Learning Collective Variables from Time-lagged Generation

Anonymous Authors[1]

Abstract

Rare events such as state transitions are difficult to observe directly with molecular dynamics simulations due to long timescales. Enhanced sampling techniques overcome this by introducing biases along carefully chosen low-dimensional features, known as collective variables (CVs), which capture the slow degrees of freedom. Machine learning approaches (MLCVs) have automated CV discovery, but existing methods typically focus on discriminating meta-stable states without fully encoding the detailed dynamics essential for accurate sampling. We propose TLC, a framework that learns CVs directly from time-lagged conditions of a generative model. Instead of modeling the static Boltzmann distribution, TLC model a time-lagged conditional distribution yielding CVs to capture the slow dynamic behavior. We validate TLC on the Alanine Dipeptide system using two CV-based enhanced sampling tasks: (i) steered molecular dynamics (SMD) and (ii) on-the-fly probability enhanced sampling (OPES), demonstrating equal or superior performance compared to existing MLCV methods in both transition path sampling and state discrimination.

1. Introduction

Understanding rare events in molecular systems, such as ligand binding in drug discovery (De Vivo et al., 2016; Abel et al., 2017), conformational changes in protein folding (Piana et al., 2012; Seong et al., 2025), and phase transformations in materials science (Lookman et al., 2019; Spotte-Smith et al., 2022) is essential in biology and chemistry. However, these transitions involve crossing high free-energy barriers between meta-stable states, making them exceedingly rare and challenging to observe directly using conventional molecular dynamics (MD) simulations.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

To accelerate this sampling challenge, numerous enhanced sampling techniques have been developed. For example, replica-exchange MD (Sugita & Okamoto, 1999) exchanges configurations between parallel simulations at different temperatures, while accelerated MD (Hamelberg et al., 2004) globally boosts the potential energy surface to overcome energy barriers. Many prominent enhanced sampling techniques, including Metadynamics (Barducci et al., 2011) and on-the-fly probability enhanced sampling (Invernizzi & Parrinello, 2020, OPES), rely on biasing the simulations along a molecular configuration projected on to a set of coordinates, known as *Collective Variables*.

Collective variables (CVs) are low-dimensional functions of atomic coordinates designed to represent the transition-relevant slow degree of freedom (Torrie & Valleau, 1977; Valsson et al., 2016). By applying biases along these CVs in simulations, enhanced sampling techniques efficiently drive the configuration over energy barriers and enable transitions between meta-stable states. For example, Metadynamics (Barducci et al., 2011) and OPES (Invernizzi & Parrinello, 2020) employ time-dependent bias potentials to progressively fill free-energy wells along the CV space, thereby accelerating transitions. Furthermore, steered molecular dynamics (Izrailev et al., 1999; Fiorin et al., 2013, SMD) adds a harmonic restraint along the CVs, pulling the molecular configuration from one state to another. Although enhanced sampling techniques can operate without well-defined CVs, their efficiency, interpretability, and effectiveness are significantly reduced.

Recently, machine learning (ML) methods have emerged as a promising approach for automating CV discovery, reducing reliance on human intuition, domain knowledge, and extensive trial and error. Supervised methods, such as DeepLDA (Bonati et al., 2020) and DeepTDA (Trizio & Parrinello, 2021), train neural networks to discriminate labeled meta-stable states. Time-lagged methods, including DeepTICA (Bonati et al., 2021) and time-lagged autoencoders (Bonati et al., 2021; Wehmeyer & Noé, 2018, TAE) explicitly incorporate temporal correlations by reconstructing or predicting time-lagged configurations.

In this work, we propose TLC, a novel framework for discovering CVs from time-lagged conditional distributions learned via generative modeling. Using the transferable

Boltzmann generators (Klein & Noé, 2024, TBG), we model the time-lagged conditional distribution $p(x_{t+\tau} \mid x_t)$ of time lag τ, rather than the equilibrium Boltzmann distribution $p(x)$. Inspired by the concept of time-lagged encoder (Wehmeyer & Noé, 2018, TAE), we encode a molecular configuration x_t into a low-dimensional condition s_t and train the generative model to predict a time-lagged configuration $x_{t+\tau}$. resulting CVs to capture slow dynamics. Furthermore, we benchmark TLC against existing MLCVs approaches with two downstream enhanced sampling techniques; on-the-fly probability enhanced sampling (Invernizzi & Parrinello, 2020, OPES) and additionally steered molecular dynamics (Izrailev et al., 1999; Fiorin et al., 2013, SMD), on the Alanine Dipeptide system without using any transition data. In short, our contributions can be summarized as follows:

- We introduce a novel framework for learning collective variables from the time-lagged conditions of a generative modeling approach.
- We demonstrate that our MLCVs captures the slow degree of freedom with two CV-based enhanced sampling techniques, achieving competitive or superior performance compared to existing methods.

2. Background

Molecular dynamics simulations. Molecular dynamics (MD) describe the temporal evolution of molecular systems by integrating stochastic differential equations (SDEs). In particular, we consider under-damped Langevin dynamics (Bussi & Parrinello, 2007), which combine deterministic forces with stochastic fluctuations as follows:

$$\mathrm{d}x_t = v_t \,\mathrm{d}t\,, \tag{1}$$

$$\mathrm{d}v_t = \frac{-\nabla U(x_t)}{m}\,\mathrm{d}t - \gamma v_t\,\mathrm{d}t + \sqrt{\frac{2\gamma k_B T}{m}}\,\mathrm{d}W_t\,. \tag{2}$$

Here, x_t and v_t denotes atomic position and velocity at time t, m the diagonal matrix consisting of the mass of the corresponding atom, $U(x)$ the potential energy function, and $\nabla U(x_t)$ its gradient concerning position, i.e., the negative force. Parameters γ, k_B, T and W_t represents the friction coefficient, the Boltzmann constant, the absolute temperature, and the standard Brownian motion, respectively. Despite their theoretical accuracy, conventional MD simulation face practical time scale limitations that hinder observation of rare events, such as transitions between meta-stable states.

Enhanced sampling. Enhanced sampling techniques play a vital role in modern simulation techniques, overcoming timescale limitations inherent to standard MD simulations enabling efficient exploration of rarely visited molecular states (Torrie & Valleau, 1977; Valsson et al., 2016; Invernizzi & Parrinello, 2020; Barducci et al., 2011; Fiorin et al., 2013). Many enhanced samplings rely on collective variables (CVs) as reaction coordinate, biasing simulations along these coordinates to facilitate transitions. For example, on-the-fly probability-enhanced simulations (Invernizzi & Parrinello, 2020, OPES) construct time-dependent bias potential on previously observed CV values, accelerating rare transitions and exploring high-energy regions.

Collective variables (CVs). CVs are low-dimensional functions of atomic coordinates designed to capture the system's slow dynamical modes and essential transition pathways (Bonati et al., 2023). Formally, given a molecular configuration $x \in \mathbb{R}^{3N}$ where N is the number of particles, CVs are defined by a small set of functions $s = (\xi_i(x))_{i=1}^{k} (k \ll 3N)$ where $\xi_i(x)$ are scalar functions. For example, the two backbone dihedral angles ϕ, ψ are optimal CVs for the Alanine Dipeptide system. Additionally, effective CVs serve as reaction coordinates for enhanced sampling techniques such as Metadynamics (Barducci et al., 2011) and umbrella sampling (Torrie & Valleau, 1977; Laio & Parrinello, 2002). Importantly, an effective CV must satisfy three key criteria as follows:

- Capable of distinguishing meta-stable states
- Limited in number, ensuring low dimensionality
- Encoding the slow degree of freedom, i.e., characterizing the correct transition state when using a biasing force or potential to overcome the energy barrier

where the third criterion is considered particularly challenging (Fu et al., 2024; Barducci et al., 2011; Bonati et al., 2023). It ensures that the CV-based biasing force or potentials will guide the system over free energy barriers via physically and realistically transition paths, resulting in lower maximum energy in transition paths.

