PHYSGNO: PHYSICS-GUIDED GRAPH NEURAL OPERATORS FOR MOLECULAR DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Molecular dynamics (MD) plays a central role in scientific applications such as drug discovery, materials design, and chemical reaction analysis. Recent advances in deep learning have further expanded the capabilities of MD modeling. In particular, graph neural network-based interatomic potentials have achieved remarkable success in modeling molecular interactions, enabling accurate prediction of molecular energies and forces while effectively preserving physical information. However, these methods still rely on a large number of iterative integration steps to simulate long-horizon trajectories. Neural operators provide a complementary paradigm by learning trajectory evolution without step-by-step integration, thereby offering significantly improved efficiency. Nevertheless, such data-driven approaches often lack sufficient physical priors, which limits their ability to produce accurate molecular trajectories. To address these challenges, we propose PhysGNO, a physics-guided neural operator for molecular trajectory learning. PhysGNO leverages physical information captured by potential energy networks to guide trajectory evolution, combining the efficiency of neural operators with physically meaningful inductive bias for accurate trajectory prediction. In addition, PhysGNO incorporates molecular flexibility and rich vector-valued features into its spatiotemporal evolution layers, making it better suited to molecular dynamics systems. Extensive experiments on MD17, rMD17, and MD22 demonstrate the effectiveness and efficiency of the proposed method.

1 INTRODUCTION

Molecular dynamics (MD) provides an atomistic view of how the three-dimensional trajectories of molecular systems evolve over time, and has become an important computational tool across chemistry, biology, and materials science (van Gunsteren and Berendsen, 1990; Hollingsworth and Dror, 2018; Allen and Tildesley, 2017). For example, in computational biology and drug discovery, MD has been widely used to investigate protein dynamics, ligand binding, and biomolecular mechanisms (Durrant and McCammon, 2011; Hollingsworth and Dror, 2018). Conventional atomistic MD often relies on a predefined potential-energy function to evaluate the energy and forces at each time step according to Newtonian mechanics. A numerical integrator, such as the Verlet integrator, is then used to iteratively update atomic positions and velocities over many small time steps to obtain a complete molecular trajectory (van Gunsteren and Berendsen, 1990; Durrant and McCammon, 2011). Consequently, despite their strong physical foundations and interpretability, conventional MD methods remain computationally demanding for long-horizon molecular trajectory simulation.

Deep learning has substantially improved the efficiency and accuracy of molecular interaction modeling. Neural potential models, such as Behler–Parrinello networks and SchNet, have demonstrated strong representation capabilities for approximating molecular energies and forces at a fraction of the cost of first-principles calculations (Behler and Parrinello, 2007; Schütt et al., 2017). Subsequent works, including PaiNN and NequIP (Schütt et al., 2021; Batzner et al., 2022b), further incorporate equivariant geometric structures into molecular models, enabling more faithful modeling of rotational and translational symmetries in dynamical systems. More recently, methods such as Allegro, MACE, and ViSNet have further improved the accuracy, scalability, and expressiveness of learned potential networks (Musaelian et al., 2023; Batatia et al., 2022; Wang et al., 2024). These advances have significantly enhanced molecular energy and force prediction while retaining useful physical and geometric inductive biases. However, when such potential models are used for trajectory simulation,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

they still rely on numerical integrators to evolve the system sequentially in a step-by-step manner, which limits their efficiency in long-horizon trajectory prediction.

Neural operators provide a complementary paradigm for modeling dynamical systems by learning mappings between function spaces rather than explicitly resolving every intermediate numerical step (Li et al., 2021; Kovachki et al., 2023). This perspective has recently been extended to physical and molecular dynamics. ESTAG reformulates physical simulation as an equivariant spatio-temporal prediction problem, using historical trajectories to capture temporal dependencies beyond frame-to-frame modeling (Wu et al., 2023). EGNO further introduces an equivariant graph neural operator that represents three-dimensional dynamics directly as trajectory functions, enabling finite-time prediction while preserving SE(3)-equivariance (Xu et al., 2024). More recently, ATOM develops a pretrained transformer neural operator for molecular dynamics and supports parallel prediction of multiple future molecular states across different molecules and time horizons (Thompson et al., 2026). These methods demonstrate the strong capability of operator learning to reduce the dependence on long sequential rollouts in molecular dynamics. Despite this progress, existing trajectory operators are primarily optimized from trajectory data alone and do not explicitly exploit useful physical information, such as molecular energies and forces, which may limit the physical plausibility and prediction accuracy of the generated molecular trajectories.

A natural question therefore arises: *can we retain the efficiency of neural operators while effectively exploiting the physical information inherent in molecular systems?* To answer this question, we propose **PhysGNO**, a physics-guided graph neural operator for effective and efficient molecular trajectory modeling. Our key insight is that pretrained potential-energy models have already learned informative representations of molecular interactions and local physical quantities, particularly potential energies and their induced forces. Rather than using the potential model as a simulator that repeatedly advances the molecular system through numerical integration, we leverage the physical information encoded in the potential model as a prior to guide the learning of a finite-time neural operator. Specifically, we first introduce a **physical conditioning** mechanism inspired by Newtonian dynamics to inject energy- and force-related information extracted from a pretrained potential model into the operator for trajectory evolution. Furthermore, molecular dynamical systems contain rich directional information, including atomic velocities and interatomic forces, while different molecules may exhibit substantially different degrees of flexibility. We therefore incorporate diverse vector-valued dynamical features into an **equivariant spatial module** and a flexibility-aware temporal module, allowing the operator to better capture spatial and temporal dependencies within molecular systems. Together, these components integrate physical guidance with spatial and temporal inductive biases tailored to molecular dynamics. Extensive experiments on MD17, rMD17, and MD22 evaluate both trajectory accuracy and computational efficiency, demonstrating the effectiveness of PhysGNO.

Our main contributions can be summarized as follows:

- We propose **PhysGNO**, a physics-guided graph neural operator that bridges physical priors and finite-time molecular trajectory modeling. Instead of relying on repeated potential evaluation and numerical integration, PhysGNO leverages the physical information encoded in pretrained potential-energy models to guide neural operator learning, combining physically meaningful guidance with efficient trajectory prediction.
- We leverage a spatial–temporal modeling framework tailored to molecular dynamical systems, which incorporates diverse vector-valued dynamical features for spatial interaction modeling and molecular flexibility information for temporal evolution.
- We conduct extensive experiments on MD17, rMD17, and MD22, demonstrating the effectiveness of PhysGNO for molecular trajectory prediction and systematically evaluating its long-horizon modeling capability and computational efficiency.

2 Related Work

Deep Learning for Molecular Dynamics. Molecular dynamics trajectories describe how the three-dimensional configurations of molecular systems evolve over time. Conventional MD generates trajectories using predefined force fields or electronic-structure calculations, such as density functional theory (DFT), followed by numerical integration, which can be computationally expensive for long-time simulations (Marx and Hutter, 2009). Deep learning has substantially accelerated this process by

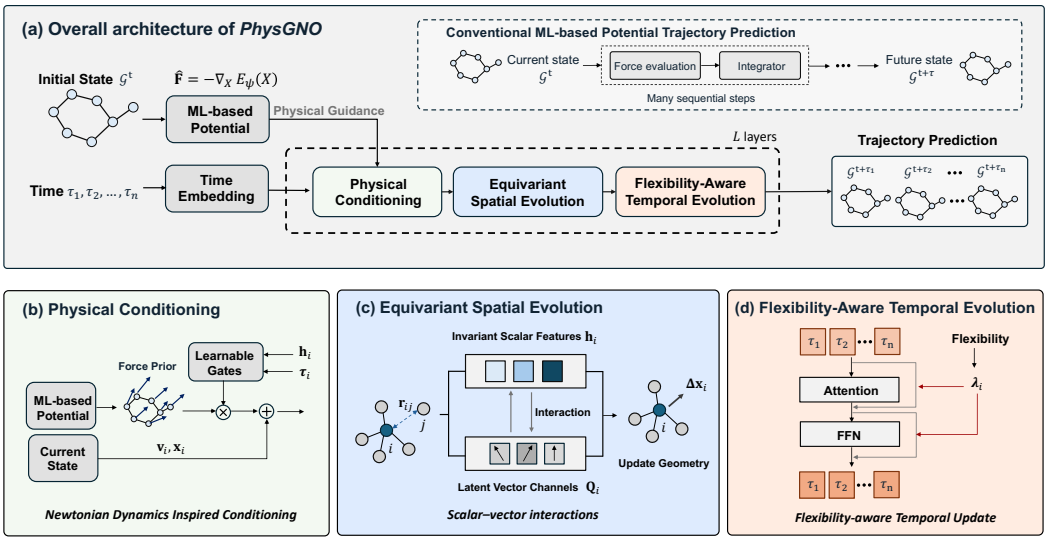

Figure 1: **Overview of PhysGNO.** (a) Overall architecture of PhysGNO and comparison with conventional ML-based potential trajectory prediction, which requires repeated force evaluations and numerical integration. PhysGNO instead performs finite-time trajectory prediction through a small number of learned evolution layers. (b)–(d) The three key components of PhysGNO: Physical Conditioning, Equivariant Spatial Evolution, and Flexibility-Aware Temporal Evolution.

