THERMO-SYMPLECTIC NEURAL LANGEVIN FLOWS

Anonymous authors
Paper under double-blind review

ABSTRACT

Molecular dynamics (MD) underlies biomolecular simulation by modeling atomistic motion over physical time. In thermostatted settings, this evolution is commonly formulated as underdamped Langevin dynamics, whose sequential force-driven integration makes long horizon simulation computationally expensive. While learning based approaches use MD data to model ensembles or state transitions, their computational primitive is defined over empirical state transitions rather than the Langevin update itself. We propose a *Thermo-Symplectic Neural Langevin Flow* (TSL-FLOW), a learnable framework that models underdamped Langevin dynamics through phase-space updates designed to preserve thermo-symplectic structure. Our approach lifts local stochastic updates into Lie algebraic summaries, allowing the learned dynamics to preserve the geometric structure of phase-space while exposing an associative composition across time. To enable accurate and efficient long horizon simulation, we use this compositional structure to define a prefix scan compatible product that propagates trajectory segments in parallel. Empirically, our method improves coordinate and physical fidelity across biomolecular benchmarks and achieves the lowest inference time. These results indicate that temporal efficiency in learned molecular simulation can be improved by representing the Langevin update as a composable stochastic physical primitive.

1 INTRODUCTION

Molecular dynamics (MD) has long served as a central computational methodology for studying molecular behavior across chemistry (Essmann et al., 1995), materials science (Plimpton, 1995), biology (Karplus & McCammon, 2002), and drug discovery (De Vivo et al., 2016). By integrating force driven equations of motion over physical time, MD generates atomistic trajectories rather than a single static configuration (Allen & Tildesley, 2017). For biomolecular simulation, this trajectory based view is especially important because proteins and protein ligand complexes evolve through coupled residue interactions, ligand contacts, and solvent effects. These coupled effects give rise to fluctuations and structural reorganizations beyond the information available from a single static structure, making the target of simulation a high dimensional temporal process rather than a single molecular structure (Henzler-Wildman & Kern, 2007; Boehr et al., 2009; Limongelli et al., 2012).

Under thermostatted conditions, this evolution is often formulated through *underdamped Langevin dynamics* (Leimkuhler & Matthews, 2015; Pavliotis, 2014), which couples interatomic forces, inertia, dissipation, and thermal noise. Under this model, the evolution of the position and momentum variables $(\mathbf{Q}_t, \mathbf{P}_t) \in \mathbb{R}^{2n}$ is given by

$$d\mathbf{X}_t = d\begin{pmatrix}\mathbf{Q}_t \\ \mathbf{P}_t\end{pmatrix} = \begin{pmatrix}\mathbf{M}^{-1}\mathbf{P}_t \\ -\nabla \mathrm{U}(\mathbf{Q}_t) - \gamma \mathbf{P}_t\end{pmatrix} dt + \begin{pmatrix}0 \\ \sqrt{2\gamma\beta^{-1}}\mathbf{I}_n\end{pmatrix} d\mathbf{B}_t, \quad \mathbf{X}_0 \sim \mathbb{P}_{\text{init}}, \tag{1}$$

where $\mathbf{M} \in \mathbb{R}^{n\times n}$ and $\mathrm{U} : \mathbb{R}^n \to \mathbb{R}$ denote the mass matrix and potential energy, respectively, and γ and β denote the friction coefficient and temperature parameter. Although this force driven time integration gives MD its physical grounding, it also ties the computational cost to the simulated physical horizon. With the fine temporal discretization required for stable atomistic simulation, extending the physical horizon entails repeated force evaluations along a causally ordered sequence of state updates.

Driven by this cost, prior work has pursued many routes to accelerating biomolecular simulation, including cheaper force evaluation (Unke et al., 2021), improved sampling efficiency (Laio &

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.

Table 1: **Comparison with existing biomolecular dynamics methods.** We compare whether each method family treats Langevin dynamics as the simulation primitive and supports accelerated mechanism.

Method family	Langevin Primitive	Learnable Pipeline	Accelerated Computation
Conventional MD simulator (Leimkuhler & Matthews, 2015)	✓	✗	△
Machine learning force fields (Schütt et al., 2018; Batzner et al., 2022)	△	✓	✗
Conformation ensemble generation (Jing et al., 2024a)	✗	✓	✗
Trajectory and transition generation (Jing et al., 2024b; Kapusniak et al., 2026)	✗	✓	✗
Sequential learned dynamics (Yu et al., 2025; 2026; Sengar et al., 2026)	✗	✓	✗
Ours (TSL-Flow)	✓	✓	✓

Parrinello, 2002), and learned surrogate dynamics from trajectory data (Jing et al., 2024b). Among these directions, recent learning-based approaches increasingly recast biomolecular dynamics as a generative modeling problem, learning conformational ensembles, trajectories, metastable transitions, or sequential state evolution from biomolecular MD simulation data (Jing et al., 2024a;b; Kapusniak et al., 2026; Yu et al., 2025; 2026; Sengar et al., 2026).

Yet, within this learning-based paradigm, the dominant modeling target remains empirical state transitions. In many cases, learned simulators connect samples $\mathbf{X}_t$ and $\mathbf{X}_{t+\delta t}$ through probabilistic update rules $p(\mathbf{X}_{t+\delta t} \mid \mathbf{X}_t)$, such as variational bridges (Yu et al., 2026) or latent sequential updates (Sengar et al., 2026). While effective for reproducing observed trajectories, these approaches typically bypass the physical update mechanism underlying the dynamics. As a result, the coupled structure of transport, dissipation, and thermal noise remains outside the computational primitive, which limits the learned simulator as a physically grounded model. Moreover, a substantial portion of these models retain an autoregressive generation scheme. This sequential dependency preserves the same causal bottleneck as conventional simulation, making long-horizon inference inherently expensive. Consequently, even when the transition model is learned, the computational structure remains poorly suited to efficient large-scale experiments.

Taken together, these limitations point to fundamental questions for learning biomolecular dynamics:

*How can thermodynamical structure be **preserved** in a learned yet **accelerable** simulator?*

This paper addresses this question through two coupled components. The structured phase-space update for Langevin dynamics is developed in **Section 2**, while the scan-based mechanism for long-horizon acceleration is developed in **Section 3**. Table 1 summarizes how this perspective relates to existing biomolecular dynamics method families. Overall, the main contributions are as follows.

