Data-driven particle dynamics: Structure-preserving coarse-graining for emergent behavior in non-equilibrium systems

Quercus Hernández
School of Engineering and Applied Science
University of Pennsylvania
Emmi AI
quercus@emmi.ai

Max Win
School of Engineering and Applied Science
University of Pennsylvania
maxw2@seas.upenn.edu

Thomas O'Connor
Department of Materials Science and Engineering
Carnegie Mellon University
thomaso@andrew.cmu.edu

Paulo E. Arratia
School of Engineering and Applied Science
University of Pennsylvania
parratia@seas.upenn.edu

Nathaniel Trask
School of Engineering and Applied Science
University of Pennsylvania
ntrask@seas.upenn.edu

Abstract

Multiscale systems are common in science and engineering but remain difficult to simulate, as fine spatiotemporal scales must be consistently linked to emergent bulk behavior. Coarse-graining high-dimensional dynamics into low-dimensional models causes entropic information loss, producing dissipative, history-dependent, and stochastic effects. We present a metriplectic bracket based framework for learning coarse-grained dynamics from particle-trajectory time series, which enforces by construction the first and second laws of thermodynamics, momentum conservation, and discrete fluctuation–dissipation balance, crucial for capturing non-equilibrium statistics. After introducing the formalism, we specialize it to particle discretizations and develop a self-supervised strategy to recover unobserved entropic state variables. The method is applied to: (i) coarse-graining star polymers at extreme resolutions while retaining non-equilibrium statistics, and (ii) learning from high-speed video of colloidal suspensions to capture coupling between local rearrangements and emergent stochasticity. Our open-source PyTorch and LAMMPS implementations enable large-scale inference and extension to diverse particle-based systems.

1 Introduction

Multiscale phenomena, from quantum materials to geophysical flows, arise from interactions across widely separated length and time scales. Quantum effects can drive ferromagnetic ordering [1], and topological entanglement can cause anomalous diffusion in polymer melts [2]. At geological scales, bulk soil evolution connects to microstructural processes such as jamming in dense suspensions [3, 4]. Bridging these scales is difficult because direct simulation would require resolving prohibitively short timescales, so coarse-graining replaces fine-scale degrees of freedom with effective macroparticles

Machine Learning and the Physical Sciences Workshop, NeurIPS 2025.

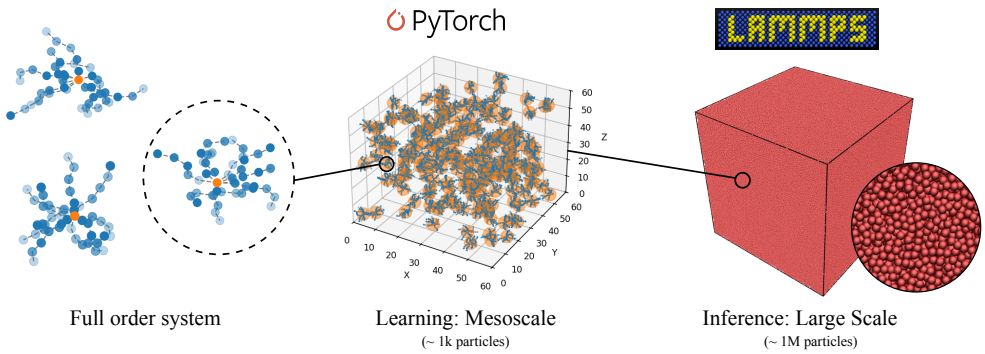

Figure 1: Data-driven particle dynamics pipeline to bridge scales in molecular systems. Detailed simulation of polymers fully resolve stochastic fluctuations (left), in-silico experiments of a small domain supervise model discovery (middle), yielding a data-driven model for non-equilibrium bulk response solved in a massive parallel simulator (right).

or filtered fields. Examples include coarse-grained MD, which preserves emergent physics while avoiding stiff dynamics, and continuum approaches such as LES or PCA-based reduced-order models. The Mori-Zwanzig (MZ) formalism [5, 6, 7, 8] gives a rigorous framework for projecting full-order dynamics onto coarse variables. This yields mean-field dynamics, a history-dependent dissipative memory kernel, and a stochastic term from unresolved physics. Even if the underlying system is reversible, coarse-graining introduces dissipation and stochasticity consistent with fluctuation-dissipation relations [9, 10]. In non-equilibrium systems, MZ offers a probabilistic description of energy transfer from unresolved to resolved variables.