learning potential-energy surfaces and predicting molecular energies and forces directly from atomic configurations. Representative neural potentials include SchNet and PhysNet (Schütt et al., 2017; Unke and Meuwly, 2019), while recent equivariant models such as PaiNN, NequIP, MACE, and ViSNet further improve the modeling of three-dimensional molecular interactions (Schütt et al., 2021; Batzner et al., 2022b; Batatia et al., 2022; Wang et al., 2024). Nevertheless, these methods typically still rely on repeated force evaluations and numerical integration to generate long trajectories. In contrast, diffusion-based generative models (Wu and Li, 2023; Jing et al., 2024) and neural operators such as EGNO and ATOM (Xu et al., 2024; Thompson et al., 2026) have explored direct trajectory modeling without explicitly resolving every microscopic integration step. Recent approaches such as GF-NODE, PAINET, and SGDD further improve long-horizon 3D dynamics modeling through multi-scale, physics-inspired, and spectral modeling, respectively (Sun et al., 2025; Yang et al., 2026; Kim et al., 2026). Compared to these advanced methods, PhysGNO combines physical priors learned by potential models with an operator specifically designed for molecular dynamics, enabling effective and efficient molecular trajectory prediction.

Neural Operators. Neural operators provide a data-driven framework for learning mappings between functions, and have been widely used to approximate the solution operators of partial and ordinary differential equations in scientific computing (Li et al., 2021; Kovachki et al., 2023). Unlike conventional numerical solvers that resolve system evolution through successive time steps, neural operators can directly learn the evolution of a dynamical system over a finite time horizon. For three-dimensional physical dynamics, ESTAG introduces equivariant spatio-temporal modeling to capture the evolution of interacting systems (Wu et al., 2023), while EGNO models time-indexed 3D trajectories through an equivariant graph neural operator (Xu et al., 2024). More recently, ATOM develops a transformer-based neural operator for molecular dynamics, enabling parallel prediction across multiple future time horizons (Thompson et al., 2026). Different from these works, we investigate how to retain the efficient finite-time modeling capability of neural operators while incorporating physical priors inherent in molecular dynamics systems.

3 PRELIMINARIES

Molecular Trajectory Prediction. We consider a molecular system of N atoms represented at time t by $\mathcal{G}^{(t)} = [\mathbf{H}, \mathbf{X}^{(t)}, \mathbf{V}^{(t)}]$, where $\mathbf{H} \in \mathbb{R}^{N \times d}$ denotes time-independent atomic features, and

$\mathbf{X}^{(t)}, \mathbf{V}^{(t)} \in \mathbb{R}^{N\times 3}$ denote atomic coordinates and velocities, respectively. A molecular trajectory is therefore a time-indexed sequence $\mathcal{T} = \{\mathcal{G}^{(t)}\}_t$. Given an observed state $\mathcal{G}^{(t)}$ and a future horizon $\tau > 0$, the trajectory prediction task is to estimate $\widehat{\mathcal{G}}^{(t+\tau)}$. For multiple horizons $\{\tau_p\}_{p=1}^{P}$, the model predicts the corresponding future states $\{\widehat{\mathcal{G}}^{(t+\tau_p)}\}_{p=1}^{P}$.

Neural Operators for Molecular Dynamics. Neural operators learn the finite-time evolution of dynamical systems directly. For molecular dynamics, we denote this mapping as $\mathcal{F}_\theta : (\mathcal{G}^{(t)}, \tau) \mapsto \widehat{\mathcal{G}}^{(t+\tau)}$, where θ denotes the learnable parameters and τ specifies the prediction horizon. This formulation avoids explicitly resolving every integration step and naturally supports multi-horizon trajectory prediction, thereby enabling efficient molecular trajectory forecasting.

$SE(3)$ **Equivariance.** Molecular dynamics describes the evolution of atomic configurations in three-dimensional Euclidean space. The underlying molecular dynamics should therefore be independent of the choice of global coordinate system. This geometric consistency is commonly characterized by $SE(3)$ equivariance. Specifically, $SE(3)$ consists of three-dimensional rotations and translations. For a rotation $\mathbf{R} \in SO(3)$ and a translation $\mathbf{b} \in \mathbb{R}^3$, the molecular coordinates and velocities transform as $\mathbf{X} \mapsto \mathbf{X}\mathbf{R}^\top + \mathbf{1}\mathbf{b}^\top$ and $\mathbf{V} \mapsto \mathbf{V}\mathbf{R}^\top$, while invariant atomic features $\mathbf{H}$ remain unchanged. A trajectory model $\mathcal{F}_\theta$ is $SE(3)$-equivariant if $\mathcal{F}_\theta(g\cdot\mathcal{G}^{(t)}, \tau) = g\cdot\mathcal{F}_\theta(\mathcal{G}^{(t)}, \tau)$ for any $g \in SE(3)$.

4 METHODS

4.1 MOTIVATION AND OVERVIEW

Motivation. Conventional potential energy networks have demonstrated strong capability in modeling molecular interactions and capturing physically meaningful information, such as interatomic force fields and potential-energy landscapes. However, they typically rely on step-by-step state propagation to generate molecular trajectories. In contrast, neural operators enable efficient finite-time trajectory prediction without explicitly resolving every microscopic integration step, but existing approaches primarily learn from observed 3D trajectories without explicitly exploiting the physical information inherent in molecular dynamics. Moreover, molecular systems exhibit rich directional features and heterogeneous flexibility, providing useful priors for spatial and temporal modeling. Motivated by these observations, we propose **PhysGNO**, which combines the efficiency of neural operators with physical and molecular priors for effective and efficient molecular trajectory forecasting.

Overview. PhysGNO integrates physical information learned by a pretrained potential model with spatiotemporal inductive biases tailored to molecular dynamics. As illustrated in Figure 1, PhysGNO consists of three main components: (1) Physical Conditioning, which uses potential-derived information to guide operator learning; (2) Equivariant Spatial Evolution Layers, which model directional spatial interactions; and (3) Flexibility-Aware Temporal Evolution Layers, which incorporate molecular flexibility to capture heterogeneous temporal dynamics.

4.2 PHYSICAL CONDITIONING

Potential energy models encode useful physical knowledge about molecular interactions through the learned energy landscape and its induced force field. We leverage such knowledge to guide molecular trajectory evolution by introducing a **Physical Conditioning** mechanism that injects the physical prior captured by a pretrained potential model into the neural operator.

Given molecular coordinates $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]^\top \in \mathbb{R}^{N\times 3}$, a pretrained potential model $\hat{E}_\psi(\mathbf{X})$ predicts the potential energy of the molecular configuration. The force acting on atom i is obtained as the negative gradient of the energy,

$$\hat{\mathbf{F}}_i = -\frac{\partial \hat{E}_\psi(\mathbf{X})}{\partial \mathbf{x}_i}, \qquad \hat{\mathbf{F}}_i \in \mathbb{R}^3, \tag{1}$$

where ψ denotes the parameters of the potential model and $\mathbf{x}_i \in \mathbb{R}^3$ is the coordinate of atom i. Compared with the scalar energy value, the induced force field provides explicit directional information about the local evolution tendency of the molecular configuration.