Machine learning CVs. DeepLDA (Bonati et al., 2020), DeepTDA (Trizio & Parrinello, 2021) have discovered CVs based on discriminant analysis methods, using binary labels dependent on ϕ. On the other hand, DeepTICA (Bonati et al., 2021), time-lagged autoencoder (Wehmeyer & Noé, 2018, TAE), and variational dynamics encoder (Hernández et al., 2018, VDE) have used time-lagged data to learn collective variables. To be specific, DeepTICA applies time-lagged independent component analysis (Molgedey & Schuster, 1994, TICA) on representation reduced by the encoder network, while TAE and VDE reconstruct a time-lagged configuration $x_{t+\tau}$ from the current configuration x_t with autoencoders (Rumelhart et al., 1985) and variational autoencoders (Kingma & Welling, 2014).

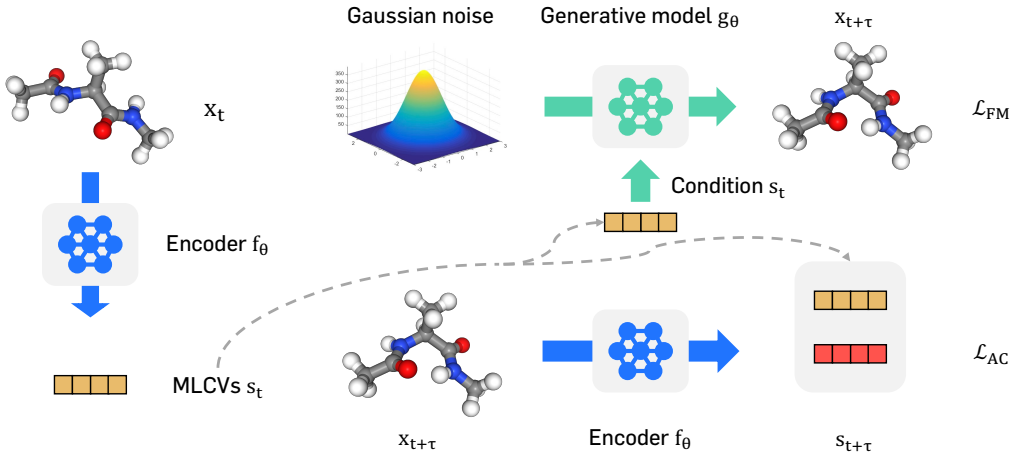


Figure 1. **Overview of our method.** We train an additional MLCV model $f_\theta()$ as conditions to a generative model $g_\theta()$ to learn the collective variables. To be specific, the MLCV model computes a reduced representation s_t from a frame x_t, while the generative model aims to construct the molecular configuration $x_{t+\tau}$ with a condition s_t.

3. Learning CVs from time-lagged conditions

In this section, we first outline our motivation, building upon prior methods and recent advances in generative models that approximate the Boltzmann distributions. We then present our proposed approach for learning collective variables from time-lagged conditions of generative models.

3.1. Generative models

Motivation. Previous works, such as TAE (Wehmeyer & Noé, 2018), utilize time-lagged data to learn collective variables. Given a molecular configuration x_t at time t, it reconstructs a time-lagged data by $x_{t+\tau} \approx h_{theta}(f_{theta}(x_t))$ where $f_theta()$ is an encoder and $h_{theta}()$ is a decoder of an autoencoder (Rumelhart et al., 1985). Stemming from this, we extend this approach using recent generative models that learn the Boltzmann distribution (Noé et al., 2019).

Continuous Normalizing Flows. We leverage generative models that learn the Boltzmann distribution (Klein & Noé, 2024, TBG) with continuous normalizing flow (Chen et al., 2018; Grathwohl et al., 2019, CNFs). CNFs map a simple prior distribution $p_0(x)$ to a target distribution $p_1(x)$, e.g., from a Gaussian noise to the Boltzmann distribution $q(x) \propto \exp(-U(x)/k_BT)$. Formally, the flow ϕ_r is defined by the ordinary differential equation (ODE) as follows:

$$\frac{\mathrm{d}}{\mathrm{d}r}\phi_r(x) = u_r\big(\phi_r(x)\big)\,, \quad \phi_0(x) \sim p_0\,, \tag{3}$$

where $u_r(x) : \mathbb{R}^n \to \mathbb{R}^n$ is a time-dependent vector field. Note that we use r for time index, to avoid confusion with the MD time step t. Generative model parameterizes the vector field $u_r(x)$ using E(3)-equivariant graph neural networks (Satorras et al., 2021, EGNN), enabling direct generation of molecular configurations in Cartesian coordinates. However, simulation-based training of CNFs is typically computationally expensive.

Flow matching. To alleviate the computational burden, textitflow matching (Lipman et al., 2023; Liu et al., 2023) is used, a simulation-free and computationally efficient training method. Specifically, flow matching directly trains a vector field by minimizing a regression between the predicted and conditional vector field $u_t(x|z)$ as follows:

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{r\sim[0,1],x\sim p_r(x|z)}\|v_\theta(x,r) - u_r(x|z)\|_2^2,. \tag{4}$$

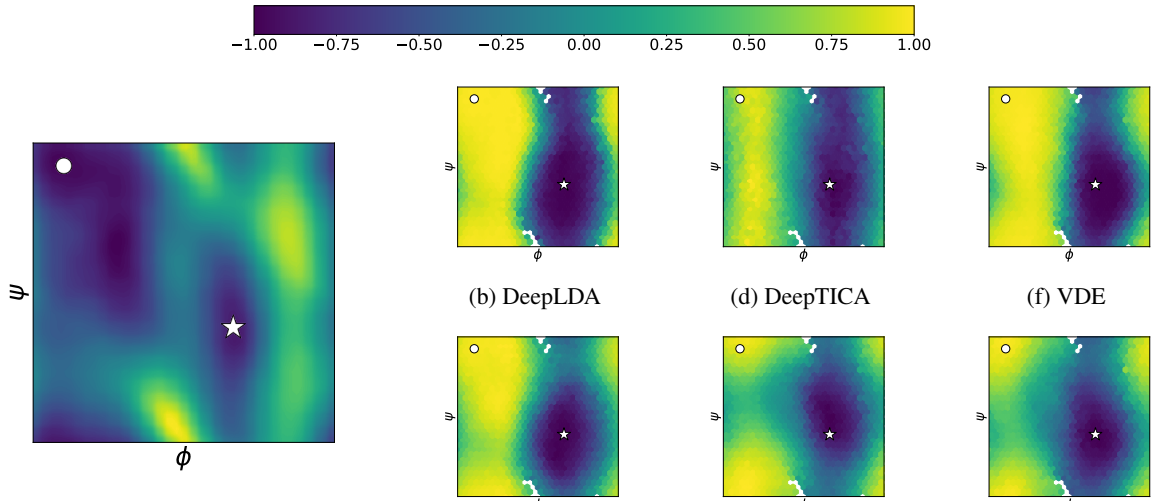
For the conditional vector field $u_t(x|z)$, a simple yet powerful parameterization as follows:

$$u_r(x|z) = x_1 - x_0\,, \tag{5}$$
$$p_r(x|z) = \mathcal{N}(x|r\cdot x_1 + (1-r)\cdot x_0, \sigma^2)\,, \tag{6}$$

where z is x_0.