2 Background

Force matching Force matching fits interatomic potentials by minimizing force errors against costly *ab initio* data. It is widely used with methods including GNNs [11, 12, 13], equivariant and invariant models [14, 15, 16, 17, 18, 19, 20], Transformers [21, 22, 23, 24], and generative diffusion models [25, 26]. Limitations include reliance on expensive quantum data, small benchmark sets, and inability to fit emergent mesoscale or continuum behaviors.

Coarse-graining models Classical particle-based coarse-graining such as dissipative particle dynamics (DPD) [27] and smoothed DPD [28] models conservative, dissipative, and noise terms from unresolved DOFs, often with MZ justification, and extends to multiphase [29], free-surface [30], and many-body flows [31]. ML analogues use autoencoders [32, 33, 34] or GNN-based graph U-Nets [35, 36], but typically fit to mean dynamics without embedding non-equilibrium structure.

Structure-preserving algorithms Other work embeds physical structure via learned Lagrangian [37] or Hamiltonian [38, 39, 40] formalisms, extended to dissipation through generalized variational principles [41, 42, 43], or Onsager brackets [44, 45]. We adopt the metriplectic (GENERIC) framework [46, 47], which augments Hamiltonian mechanics with an entropy-producing bracket. Although derivable from MZ [48], its algebra maps naturally to neural architectures. Prior ML metriplectic models has seen limited adoption: existing methods either rely on penalty-based constraints that fail to preserve structure exactly [49, 50], or require expensive bracket construction that limits scalability to small toy systems [51, 52, 53].

We propose a scalable $O(N)$ metriplectic particle architecture with localized pairwise interactions, integrable into large-scale simulators (for example LAMMPS), and released open source. The model is interpretable, with subnetworks predicting physically meaningful quantities such as temperature, pressure, energy, and entropy, and self-supervised, inferring unmeasurable internal variables consistently with thermodynamics [54, 32]. This enables learning emergent physics driven by stochastic unresolved processes.

3 Method

Let $\boldsymbol{x}$ be a set of a dynamical system state variables. In metriplectic dynamics [46, 47], the evolution of $\boldsymbol{x}$ is given by

$$d\boldsymbol{x} = \left[\boldsymbol{L}\frac{\partial E}{\partial \boldsymbol{x}} + \boldsymbol{M}\frac{\partial S}{\partial \boldsymbol{x}} + k_B\frac{\partial}{\partial \boldsymbol{x}}\cdot\boldsymbol{M}\right]dt + d\tilde{\boldsymbol{x}}. \tag{1}$$

where $E(\boldsymbol{x})$ and $S(\boldsymbol{x})$ are functionals prescribing generalized energy and entropy of the system, $\boldsymbol{L}(\boldsymbol{x})$ is a skew-symmetric Poisson matrix and $\boldsymbol{M}(\boldsymbol{x})$ is a symmetric positive semi-definite friction matrix. These two terms prescribe the reversible and irreversible dynamics of the system, respectively. By imposing the degeneracy conditions, $\boldsymbol{L}\boldsymbol{\nabla}S = \boldsymbol{M}\boldsymbol{\nabla}E = \boldsymbol{0}$, the dynamics satisfy the first ($\dot{E} = 0$) and second ($\dot{S} \geq 0$) laws of thermodynamics independent of the choice of state variables. The stochastic effects $d\tilde{\boldsymbol{x}}$ are directly related to the dissipative effects via the fluctuation dissipation theorem [9, 10]: $d\tilde{\boldsymbol{x}}d\tilde{\boldsymbol{x}}^T = 2k_B\boldsymbol{M}dt$, where k_B is the Boltzmann constant.