Inspired by Newtonian dynamics, we incorporate such physical knowledge into the neural operator. According to Newton's second law, $\mathbf{F}_i = m_i\, d\mathbf{v}_i/dt$, force governs the local evolution of atomic motion. In particular, the force-induced variations in velocity and position satisfy

$$\underbrace{\begin{cases} \Delta\mathbf{v}_i \propto \dfrac{\mathbf{F}_i}{m_i}\Delta t, \\ \Delta\mathbf{x}_i^F \propto \dfrac{\mathbf{F}_i}{m_i}\Delta t^2 \end{cases}}_{\text{Newtonian dynamics}} \quad\Longrightarrow\quad \underbrace{\begin{cases} \widetilde{\mathbf{v}}_i = \mathbf{v}_i + g_{v,i}(\mathbf{h}_i,\tau)\,\hat{\mathbf{F}}_i, \\ \widetilde{\mathbf{x}}_i = \mathbf{x}_i + g_{x,i}(\mathbf{h}_i,\tau)\,\hat{\mathbf{F}}_i \end{cases}}_{\text{Physical Conditioning}}, \tag{2}$$

where $\Delta\mathbf{x}_i^F$ denotes the force-induced component of the coordinate variation, $\mathbf{h}_i$ is the invariant atomic representation, and τ denotes the target prediction horizon.

Rather than explicitly performing numerical integration, we use the force predicted by the pretrained potential model as a directional signal and modulate its effect through learnable gates. Specifically, instead of using fixed coefficients determined by atomic mass and a microscopic time step, we define

$$g_{r,i}(\mathbf{h}_i,\tau) = \phi_r\left([\mathbf{h}_i, \mathrm{PE}(\tau)]\right), \qquad r \in \{x, v\}, \tag{3}$$

where $\mathrm{PE}(\tau)$ denotes the positional encoding of the prediction horizon and ϕ_x, ϕ_v are learnable gating functions. The atomic representation $\mathbf{h}_i$ contains atom-specific information, including atomic identity and implicitly encoded mass-related properties, while $\mathrm{PE}(\tau)$ characterizes the temporal scale. Hence, the gates learn the effective atom- and horizon-dependent strength of the physical guidance.

Overall, Physical Conditioning combines two complementary sources of information: the potential-derived force provides a physically meaningful *direction* for molecular evolution, while the learnable gates determine *how strongly* this local physical tendency influences the finite-horizon dynamics.

In practice, we adopt a gradient scaling strategy so that the energy model $\hat{E}_\psi(\cdot)$ can adapt to the downstream trajectory prediction task while preserving most of its pretrained representation capacity:

$$\bar{\mathbf{F}}_i = \alpha\,\hat{\mathbf{F}}_i + (1-\alpha)\,\mathrm{sg}\Big(\hat{\mathbf{F}}_i\Big), \qquad \alpha \in [0,1], \tag{4}$$

where $\mathrm{sg}(\cdot)$ denotes the stop-gradient operation. We replace $\hat{\mathbf{F}}_i$ in (2) with $\bar{\mathbf{F}}_i$. This leaves the forward-pass physical guidance unchanged, while α controls the extent to which the trajectory prediction loss refines the pretrained potential model during training.

4.3 Equivariant Spatial Evolution

Molecular dynamics contains diverse directional information, including velocities, forces, relative displacements, and local interaction directions. Inspired by PaiNN (Schütt et al., 2021), we maintain an invariant scalar representation $\mathbf{h}_i^{(\ell)} \in \mathbb{R}^{d_s}$ together with a latent equivariant vector representation $\mathbf{Q}_i^{(\ell)} \in \mathbb{R}^{d_v\times 3}$. The d_v three-dimensional channels provide a latent space for encoding different directional dynamics. The spatial evolution jointly models invariant molecular interactions and directional vector interactions, while allowing information to exchange between the two representations.

For clarity, we omit the layer index below. For neighboring atoms i and j, let $\mathbf{r}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ and $d_{ij} = \|\mathbf{r}_{ij}\|_2$. We construct the pairwise message as

$$(\Delta\mathbf{h}_{ij}, \mathbf{a}_{ij}, \mathbf{b}_{ij}) = \mathrm{MLP}\left([\mathbf{h}_i, \mathbf{h}_j, \mathrm{RBF}(d_{ij}), \mathbf{e}_{ij}]\right), \tag{5}$$

where $\Delta\mathbf{h}_{ij} \in \mathbb{R}^{d_s}$ and $\mathbf{a}_{ij}, \mathbf{b}_{ij} \in \mathbb{R}^{d_v}$. Here, the outputs of MLP are split into an invariant message and two invariant vector-channel coefficients. We then propagate scalar and vector information by

$$\begin{aligned} \mathbf{h}_i' &= \mathbf{h}_i + \sum_{j\in\mathcal{N}(i)} \Delta\mathbf{h}_{ij}, \\ \mathbf{Q}_i' &= \mathbf{Q}_i + \sum_{j\in\mathcal{N}(i)} \left[\mathbf{a}_{ij}\mathbf{r}_{ij}^\top + \mathrm{Diag}(\mathbf{b}_{ij})\mathbf{Q}_j\right]. \end{aligned} \tag{6}$$

The scalar update aggregates invariant molecular interactions, while the vector update combines relative geometric directions with directional features propagated from neighboring atoms.

We further model the interaction between scalar and vector features. Let $\mathbf{U}_i = \mathbf{W}_U \mathbf{Q}'_i$ and $\mathbf{Z}_i = \mathbf{W}_Z \mathbf{Q}'_i$. Their channel-wise inner product $\mathbf{c}_i = \langle \mathbf{U}_i, \mathbf{Z}_i \rangle_3 \in \mathbb{R}^{d_v}$ summarizes directional interactions into invariant features. We then compute $(\mathbf{g}_i^h, \mathbf{g}_i^q) = \mathrm{MLP}(\mathbf{h}'_i)$ and update

$$\mathbf{h}_i^{(\ell+1)} = \mathbf{h}'_i + \mathbf{g}_i^h \odot \mathbf{c}_i, \qquad \mathbf{Q}_i^{(\ell+1)} = \mathbf{Q}'_i + \mathbf{g}_i^q \odot \mathbf{U}_i. \tag{7}$$

Thus, directional information contributes to the invariant representation, while the scalar state adaptively modulates different latent vector channels. Finally, the vector channels are combined to evolve the coordinates,

$$\mathbf{x}_i^{(\ell+1)} = \mathbf{x}_i^{(\ell)} + \sum_{k=1}^{d_v} g_{i,k}^x \, \mathbf{Q}_{i,k}^{(\ell+1)}, \qquad \mathbf{g}_i^x = \mathrm{MLP}(\mathbf{h}_i^{(\ell+1)}). \tag{8}$$

The resulting spatial evolution is designed to preserve $SE(3)$ equivariance; as provided in Theorem 1.

4.4 Flexibility-Aware Temporal Evolution

Neural operators for molecular trajectory modeling generally introduce temporal layers to capture dependencies across different prediction horizons. However, such temporal modeling often overlooks the intrinsic flexibility of molecular structures, which provides a useful inductive bias for the degree of conformational change. Relatively rigid regions tend to exhibit smaller structural variations, whereas flexible regions may undergo larger motions.

To characterize molecular flexibility without introducing additional test-time observations, we compute a static flexibility prior using only the training set. Specifically, we measure the normalized variation of pairwise distances across training configurations:

$$\phi_{ij} = \frac{\mathrm{Var}_{\mathbf{X} \in \mathcal{D}_{\mathrm{train}}} \left(d_{ij}(\mathbf{X}) \right)}{\mathrm{Mean}_{\mathbf{X} \in \mathcal{D}_{\mathrm{train}}} \left(d_{ij}(\mathbf{X}) \right)^2 + \epsilon}, \qquad d_{ij}(\mathbf{X}) = \|\mathbf{x}_i - \mathbf{x}_j\|_2, \tag{9}$$

and summarize the local flexibility of atom i by

$$\bar{\phi}_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \phi_{ij}. \tag{10}$$

The flexibility statistics are precomputed exclusively from training configurations and kept fixed during validation and testing. Therefore, no validation/test configurations, future target states, or additional historical frames are used, avoiding information leakage.