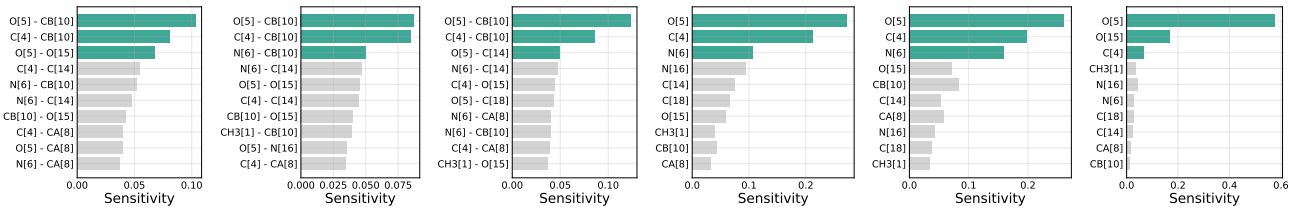
3.2. Conditional Boltzmann distribution

Time-lagged conditions. We now extend the generative backbone to model the conditional Boltzmann distribution $q(x_{t+\tau}|s_t)$, given a time-lagged molecular configuration pair $(x_t, x_{t+\tau})$ where t denotes the timestep in a simulation, τ denotes a fixed lag time. While one can consider learning conditions from a pre-trained model (Zhang et al., 2023), we train both the generative backbone and the MLCV encoder from scratch. Our MLCV encoder f_theta implemented as a simple MLP, compresses the current molecular configuration into a low-dimensional condition MLCV s_t. We concatenate s_t with the initial node features of the EGNN, thus conditioning the generative flow on this low-dimensional



(a) Energy landscape (b) DeepLDA (c) DeepTDA (d) DeepTICA (e) TAE (f) VDE (g) TLC

Figure 2. **Ramachandran plot of MLCVs** marginalized over the two dihedral angles and the energy landscape. Two meta-stable states $C5$ and $C7_{ax}$ are each denoted by a white circle and a star. For visualization and simplicity, collective variables are normalized from $(1, -1)$ based on Metadynamics samples, and the collective variables of the meta-stable state $C5$ are set to positive. While all methods discriminate two meta-stable states, DeepLDA, DeepTICA, and VDE fail to show the slow degree of freedom visualized in (a).



(a) DeepLDA (b) DeepTDA (c) DeepTICA (d) TAE (e) VDE (f) TLC

Figure 3. **MLCVs sensitivity** against the top ten input features, i.e., heavy atom distance or RMSD aligned heavy atom coordinates. Sensitivity is computed as the gradients of MLCVs against input features and averaged over the projection dataset.

representation. The conditional flow matching loss from Equation (4) incorporates these conditions as follows:

$$\mathcal{L}_{\text{TLC}}(\theta) = \mathbb{E}_{t,(x_t,x_{t+\tau})}\left[\|v_\theta(x_t, t|s_t) - u_t(x|z)\|^2\right], \quad (7)$$

$$s_t = f_\theta(x_t), \quad (8)$$

Intuitively, the MLCV encoder is encouraged to encode information capturing the slow degree of freedom in the molecular system, as it learns the distribution of future molecular configurations from the current state. We provide an overview of our method in Figure 1.

Autocorrelation loss. Inspired by the Variational Dynamics Encoders (Hernández et al., 2018, VDE), we further propose an additional autocorrelation loss to ensure temporal consistency in the learned CVs $s_t, s_{t+\tau}$. Maximizing the autocorrelation results in CVs to remain similar over the time lag τ, highlighting the slow degree of freedom. Formally, we define the autocorrelation loss as follows:

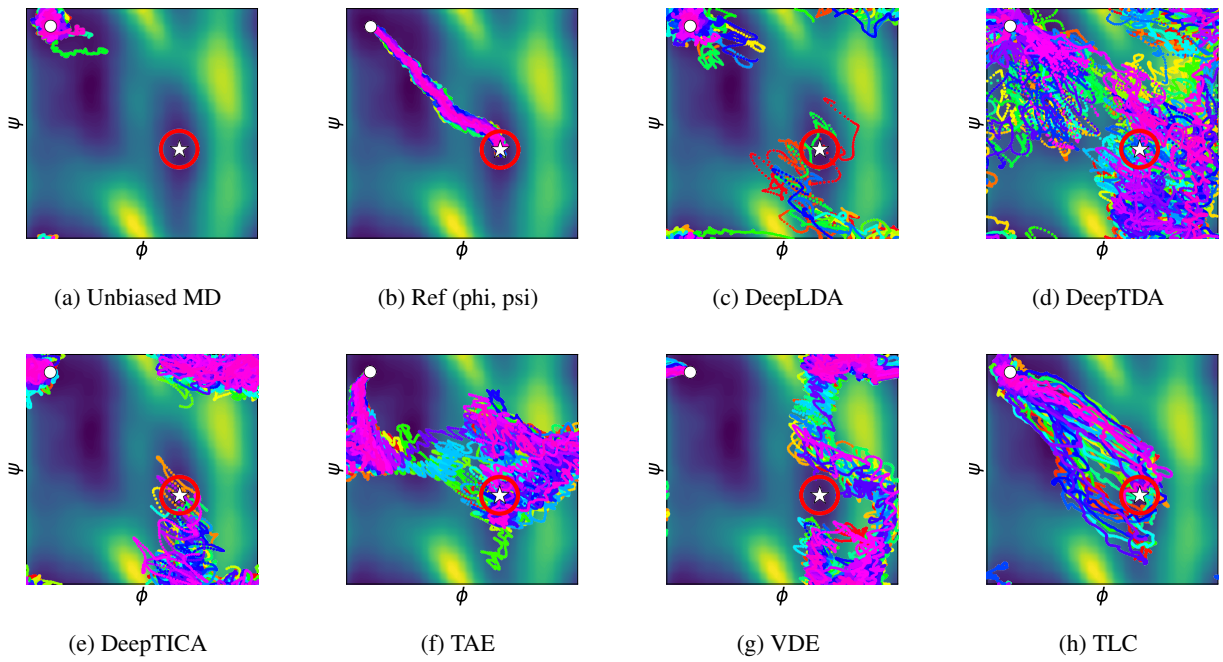
$$\mathcal{L}_{\text{AC}}(\theta) = -\frac{\mathbb{E}[(s_t - \overline{s}_t)(s_{t+\tau} - \overline{s}_{t+\tau})]}{\sigma_{s_t}\sigma_{s_{t+\tau}}}, \quad (9)$$

where $\overline{s}_t$ and σ_{s_t} denote the mean and standard deviation of encoded collective variables for a batch of data. Eventually, we combine this loss with the conditional flow matching loss into the following:

$$\mathcal{L}_{\text{total}}(\theta) = \mathcal{L}_{\text{TLC}}(\theta) + \lambda\mathcal{L}_{\text{AC}}(\theta) \quad (10)$$

where λ is a scaling factor for the autocorrelation loss. We provide ablation studies demonstrating the benefit of the autocorrelation loss in Appendix A.

Invariant representations. While prior generative models utilize SE(3)-equivariant flows to generate Cartesian coordinates of molecular configurations, CVs should remain invariant under rotations and translations. Unlike prior works using invariant features such as heavy atom distances



(a) Unbiased MD (b) Ref (phi, psi) (c) DeepLDA (d) DeepTDA

(e) DeepTICA (f) TAE (g) VDE (h) TLC

Figure 4. **Ramachandran plot of 64 trajectories** of length 1000 fs by unbiased MD and steered MD simulations using MLCVs. Initial state $C5$ and the target state $C7_{ax}$ are each denoted as white circles and stars. The red circle indicates the target hit region.

(Bonati et al., 2020; Trizio & Parrinello, 2021; Bonati et al., 2021), we retain raw Cartesian coordinates but enforce invariance through rigid-body alignment to a reference configuration, minimizing the Euclidean distance. Specifically, we align every configuration to the $C5$ meta-stable state via the Kabsch algorithm (Kabsch, 1976). This RMSD-based alignment significantly enhanced the efficacy of our learned CVs, as we experimentally shwosn in Appendix A.

