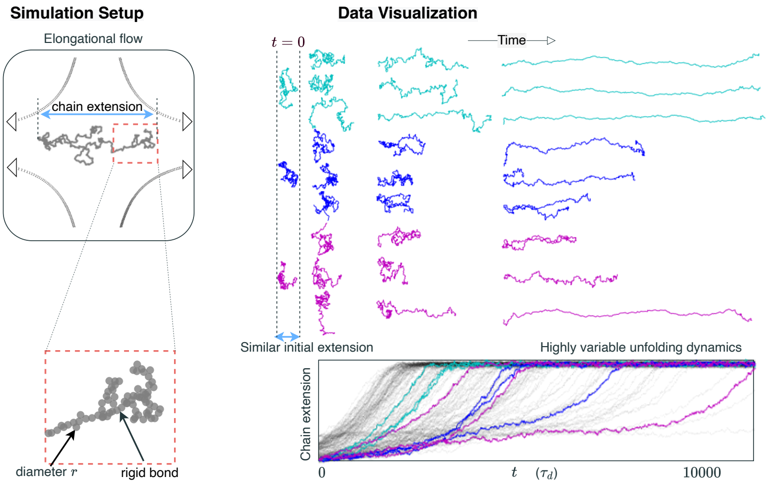
The closure problem for polymer dynamics



Key questions:

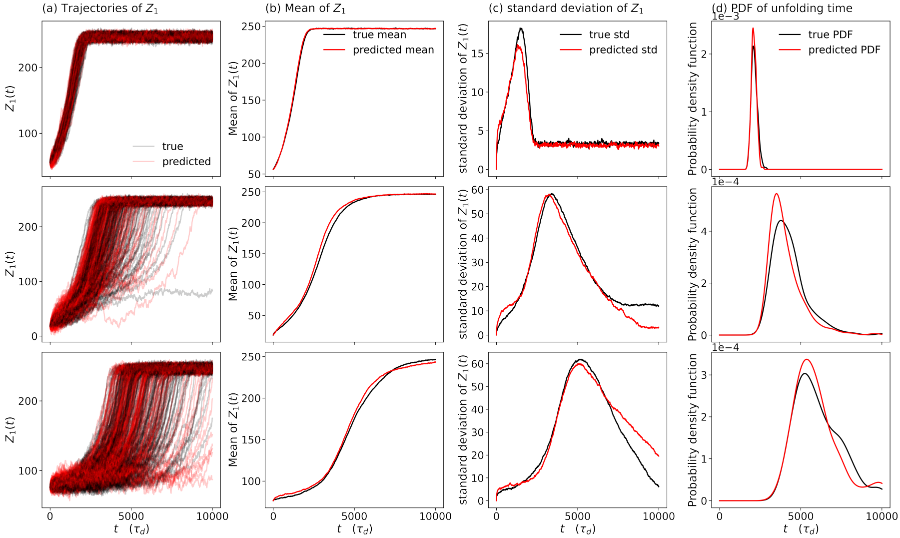
- Accurate dynamical model for the extension $Z^*(t)$?
- What are the closure coordinates that drive the dynamics?
- What can we learn about the dynamical landscape? Can we interact with it?