For P queried horizons, we stack the scalar representations of atom i as $\mathbf{H}_i \in \mathbb{R}^{P \times d_s}$, where the p-th row corresponds to prediction horizon τ_p. We use a standard temporal Transformer block, where $\mathrm{Attn}(\cdot)$ denotes self-attention along the horizon dimension and $\mathrm{FFN}(\cdot)$ denotes the standard position-wise feed-forward network. To incorporate flexibility, we compute a horizon-dependent gate

$$\boldsymbol{\lambda}_i = \sigma \left(G \left(\bar{\phi}_i, \mathrm{PE}(\tau_{1:P}) \right) \right) \in \mathbb{R}^P, \tag{11}$$

where the gate is broadcast over feature channels. The temporal evolution is then

$$\mathbf{H}'_i = \mathbf{H}_i + \boldsymbol{\lambda}_i \odot \mathrm{Attn}(\mathbf{H}_i), \qquad \mathbf{H}''_i = \mathbf{H}'_i + \boldsymbol{\lambda}_i \odot \mathrm{FFN}(\mathbf{H}'_i). \tag{12}$$

The same temporal evolution is applied to the latent vector representations $\mathbf{Q}_i$ across the horizon dimension. The attention weights and flexibility-conditioned gates are shared across the three Cartesian components, preserving the equivariant vector structure.

4.5 Training Objective

Given the observed molecular state at time t, PhysGNO predicts molecular coordinates at a set of future horizons $\{\tau_p\}_{p=1}^P$. Let $\widehat{\mathbf{X}}^{(p)} \in \mathbb{R}^{N \times 3}$ and $\mathbf{X}^{(p)} \in \mathbb{R}^{N \times 3}$ denote the predicted and ground-truth coordinates at horizon τ_p, respectively. We train the model primarily with the trajectory prediction objective

$$\mathcal{L}_{\mathrm{traj}} = \frac{1}{PN} \sum_{p=1}^{P} \sum_{i=1}^{N} \left\| \widehat{\mathbf{x}}_i^{(p)} - \mathbf{x}_i^{(p)} \right\|_2^2. \tag{13}$$

5 EXPERIMENTS

We organize our experiments around five research questions:

- **RQ1: Overall Effectiveness.** How does PhysGNO compare with advanced equivariant dynamics models and neural trajectory operators for molecular trajectory prediction?
- **RQ2: Component Contribution.** How much do the different Components of PhysGNO contribute to the final performance?
- **RQ3: Effect of the Physical Prior.** Does potential-derived physical information improve trajectory prediction, and how should such information be incorporated into the operator?
- **RQ4: Efficiency.** What is the computational efficiency of PhysGNO?
- **RQ5: Physical rationality of the generated trajectory.** Can PhysGNO produce physically faithful trajectories that better match the ground-truth molecular dynamics?

5.1 EXPERIMENTAL SETUP

Datasets and Tasks. We evaluate PhysGNO on MD17 (Christensen and von Lilienfeld, 2020), rMD17 (Christensen and von Lilienfeld, 2020) and MD22 (Chmiela et al., 2023) MD17/rMD17 contain small organic molecules, while MD22 comprises large molecular systems with substantially more atoms and greater structural complexity. We follow the MD17 evaluation protocol used by EGNO and PAINET (Xu et al., 2024; Yang et al., 2026), and the MD22 protocol of PAINET. For each molecule, we use 500/2000/2000 samples for training/validation/testing. Following PAINET, we consider state-to-state (S2S) prediction with a single query at the final horizon ($T = 1$), and state-to-trajectory (S2T) prediction with $T = 8$ queried states on MD17 and $T = 5$ on MD22, where T denotes the number of predicted states.

Baselines. We compare PhysGNO against advanced models, including RF (Köhler et al., 2019), EGNN (Satorras et al., 2021), ClofNet (Du et al., 2022), EGNO (Xu et al., 2024), GF-NODE (Sun et al., 2025), HEGNN (Cen et al., 2024), PAINET (Yang et al., 2026), and ATOM (Thompson et al., 2026). We follow ATOM and use two EGNN baselines for S2T prediction: EGNN-Rollout (EGNN-R), which predicts timesteps autoregressively, and EGNN-Sequential (EGNN-S), which uses the output of each GNN as the prediction for a given frame.

Metrics and Implementation. Following EGNO, we report final mean squared error (F-MSE) for S2S prediction and average mean squared error (A-MSE) over all queried states for S2T prediction. Detailed implementation and experimental settings can be found in the Appendix B.

5.2 RQ1: OVERALL EFFECTIVENESS AND LONG-HORIZON PERFORMANCE.

We first compare PhysGNO with advanced methods on molecular trajectory prediction. Results on MD17 are reported in Tables 1 and 2, while Tables 5 summarizes the results on the larger MD22 systems. Results on rMD17 are provided in Appendix C.1.

Overall, PhysGNO achieves strong performance across all benchmarks. By leveraging physical priors from the pretrained potential model, PhysGNO produces more accurate and physically informed trajectory predictions than prior approaches. The improvement is particularly notable on molecules such as Ethanol and Toluene, where PhysGNO shows clear advantages over competing methods. Moreover, these performance gains remain evident on the larger MD22 molecules, further confirming the applicability of our method to more complex molecular systems. We also provide visualization results to further support these findings, as shown in Section 5.6. To further illustrate the effectiveness of our methods in long-horizon scenarios, we include additional experiments in the Appendix C.4.

5.3 RQ2: COMPONENT CONTRIBUTION

To evaluate the impact of each component on overall performance, we further conduct ablation studies on different components of PhysGNO. We observe that introducing Physical Conditioning effectively provides informative guidance, which facilitates more accurate prediction. A more detailed analysis of Physical Conditioning is provided in Section 5.4.

Table 1: State-to-state (S2S) position F-MSE ($\times 10^{-2}$) on MD17. Lower is better.

Method	Aspirin	Ethanol	Malonaldehyde	Naphthalene	Salicylic	Toluene	Uracil
RF	10.90 ± 0.003	4.64 ± 0.001	12.90 ± 0.060	0.40 ± 0.0029	1.24 ± 0.0105	10.90 ± 0.010	0.62 ± 0.0050
EGNN	10.20 ± 0.158	4.65 ± 0.005	12.80 ± 0.010	0.39 ± 0.0042	1.04 ± 0.0154	10.70 ± 0.150	0.58 ± 0.0082
ClofNet	9.46 ± 0.029	4.62 ± 0.006	12.80 ± 0.010	0.39 ± 0.0066	1.03 ± 0.0581	10.30 ± 0.020	0.62 ± 0.0415
EGNO	9.44 ± 0.122	4.63 ± 0.004	12.80 ± 0.010	0.38 ± 0.0077	0.90 ± 0.0641	10.60 ± 0.050	0.55 ± 0.0149
GF-NODE	9.15 ± 0.030	4.63 ± 0.001	12.80 ± 0.000	0.37 ± 0.0166	0.86 ± 0.0254	10.40 ± 0.020	0.58 ± 0.0017
HEGNN	9.38 ± 0.056	4.64 ± 0.001	12.90 ± 0.020	0.37 ± 0.0037	0.84 ± 0.0021	10.50 ± 0.040	0.53 ± 0.0047
PAINET	8.98 ± 0.040	4.31 ± 0.004	12.60 ± 0.010	**0.34 ± 0.0008**	**0.78 ± 0.0007**	10.00 ± 0.000	0.51 ± 0.0004
ATOM	6.82 ± 0.060	3.52 ± 0.040	14.72 ± 0.010	0.50 ± 0.000	0.88 ± 0.010	4.66 ± 0.210	0.63 ± 0.000
PhysGNO	**6.33 ± 0.040**	**2.98 ± 0.008**	**12.22 ± 0.002**	**0.34 ± 0.004**	0.80 ± 0.003	**3.96 ± 0.001**	**0.50 ± 0.006**

Table 2: State-to-trajectory (S2T) position A-MSE ($\times 10^{-2}$) on MD17. Lower is better.