4. Experiments

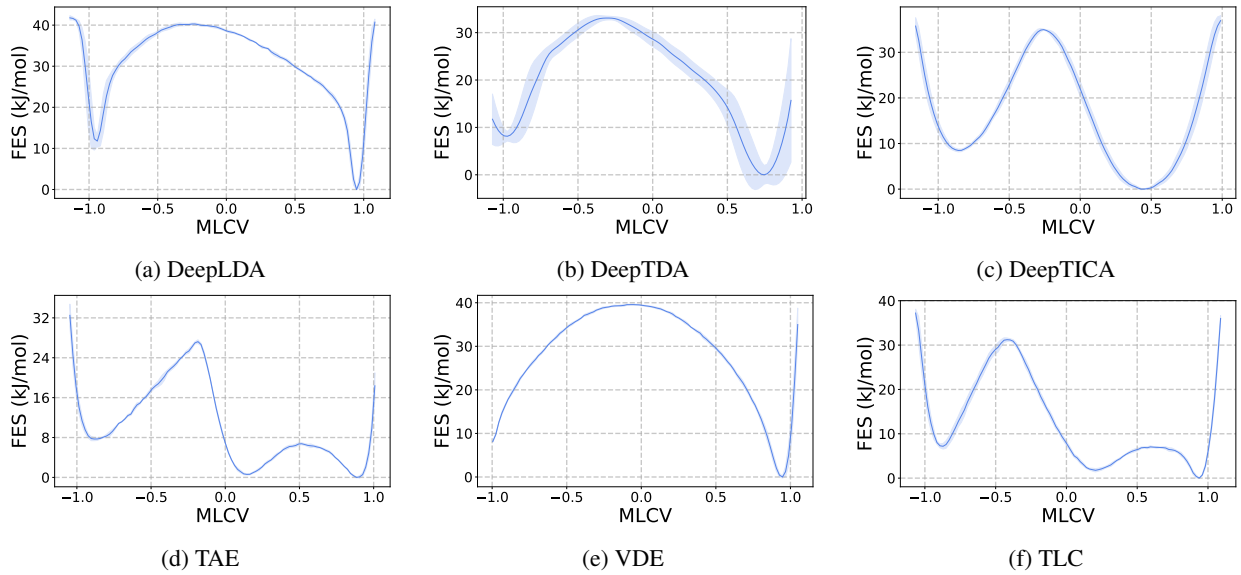
In this section, we evaluate how well our machine-learned collective variables (MLCVs) capture the system's slow degrees of freedom. First, steered molecular dynamics (Izrailev et al., 1999; Fiorin et al., 2013, SMD), applying a biasing force based on MLCVs to measure its ability to drive transitions along slow modes. Additionally, we use on-the-fly probability-enhanced sampling (Invernizzi & Parrinello, 2020, OPES) to compare the distributions sampled by MLCVs against those from known, optimal CVs (dihedral angles) as in Bonati et al. (2020). All experiments are conducted on the Alanine Dipeptide system. Note that the optimal CVs for the two enhanced sampling techniques are different, where SMD requires a slow degree of freedom related to transitions, while OPES requires the slowest degree of freedom between two meta-stable states.

Alanine Dipeptide. Alanine Dipeptide is a widely studied molecular system consisting of 22 atoms, where the backbone dihedral angles ϕ and ψ are known to be the optimal collective variables. We use two meta-stable states defined by these angles: $C5$ at $(-2.49,\ 2.67)$ and $C7_{ax}$ at $(1.02,\ -0.70)$ in the (ϕ, ψ) space. While we do not use these angles directly during training, we use them for ground-truth references and visualization purposes.

Simulation data. To ensure a fair comparison, all models were trained on identical datasets, with the MLCV dimension fixed to one. We generate ten 10 ns trajectories using OpenMM (Eastman et al., 2023), initializing five trajectories each in the $C5$ and $C7_{ax}$ meta-stable states. Training data were then randomly extracted from these trajectories, explicitly excluding transition events, i.e., the sign of ϕ remains consistent between any paired time-lagged data x_t and $x_{t+\tau}$. We provide more details in Appendix B.

Baselines. We compare our approach TLC with both supervised and time-lagged methods. Supervised baselines include DeepLDA (Bonati et al., 2020) and DeepTDA (Trizio & Parrinello, 2021), which rely on ϕ-based binary labels. Time-lagged approaches include DeepTICA (Bonati et al., 2021), time-lagged autoencoder (Wehmeyer & Noé, 2018, TAE), and variational dynamics encoder (Hernández et al., 2018, VDE). For additional details, refer to Appendix C.

Visualization. Also, we visualize the MLCVs in Figure 2. To be specific, we collect diverse configurations with Metadynamics and marginalize the values over the dihedral angles. All methods distinguish the two meta-stable states $C5$



(a) DeepLDA (b) DeepTDA (c) DeepTICA

(d) TAE (e) VDE (f) TLC

Figure 5. **Free energy surface (FES)** of MLCVs, averaged over four OPES simulations. Samples from OPES simulations are reweighted to compute the free energy for each MLCVs value, and local minima in each FES curve refer to capturing the meta-stable state basin. DeepLDA, DeepTDA, and DeepTICA show two meta-stable state basins, while TAE and TLC show three meta-stable state basins. However, VDE only captures one meta-stable state basin.

and $C7_{ax}$, while the detailed slow degree of freedom differs. For details on Metadynamics, refer to Appendix D.

Sensitivity analysis. Finally, we present the sensitivity of MLCVs against the input (Bonati et al., 2023) in Figure 3. Among input features, we plot the top ten features and color-highlight the top three. Y-axis denotes the atom type and index in the Alanine Dipeptide system. Dihedral angles are computed from the 4, 6, 8, 14, and 16th atoms, MLCVs show high correlation with the dihedral angles.

4.1. Steered Molecular Dynamics (SMD)

SMD (Izrailev et al., 1999; Fiorin et al., 2013) is an enhanced sampling technique that steers the molecular configuration to a target state with a time-dependent bias. It requires the CV to encode the system's slow degree of freedom not only for distinguishing the meta-stable states, but also for transition paths between the meta-stable states. To be specific, the bias is computed as the time interpolation of the initial and target state CVs as follows:

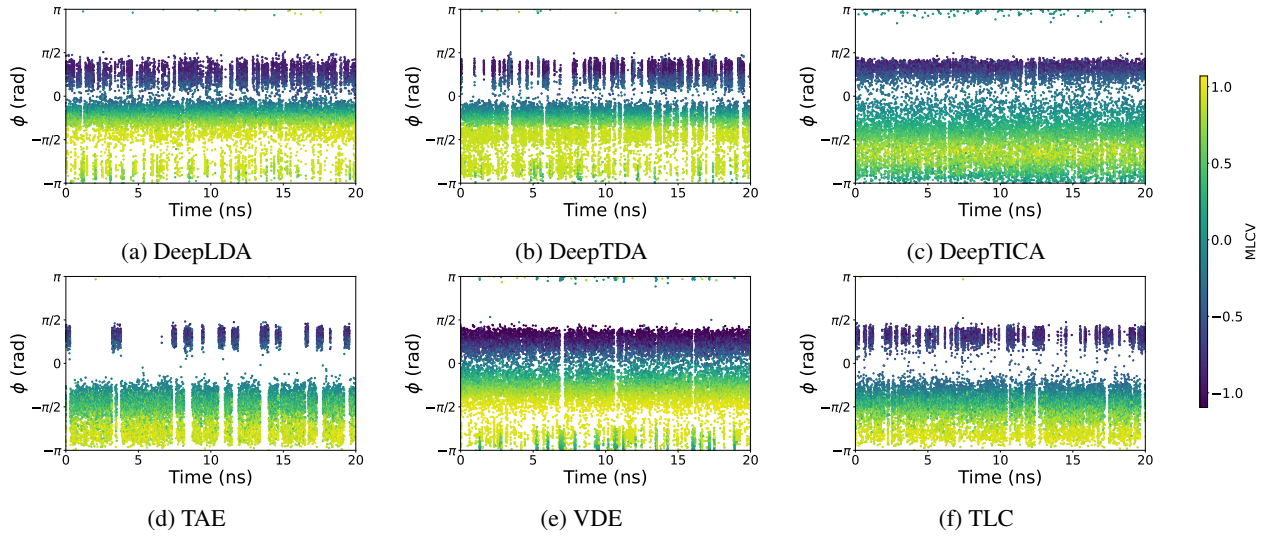
$$U(x,t) = \frac{k}{2}\left(\frac{ts_{\text{target}} + (T-t)s_{\text{initial}}}{T} - f_\theta(x)\right)^2, \quad (11)$$

where t denotes the current time step, T the simulation time horizon, k the force constant, x the current molecular configuration, and $s_{\text{target}}, s_{\text{initial}}$ each denotes the CVs of the initial and target meta-stable state. Intuitively, the bias potential of Equation (11) encourages the CVs to linearly evolve towards the target meta-stable state value starting from the initial meta-stable state value. If machine learned CVs well reflects the slow degree of freedom, the system will transition smoothly with a minimum energy penalty.