Consider now a collection of N particles with positions $\boldsymbol{r}_i$ and velocities $\boldsymbol{v}_i = \frac{d\boldsymbol{r}_i}{dt}$ for $i \in V = \{1, ..., N\}$. We introduce energy and entropy functionals

$$E = \sum_i \left[\frac{1}{2}m_i\boldsymbol{v}_i^2 + U_i\right], \quad S = \sum_i S_i, \tag{2}$$

where: $m_i = m$ is a particle mass, assumed to be constant; S_i is a per particle entropy; $U_i = U(\mathcal{V}_i, S_i)$ is a function prescribing per particle internal energy; and $\mathcal{V}_i(\boldsymbol{r}_1, ..., \boldsymbol{r}_N)$ is a per particle volume. We select position and velocity as reversible variables and entropy as irreversible variable such that $\boldsymbol{x}_i = (\boldsymbol{x}_{\text{rev}}, \boldsymbol{x}_{\text{irr}}) = (\boldsymbol{r}_i, \boldsymbol{v}_i, S_i)$. Substituting these expressions into Eq. 1 and making use of the first law of thermodynamics $dU = TdS - Pd\mathcal{V}$ we obtain per particle dynamics

$$\underbrace{\begin{bmatrix} d\boldsymbol{r}_i \\ d\boldsymbol{v}_i \\ dS_i \end{bmatrix}}_{d\boldsymbol{x}_i} = \sum_j \left(\underbrace{\frac{1}{m}\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}\delta_{ij} & \boldsymbol{0} \\ -\boldsymbol{I}\delta_{ij} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{L}_{ij}} \underbrace{\begin{bmatrix} \frac{\partial U}{\partial \boldsymbol{r}_j} \\ m\boldsymbol{v}_j \\ T_j \end{bmatrix}}_{\frac{\partial E}{\partial \boldsymbol{x}_j}} + \boldsymbol{M}_{ij} \underbrace{\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ 1 \end{bmatrix}}_{\frac{\partial S}{\partial \boldsymbol{x}_j}} + k_B\frac{\partial}{\partial \boldsymbol{x}_j}\cdot\boldsymbol{M}_{ij} \right)dt + \underbrace{\begin{bmatrix} d\tilde{\boldsymbol{r}}_i \\ d\tilde{\boldsymbol{v}}_i \\ d\tilde{S}_i \end{bmatrix}}_{d\tilde{\boldsymbol{x}}_i}. \tag{3}$$

To close this system of equations, we will introduce GNN-style node and edge-wise neural networks shared across all particles to define:

Particle volume To obtain a method which scales as $O(N)$ in the number of particles, we prescribe $\mathcal{V}$ to include only contributions from particles within a neighborhood of radius h, denoted as $\mathcal{N}_i = \{j \in V, |\boldsymbol{r}_{ij}| < h\}$ and parameterize the volume

$$\mathcal{V}_i^{-1} = d_i = \sum_{j\in\mathcal{N}_i} W_{ij} = \sum_{j\in\mathcal{N}_i} \exp\left[\text{MLP}\left(|\boldsymbol{r}_{ij}|/h\right)\right]\left(1 - |\boldsymbol{r}_{ij}|^2/h^2\right), \tag{4}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$ denotes the relative positions between particle i and neighboring particles $j \in \mathcal{N}_i$, and $|\cdot|$ denotes the Euclidean norm.

Per particle internal energy In classical variational mechanics, Gibbs relations require internal energy to satisfy convexity and monotonicity constraints for hyperbolicity. Physically, pressure and temperature must be positive, implying $\partial_{\mathcal{V}}U \leq 0$ and $\partial_S U \geq 0$, with positive second derivatives (isothermal compressibility and specific heat), so U is convex in both variables. We use the Constrained Monotonic Neural Network [55], which enforces these monotonicities by construction: $U_i = \text{CMNN}(S_i, \mathcal{V}_i)$.