Method	Aspirin	Ethanol	Malonaldehyde	Naphthalene	Salicylic	Toluene	Uracil
EGNN-R	7.35 ± 0.190	3.21 ± 0.000	10.75 ± 0.040	0.34 ± 0.060	1.09 ± 0.120	4.53 ± 0.080	0.89 ± 0.020
EGNN-S	9.01 ± 0.340	3.21 ± 0.000	11.20 ± 0.030	0.42 ± 0.010	1.41 ± 0.000	4.86 ± 0.040	0.65 ± 0.010
EGNO	9.64 ± 0.150	4.57 ± 0.010	12.92 ± 0.000	0.39 ± 0.000	0.90 ± 0.010	10.99 ± 0.000	0.58 ± 0.020
GF-NODE	9.25 ± 0.017	4.63 ± 0.006	12.80 ± 0.000	0.46 ± 0.007	0.86 ± 0.001	10.30 ± 0.020	0.58 ± 0.001
HEGNN	9.54 ± 0.024	4.67 ± 0.001	12.90 ± 0.010	0.43 ± 0.000	0.90 ± 0.002	10.60 ± 0.070	0.61 ± 0.006
PAINET	8.84 ± 0.045	4.55 ± 0.001	12.60 ± 0.010	0.32 ± 0.012	0.79 ± 0.014	10.00 ± 0.020	0.51 ± 0.001
ATOM	5.62 ± 0.050	2.62 ± 0.040	12.49 ± 0.010	0.43 ± 0.000	0.86 ± 0.010	**2.27 ± 0.100**	0.61 ± 0.000
PhysGNO	**5.51 ± 0.041**	**2.22 ± 0.010**	**10.43 ± 0.020**	**0.28 ± 0.007**	**0.67 ± 0.013**	2.32 ± 0.020	**0.41 ± 0.010**

Besides, we examine variants that remove the flexibility-aware temporal evolution module or replace the proposed spatial evolution module with an EGNN-based alternative. As shown in Table 3, the full model achieves the best performance among all variants. This result demonstrates the effectiveness of our neural operator design tailored for molecular dynamics. More ablation studies, including the gradient scaling strategy, can be found in the Appendix C.2.

Table 3: Component ablation of PhysGNO on representative MD17 molecules ($\times 10^{-2}$). PC, FTE, and SE denote Physical Conditioning, Flexibility-aware Temporal Evolution, and Spatial Evolution, respectively.

Component			Aspirin		Ethanol		Toluene	
PC	FTE	SE	S2S	S2T	S2S	S2T	S2S	S2T
✗	✗	✗	8.96	8.62	4.62	4.64	9.95	9.84
✗	✓	✗	8.54	8.40	4.24	4.12	6.24	6.20
✗	✗	✓	7.24	6.78	3.44	3.56	8.48	8.14
✗	✓	✓	6.84	6.33	3.67	3.42	5.14	5.05
✓	✓	✓	**6.33**	**5.51**	**2.98**	**2.22**	**3.96**	**2.32**

5.4 RQ3: Effect of Physical Conditioning

To further investigate the effectiveness of the physical information provided by ML-based potentials, we conduct ablation studies on different conditioning strategies. Unless otherwise specified, we use PaiNN as the default ML potential. As shown in Table 4, incorporating Physical Conditioning consistently improves trajectory prediction, indicating that the force information encoded by the pretrained potential provides useful local dynamical guidance for finite-time molecular evolution.

Table 4: Effect of different ML potentials on physical conditioning for MD17. We report S2S and S2T position F-MSE and A-MSE, respectively ($\times 10^{-2}$).

Potential	Aspirin		Ethanol		Toluene	
	S2S	S2T	S2S	S2T	S2S	S2T
None	6.84	6.33	3.67	3.42	5.14	5.05
PaiNN (default)	6.33	5.51	2.98	2.22	3.96	2.32
MACE	6.58	5.96	2.64	2.25	4.42	3.87
NequIP	6.69	5.83	3.06	2.45	4.12	4.10

We further investigate the robustness of physical conditioning to the choice of potential architecture by replacing the default PaiNN potential with alternative pretrained ML potentials, including MACE (Batatia et al., 2022) and NequIP (Batzner et al., 2022a). As shown in Table 4, Physical Conditioning consistently improves trajectory prediction across different potential architectures, suggesting that PhysGNO can effectively leverage meaningful information from diverse ML potentials.

Table 5: Test MSE (↓) on MD22 ($\times 10^{-1}$). F-MSE is reported for S2S ($T = 1$) and A-MSE for S2T ($T = 5$). DHA results are reproduced by running the official code. Lower is better.

Method	Stachyose		Ac-Ala3-NHMe		DHA	
	F-MSE	A-MSE	F-MSE	A-MSE	F-MSE	A-MSE
EGNO	3.33 ± 0.05	4.43 ± 0.11	9.11 ± 0.17	8.92 ± 0.12	15.25 ± 0.10	9.16 ± 0.02
HEGNN	3.34 ± 0.14	3.81 ± 0.20	12.41 ± 0.53	11.63 ± 1.07	12.57 ± 0.11	9.43 ± 0.00
GF-NODE	3.11 ± 0.27	2.54 ± 0.17	9.03 ± 0.31	8.52 ± 0.07	11.58 ± 0.58	7.65 ± 0.28
PAINET	3.00 ± 0.06	2.40 ± 0.11	8.83 ± 0.18	8.35 ± 0.27	8.46 ± 0.53	7.28 ± 0.04
PhysGNO	**2.65** ± 0.06	**2.23** ± 0.14	**6.42** ± 0.08	**6.21** ± 0.02	**1.89** ± 0.03	**1.34** ± 0.01

5.5 RQ4: Efficiency

To further evaluate the practical efficiency of PhysGNO, we measure the average inference time on the MD22 S2T task over six molecular systems: Ac-Ala3-NHMe, DHA, stachyose, etc. We compare against advanced methods, including ATOM, PAINET, EGNN-R, and a PaiNN-based rollout baseline that generates trajectories through step-by-step velocity-Verlet integration. As shown in Table 6, PhysGNO predicts a full trajectory in 12.10 ms on average, achieving inference efficiency comparable to PAINET while remaining considerably faster than EGNN-R. Together with the strong predictive performance, these results demonstrate that PhysGNO achieves a favorable balance between prediction accuracy and computational efficiency.

Table 6: Average inference time on the MD22 S2T task.

Method	Time (ms / traj.) ↓
ATOM	9.78
PAINET	14.00
EGNN-R	21.22
PaiNN rollout	199.46
PhysGNO	**12.10**

5.6 RQ5: Physical rationality of the generated trajectory.

To better illustrate the prediction accuracy of PhysGNO, we visualize the molecular trajectories generated by EGNO and PhysGNO on DHA. As shown in Figure 2, PhysGNO more accurately follows the ground-truth physical trajectories than EGNO, further demonstrating its effectiveness. To more clearly demonstrate the physical plausibility of the trajectories generated by PhysGNO, we provide additional physics-related evaluation metrics in the Appendix C.5.

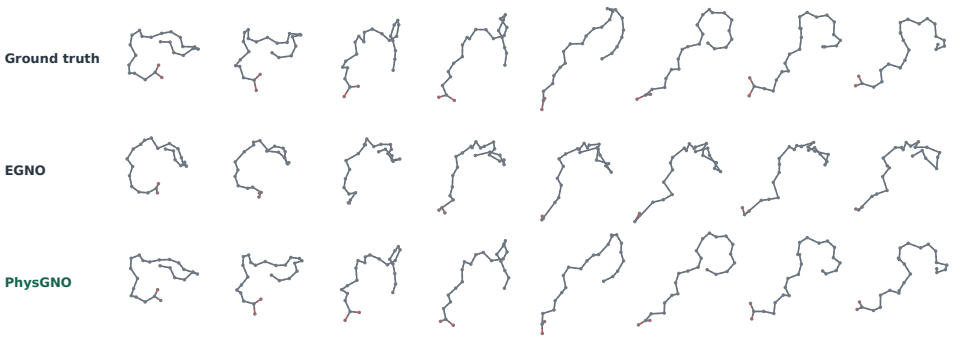

Figure 2: Trajectory comparison on DHA. We use $T = 8$ for better visualization.