Simulation setup and training

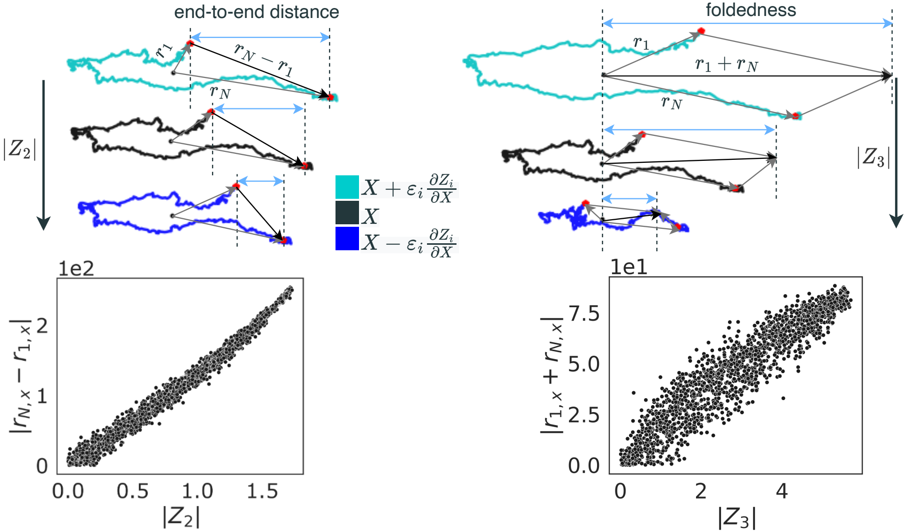


We train a S-OnsagerNet with 2 closure variables, leading to a 3D dynamics

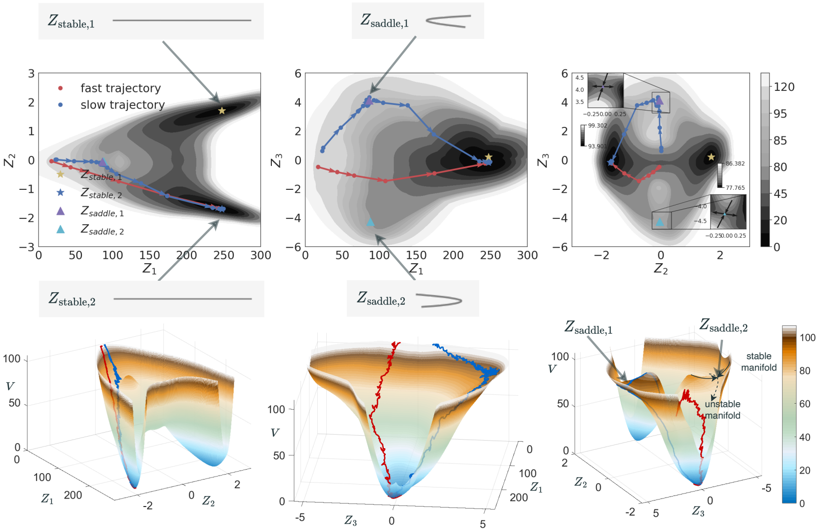
Capturing unfolding statistics



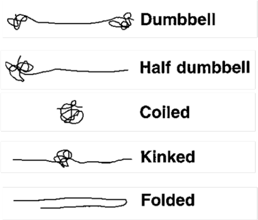
Finding sufficient descriptors for the stretching dynamics

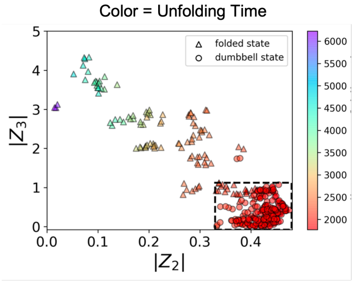


Visualising the energy landscape



Advancing the classification of polymer behaviour





Classical
(Qualitative)

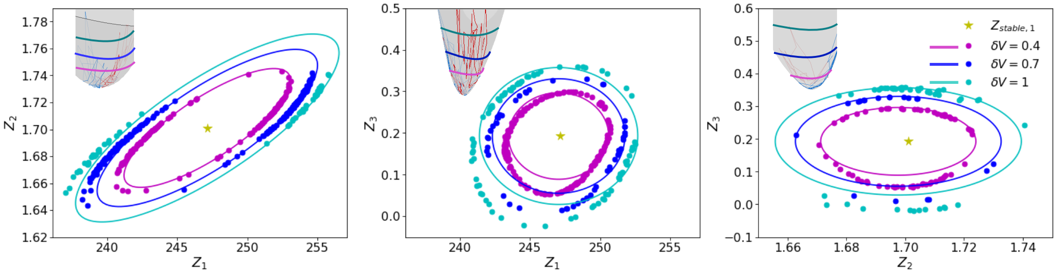
Stochastic OnsagerNet
(Quantitative)

Near Z_{stable} (stretched state), denote by $\delta V \sim k_B T$ the typical energy fluctuation. How does Z_1, Z_2, Z_3 fluctuate?

We obtain via automatic differentiation the expansion

$$k_B T \sim \delta V \approx a_1 (\delta Z_1 - a_4 \, \delta Z_2)^2 + a_2 \, \delta Z_2^2 + a_3 \, \delta Z_3^2, \qquad \delta Z_i = Z_i - [Z_{\text{stable}}]_i,$$

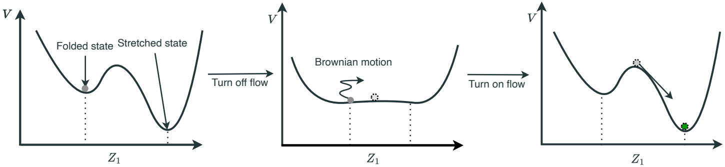
which can be viewed as a partial “equation of state”