Table 1. **Quantitative metrics** of MLCV SMD simulations. RMSD and target hit percentage (THP) are averaged over 256 trajectories, while max energy (E_{TS}) is averaged over trajectories only hitting the target state. Best results are highlighted in **bold** and second in underline, excluding the reference simulation.

Method	k	RMSD (↓) Å	THP (↑) %	E_{TS} (↓) kJmol^{-1}
Ref (phi, psi)	200	1.0640	100.00	-3.89 ± 5.80
DeepLDA	600	1.1678	3.90	887.50 ± 211.36
DeepTDA	500	1.1043	48.04	904.06 ± 261.26
DeepTICA	400	0.9729	8.59	814.52 ± 115.74
TAE	1200	1.0086	58.59	755.41 ± 92.30
VDE	700	**0.8582**	5.08	901.69 ± 115.59
TLC (Ours)	300	0.9593	**60.93**	**33.58 ± 15.19**

Metrics. We quantitatively evaluate transition path from MLCVs steered MD of length 1000 fs with three metrics (Seong et al., 2025; Holdijk et al., 2023): (i) root mean square distance (RMSD), (ii) target hit percentage (THP), and (iii) transition state energy (E_{TS}). RMSD computes the Euclidean distances between atoms of the closest state in the transition path to the target state, aligning the states using the Kabsch algorithm (Kabsch, 1976). Next, THP measures the number of paths that arrive near the target meta-stable state with a dihedral angle threshold, i.e., L2-distance smaller than $0.5°$ for the two dihedral angles ϕ and ψ. Finally, the transition state energy measures the ability



(a) DeepLDA (b) DeepTDA (c) DeepTICA

(d) TAE (e) VDE (f) TLC

Figure 6. **ϕ-angle distribution** of a single OPES Metadynamics simulation. ϕ distribution close to a uniform indicates that the MLCVs successfully mimic the effect of the known optimal CVs, i.e., the backbone dihedral angle ϕ. While most methods show a uniform ϕ distribution, TAE fails to show frequent transitions between the two meta-stable states.

to identify the transition state in terms of energy, a lower energy would refer to more physically realistic transition paths. Additionally, we sweep the force constant k since there exists a tradeoff between the *target hit* probability with *max energy*, and report the highest success rate under a maximum energy threshold of 1000 kJmol^{-1}.

Transition paths. In Table 1, one can see that TLC outperforms baselines in for THP and E_{TS}, while DeepTICA shows the smallest RMSD. Surprisingly, TLC yields much lower values in E_{TS} compared to prior works, implying TLC generates a realistic transition path from meta-stable state $C5$ to meta-stable state $C7_{ax}$. Additionally, in Figure 4, prior works mostly reach the target meta-stable state, ignoring the energy landscape. In contrast, TLC crosses the low energy points in the energy barrier located at $\phi = 0$ and reaches the target meta-stable state with high probability.

4.2. On-the-fly Probability-Enhanced Sampling (OPES)

OPES (Invernizzi & Parrinello, 2020) is an enhanced sampling technique that adaptively constructs a bias potential to accelerate exploration in the CV space. It aims for an equilibrium sampling of the molecular configuration. To be specific, the probability distribution at the n-th iteration is as follows:

$$P_n(s) = \frac{\sum_k^n w_k G(s, s_k)}{\sum_k^n w_k}, \quad w_k = e^{\beta V_{k-1}(s_k)}, \qquad (12)$$

where w_k denotes the bias potential of the previous iteration, β the inverse temperature, and $G(s, s_k)$ the multivariate Gaussian. Additionally, the bias potential $V_n(s)$ in Equation (12) is computed as follows:

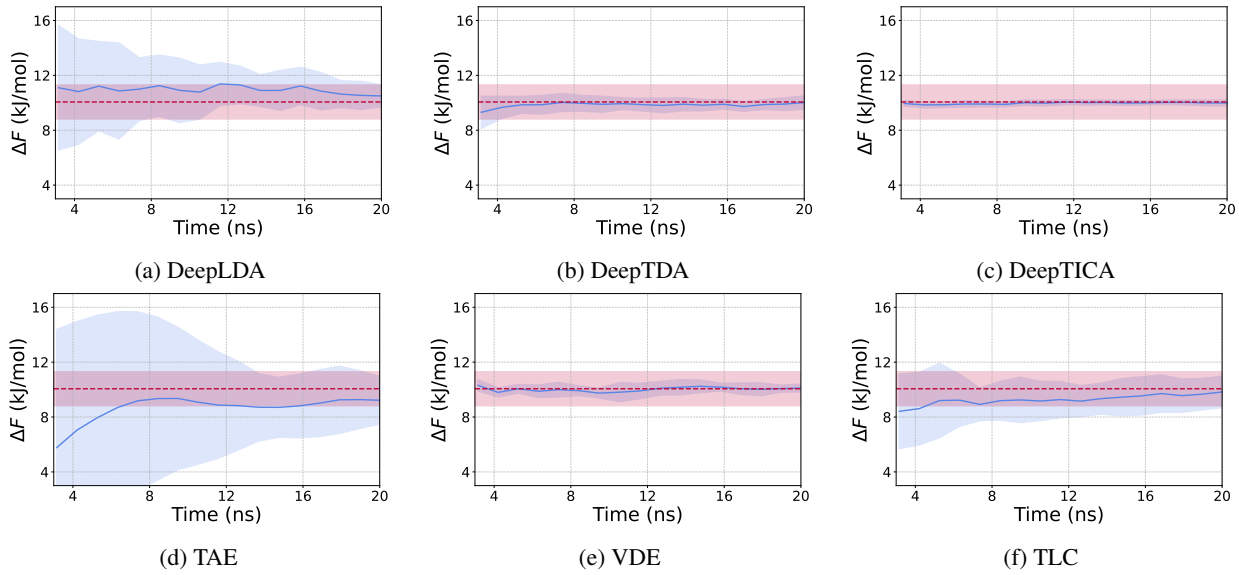
$$V_n(s) = \left(1 - \frac{1}{\gamma}\right)\frac{1}{\beta}\log\left(\frac{P_n(s)}{Z_n} + \epsilon\right) \qquad (13)$$

where Z_n denotes the normalization factor, γ the broadening of the base distribution, and ϵ is a regularization term limiting the maximum value of the bias for the exploration of higher free-energy regions. Intuitively, OPES adds bias in the CV-space targeting a uniform distribution, where CVs encoding the slow degree of freedom would result in better exploration. Results are averaged over four independent simulations, and a 100 ns OPES reference simulation using the dihedral angles ϕ and ψ serves as the ground truth.

Table 2. **Free-energy difference** ΔF between two meta-stable state $C5$ and $C7_{ax}$, averaged over four OPES simulations. Free energy difference values within the range of $0.5\ k_BT \approx 1.25\ \mathrm{kJmol}^{-1}$ from the value of reference OPES simulations are considered to capture the slow degree of freedom.

Method	SIGMA	ΔF
Ref (ϕ, ψ)	0.05	10.06 ± 0.22
DeepLDA	0.05	10.50 ± 0.80
DeepTDA	0.20	10.01 ± 0.49
DeepTICA	0.10	9.99 ± 0.21
TAE	0.05	9.22 ± 1.74
VDE	0.05	10.11 ± 0.28
TLC (Ours)	0.05	9.83 ± 1.15

Free energy surface. First of all, we plot the free energy surface (FES) along the MLCVs in Figure 5. The FES is



(a) DeepLDA (b) DeepTDA (c) DeepTICA

(d) TAE (e) VDE (f) TLC

Figure 7. **Free energy difference** between two basins averaged over four OPES Metadynamics simulations. The first 3 ns have been discarded, and ΔF is updated every 1 ns. Convergence to the reference value within $0.5\ k_BT \approx 1.25\text{kJmol}^{-1}$, i.e, the red region, is considered to reproduce the known CVs (two dihedral angles).

computed by binning the MLCV values sampled during the OPES simulations and applying Boltzmann inversion to estimate free energy. These plots illustrate different methods' ability to recover the system's slow degree of freedom. DeepLDA, DeepTDA, and DeepTICA capture two metastable states basins, while TAE and TLC identify three metastable states. However, VDE falls short on the FES, where it only shows one metastable state basin.