6 Conclusion

We propose PhysGNO, a physics-guided graph neural operator for molecular trajectory modeling. By leveraging force information from pretrained ML potentials as physical guidance, PhysGNO enables efficient finite-time trajectory prediction with improved physical fidelity. To better characterize molecular dynamics, we further design spatial-temporal evolution layers tailored to molecular dynamical systems, enabling effective modeling of coupled spatial interactions and temporal evolution. Extensive experiments on MD17/rMD17 and MD22 show that PhysGNO achieves strong performance across diverse molecular systems.

AI USAGE STATEMENT

We used generative AI tools primarily for language polishing, manuscript editing, and code optimization. The authors take full responsibility for the accuracy and integrity of the final manuscript, code, and reported results.

REPRODUCIBILITY STATEMENT

Appendix B provides the implementation details. All experiments were conducted with different random seeds to ensure the reliability of the results. We will release the code upon acceptance of the paper.

REFERENCES

Michael P Allen and Dominic J Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 2 edition, 2017.

Ilyes Batatia, Dávid P. Kovács, Gregor N. C. Simm, Christoph Ortner, and Gábor Csányi. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):2453, 2022a.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13:2453, 2022b.

Jorg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 98:146401, 2007. doi: 10.1103/PhysRevLett.98.146401.

Jiacheng Cen, Anyi Li, Ning Lin, Yuxiang Ren, Zihe Wang, and Wenbing Huang. Are high-degree representations really unnecessary in equivariant graph neural networks? In *Advances in Neural Information Processing Systems*, volume 37, pages 26238–26266, 2024. doi: 10.52202/079017-0826. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/2e4531a45b99f61947b23ccdd608303b-Abstract-Conference.html.

Stefan Chmiela, Valentin Vassilev-Galindo, Oliver T. Unke, Adil Kabylda, Huziel E. Sauceda, Alexandre Tkatchenko, and Klaus-Robert M"uller. Accurate global machine learning force fields for molecules with hundreds of atoms. *Science Advances*, 9(2):eadf0873, 2023. doi: 10.1126/sciadv.adf0873.

Anders S. Christensen and O. Anatole von Lilienfeld. On the role of gradients for machine learning of molecular energies and forces. *Machine Learning: Science and Technology*, 1(4):045018, 2020.

Weitao Du, He Zhang, Yuanqi Du, Qi Meng, Wei Chen, Nanning Zheng, Bin Shao, and Tie-Yan Liu. SE(3) equivariant graph neural networks with complete local frames. In *International Conference on Machine Learning*, volume 162, pages 5583–5608, 2022.

Jacob D. Durrant and J. Andrew McCammon. Molecular dynamics simulations and drug discovery. *BMC Biology*, 9:71, 2011. doi: 10.1186/1741-7007-9-71.

Scott A. Hollingsworth and Ron O. Dror. Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143, 2018. doi: 10.1016/j.neuron.2018.08.011.

Bowen Jing, Hannes Stark, Tommi Jaakkola, and Bonnie Berger. Generative modeling of molecular dynamics trajectories. In *Advances in Neural Information Processing Systems*, 2024.

Youjin Kim, Dagyeong Na, Jae Yong Lee, and Junseok Kwon. Spectral-guided physical dynamics distillation. In *International Conference on Learning Representations*, 2026.

Jonas Köhler, Leon Klein, and Frank Noé. Equivariant flows: sampling configurations for multi-body systems with symmetric energies. *arXiv preprint arXiv:1910.00753*, 2019.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89):1–97, 2023.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

Dominik Marx and Jürg Hutter. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press, 2009.

Albert Musaelian, Simon Batzner, Anders Johansson, Lixin Sun, Cameron J. Owen, Mordechai Kornbluth, and Boris Kozinsky. Learning local equivariant representations for large-scale atomistic dynamics. *Nature Communications*, 14:579, 2023. doi: 10.1038/s41467-023-36329-y.

Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *International Conference on Machine Learning*, pages 9323–9332, 2021.

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.

Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International conference on machine learning*, pages 9377–9388. PMLR, 2021.

Fang Sun, Zijie Huang, Haixin Wang, Huacong Tang, Xiao Luo, Wei Wang, and Yizhou Sun. Graph fourier neural odes: Modeling spatial-temporal multi-scales in molecular dynamics. *Transactions on Machine Learning Research*, 2025.

Luke Thompson, Davy Guan, Slade Matthews, Dai Shi, Junbin Gao, and Andi Han. ATOM: A pretrained neural operator for multitask molecular dynamics. In *International Conference on Learning Representations*, 2026.

Oliver T. Unke and Markus Meuwly. Physnet: A neural network for predicting energies, forces, dipole moments, and partial charges. *Journal of Chemical Theory and Computation*, 15(6):3678–3693, 2019. doi: 10.1021/acs.jctc.9b00181.

Wilfred F van Gunsteren and Herman J C Berendsen. Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angewandte Chemie International Edition in English*, 29(9):992–1023, 1990.

Yusong Wang, Tong Wang, Shaoning Li, Xinheng He, Mingyu Li, Zun Wang, Nanning Zheng, Bin Shao, and Tie-Yan Liu. Enhancing geometric representations for molecules with equivariant vector-scalar interactive message passing. *Nature Communications*, 15:313, 2024. doi: 10.1038/s41467-023-43720-2.

Fang Wu and Stan Z. Li. Diffmd: A geometric diffusion model for molecular dynamics simulations. *AAAI*, 2023.

Liming Wu, Zhichao Hou, Jirui Yuan, Yu Rong, and Wenbing Huang. Equivariant spatio-temporal attentive graph networks to simulate physical dynamics. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Minkai Xu, Jiaqi Han, Aaron Lou, Jean Kossaifi, Arvind Ramanathan, Kamyar Azizzadenesheli, Jure Leskovec, Stefano Ermon, and Anima Anandkumar. Equivariant graph neural operator for modeling 3d dynamics. In *International Conference on Machine Learning*, volume 235, pages 55015–55032, 2024.

Kai Yang, Yuqi Huang, Junheng Tao, Wanyu Wang, and Qitian Wu. PAINET: A principled efficient transformer for 3d dynamics modeling. In *International Conference on Learning Representations*, 2026.

A Equivariance of the Spatial Evolution Layer

This section supplies the formal statement referenced in Section 4.3.

Theorem 1 (Equivariance of the spatial evolution)**.** *Assume that (i) the scalar node features $\mathbf{h}_i$ and edge features $\mathbf{e}_{ij}$ are invariant under global rotations and translations; (ii) neighborhood construction is unchanged by a global rigid transformation; and (iii) the channel-mixing matrices act only on feature channels and not on Cartesian axes. Then the spatial updates in Eqs.* (5)–(8) *preserve invariant scalar features and transform the vector features and coordinates equivariantly under $SE(3)$.*

Proof. Let a global rigid transformation act as $\mathbf{x}_i \mapsto \mathbf{x}_i\mathbf{R}^\top + \mathbf{b}^\top$ with $\mathbf{R} \in SO(3)$. Then $\mathbf{r}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ transforms as $\mathbf{r}_{ij} \mapsto \mathbf{r}_{ij}\mathbf{R}^\top$, while $d_{ij} = \|\mathbf{r}_{ij}\|_2$ is invariant. Therefore every input to the MLP in Eq. (5) is invariant, and its outputs $\Delta\mathbf{h}_{ij}$, $\mathbf{a}_{ij}$, and $\mathbf{b}_{ij}$ are invariant coefficients.

Suppose the incoming vector state satisfies $\mathbf{Q}_i \mapsto \mathbf{Q}_i\mathbf{R}^\top$. In Eq. (6), $\mathbf{a}_{ij}\mathbf{r}_{ij}^\top$ and $\mathrm{Diag}(\mathbf{b}_{ij})\mathbf{Q}_j$ both transform with the same right multiplication by $\mathbf{R}^\top$. Hence $\mathbf{Q}'_i$ is equivariant, whereas $\mathbf{h}'_i$ remains invariant.