1. **Thermo-Symplectic Langevin Dynamics.** We introduce a stochastic Lie-conformal framework that represents underdamped Langevin dynamics as a learnable phase-space update with thermo-symplectic symmetry, using an inhomogeneous conformal symplectic transformation group to encode Hamiltonian transport, dissipation, and stochastic forcing within a single structured update.
2. **Prefix-Scan Acceleration for Trajectory Generation.** We develop a Lie-algebraic signature scan that makes this structured update composable across time through an associative product, enabling efficient long-horizon trajectory generation. Empirically, our method achieves strong coordinate fidelity on mdCATH, MISATO and ATLAS datasets while substantially reducing inference time.

2 Thermo-Symplectic Langevin Dynamics of Biomolecular System

In this section, we develop the thermodynamical formulation underlying our method. For omitted definitions and notation, we refer the reader to Section A.1 and Section A.4.

Motivation. A general way to model stochastic motion is to describe evolution in state space through Stratonovich dynamics. When the state space is equipped with a structured transformation mechanism, such as the action of a *Lie group* (Hall, 2013), the resulting trajectory can be viewed as being generated by a group-valued stochastic flow acting on the initial state (Hsu, 2002; Malham & Wiese, 2008). This viewpoint replaces the black-box state-to-state transition picture with one in which dynamics is built from structured infinitesimal transformations. For our purposes, this provides a physically meaningful foundation for the Lie-Stratonovich formulation developed below.

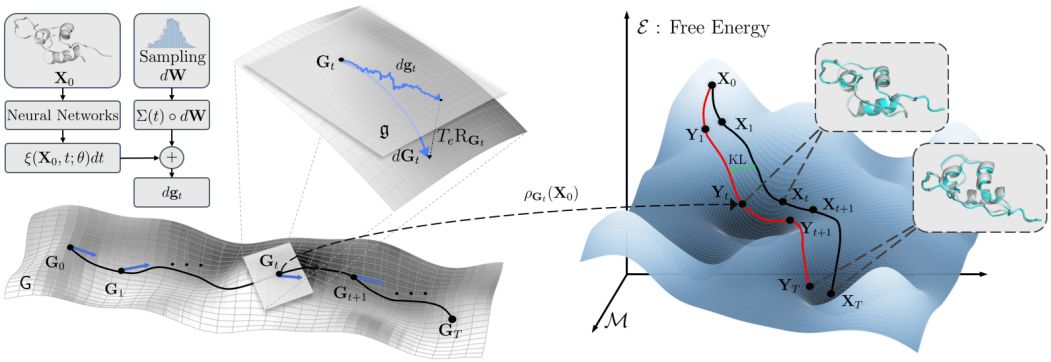


Figure 1: **Overall pipeline.** *(Left)* Starting from an initial molecular phase-space state $\mathbf{Y}_0$ and sampled noise $d\mathbf{W}$, TSL-FLOW predicts a Lie-algebraic drift term, which is combined with the noise input to form the local driver $d\mathbf{g}_t$. This driver is lifted to the group-valued increment $d\mathbf{G}_t$ and composed into the stochastic flow $\mathbf{G}_t$. *(Right)* Through the group action $\rho_{\mathbf{G}_t}$, the learned flow acts on the initial state, producing the phase-space trajectory and molecular conformations along the trajectory, while KL indicates the trajectory-level matching objective between the learned and reference dynamics, with (gray) as ground truth and (cyan) as our prediction.

Stochastic Phase-Space Lie-Stratonovich Flows. Let us consider the atomistic physical configuration space $\mathcal{M} \coloneqq T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ and let $(\mathsf{G}, \mathfrak{g})$ be a finite-dimensional Lie group and its Lie algebra acting smoothly on $\mathcal{M}$ through *group representation* $\rho : \mathsf{G} \times \mathcal{M} \to \mathcal{M}$ with $(g, y) \mapsto \rho_g(y)$. In this paper, we first consider a Lie-group valued stochastic process $\mathbf{G}_t \in \mathsf{G}$ that solves the following right-invariant *neural Stratonovich dynamics* (Hsu, 2002; Malham & Wiese, 2008):

$$\textbf{(Blue)} \quad d\mathbf{G}_t = T_e\mathrm{R}_{\mathbf{G}_t}\Big(\underbrace{\xi(\mathbf{X}_0, t; \theta)dt + \sum_{i=1} \Sigma_i(t) \circ d\mathbf{W}_t^i}_{\coloneqq d\mathbf{g}_t}\Big), \quad \mathbf{G}_0 = e, \quad \mathrm{R}_{\mathbf{G}_t}(h) \coloneqq h \cdot \mathbf{G}_t,$$

where $\xi : \mathcal{M} \times [0, T] \times \Theta \to \mathfrak{g}$ and $\Sigma : [0, T] \to \mathfrak{g}$ denote the Lie algebra-valued drift and diffusion functions parameterized by neural networks $\theta \in \Theta$. By construction, the Lie-group-valued process $\mathbf{G}_t$ plays the role of a stochastic evolution operator on phase space: it accumulates structured transformations over time, and the observable molecular trajectory is obtained only after this operator is applied to the initial state through the action of group representation. Specifically, the stochastic flow of physical state can be described as $\mathcal{M} \ni \mathbf{Y}_t = \rho_{\mathbf{G}_t}(\mathbf{Y}_0 \coloneqq \mathbf{X}_0)$. Following Proposition A.1, we apply the chain rule of Stratonovich dynamics to describe phase-space evolution of molecules:

$$\textbf{(Red)} \quad d\mathbf{Y}_t \coloneqq d\begin{pmatrix}\mathbf{q}_t \\ \mathbf{p}_t\end{pmatrix} = V_{\xi(t)}(\mathbf{Y}_t)dt + \sum_{i=1} V_{\Sigma_i(t)}(\mathbf{Y}_t) \circ d\mathbf{W}_t^i, \quad V_\xi(y) \coloneqq \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \rho_{\exp(\varepsilon\xi)}(y).$$