Phi distribution. Next, we compare the ϕ distribution of the OPES simulations as in Bonati et al. (2020); Trizio & Parrinello (2021). CVs capturing the slowest degree of freedom will distinguish and keep drive transitions between the two meta-stable states throughout the OPES simulation, mimicking the effect of dihedral angles. In Figure 6, all methods except TAE effectively drive transitions between the two meta-stable states, validating their ability to capture the slow degree of freedom as in dihedral angles.

Free energy convergence. Finally, we monitor the convergence of the free energy difference between two basins. CVs capturing the slow degree of freedom will effectively drive the transition between meta-stable states and result in a similar free energy difference. Additionally, free energy differences falling within the range of $0.5\ k_BT \approx 1.25\ \text{kJmol}^{-1}$ from the reference value are considered to reproduce the slowest degree of freedom (Invernizzi & Parrinello, 2020; Bonati et al., 2020). Formally, the free energy difference between two basins is defined as follows:

$$\Delta F = \frac{1}{\beta} \log \frac{\int_A e^{-\beta F(\phi)} \mathrm{d}\phi}{\int_B e^{-\beta F(\phi)} \mathrm{d}\phi}, \tag{14}$$

where β denotes the inverse temperature, $F(\phi)$ the reweighted free energy, A and B each the regions corresponding to $\phi > 0$ and $\phi < 0$. The first 3 ns of the OPES simulations are discarded, and ΔF is updated every 1 ns (Bonati et al., 2020). In Figure 7 and Table 2, most methods quickly converge to the reference free energy difference value, while TAE exhibits high variance.

5. Conclusion

We present a framework for learning collective variables from the time-lagged conditions in a generative model, capturing the slow degree of freedom. While VDE has first applied generative models to learn collective variables, it shows poor performance for steered MD. On the other hand, TLC shows superior performance in steered molecular dynamics tasks and competitive performance in free energy convergence for OPES simulations. An interesting future work would be investigating which collective variables are optimal for each enhanced sampling tasks.

Impact Statement

This work advances machine learning methods for molecular simulation by improving the automated discovery of collective variables, which may accelerate research in drug design and materials science. While our methods pose no direct ethical risks, they contribute to broader capabilities in modeling complex chemical and biological systems.

References

Abel, R., Wang, L., Harder, E. D., Berne, B., and Friesner, R. A. Advancing drug discovery through enhanced free energy calculations. *Accounts of chemical research*, 50 (7):1625–1632, 2017. 1

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1: 19–25, 2015. 12

Barducci, A., Bonomi, M., and Parrinello, M. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5):826–843, 2011. 1, 2

Bonati, L., Rizzi, V., and Parrinello, M. Data-driven collective variables for enhanced sampling. *The journal of physical chemistry letters*, 11(8):2998–3004, 2020. 1, 2, 5, 8, 13

Bonati, L., Piccini, G., and Parrinello, M. Deep learning the slow modes for rare events sampling. *Proceedings of the National Academy of Sciences*, 118(44):e2113533118, 2021. 1, 2, 5, 13

Bonati, L., Trizio, E., Rizzi, A., and Parrinello, M. A unified framework for machine learning collective variables for enhanced sampling simulations: mlcolvar. *The Journal of Chemical Physics*, 159(1), 2023. 2, 6

Bussi, G. and Parrinello, M. Accurate sampling using langevin dynamics. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 75(5):056707, 2007. 2

Chen, R. T., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018. 3

De Vivo, M., Masetti, M., Bottegoni, G., and Cavalli, A. Role of molecular dynamics and related methods in drug discovery. *Journal of medicinal chemistry*, 59(9):4035–4061, 2016. 1

Dolezal, R., Fronckova, K., Kirimtat, A., and Krejcar, O. Computational complexity of kabsch and quaternion based algorithms for molecular superimposition in computational chemistry. In *Proceedings of the 21st EANN (Engineering Applications of Neural Networks) 2020 Conference: Proceedings of the EANN 2020 21*, pp. 473–486. Springer, 2020. 11

Eastman, P., Galvelis, R., Peláez, R. P., Abreu, C. R., Farr, S. E., Gallicchio, E., Gorenko, A., Henry, M. M., Hu, F., Huang, J., et al. Openmm 8: molecular dynamics simulation with machine learning potentials. *The Journal of Physical Chemistry B*, 128(1):109–116, 2023. 5, 12

Fiorin, G., Klein, M. L., and Hénin, J. Using collective variables to drive molecular dynamics simulations. *Molecular Physics*, 111(22-23):3345–3362, 2013. 1, 2, 5, 6

Fu, H., Bian, H., Shao, X., and Cai, W. Collective variable-based enhanced sampling: From human learning to machine learning. *The journal of physical chemistry letters*, 15(6):1774–1783, 2024. 2

Grathwohl, W., Chen, R. T., Bettencourt, J., and Duvenaud, D. Scalable reversible generative models with free-form continuous dynamics. In *International Conference on Learning Representations*, volume 3, 2019. 3

Hamelberg, D., Mongan, J., and McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *The Journal of chemical physics*, 120(24):11919–11929, 2004. 1

Hernández, C. X., Wayment-Steele, H. K., Sultan, M. M., Husic, B. E., and Pande, V. S. Variational encoding of complex dynamics. *Physical Review E*, 97(6):062412, 2018. 2, 4, 5, 13

Holdijk, L., Du, Y., Hooft, F., Jaini, P., Ensing, B., and Welling, M. Stochastic optimal control for collective variable free sampling of molecular transition paths. *Advances in Neural Information Processing Systems*, 36: 79540–79556, 2023. 6

Invernizzi, M. and Parrinello, M. Rethinking metadynamics: from bias potentials to probability distributions. *The journal of physical chemistry letters*, 11(7):2731–2736, 2020. 1, 2, 5, 7, 8

Izrailev, S., Stepaniants, S., Isralewitz, B., Kosztin, D., Lu, H., Molnar, F., Wriggers, W., and Schulten, K. Steered molecular dynamics. In *Computational Molecular Dynamics: Challenges, Methods, Ideas: Proceedings of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin, May 21–24, 1997*, pp. 39–65. Springer, 1999. 1, 2, 5, 6

Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Foundations of Crystallography*, 32(5): 922–923, 1976. 5, 6, 13

Kingma, D. P. and Welling, M. Auto-encoding variational bayes. *International conference on Learning representations*, 2014. 2

Klein, L. and Noé, F. Transferable boltzmann generators. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*. PMLR, 2024. 2, 3

Laio, A. and Parrinello, M. Escaping free-energy minima. *Proceedings of the national academy of sciences*, 99(20): 12562–12566, 2002. 2

Lipman, Y., Chen, R. T., Ben-Hamu, H., Nickel, M., and Le, M. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023. 3

Liu, X., Gong, C., and Liu, Q. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations*, 2023. 3

Lookman, T., Balachandran, P. V., Xue, D., and Yuan, R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials*, 5(1):21, 2019. 1

Molgedey, L. and Schuster, H. G. Separation of a mixture of independent signals using time delayed correlations. *Physical review letters*, 72(23):3634, 1994. 2, 13

Noé, F., Olsson, S., Köhler, J., and Wu, H. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457):eaaw1147, 2019. 3

Piana, S., Lindorff-Larsen, K., and Shaw, D. E. Protein folding kinetics and thermodynamics from atomistic simulation. *Proceedings of the National Academy of Sciences*, 109(44):17845–17850, 2012. 1

Rumelhart, D. E., Hinton, G. E., Williams, R. J., et al. Learning internal representations by error propagation, 1985. 2, 3

Satorras, V. G., Hoogeboom, E., and Welling, M. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021. 3

Seong, K., Park, S., Kim, S., Kim, W. Y., and Ahn, S. Transition path sampling with improved off-policy training of diffusion path samplers. *arXiv preprint arXiv:2405.19961*, 2025. 1, 6

Spotte-Smith, E. W. C., Kam, R. L., Barter, D., Xie, X., Hou, T., Dwaraknath, S., Blau, S. M., and Persson, K. A. Toward a mechanistic model of solid–electrolyte interphase formation and evolution in lithium-ion batteries. *ACS Energy Letters*, 7(4):1446–1453, 2022. 1

Sugita, Y. and Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters*, 314(1-2):141–151, 1999. 1

Torrie, G. M. and Valleau, J. P. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of computational physics*, 23(2): 187–199, 1977. 1, 2

Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. Plumed 2: New feathers for an old bird. *Computer physics communications*, 185(2):604–613, 2014. 12

Trizio, E. and Parrinello, M. From enhanced sampling to reaction profiles. *The Journal of Physical Chemistry Letters*, 12(35):8621–8626, 2021. 1, 2, 5, 8, 13

Valsson, O., Tiwary, P., and Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annual review of physical chemistry*, 67(1):159–184, 2016. 1, 2

Wehmeyer, C. and Noé, F. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. *The Journal of chemical physics*, 148(24), 2018. 1, 2, 3, 5, 13

Zhang, L., Rao, A., and Agrawala, M. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3836–3847, 2023. 3

A. Ablation studies

We conduct ablation studies to empirically verify two components of our framework, and report results in Table 3 and Figure 8.