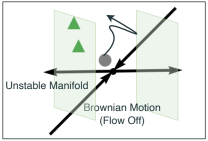


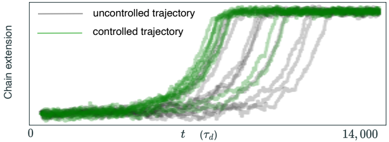
Another expansion near a saddle (folded state) gives

$$\delta V \approx b_1\,\delta Z_1^2 - b_2\,(\delta Z_2 - b_4\,\delta Z_3)^2 + b_3\,\delta Z_3^2.$$

Escape from folded state by increasing $|\delta Z_2 - b_4 \delta Z_3|$ – a control protocol!

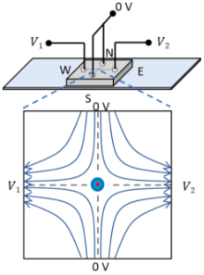




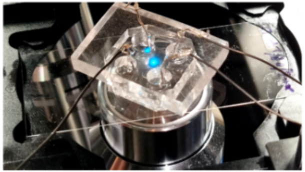


Extension to experimental data

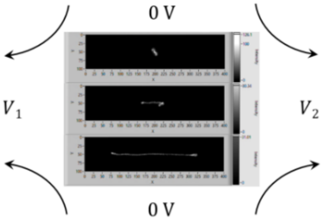
Experimental Setup



Microfluidic Device

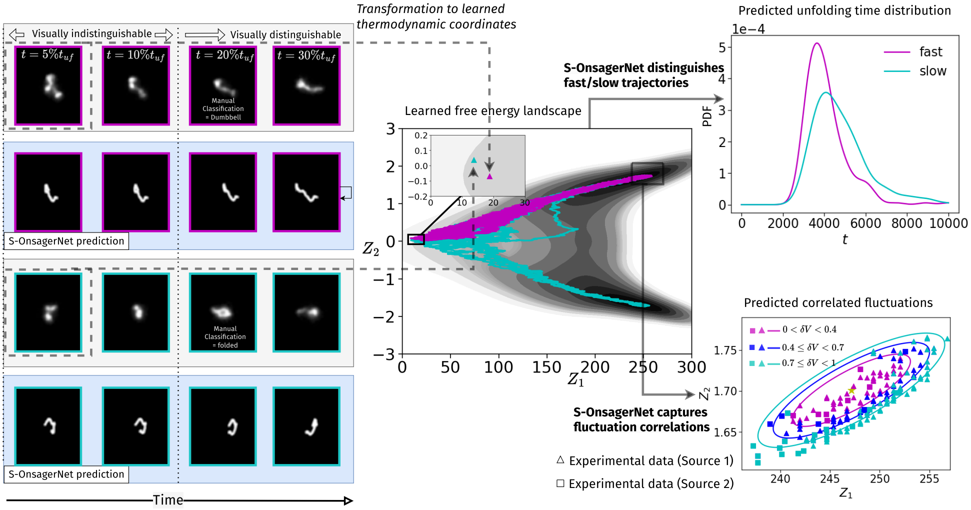


DNA Molecules Stretching



Dancing DNAs

Validation on experimental data



Summary

In this talk we discussed how to build a custom thermodynamic description of a dynamical system via machine learning that allows for Interpretation, analysis and control

http://arxiv.org/abs/2308.04119

Other applications of this idea

- Rayleigh-Bénard convection[8]
- Dynamics of intrinsic self-healing materials[9]

---

[8] H. Yu, X. Tian, W. E, Q. Li, *Physical Review Fluids* **6**, 114402 (Nov. 23, 2021).

[9] H. P. Anwar Ali, Z. Zhao, Y. J. Tan, W. Yao, Q. Li, B. C. K. Tee, *ACS Applied Materials & Interfaces* **14**, 52486–52498 (Nov. 23, 2022).

Outlook

Many methodological improvements

- Systematic identification of latent dimensions
- Parametric or controlled dynamics
- More general structures

More generally

- Mathematics of unstructured vs structured models
- Flexibility $\longleftrightarrow$ Interpretability

Thank you!

Joint work with:

Xiaoli Chen, Beatrice Soh, Eleonore Vissol-Gaudin, Zi En Ooi
Haijun Yu, Kedar Hippalgaonkar, Kostya Novoselov

X. Tian, Weinan E, Hashina P. Anwar Ali, Zichen Zhao, Benjamin C. K. Tee