The fundamental vector field V_ξ acts as a decoder from Lie-algebra directions to state-space velocities, translating an infinitesimal transformation $\xi \in \mathfrak{g}$ into the local motion $d\mathbf{Y}_t$. In light of Proposition A.4, we introduce the objective function that measures the distributional discrepancy (*i.e.*, KL divergence) between the **target** and the **learned** phase-space dynamics:

$$\mathcal{J}(\theta) \coloneqq \mathrm{KL}\left(\mathbb{P}_{\mathbf{Y}}(\theta) \| \mathbb{P}_{\mathbf{X}}\right) = \mathrm{KL}\left(\mathrm{Law}(\mathbf{q}_{[0,T]}, \mathbf{p}_{[0,T]}) \parallel \mathrm{Law}(\mathbf{Q}_{[0,T]}, \mathbf{P}_{[0,T]})\right)$$

$$= \frac{\beta}{4\gamma}\mathbb{E}_{\mathbb{P}_{\mathbf{Y}}} \int_0^T \left\| (0 \quad \mathbf{I}_n) \left(\begin{pmatrix} \mathbf{M}^{-1}\mathbf{p}_t \\ -\nabla \mathrm{U}(\mathbf{q}_t) - \gamma \mathbf{p}_t \end{pmatrix} - V_{\xi(t)}(\mathbf{Y}_t) - \frac{1}{2}\sum_{i=1}^n D\big(V_{\Sigma_i(t)}\big)(\mathbf{Y}_t) V_{\Sigma_i(t)}(\mathbf{Y}_t) \right) \right\|^2 dt,$$

where D denotes the differential with respect to the state variable. In the definition, two representations $\mathbb{P}_{\mathbf{Y}}(\theta)$ and $\mathbb{P}_{\mathbf{X}}$ denote the path measures on $C([0, T], \mathcal{M})$ induced by the learned and reference dynamics, respectively. The objective is therefore posed at the level of entire phase-space trajectories, rather than pointwise state comparisons, so that learning is formulated as probabilistic matching of the full physical evolution law over time.

Fig. 1 illustrates the three coupled dynamics (*e.g.*, $d\mathbf{g}_t, d\mathbf{G}_t, d\mathbf{Y}_t$), and how they are linked across representation levels to propagate biomolecular motion, from local Lie-algebraic updates to group-valued stochastic flow and finally to the observable phase-space trajectory of the protein structures.

Thermo-Symplectic Symmetry. A discrepancy-based objective alone is not sufficient to determine the geometric and thermodynamic structure governing the learned update. To impose this missing

physical constraint, we endow G with a geometric structure that restricts the learned evolution to physically admissible phase-space transformations. This structural principle will be referred to as *thermo-symplectic symmetry*, and its formal statement is given below.

Definition 2.1 (Thermo-Symplectic Symmetry)**.** *We say that the Stratonovich dynamics possesses* ***thermo-symplectic symmetry*** *if the following structural conditions hold:*

$$\mathcal{L}_{V_{\xi(t)}}\omega_J = -\gamma\omega_J, \quad \mathcal{L}_{V_{\Sigma_i(t)}}\omega_J = 0, \quad \forall t \in [0,T],$$

where the Lie derivative $\mathcal{L}_\xi\omega_J$ of the symplectic form ω_J along a vector field ξ is given as

$$(\mathcal{L}_\xi\omega_J)(U,V) = \xi[\omega_J(U,V)] - \omega_J([\xi,U],V) - \omega_J(U,[\xi,V]), \quad U,V \in \xi(\mathcal{M}).$$

Under this requirement, the Stratonovich dynamics is constrained to preserve the same physical organization as the reference Langevin dynamics. In particular, its deterministic component remains compatible with Hamiltonian phase-space evolution up to the damping induced by friction (*i.e.*, γ). Consequently, the dynamics is restricted to a physically admissible class of phase-space transformations rather than an arbitrary data-driven transition rule. As shown in Proposition A.3, the physical-oracle Langevin dynamics in Eq. 1 already satisfies this property, and our goal is therefore to identify a structure on G that allows the learned dynamics to inherit it.

To realize the introduced symmetry concretely, we model the learned evolution through transformations that preserve the canonical symplectic form up to conformal scaling while allowing phase-space translations. This leads to the so-called *conformal symplectic group* (*i.e.*, CSp) (Wojtkowski & Liverani, 1998) and its inhomogeneous extension (*i.e.*, $\mathsf{CSp} \ltimes \mathbb{R}^{2n}$) as our algebraic primitives:

$$\begin{aligned} \mathsf{CSp}(2n) &:= \left\{A \in \mathrm{GL}(2n) : A^\top \mathbf{J} A = \lambda_A \mathbf{J},\ \lambda_A > 0\right\}, & \mathsf{G} &:= \mathsf{CSp}(2n) \ltimes \mathbb{R}^{2n}, \\ \mathfrak{csp}(2n) &:= \left\{X \in \mathfrak{gl}(2n) : X^\top \mathbf{J} + \mathbf{J} X = \lambda_X \mathbf{J},\ \lambda_X \in \mathbb{R}\right\}, & \mathfrak{g} &:= \mathfrak{csp}(2n) \ltimes \mathbb{R}^{2n}, \end{aligned} \tag{2}$$

where $\ltimes$ denotes the semidirect product, satisfying $(A,\mathbf{a})\cdot(B,\mathbf{b}) = (AB, A\mathbf{b}+\mathbf{a})$. With this choice of G and $\mathfrak{g}$, the learned evolution is placed in the geometric class compatible with thermo-symplectic structure. The following proposition makes this point explicit.

Proposition 2.2 (informal)**.** *When $(\mathsf{G},\mathfrak{g})$ is chosen as the inhomogeneous conformal symplectic group in Eq. 2, the proposed learned dynamics $\mathbf{Y}_t$ approximates the thermo-symplectic symmetry.*

Proposition 2.2 shows that the proposed parameterization enables the learned dynamics to approximate the symplectic structure of the reference Langevin dynamics at the level of phase-space geometry. In contrast, prior learned simulators generally almost fail to encode this structure at the level of their update rules. In Section 5, we further assess this property through physical-structure fidelity metrics that measure consistency with the corresponding conformal symplectic invariants in practice.