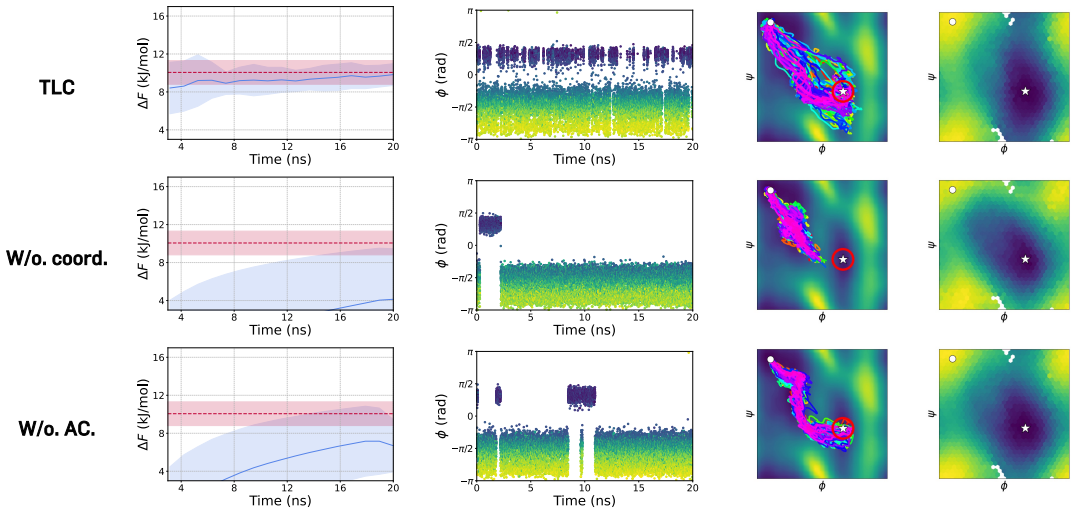


Figure 8. **Enhanced sampling results on component ablation studies**. Each column from the left refers to the free energy difference convergence, the phi distribution of OPES simulations, transition paths from SMD simulations, and MLCV Ramachandran plots.

Table 3. **Ablation studies** on the components of TBGCV. *w/o coord.* refers to using heavy atom distances instead of RMSD-aligned heavy atom coordinates. Best results are highlighted in **bold**.

	TLC	W/o coord.	W/o ac. loss
ΔF	**9.83 ± 1.15**	4.13 ± 5.32	9.15 ± 1.66
RMSD (↓)	**0.9593**	4.8773	3.4465
THP (↑)	60.93	7.81	**97.27**
E_{max} (↓)	**33.58. ± 15.19**	195.63 ± 19.55	116.58 ± 24.02

Autocorrelation loss. We introduced an additional autocorrelation loss term for our framework, which results in better convergence in the free energy difference. In Figure 8, one can see that autocorrelation loss results in better free energy difference convergence and a uniform phi distribution in OPES simulations. Nevertheless, there exists a minor tradeoff in the performance of SMD simulations, where the autocorrelation loss degrades performance.

Input representation. While prior works mainly use heavy atom distance as input representation, we instead propose to use Cartesian coordinates with RMSD aligned to a reference state, e.g., the $C5$ meta-stable state. Consequently, we validate the effectiveness of RMSD-aligned Cartesian coordinates against heavy atom distance. In Figure 8 and Table 3, one can see that using RMSD-aligned coordinates clearly shows better performance compared to heavy atom distance for both OPES and SMD simulations. We also note that the Kabsch algorithm operates in $O(n)$ where n denotes the number of atoms (Dolezal et al., 2020).

B. Dataset details

Table 4. Simulation details for collecting training data and projection data.

	Engine	Time horizon	Time step	Force field	Solvent	temperature (K)
Training	OpenMM	10 ns	1 fs	amber99sbildn	tip3p	300
Projection	GROMACS	100 ns	2 fs	amber99sbildn	tip3p	300

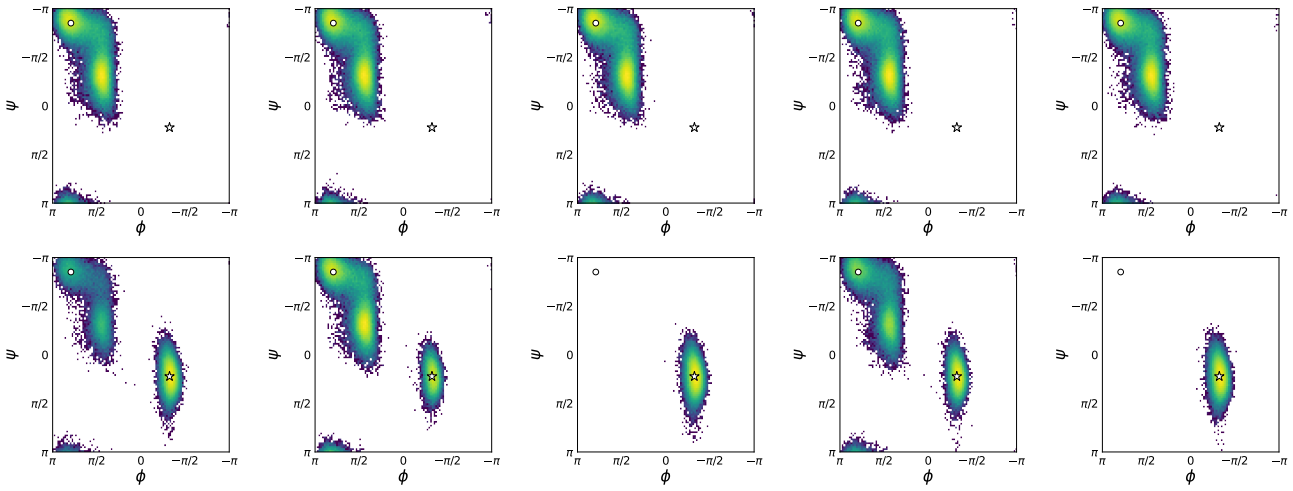


Figure 9. **10ns simulation trajectories** plotted on the Ramachandran plot. White circle and star each indicate meta-stable states $C5$ and $C7_{ax}$, respectively. The top and bottom row each refers to simulations starting from $C5$ and $C7_{ax}$ meta-stable states.

Training dataset. Data used for training models were all collected from simulations run by OpenMM (Eastman et al., 2023). From each meta-stable state $C5$ and $C7_{ax}$, we run 5 10 ns simulations with a record frequency of 100 fs. Afterwards, we randomly sample configurations from the trajectory. For the case of time-lagged data, we sample configurations with a time lag of 1000 fs. No transition data, i.e., time-lagged data where the sign of ϕ is opposite, were included in the dataset.

Projection dataset. Alanine Dipeptide configurations used for projection and normalizing MLCVs were collected from $100ns$ Metadynamics simulations by GROMACS (Abraham et al., 2015) and PLUMED (Tribello et al., 2014). Coordinates were recorded at a 100 fs frequency.