Thermal fluctuations Motivated by [28], we assume a functional form $d\tilde{\boldsymbol{x}}_i = (\boldsymbol{0}, d\tilde{\boldsymbol{v}}_i, d\tilde{S}_i)$ for which random forcing in $d\tilde{\boldsymbol{v}}$ consists only of pairwise interactions, from which the degeneracy conditions enforce as a consequence. In this work, instead of choosing the noise amplitudes A, B, and C to arrive at an accurate discretization of the fluctuating Navier-Stokes equations, we machine learn them to identify state-dependent nonlinear scalings of the fluctuations, ensuring the symmetry over particle pair indices by construction and the functional dependency of viscosity with temperature: $A_{ij}, B_{ij}, C_{ij} = \text{MLP}\left(|\boldsymbol{r}_{ij}|/h, T_i\right) \cdot \text{MLP}\left(|\boldsymbol{r}_{ij}|/h, T_j\right)$.

Entropy self-supervision While position and velocity are easily measurable either in experimental or by post-processed synthetic data, calculating entropy would require an enumeration of microstates which is intractable for the vast majority of cases. To fully supervise the system, we therefore need to generate time-series labels of S_i. In this work we adopt a self-supervised approach to generate labels for S_i, based on an edge convolutional operator [56] which constructs a closure for the entropy in terms of position and velocity: $S_i = \text{EdgeConv}\left(|\boldsymbol{r}_{ij}|/h, \boldsymbol{v}_{ij}\right)$.

Training procedure Let $\mathcal{D} = \{(\boldsymbol{x}_{\text{rev}}^t, \boldsymbol{x}_{\text{rev}}^{t+1})\}_{t=0}^{N_{\text{train}}}$ be a dataset consisting of position-velocity pairs. To train the system, we construct a joint probability distribution $p(\boldsymbol{x}_{\text{rev}}^{t+1}, \boldsymbol{x}_{\text{irr}}^{t+1} | \boldsymbol{x}_{\text{rev}}^t, \boldsymbol{x}_{\text{irr}}^t)$ corresponding to a numerical integration of Equation 1 and train the architectures with maximum likelihood. We denote all trainable parameters via Θ and define the maximum log-likelihood objective

$$\Theta^* = \arg\min_{\Theta} \text{NLL}(x^1, ..., x^T). \tag{5}$$

By marginalizing over the rest of the particles, we can split the problem as per particle-wise loss corresponding to the multivariate Gaussian distribution:

$$\boldsymbol{\mu}_i^{t+1} = \begin{bmatrix} \boldsymbol{v}_i^t + \frac{d\boldsymbol{v}_i^t}{dt}\Delta t \\ S_i^t + \frac{dS_i^t}{dt}\Delta t \end{bmatrix}, \quad \boldsymbol{\Sigma}_{ii}^{t+1} = \begin{bmatrix} \Delta\tilde{\boldsymbol{v}}_i \Delta\tilde{\boldsymbol{v}}_i^T & \Delta\tilde{\boldsymbol{v}}_i \Delta\tilde{S}_i \\ \Delta\tilde{S}_i \Delta\tilde{\boldsymbol{v}}_i^T & \Delta\tilde{S}_i \Delta\tilde{S}_i \end{bmatrix}. \tag{6}$$

where Δt is the simulation time step and $\Delta\tilde{\boldsymbol{v}}_i, \Delta\tilde{S}_i$ correspond to the discretized Wiener processes of velocity and entropy respectively. Thus the final loss function is defined as

$$\mathcal{L} = \frac{1}{N_{\text{train}}} \sum_{t=0}^{N_{\text{train}}} \frac{1}{N} \sum_{i=0}^{N} \Big[\frac{1}{2} \ln |\boldsymbol{\Sigma}_{ii}^{t+1}| + \frac{1}{2} (\boldsymbol{x}_i^{t+1} - \boldsymbol{\mu}_i^{t+1})^T \left(\boldsymbol{\Sigma}_{ii}^{t+1}\right)^{-1} (\boldsymbol{x}_i^{t+1} - \boldsymbol{\mu}_i^{t+1}) \Big], \tag{7}$$

which can be minimized using standard batch gradient descent. The proposed architecture and training procedure has the following desirable properties:

Theorem 3.1. *The equations of motion in Eq. 3: (i) satisfy energy conservation and entropy inequality by construction, (ii) satisfy momentum conservation, (iii) enforce the fluctuation-dissipation theorem exactly, and (iv) the addition of new types of physics is trivial by adding additional terms to the energy and entropy functionals in Eq. 2. Furtheremore, the training procedure in Eq. 7: (v) can be supervised by only using position/velocity information, and (vi) scales linearly $O(N)$ with the number of particles, being able to simulate massive-scale systems.*

Proof. The complete derivations can be found in [57]. □

4 Experiments

Molecular dynamics: Coarse-graining of star polymers First, we consider a star polymer melt in a 3D periodic box, where each polymer has a core atom bonded to 10 arms. Two internal configurations are tested: 1 and 5 beads per arm, corresponding to lower and higher coarse-graining levels. Fully resolved datasets are generated with LAMMPS [58] and coarse-grained via the center-of-mass position and velocity of each polymer. The method is benchmarked against a Graph Network-based Solver [59] (GNS), its stochastic variant (GNS-SDE), and a calibrated DPD model [27] in extrapolation rollouts 25 times the training horizon.

Table 1 shows good agreement with spatial (RDF) and dynamical (MSD) statistics, while GNS and GNS-SDE fail to match flow statistics or outperform DPD, lacking spatial structure and correlations. This underscores the qualitative impact of preserving physical structure, even in simple systems, and shows that "black-box" approaches are poor candidates for coarse-graining.

Extension to general physics: Experimental colloidal system The above energy choice suits fluid-like systems that store energy only via dilatation. To extend this to broader physics, we redefine the potential energy to include a tensorial strain dependence appropriate for solids, $U_i = U_i^{\text{vol}} + U_i^{\text{dev}} = \text{CMNN}(S_i, \mathcal{V}_i) + \text{CMNN}(S_i, \bar{\boldsymbol{\varepsilon}}_i)$, where $\bar{\boldsymbol{\varepsilon}}_i$ is the traceless strain tensor estimated similarly to $\mathcal{V}_i$. This addition accounts for reversible shear stress storage absent in fluids. Dissipative and noise

Table 1: L2 relative error (↓) of correlation metrics for the STAR POLYMER and COLLOIDS examples.

	STAR POLYMER 11				STAR POLYMER 51				COLLOIDS
	GNS	GNS-SDE	DPD	Ours	GNS	GNS-SDE	DPD	Ours	Ours
RDF	4.3e-1	4.3e-1	2.3e-1	**2.0e-2**	2.1e-1	2.2e-1	9.6e-2	**5.3e-2**	**1.15e-1**
MSD	2.1e6	6.5e5	9.5e-1	**4.7e-2**	2.1e1	1.6e7	7.6e0	**5.3e-2**	-

terms are unaffected; the case illustrates how to incorporate additional reversible physics into the framework.

We apply the method to measured particle trajectories from a jammed 2D monolayer of repulsive polymer microspheres under cyclic shear [60]. Shearing is performed with an interfacial stress rheometer [60, 61], and high-resolution optical microscopy yields particle positions and finite-difference velocities. The colloidal particles themselves act as the coarse-grained degrees of freedom, representing a system that combines fluid, electrostatic, elastic, and contact effects, demonstrating the framework's broad applicability.

Table 1 shows the model accurately predicts RDF statistics. We also predict emerging behavior such as local particle rearrangements, and dislocation distribution trends over time [57]. Given the challenge of predicting these kind of dynamics mechanisms [61], it is notable that this rare-event mechanism emerges from top-down training.

5 Conclusions

We have introduced a deep learning approach to model conservative, dissipative, and stochastic effects in compliance with thermodynamics and the fluctuation-dissipation theorem. Supervised only by position and velocity time series, the method uses a self-supervised scheme to identify emergent entropic variables governing dissipation. It reproduces coarse-grained molecular responses from top-down information alone, offering an alternative to bottom-up models from density functional theory. By modifying the internal energy functional, the framework accommodates diverse physics, including viscoelasticity and electrokinetic suspensions. The methodology is largely agnostic to the architecture of the learnable components. It could be enhanced through more advanced designs such as cross-attention [62], parametric conditioning, or graph autoencoders mapping microstructure to entropy variables [63, 64].