3 Prefix-Scan Acceleration of Lie-Stratonovich Flow

The cost of simulating the full phase-space trajectory $\{\mathbf{Y}_t\}$ depends largely on how efficiently we can represent and propagate the underlying group-valued flow, *i.e.*, **(Blue)**. Motivated by recent machine learning models such as Mamba (Gu & Dao, 2024), which reduce sequential depth through a *parallel prefix-scan* (Blelloch, 1990), we seek a representation of $\mathbf{G}_t$, *i.e.*, **(Green)**, whose temporal evolution can be composed through a prefix-scan over local updates. In particular, our goal is to replace the $\mathcal{O}(M)$ sequential depth by a scan-compatible sequential depth $\mathcal{O}(\log M)$.

Rough-Path Theory for Prefix-Scan Acceleration. In the discretized time interval $0 = t_0 < t_1 < \cdots < t_M = T$, let $\mathrm{I}_m := [t_{m-1}, t_m]$ be a temporal partition. Guided by *rough path theory* (Friz & Hairer, 2014), we encode the local driver on each interval through iterated Stratonovich integrals to obtain an algebraic representation that can be composed across time. Specifically, we represent the local driver (*i.e.*, $d\mathbf{g}_t$) on each interval by its algebraic extension called a *signature* (*i.e.*, Sig):

$$\mathrm{Sig}_m(\mathbf{g}) := 1 + \sum_{k, a_1,\ldots,a_k \in 1,\ldots,d} \left(\int_{t_{m-1}<t_1<\cdots<t_k\le t} \circ d\mathbf{g}_{t_1}^{a_1} \cdots \circ d\mathbf{g}_{t_k}^{a_k} \right) \mathrm{E}_{a_1} \otimes \cdots \otimes \mathrm{E}_{a_k}, \tag{3}$$

where $\{\mathrm{E}_a\}_{a=1}^d$ is a basis of the Lie algebra $\mathfrak{g}$, and $\circ d\mathbf{g}_t = \sum_{a=1}^d \mathrm{E}_a \circ d\mathbf{g}_t^a$. Each tensor term $\mathrm{E}_{a_1} \otimes \cdots \otimes \mathrm{E}_{a_k}$ records the ordered sequence of directions taken by the local signal, so that the

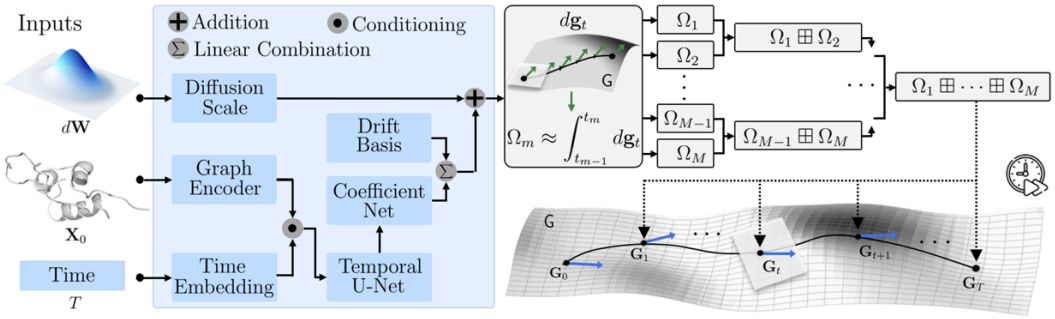


Figure 2: **Lie-Stratonovich flows with BCH prefix-scan.** *(Left)* TSL-FLOW parameterizes the local Lie-algebraic driver $d\mathbf{g}_t$ using a diffusion branch and coefficient-weighted drift basis conditioned on molecular and temporal features. *(Right)* For each temporal segment, the local driver induces a logarithmic coordinate Ω_m, which the BCH prefix-scan associatively composes across time to construct the group-valued flow $\mathbf{G}_t$.

signature becomes a compact summary of the driver in a form that can be composed through prefix-scan across time. To turn this algebraic summary into a local coordinate for closed-form computation, we next pass to the logarithmic chart associated with the matrix exponential map and its inverse ($i.e.$, Exp, Log). Consecutively, Proposition A.5 elucidates the following segment-level relation:

$$\textbf{(Green)}^1 \quad \mathrm{H}_t \sim \int_{t_{m-1}}^{t} \dot{\mathbf{G}}_s \mathbf{G}_s^{-1} ds, \quad \Omega_m \approx \int_{t_{m-1}}^{t_m} d\mathbf{g}_t \sim \int_{\mathrm{I}_m} \mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_m(\mathbf{g})\mathrm{Sig}_m(\mathbf{g})^{-1}\right).$$

We then extend this idea to the full time horizon $[0, T]$ by composing these local updates in order to represent the cumulative evolution over $[0, t_m] \subseteq [0, T]$ by $\Omega_{1:m}$:

$$\Omega_{1:m} := \Omega_1 \boxplus \Omega_2 \boxplus \cdots \boxplus \Omega_m, \quad \mathrm{A} \boxplus \mathrm{B} := \log\big(\exp(\mathrm{B})\exp(\mathrm{A})\big), \quad \mathrm{A}, \mathrm{B} \in \mathfrak{g}, \tag{4}$$

Here, $\boxplus$ denotes the *Baker-Campbell-Hausdorff* (BCH) (Hall, 2013) product, which simply combines two local updates into a single update while keeping the same logarithmic-coordinate form. In this form, $\Omega_{1:m}$ is a single Lie-algebraic representation of the composed evolution up to t_m, written so that it can be combined further in the same way. By the definition, on the local coordinate domain where the exponential ($i.e.$, Exp), logarithm ($i.e.$, Log), and all intermediate compositions are well defined, this composition satisfies *commutative relation*:

$$(\Omega_m \boxplus \Omega_l) \boxplus \Omega_r = \Omega_m \boxplus (\Omega_l \boxplus \Omega_r), \qquad \forall (m, l, r) \in [1, M]^3. \tag{5}$$

Since the BCH product in Eq. 5 provides an associative composition rule, the cumulative quantities $\{\Omega_{1:m}\}_{m=1}^{M}$ can be computed by a prefix-scan over the segment-wise coordinates $\{\Omega_m\}_{m=1}^{M}$. Overall, the time-discretized approximation of the Lie group-valued flow $\mathbf{G}_t$ and the corresponding accelerated state trajectory induced by our model are given by

$$\mathbf{G}_{t_m} \approx \exp(\Omega_m) \cdots \exp(\Omega_1) = \exp(\Omega_{1:m}), \quad \mathbf{Y}_{t_m} := \rho_{\mathbf{G}_{t_m}}(\mathbf{Y}_0) \approx \rho_{\exp(\Omega_{1:m})}(\mathbf{Y}_0). \tag{6}$$

As a result, the inferred dynamics can be carried out with sequential depth $\mathcal{O}(\log M)$ rather than $\mathcal{O}(M)$, yielding a substantial gain in efficiency for long-horizon simulation.