C. Experimental details

C.1. Baselines

In this section, we provide the details of experiments and baselines. We report the detailed model configuration in Table 5. For fair comparison, we used 100 for the hidden dimension across all models.

Input representation. We use heavy-atom-related information for input descriptors. For DeepLDA (Bonati et al., 2020), DeepTDA (Trizio & Parrinello, 2021), and DeepTICA (Bonati et al., 2021), we use heavy atom distance as denoted, i.e., distances between atoms excluding Hydrogen. For the Alanine Dipeptide system, ten heavy atoms exist, resulting in 45 input descriptors. For other models, we use heavy atom coordinates by aligning the configuration to the $C5$ meta-stable state with the Kabsch algorithm (Kabsch, 1976).

Time-lag. We fix the time-lag tau at 1000 fs for all time-lagged methods. Importantly, no true transition events, i.e., crossing between $C5$ and $C7_{ax}$, are included in the training pairs $(x_t, x_{t+\tau})$, ensuring that models do not simply memorize completed transitions.

DeepLDA, DeepTDA. Both are supervised, discriminant-analysis approaches, where an encoder network maps the input descriptors to an MLCV. Binary labels are used, dependent on the ϕ sign.

DeepTICA. DeepTICA combines a neural encoder with Time-lagged Independent Component Analysis (TICA) (Molgedey & Schuster, 1994). It maximizes the autocovariance of the learned one-dimensional CV at lag τ, capturing the slowest linear combination of features.

Time-lagged autoencoder (TAE). TAE is an unsupervised, reconstruction-based method (Wehmeyer & Noé, 2018). Its encoder–decoder architecture is trained to reconstruct the future configuration $x_{t+\tau}$ from x_t via a low-dimensional bottleneck CV, encouraging that CV to encode predictive, slow-varying information.

Variational Dynamics Encoder (VDE). VDE (Hernández et al., 2018) extends the TAE with a variational autoencoder, framing future-frame prediction as a latent-variable model. While C_α contact distances were used at Hernández et al. (2018), we use RMSD aligned heavy atom distances since only two alpha carbons exist in the Alanine Dipeptide system.

Table 5. **Details on model configurations.** H.A. refers to heavy atoms.

Model	Layers	Input Representation	Equi/in-variance
DeepLDA	[45, 100, 100, 100, 1]	H.A. distance	Invariance
DeepTDA	[45, 100, 100, 1]	H.A. distance	Invariance
DeepTICA	[45, 100, 100, 3]	H.A. distance	Invariance
TAE	[30, 100, 100, 1]	H.A. coordinate	Invariance (RMSD align)
TLC	[30, 100, 100, 1]	H.A. coordinate	Invariance (RMSD align)

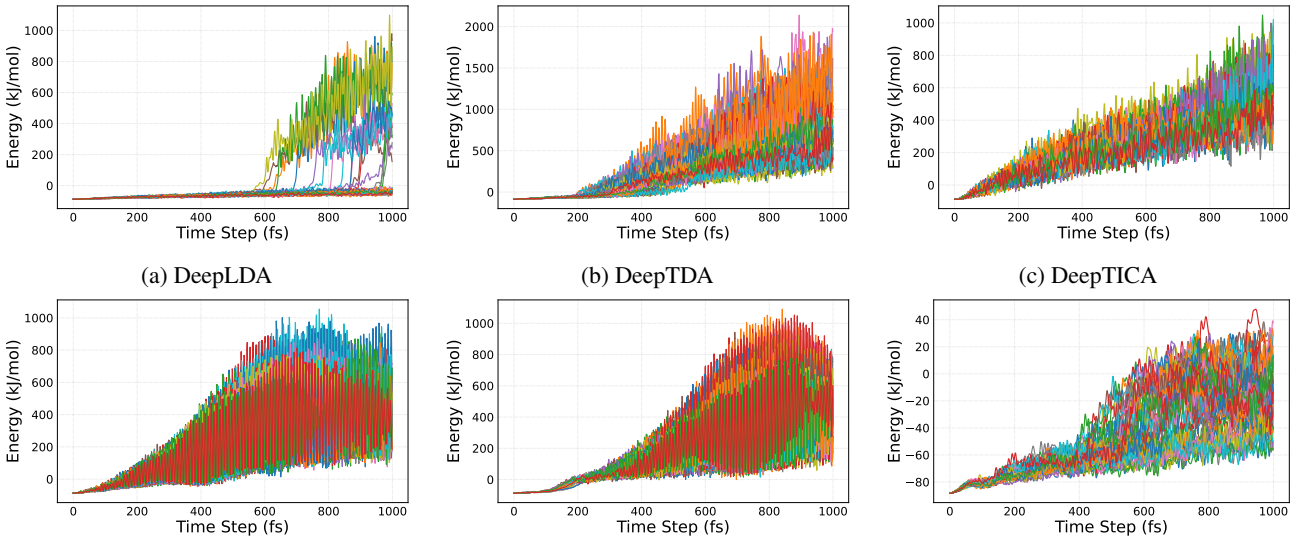
C.2. Enhanced samplings

OPES. We all use PACE of 500 and BARRIER of $30\mathrm{kJmol}^{-1}$, with record frequency of 500.

SMD. We search force constant ranging from 100 to 1000, with steps of 100.

D. Additional results

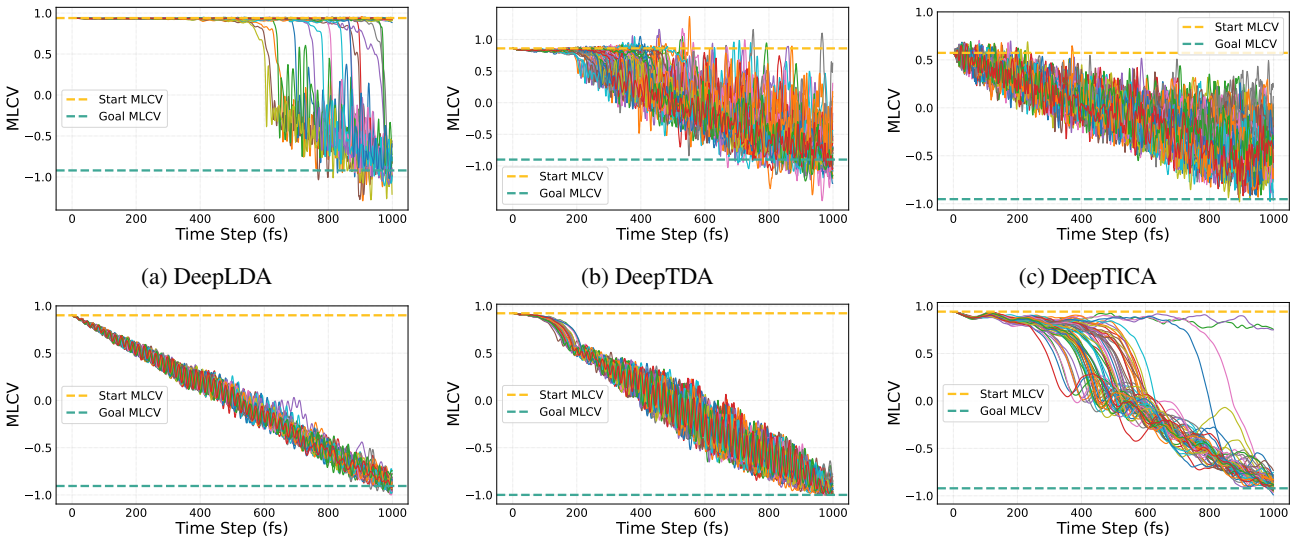
For steered MD simulations, we present additional figures on the energy and MLCV values during the transition paths.



(a) DeepLDA (b) DeepTDA (c) DeepTICA

(d) TAE (e) VDE (f) TLC

Figure 10. **Energy along the transition path** of 64 trajectories in Steered MD simulations.



(a) DeepLDA (b) DeepTDA (c) DeepTICA

(d) TAE (e) VDE (f) TLC

Figure 11. **MLCV along the transition path** of 64 trajectories from Steered MD simulations. Initial and goal states, i.e., meta-stable states $C5$ and $C7_{ax}$, are each denoted in green and yellow horizontal lines.