Acknowledgments and Disclosure of Funding

We thank Pep Español and George Karniadakis for their insightful feedback and valuable discussions throughout the development of this work. We also acknowledge Joel T. Clemmer for his assistance with the LAMMPS code implementation and Aleksandra Vojvodic for the computational resources.

References

[1] Jin Soo Lim, Diomedes Saldana-Greco, and Andrew M Rappe. Improper magnetic ferroelectricity of nearly pure electronic nature in helicoidal spiral camn 7 o 12. *Physical Review B*, 97(4):045115, 2018.

[2] Thomas C O'Connor, Ting Ge, Michael Rubinstein, and Gary S Grest. Topological linking drives anomalous thickening of ring polymers in weak extensional flows. *Physical review letters*, 124(2):027801, 2020.

[3] Douglas J Jerolmack and Karen E Daniels. Viewing earth's surface as a soft-matter landscape. *Nature Reviews Physics*, 1(12):716–730, 2019.

[4] Robert Kostynick, Hadis Matinpour, Shravan Pradeep, Sarah Haber, Alban Sauret, Eckart Meiburg, Thomas Dunne, Paulo Arratia, and Douglas Jerolmack. Rheology of debris flow materials is controlled by the distance from jamming. *Proceedings of the National Academy of Sciences*, 119(44):e2209109119, 2022.

[5] Robert Zwanzig. Ensemble method in the theory of irreversibility. *The Journal of Chemical Physics*, 33(5):1338–1341, 1960.

[6] Hazime Mori. Transport, collective motion, and brownian motion. *Progress of theoretical physics*, 33(3):423–455, 1965.

[7] Hermann Grabert. *Projection operator techniques in nonequilibrium statistical mechanics*, volume 95. Springer, 2006.

[8] Robert Zwanzig. *Nonequilibrium statistical mechanics*. Oxford university press, 2001.

[9] Herbert B Callen and Theodore A Welton. Irreversibility and generalized noise. *Physical Review*, 83(1):34, 1951.

[10] Rep Kubo. The fluctuation-dissipation theorem. *Reports on progress in physics*, 29(1):255, 1966.

[11] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. In *International conference on machine learning*, pages 1263–1272. PMLR, 2017.

[12] Anuroop Sriram, Abhishek Das, Brandon M Wood, Siddharth Goyal, and C Lawrence Zitnick. Towards training billion parameter graph neural networks for atomic simulations. *arXiv preprint arXiv:2203.09697*, 2022.

[13] Linfeng Zhang, Jiequn Han, Han Wang, Roberto Car, and Weinan E. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Physical review letters*, 120(14):143001, 2018.

[14] Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219*, 2018.

[15] Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco S Cohen. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. *Advances in Neural information processing systems*, 31, 2018.

[16] Risi Kondor, Zhen Lin, and Shubhendu Trivedi. Clebsch–gordan nets: a fully fourier space spherical convolutional neural network. *Advances in Neural Information Processing Systems*, 31, 2018.

[17] Johannes Brandstetter, Rob Hesselink, Elise van der Pol, Erik J Bekkers, and Max Welling. Geometric and physical quantities improve e (3) equivariant message passing. *arXiv preprint arXiv:2110.02905*, 2021.

[18] Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems*, 30, 2017.

[19] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123*, 2020.

[20] Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations (ICLR)*, 2022.

[21] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[22] Philipp Thölke and Gianni De Fabritiis. Torchmd-net: equivariant transformers for neural network based molecular potentials. *arXiv preprint arXiv:2202.02541*, 2022.

[23] Tuan Le, Frank Noé, and Djork-Arné Clevert. Equivariant graph attention networks for molecular property prediction. *arXiv preprint arXiv:2202.09891*, 2022.

[24] Yi-Lun Liao and Tess Smidt. Equiformer: Equivariant graph attention transformer for 3d atomistic graphs. *arXiv preprint arXiv:2206.11990*, 2022.