4 Related Work

Conformation Ensemble Generation. Recent approaches in protein modeling have expanded from predicting single static structures to generating conformational ensembles that reflect thermal fluctuations. Within protein structure predictors, AlphaFlow (Jing et al., 2024a) adapts flow matching to generate sequence conditioned conformational ensembles, offering a direct way to learn structural distributions represented by MD samples. As a representative approach, it effectively captures the conformational variability of protein dynamics without specifying time-ordered transitions. Unlike this distributional focus, we explicitly model the temporal Langevin evolution through which such conformational ensembles arise over time.

Trajectory and Metastable Transition Generation. Beyond conformational ensembles, several generative models describe molecular motion by focusing on continuous trajectories or state-to-state transitions. For trajectory generation, molecular motion can be represented as spatiotemporal

[1]Since the expression is abstract, Proposition A.6 provides an explicit local coordinate expression of Ω_m.

Table 2: **Quantitative results on mdCATH solvated protein systems.** We report coordinate fidelity, physical fidelity, and computational efficiency for long-horizon trajectory reproduction. The best and second-best results are highlighted in **bold** and **blue**, respectively.

Method	Coordinate Fidelity		Physical Fidelity		Computational Efficiency	
	Mean MSE ↓	Mean RMSD ↓	Pair JSD ↓	CSP ↓	Inference Time ↓	Train Time ↓
(A) Conformation Ensemble Generation						
AlphaFlow (Jing et al., 2024a)	6.01×10^{-2}	3.05×10^{-1}	1.56×10^{-2}	4.92×10^{-1}	4.17×10^{2}	1.05×10^{2}
(B) Trajectory and Metastable Transition Generation						
MDGen (Jing et al., 2024b)	7.47×10^{-2}	4.51×10^{-1}	2.12×10^{-3}	**4.36×10^{-2}**	5.35×10^{0}	**2.52×10^{1}**
MarS-FM (Kapusniak et al., 2026)	4.07×10^{-2}	2.63×10^{-1}	1.36×10^{-3}	4.55×10^{-2}	2.60×10^{1}	8.21×10^{1}
(C) Sequential Learned Dynamics						
UniSim (Yu et al., 2025)	2.11×10^{-2}	2.11×10^{-1}	2.13×10^{-3}	7.71×10^{-1}	1.19×10^{2}	7.35×10^{2}
PVB (Yu et al., 2026)	6.15×10^{-2}	3.62×10^{-1}	7.03×10^{-3}	9.06×10^{-1}	7.61×10^{2}	3.86×10^{2}
GLDP (Sengar et al., 2026)	**5.14×10^{-3}**	**1.14×10^{-1}**	**8.49×10^{-4}**	6.93×10^{-2}	**4.77×10^{0}**	1.67×10^{2}
Ours (TSL-Flow)	**1.64×10^{-3}**	**6.21×10^{-2}**	**5.45×10^{-4}**	**1.66×10^{-2}**	**4.09×10^{0}**	**8.16×10^{0}**

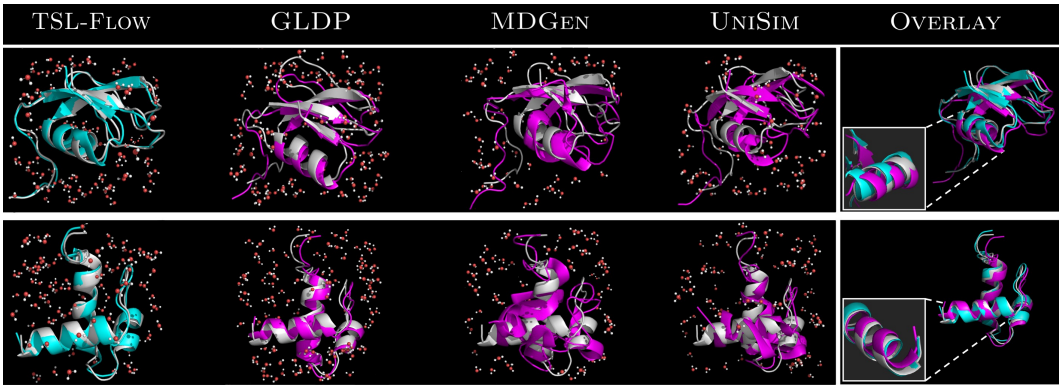


Figure 3: **Qualitative results for 2LSM and 1BA5.** We visualize the final frames of two protein trajectories from the mdCATH benchmark, with 2LSM on the top and 1BA5 on the bottom. *(Left)* We overlay each predicted final frame on the ground truth (gray), comparing our framework (cyan) with the prior works (magenta). *(Right)* We further overlay the final conformations of our framework and GLDP directly against the ground truth.

coordinate sequences, as in MDGen (Jing et al., 2024b), which generates complete trajectories within a unified conditional framework. To describe metastable transitions, Markov state models can be coupled with flow matching, as in MarS-FM (Kapusniak et al., 2026), which samples transitions across metastable states. Despite their different constructions, these methods provide effective baselines for capturing long-range molecular motion. Distinct from these approaches, we model underdamped Langevin dynamics to propagate molecular states through stepwise integration.