The channel transforms $\mathbf{U}_i = \mathbf{W}_U\mathbf{Q}'_i$ and $\mathbf{Z}_i = \mathbf{W}_Z\mathbf{Q}'_i$ remain equivariant because $\mathbf{W}_U$ and $\mathbf{W}_Z$ mix feature channels only. Their Cartesian inner product $\mathbf{c}_i = \langle \mathbf{U}_i, \mathbf{Z}_i \rangle_3$ is invariant. Consequently, the scalar update in Eq. (7) is invariant and the vector update is equivariant. Finally, the coefficients $g^x_{i,k}$ in Eq. (8) depend only on invariant scalar features, so the predicted displacement is an invariant linear combination of equivariant vectors. Adding this displacement to the translated coordinate gives the required $SE(3)$-equivariant coordinate update. □

B Detailed Experimental and Implementation Settings

B.1 Datasets and prediction tasks

For MD17 and MD22, we follow the trajectory construction used by PAINET (Yang et al., 2026). Each molecule is split into 500 training, 2000 validation, and 2000 test trajectory samples. For MD17, the PAINET protocol uniformly subsamples the source trajectory with interval $\Delta T = 3000$ and discards the first and last 10,000 snapshots before constructing samples. For MD22, the corresponding interval is $\Delta T = 1000$, again discarding the first and last 10,000 snapshots. Each sample is defined by an initial molecular state containing coordinates and velocities and by one or more future coordinate targets. For rMD17 (Christensen and von Lilienfeld, 2020), we use the same train/validation/test sample counts and the same S2S/S2T query convention as in the MD17 experiments. All the experiments are conducted on the NVIDIA RTX 4090 GPUs.

Pretraining of ML Potentials. All ML potentials used in PhysGNO, including PaiNN, MACE, and NequIP, are trained exclusively on the training split of each molecular dataset using the corresponding DFT energies and atomic forces as supervision. No configurations, energies, or forces from the validation or test splits are used during potential training. The pretrained potentials are then used to provide physical guidance for trajectory prediction. This protocol ensures that the physical prior is derived solely from training data and prevents information leakage from the validation or test sets.

B.2 Metrics

Let $\widehat{\mathbf{X}}^{(p)}$ and $\mathbf{X}^{(p)}$ denote the predicted and ground-truth coordinates at query p. Consistent with the main paper, S2S uses the final-state mean squared error

$$\text{F-MSE} = \frac{1}{N}\sum_{i=1}^{N}\left\|\widehat{\mathbf{x}}_i^{(P)} - \mathbf{x}_i^{(P)}\right\|_2^2, \tag{14}$$

whereas S2T uses the average error over all queried states,

$$\text{A-MSE} = \frac{1}{PN}\sum_{p=1}^{P}\sum_{i=1}^{N}\left\|\widehat{\mathbf{x}}_i^{(p)} - \mathbf{x}_i^{(p)}\right\|_2^2. \tag{15}$$

B.3 Model and optimization details

The details of the model and optimization can be found in Table 7.

Table 7: Proposed implementation settings for PhysGNO.

Setting	Value
Optimizer	AdamW, $\beta_1 = 0.9$, $\beta_2 = 0.999$
Learning rate	$[10^{-3}, 5 \times 10^{-4}, 10^{-4}]$
Weight decay	10^{-5}
Batch size (MD17/rMD17)	100
Batch size (MD22)	32
Training epochs / stopping rule	At most 1000 epochs; early stopping after 100 epochs without validation improvement
Random seeds	42, 43, 44
Scalar hidden dimension d_s	[32, 128]
Vector hidden dimension d_v	[32, 128] vector channels
Number of spatial evolution layers	[2,4]
Number of temporal evolution layers	[2,4]
Radial basis specification	32 Gaussian radial basis functions with uniformly spaced centers on $[0, 5]$ Å; cosine cutoff envelope
Graph cutoff / connectivity	Radius graph with $r_c = 5$ Å; bidirectional edges, excluding self-loops
Gradient-scaling coefficient α	0.1

C Additional Experiments

C.1 Additional Results on rMD17

To further illustrate the effectiveness of the proposed method, we also evaluate PhysGNO on the rMD17 dataset. Compared with the original MD17 dataset, rMD17 provides more stable and reliable molecular dynamics trajectories, making it a more robust benchmark for assessing molecular trajectory prediction models. As shown in Table 8, PhysGNO also achieves advanced performance on the rMD17 dataset, further demonstrating the effectiveness and generalization ability of our method.

Table 8: Experiments on the rMD17 datasets ($\times 10^{-2}$). Upper part: S2S MSE. Lower part: S2T MSE.

	Aspirin	Ethanol	Malonaldehyde	Naphthalene	Paracetamol
EGNO	9.66	4.57	12.92	0.39	12.93
PAINET	8.43	4.43	11.85	0.42	12.78
PhysGNO	**6.06**	**3.52**	**11.62**	**0.38**	**8.49**
EGNO	**9.23**	4.57	12.81	0.39	12.35
PAINET	9.24	4.14	15.83	**0.36**	12.28
PhysGNO	**5.43**	**3.38**	**12.34**	0.42	**8.18**

C.2 Gradient-scaling strategy

Equation (4) leaves the forward force used by Physical Conditioning unchanged while controlling how much gradient from the trajectory objective reaches the pretrained potential. We evaluate this mechanism by varying $\alpha \in [0, 1]$. The endpoint values have clear optimization interpretations: $\alpha = 0$ fully stops the downstream gradient through the force prior, whereas $\alpha = 1$ allows the full downstream gradient. Intermediate values interpolate the backward signal without changing the forward force value. As shown in Table 9, when $\alpha = 0$, the pretrained ML potential cannot receive gradient feedback from the ground-truth trajectories, leading to inferior performance compared with settings that allow gradient propagation. Meanwhile, a small gradient scale such as $\alpha = 0.1$ effectively controls the magnitude of gradient updates, which helps preserve the useful physical

knowledge encoded in the pretrained potential while still enabling task-specific adaptation. Therefore, we adopt $\alpha = 0.1$ as the default setting in practice.

Table 9: Ablation of the gradient-scaling coefficient α in Eq. (4).

α	Aspirin		Ethanol		Toluene	
	S2S	S2T	S2S	S2T	S2S	S2T
0	6.62	6.16	3.25	2.52	4.43	4.26
0.1	**6.33**	**5.51**	2.98	**2.22**	**3.96**	**2.32**
1	6.68	6.25	**2.85**	2.44	4.87	4.50

C.3 Effectiveness of the PhysGNO backbone.

To further demonstrate the effectiveness of the PhysGNO backbone, we apply the same physical conditioning mechanism to the EGNO and PAINET backbones. The corresponding results are shown in Table 10. From the table, physical conditioning consistently improves the performance of these diverse backbones to a certain extent. However, the combination of PhysGNO and physical conditioning achieves the best overall performance, which further validates the effectiveness of both the physical conditioning mechanism and the PhysGNO backbone.

Table 10: Results of A-MSE ($\times 10^{-2}$) on MD17 S2T tasks with different backbones. Physical Conditioning is abbreviated as pc.

Backbone	**Setting**	**Aspirin**	**Ethanol**	**Toluene**	**Malonaldehyde**
EGNO	w/o pc	9.64	**4.57**	10.99	12.92
	+ pc	**8.27**	4.85	**6.83**	**10.56**
PAINET	w/o pc	8.84	4.55	10.00	12.60
	+ pc	**8.46**	**4.28**	**6.43**	**10.42**
PhysGNO	w/o pc	6.33	3.42	5.05	11.50
	+ pc	**5.51**	**2.22**	**2.32**	**10.43**

C.4 Experiments on Long-horizon Setting

Long-Horizon Evaluation. In the long-horizon setting, we fix the temporal interval between adjacent queried states to be the same as in the main S2T task. We then extend the maximum prediction horizon from $8\Delta\tau$ to $32\Delta\tau$ and $64\Delta\tau$, yielding $T = 32$ and $T = 64$ queried future states, respectively. The results are presented in Table 11, where PhysGNO exhibits competitive performance compared with the advanced PAINET and EGNO baselines, further confirming the effectiveness of our method under long-horizon scenarios.