[25] Bowen Jing, Hannes Stärk, Tommi Jaakkola, and Bonnie Berger. Generative modeling of molecular dynamics trajectories. *arXiv preprint arXiv:2409.17808*, 2024.

[26] Majdi Hassan, Nikhil Shenoy, Jungyoon Lee, Hannes Stärk, Stephan Thaler, and Dominique Beaini. Et-flow: Equivariant flow-matching for molecular conformer generation. *Advances in Neural Information Processing Systems*, 37:128798–128824, 2024.

[27] Robert D Groot and Patrick B Warren. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *The Journal of chemical physics*, 107(11):4423–4435, 1997.

[28] Pep Espanol and Mariano Revenga. Smoothed dissipative particle dynamics. *Physical Review E*, 67(2):026705, 2003.

[29] Domenico Davide Meringolo, Agostino Lauria, Francesco Aristodemo, and Pasquale Fabio Filianoti. Large eddy simulation within the smoothed particle hydrodynamics: Applications to multiphase flows. *Physics of Fluids*, 35(6), 2023.

[30] Jiannong Fang, Aurèle Parriaux, Martin Rentschler, and Christophe Ancey. Improved sph methods for simulating free surface flows of viscous fluids. *Applied numerical mathematics*, 59(2):251–271, 2009.

[31] Yining Han, Jaehyeok Jin, and Gregory A Voth. Constructing many-body dissipative particle dynamics models of fluids from bottom-up coarse-graining. *The Journal of Chemical Physics*, 154(8), 2021.

[32] Wujie Wang and Rafael Gómez-Bombarelli. Coarse-graining auto-encoders for molecular dynamics. *npj Computational Materials*, 5(1):125, 2019.

[33] Wujie Wang and Rafael Gómez-Bombarelli. Learning coarse-grained particle latent space with auto-encoders. *Adv. Neural Inf. Process. Syst*, 1, 2019.

[34] Agnieszka Ilnicka and Gisbert Schneider. Designing molecules with autoencoder networks. *Nature Computational Science*, 3(11):922–933, 2023.

[35] Brooke E Husic, Nicholas E Charron, Dominik Lemm, Jiang Wang, Adrià Pérez, Maciej Majewski, Andreas Krämer, Yaoyi Chen, Simon Olsson, Gianni De Fabritiis, et al. Coarse graining molecular dynamics with graph neural networks. *The Journal of chemical physics*, 153(19), 2020.

[36] Xiang Fu, Tian Xie, Nathan J Rebello, Bradley D Olsen, and Tommi Jaakkola. Simulate time-integrated coarse-grained molecular dynamics with multi-scale graph networks. *arXiv preprint arXiv:2204.10348*, 2022.

[37] Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks. *arXiv preprint arXiv:2003.04630*, 2020.

[38] Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. *Advances in neural information processing systems*, 32, 2019.

[39] Marco David and Florian Méhats. Symplectic learning for hamiltonian neural networks. *Journal of Computational Physics*, 494:112495, 2023.

[40] Zhengdao Chen, Jianyu Zhang, Martin Arjovsky, and Léon Bottou. Symplectic recurrent neural networks. *arXiv preprint arXiv:1909.13334*, 2019.

[41] Yaofeng Desmond Zhong, Biswadip Dey, and Amit Chakraborty. Dissipative symoden: Encoding hamiltonian dynamics with dissipation and control into deep learning. *arXiv preprint arXiv:2002.08860*, 2020.

[42] Andrew Sosanya and Sam Greydanus. Dissipative hamiltonian neural networks: Learning dissipative and conservative dynamics separately. *arXiv preprint arXiv:2201.10085*, 2022.

[43] Shanshan Xiao, Jiawei Zhang, and Yifa Tang. Generalized lagrangian neural networks. *arXiv preprint arXiv:2401.03728*, 2024.

[44] Haijun Yu, Xinyuan Tian, Weinan E, and Qianxiao Li. Onsagernet: Learning stable and interpretable dynamics using a generalized onsager principle. *Physical Review Fluids*, 6(11):114402, 2021.