Sequential Learned Dynamics. For learned molecular simulation, another line of work advances molecular states sequentially through data driven dynamics. With stochastic interpolants and force guidance, UniSim (Yu et al., 2025) models coarse time biomolecular evolution. Through pre-trained variational bridge dynamics, PVB (Yu et al., 2026) propagates molecular states, whereas in GLDP (Sengar et al., 2026), compressed latent representations evolve before being decoded into atomistic coordinates. Taken together, these methods establish strong baselines for sequential molecular simulation. In contrast, we explicitly evolve the full atomistic phase space via Lie group representations, thereby preserving all-atom fidelity and adhering to local physical rules.

5 Experiment

In this section, we evaluate our framework on molecular dynamics prediction across multiple protein systems. Our goal is to examine whether predicted trajectories capture the structural evolution observed in reference simulations while enabling efficient long-horizon trajectory generation. We measure performance in terms of consistency with reference biomolecular dynamics, physical fidelity, and computational efficiency. For each selected protein system, we generate an OpenMM (Eastman et al., 2017) reference trajectory using the AMBER19 force field (Tian et al., 2019) and the TIP3P-FB

Table 3: **Quantitative results on ATLAS solvated protein systems.** We report coordinate fidelity, physical fidelity, and computational efficiency for long-horizon trajectory reproduction. The best and second-best results are highlighted in **bold** and **blue**, respectively.

Method	Coordinate Fidelity		Physical Fidelity		Computational Efficiency	
	Mean MSE ↓	Mean RMSD ↓	Pair JSD ↓	CSP ↓	Inference Time ↓	Train Time ↓
(A) Conformation Ensemble Generation						
AlphaFlow (Jing et al., 2024a)	5.60×10^{-2}	2.56×10^{-1}	5.75×10^{-3}	5.40×10^{-1}	3.63×10^{2}	1.37×10^{2}
(B) Trajectory and Metastable Transition Generation						
MDGen (Jing et al., 2024b)	7.56×10^{-2}	4.55×10^{-1}	1.38×10^{-3}	5.80×10^{-2}	5.57×10^{1}	**7.90×10^{1}**
MarS-FM (Kapusniak et al., 2026)	3.50×10^{-2}	2.34×10^{-1}	1.89×10^{-3}	**2.96×10^{-2}**	2.95×10^{1}	8.80×10^{1}
(C) Sequential Learned Dynamics						
UniSim (Yu et al., 2025)	1.61×10^{-2}	2.12×10^{-1}	1.41×10^{-3}	8.87×10^{-1}	1.36×10^{2}	7.48×10^{2}
PVB (Yu et al., 2026)	5.45×10^{-2}	3.22×10^{-1}	1.23×10^{-2}	7.56×10^{-1}	7.99×10^{2}	4.02×10^{2}
GLDP (Sengar et al., 2026)	**4.36×10^{-3}**	**1.09×10^{-1}**	**7.41×10^{-4}**	7.48×10^{-2}	**4.59×10^{0}**	1.75×10^{2}
Ours (TSL-Flow)	**1.48×10^{-3}**	**6.00×10^{-2}**	**4.58×10^{-4}**	**1.19×10^{-2}**	**4.01×10^{0}**	**8.77×10^{0}**

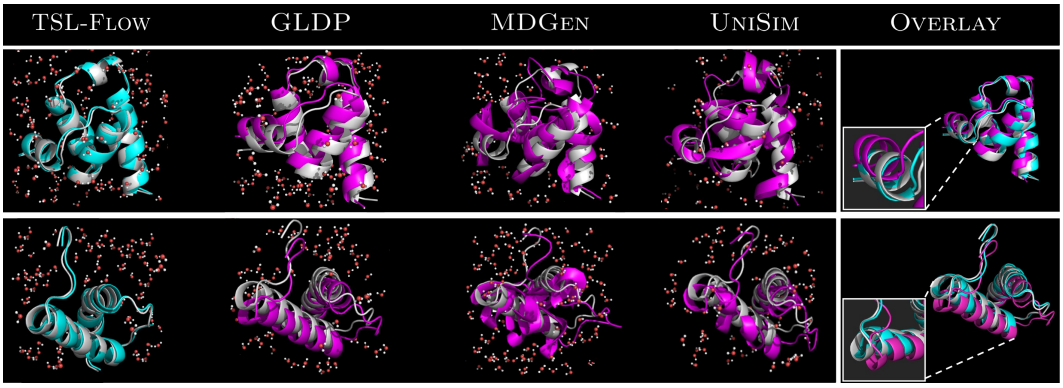


Figure 4: **Qualitative results for 5l87 and 3u5v.** We visualize the final frames of two protein trajectories from the ATLAS benchmark, with 5l87 on the top and 3u5v on the bottom. *(Left)* We overlay each predicted final frame on the ground truth (gray), comparing our framework (cyan) with the prior works (magenta). *(Right)* We further overlay the final conformations of our framework and GLDP directly against the ground truth.

explicit water model (Wang et al., 2014). Simulations use particle mesh Ewald electrostatics (Essmann et al., 1995) and a Langevin integrator in the NVT ensemble at 400 K with a 2 fs timestep. From each simulation, we construct 10,240 frame trajectories for training and long-horizon evaluation.

Datasets. We formulate our evaluation setting using proteins from mdCATH (Mirarchi et al., 2024) and ATLAS (Vander Meersche et al., 2024). mdCATH provides trajectories for protein domains sampled across diverse CATH folds, whereas ATLAS provides standardized simulations of representative proteins and their conformational flexibility. We utilize both benchmarks to validate performance consistency in long horizon trajectory prediction across distinct solvated protein systems. To further examine protein ligand dynamics, we also include a protein ligand complex from MISATO (Siebenmorgen et al., 2024). Details of evaluated systems are provided in the Appendix A.2.