Table 11: Long-horizon prediction experiments on rMD17.

Methods	Ethanol		Toluene		Naphthalene	
	$T = 32$	$T = 64$	$T = 32$	$T = 64$	$T = 32$	$T = 64$
EGNO	7.85	6.47	2.82	2.89	0.17	0.28
PAINET	7.90	6.49	**2.69**	2.84	0.17	0.27
PhysGNO	**7.16**	**6.23**	2.73	**2.68**	**0.14**	**0.24**

Table 12: Physical evaluation on MD17 molecules under the S2T setting. We report bond-length MAE (Å), bond-angle MAE (°), dihedral-angle MAE (°), and dRMSD (Å), averaged over all queried future states. Lower is better.

Molecule	Method	Bond Length MAE (Å) ↓	Bond Angle MAE (°) ↓	Dihedral MAE (°) ↓	dRMSD (Å) ↓
Aspirin	EGNO	0.089	5.499	15.689	0.241
	PAINet	0.088	5.770	15.477	0.242
	PhysGNO	**0.069**	**4.849**	**13.276**	**0.198**
Benzene	EGNO	0.215	8.999	13.877	0.359
	PAINet	0.215	9.877	16.416	0.362
	PhysGNO	**0.138**	**7.268**	**11.872**	**0.237**
Malonaldehyde	EGNO	0.313	41.862	76.910	0.323
	PAINet	0.312	42.089	77.274	0.319
	PhysGNO	**0.262**	**28.219**	**68.926**	**0.284**
Naphthalene	EGNO	0.033	2.779	6.511	0.057
	PAINet	0.032	2.704	6.526	0.055
	PhysGNO	**0.030**	**2.583**	**5.621**	**0.047**
Salicylic acid	EGNO	0.035	3.092	9.562	0.074
	PAINet	0.038	3.223	9.219	0.078
	PhysGNO	**0.031**	**2.829**	**7.589**	**0.066**
Uracil	EGNO	0.034	3.243	9.787	0.061
	PAINet	0.034	3.225	9.939	0.060
	PhysGNO	**0.032**	**3.034**	**9.691**	**0.057**

C.5 Physical plausibility Evaluation

Physical Evaluation. To further evaluate the structural and conformational plausibility of predicted trajectories, we report several geometry-based metrics in addition to position errors. Let $\hat{\mathbf{x}}_i$ and $\mathbf{x}_i$ denote the predicted and ground-truth coordinates of atom i, respectively.

Bond Length MAE. For the set of molecular bonds $\mathcal{B}$, we measure the deviation of predicted bond lengths as

$$\mathrm{BondMAE} = \frac{1}{|\mathcal{B}|} \sum_{(i,j)\in\mathcal{B}} \left| \hat{d}_{ij} - d_{ij} \right|, \qquad d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|_2, \quad \hat{d}_{ij} = \|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|_2. \tag{16}$$

Bond Angle MAE. For the set of connected atom triplets $\mathcal{A}$, the bond-angle error is defined as

$$\mathrm{AngleMAE} = \frac{1}{|\mathcal{A}|} \sum_{(i,j,k)\in\mathcal{A}} \left| \hat{\theta}_{ijk} - \theta_{ijk} \right|, \tag{17}$$

where

$$\theta_{ijk} = \arccos\left(\frac{(\mathbf{x}_i - \mathbf{x}_j)^\top (\mathbf{x}_k - \mathbf{x}_j)}{\|\mathbf{x}_i - \mathbf{x}_j\|_2 \|\mathbf{x}_k - \mathbf{x}_j\|_2} \right), \tag{18}$$

and $\hat{\theta}_{ijk}$ is computed analogously from the predicted coordinates.

Dihedral Angle MAE. For the set of bonded atom quadruplets $\mathcal{T}$, we compute the periodic angular error as

$$\mathrm{DihedralMAE} = \frac{1}{|\mathcal{T}|} \sum_{(i,j,k,l)\in\mathcal{T}} \arccos\left[\cos\left(\hat{\psi}_{ijkl} - \psi_{ijkl} \right) \right], \tag{19}$$

where ψ_{ijkl} and $\hat{\psi}_{ijkl}$ denote the ground-truth and predicted dihedral angles, respectively. The periodic formulation accounts for the 2π periodicity of dihedral angles.

Distance RMSD. We further evaluate the overall molecular geometry using the distance root-mean-square deviation (dRMSD), defined over all atom pairs as

$$\mathrm{dRMSD} = \sqrt{\frac{2}{N(N-1)} \sum_{i<j} \left(\hat{d}_{ij} - d_{ij} \right)^2}. \tag{20}$$

Lower values of all metrics indicate better preservation of local molecular geometry and overall molecular conformation.

Physical Evaluation on MD17. To further evaluate the structural and conformational quality of the predicted trajectories, we report additional geometry-based metrics under the MD17 state-to-trajectory (S2T) setting. Specifically, we compute the average bond-length MAE, bond-angle MAE, dihedral-angle MAE, and dRMSD over all queried future states. The results are summarized in Table 12.

D LIMITATION

Our experiments primarily focus on publicly available molecular dynamics datasets. We have not yet evaluated our method on large-scale protein systems, which often exhibit higher structural complexity and more diverse dynamical behaviors. We leave this important direction for future work and plan to further examine the scalability and generalization of our method in such challenging settings.

E VISUALIZATION RESULTS

To more clearly demonstrate the effectiveness of our method, we provide visualization results in Figure 3, Figure 4, and Figure 5.

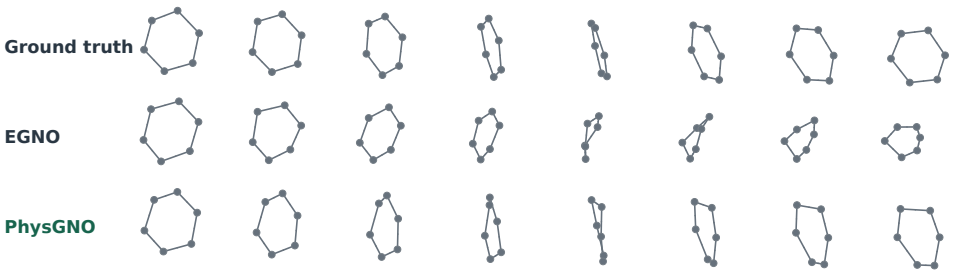

Figure 3: Trajectory comparison on Benzene.

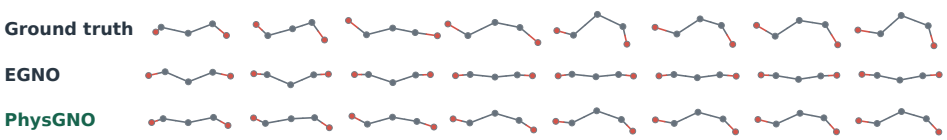

Figure 4: Trajectory comparison on Malonaldehyde.

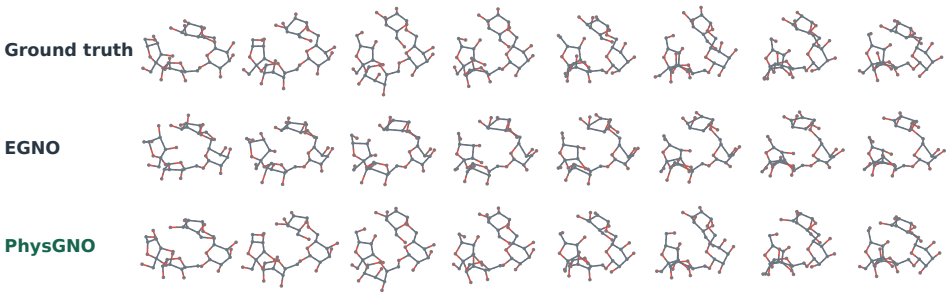

Figure 5: Trajectory comparison on Stachyose.