[45] Shenglin Huang, Zequn He, and Celia Reina. Variational onsager neural networks (vonns): A thermodynamics-based variational learning strategy for non-equilibrium pdes. *Journal of the Mechanics and Physics of Solids*, 163:104856, 2022.

[46] Miroslav Grmela and Hans Christian Öttinger. Dynamics and thermodynamics of complex fluids. i. development of a general formalism. *Physical Review E*, 56(6):6620, 1997.

[47] Hans Christian Öttinger and Miroslav Grmela. Dynamics and thermodynamics of complex fluids. ii. illustrations of a general formalism. *Physical Review E*, 56(6):6633, 1997.

[48] Alexander Mielke, Mark A Peletier, and Johannes Zimmer. Deriving a generic system from a hamiltonian system. *arXiv preprint arXiv:2404.09284*, 2024.

[49] Quercus Hernández, Alberto Badías, David González, Francisco Chinesta, and Elías Cueto. Structure-preserving neural networks. *Journal of Computational Physics*, 426:109950, 2021.

[50] Quercus Hernández, Alberto Badías, Francisco Chinesta, and Elías Cueto. Thermodynamics-informed graph neural networks. *IEEE Transactions on Artificial Intelligence*, 5(3):967–976, 2022.

[51] Zhen Zhang, Yeonjong Shin, and George Em Karniadakis. Gfinns: Generic formalism informed neural networks for deterministic and stochastic dynamical systems. *Philosophical Transactions of the Royal Society A*, 380(2229):20210207, 2022.

[52] Kookjin Lee, Nathaniel Trask, and Panos Stinis. Machine learning structure preserving brackets for forecasting irreversible processes. *Advances in Neural Information Processing Systems*, 34:5696–5707, 2021.

[53] Anthony Gruber, Kookjin Lee, and Nathaniel Trask. Reversible and irreversible bracket-based dynamics for deep graph neural networks. *Advances in Neural Information Processing Systems*, 36:38454–38484, 2023.

[54] Thomas N Thiem, Mahdi Kooshkbaghi, Tom Bertalan, Carlo R Laing, and Ioannis G Kevrekidis. Emergent spaces for coupled oscillators. *Frontiers in computational neuroscience*, 14:36, 2020.

[55] Davor Runje and Sharath M Shankaranarayana. Constrained monotonic neural networks. In *International Conference on Machine Learning*, pages 29338–29353. PMLR, 2023.

[56] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, 38(5):1–12, 2019.

[57] Quercus Hernandez, Max Win, Thomas C. O'Connor, Paulo E. Arratia, and Nathaniel Trask. Data-driven particle dynamics: Structure-preserving coarse-graining for emergent behavior in non-equilibrium systems, 2025.

[58] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.

[59] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to simulate complex physics with graph networks. In *International conference on machine learning*, pages 8459–8468. PMLR, 2020.

[60] Nathan C Keim and Paulo E Arratia. Mechanical and microscopic properties of the reversible plastic regime in a 2d jammed material. *Physical review letters*, 112(2):028302, 2014.

[61] K Lawrence Galloway, Xiaoguang Ma, Nathan C Keim, Douglas J Jerolmack, Arjun G Yodh, and Paulo E Arratia. Scaling of relaxation and excess entropy in plastically deformed amorphous solids. *Proceedings of the National Academy of Sciences*, 117(22):11887–11893, 2020.

[62] A Waswani, N Shazeer, N Parmar, J Uszkoreit, L Jones, A Gomez, L Kaiser, and I Polosukhin. Attention is all you need. In *NIPS*, 2017.

[63] Junhyun Lee, Inyeop Lee, and Jaewoo Kang. Self-attention graph pooling. In *International conference on machine learning*, pages 3734–3743. pmlr, 2019.

[64] Zhen Zhang, Jiajun Bu, Martin Ester, Jianfeng Zhang, Chengwei Yao, Zhi Yu, and Can Wang. Hierarchical graph pooling with structure learning. *arXiv preprint arXiv:1911.05954*, 2019.