Biomolecular Trajectory Reconstruction. In Tables 2 and 3, we summarize the quantitative comparison on mdCATH and ATLAS. We group prior methods into three categories based on their modeling approaches. (A) *Conformation Ensemble Generation*, which targets static conformational distributions, (B) *Trajectory and Metastable Transition Generation*, which models coordinate paths or metastable transitions, and (C) *Sequential Learned Dynamics*, which advances molecular states through learned temporal updates. Following the table axes, we compare all methods using coordinate fidelity, physical fidelity, and computational efficiency. We evaluate correspondence with reference trajectories using mean MSE and mean RMSD. We assess pairwise distance statistics using pair JSD, where the pairs denote heavy-atom pairs within each trajectory frame. Additionally, we examine consistency with the conformal symplectic structure of the learned update via CSP. Specifically, following the Lie-algebraic conformal symplectic condition in Eq. 2, we define CSP as $\left\|A^{\top}\mathbf{J} + \mathbf{J}A + \gamma\mathbf{J}\right\|_F^2$, where A denotes the linear generator of the learned phase-space update and γ is the friction coefficient. Furthermore, for computational comparison, we measure inference and

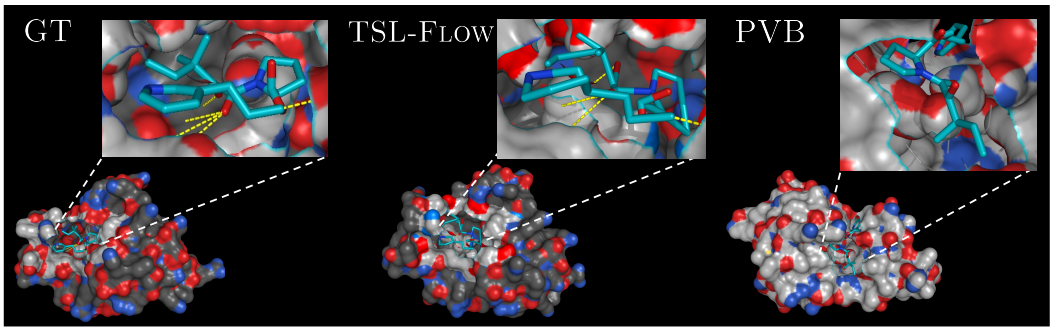


Figure 5: **Visualization of ligand pocket contact geometry on MISATO.** We visualize the bound protein–ligand complex for the ground truth, our framework, and PVB from left to right. The ligand is shown in sticks (cyan), and the binding pocket is shown as an atom colored molecular surface. We identify ligand pocket contacts (yellow) from trajectory level contact occupancy.

training time to quantify computational efficiency under the shared trajectory generation protocol. We visualize qualitative reconstructions in Figures 3 and 4, where final-frame overlays visualize how each method preserves the reference structure across representative mdCATH and ATLAS.

Across both benchmarks, the proposed framework achieves the highest trajectory fidelity among the evaluated methods. Compared with GLDP, the leading prior method in coordinate fidelity, our framework reduces mean MSE by approximately $\mathbf{3.1}\times$ on mdCATH and $\mathbf{2.9}\times$ on ATLAS. Mean RMSD is also reduced by approximately $\mathbf{1.8}\times$ on both benchmarks. Beyond coordinate accuracy, our framework obtains the lowest Pair JSD and CSP across both datasets, indicating that the generated trajectories better preserve structural distance statistics and conformal phase space consistency. This improvement is consistent with the conformal symplectic parameterization of our phase space update. By aligning the learned dynamics with the geometry induced by underdamped Langevin dynamics, the proposed framework reduces the CSP residual rather than learning an unconstrained coordinate transition. As a result, our model remains more consistent with the phase space geometry of the target dynamics and improves stability in long horizon trajectory reconstruction.

In addition to fidelity gains, the proposed framework substantially reduces computational cost compared with prior methods. On mdCATH, training time is reduced by approximately $\mathbf{3.1}\times$ compared with the fastest prior result, and on ATLAS it is reduced by approximately $\mathbf{9.0}\times$. The proposed framework also achieves the lowest inference time on both benchmarks. These results indicate that the scan compatible Langevin formulation improves long horizon trajectory reconstruction.

Protein-ligand Interaction Modeling. To examine whether the proposed formulation extends beyond protein trajectory reconstruction, we evaluate protein-ligand trajectories on MISATO. This setting is more demanding than protein reconstruction alone because the generated dynamics must capture both protein conformational evolution and the relative motion of a ligand inside a local binding pocket. Since ligand-protein interactions are localized, small errors in ligand placement can lead to changes in contact geometry or steric validity that may not be fully captured by global coordinate metrics. This setting therefore provides a complementary evaluation of local interaction geometry, which is not fully characterized by protein-level trajectory reconstruction alone.

Table 4: **Quantitative results on MISATO protein-ligand systems.** We report five protein-ligand interaction fidelity metrics. The best and second-best results are highlighted in **bold** and **blue**, respectively.

Metric	TSL-FLOW	UNISIM	PVB
EMD-lig. ↓	$\mathbf{2.80\times10^{-2}}$	$\mathbf{2.26\times10^{-1}}$	5.57×10^{-1}
EMD-CoM ↓	$\mathbf{9.00\times10^{-3}}$	$\mathbf{3.80\times10^{-2}}$	1.29×10^{-1}
RMSE-cont. ↓	$\mathbf{3.60\times10^{-2}}$	8.30×10^{-2}	$\mathbf{5.90\times10^{-2}}$
RMSD (Å) ↓	$\mathbf{8.02\times10^{-1}}$	$\mathbf{3.17\times10^{0}}$	6.60×10^{0}
Clash ↓	$\mathbf{1.60\times10^{-1}}$	$\mathbf{1.21\times10^{1}}$	N/A

Consequently, we evaluate generated protein-ligand trajectories in terms of ligand pose distributions, pocket relative motion, intermolecular contact geometry, and steric validity within the binding pocket, using metrics from prior protein-ligand trajectory evaluation (Yu et al., 2026). After aligning trajectory to the protein backbone, EMD-ligand compares ligand motion distributions using the Wasserstein-1 distance (Villani, 2009), while EMD-CoM measures the discrepancy between ligand-pocket center-of-mass distance distributions. RMSE-CONTACT evaluates ligand-pocket heavy atom contact occupancy recovery. We further report ligand RMSD under the same protein alignment for pose accuracy, and count ligand-pocket heavy atom pairs within 2Å to quantify steric validity.

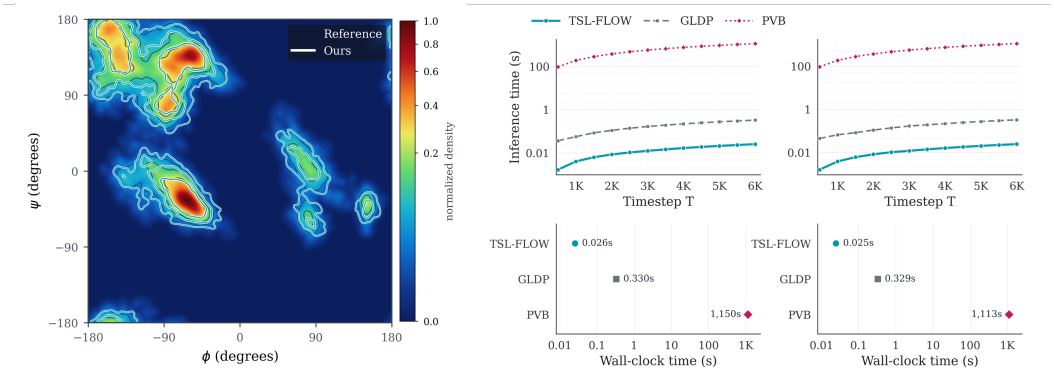


Figure 6: **Ablation study on inference scaling and backbone conformation.** *(Left)* Ramachandran plot showing that our generated backbone conformations closely follow the reference dihedral-angle distribution. *(Top)* Inference time as a function of trajectory length for TSL-FLOW (cyan), GLDP (gray), and PVB (magenta). *(Bottom)* Wall-clock inference time, where our framework remains faster than both baselines.

As can be seen in Table 4, the proposed framework improves bound state protein–ligand interaction geometry across distributional, contact, and pose based measures. It reduces EMD-ligand and EMD-CoM by $\underline{\mathbf{8.1}\times}$ and $\underline{\mathbf{4.2}\times}$ compared with the strongest baseline for each metric. It also reduces RMSE-CONTACT by $\underline{\mathbf{1.6}\times}$ and ligand RMSD by $\underline{\mathbf{4.0}\times}$. The gain is particularly pronounced in pose accuracy, where the predicted ligand remains below 1 Å RMSD on average after protein alignment. This improvement is consistent with the Langevin phase space parameterization, which couples protein and ligand motion through a structured update rather than treating ligand displacement as an unconstrained coordinate transition. Although PVB reports zero close contact clashes, its larger EMD-CoM and ligand RMSD suggest that clash count should be interpreted together with ligand localization and pose accuracy. Taken together, these results show that the structured Langevin update preserves binding pocket localization, contact occupancy, and bound state geometry, which are structural components associated with stable protein–ligand binding.

Ablation Study. We compare the inference time across increasing prediction horizons in Figure 6. This analysis quantifies how the computational cost grows with the increasing number of timesteps, and whether the proposed scan-compatible formulation avoids the temporal cost accumulation observed in prior works. Across both ATLAS and mdCATH, our method shows consistently lower wall-clock cost than GLDP and PVB. The gradual increase of inference-time curves indicates that our method alleviates the temporal bottleneck in long horizon trajectory generation.

Our method requires only 2.55×10^{-2} s on ATLAS and 2.46×10^{-2} s on mdCATH, whereas GLDP requires 3.30×10^{-1} s on ATLAS and 3.29×10^{-1} s on mdCATH, making it $\underline{\mathbf{13.0}\times}$ and $\underline{\mathbf{13.4}\times}$ slower than our method at the same horizon. We also show that PVB yields a larger gap, making it $\underline{\mathbf{44.0}\times}$ slower than our method on both datasets. This gap shows that our method holds a clear inference time advantage even against the stronger prior baseline. We verify through a Ramachandran analysis (Ramachandran et al., 1963) that we achieve the efficiency while preserving backbone statistics. For the 2LSM, the generated ensemble closely matches the reference distribution, with a Ramachandran Jensen-Shannon divergence of 0.02 bits computed from 1.21M ϕ/ψ samples per ensemble, confirming physically plausible backbone geometry.

6 Conclusion

In this work, we studied how learned biomolecular simulation can preserve the thermodynamic structure of Langevin dynamics while reducing the temporal bottleneck of long horizon trajectory generation. Our results show that these two goals can be connected by representing local stochastic updates as composable geometric transformations with an explicit temporal composition structure. Across biomolecular benchmarks, our formulation improves coordinate and physical fidelity while reducing inference cost. These findings indicate that acceleration in learned molecular simulation can be achieved not only by fitting empirical state transitions, but also by restructuring the physical update into a scan compatible form. Overall, composable Langevin learning provides a practical route toward learned biomolecular simulators that are both physically structured and temporally scalable.

AI USE STATEMENT

We used generative AI tools to refine the writing in the manuscript and support the development of mathematical proofs. We have reviewed all AI-assisted content and checked the mathematical arguments for correctness. We take full responsibility for the final manuscript.

REFERENCES

Michael P. Allen and Dominic J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 06 2017. ISBN 9780198803195. doi: 10.1093/oso/9780198803195.001.0001. URL https://doi.org/10.1093/oso/9780198803195.001.0001.

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, 2022.

Guy E Blelloch. Prefix sums and their applications. (CMU-CS-90-190), 1990.

David D Boehr, Ruth Nussinov, and Peter E Wright. The role of dynamic conformational ensembles in biomolecular recognition. *Nature chemical biology*, 5(11):789–796, 2009.

Marco De Vivo, Matteo Masetti, Giovanni Bottegoni, and Andrea Cavalli. Role of molecular dynamics and related methods in drug discovery. *Journal of medicinal chemistry*, 59(9):4035–4061, 2016. doi: 10.1021/acs.jmedchem.5b01684.

Peter Eastman, Jason Swails, John D. Chodera, Robert T. McGibbon, Yutong Zhao, Kyle A. Beauchamp, Lee-Ping Wang, Andrew C. Simmonett, Matthew P. Harrigan, Chaya D. Stern, Rafal P. Wiewiora, Bernard R. Brooks, and Vijay S. Pande. Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*, 13(7):1–17, 2017.

Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995. doi: 10.1063/1.470117.

P.K. Friz and M. Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Universitext. Springer International Publishing, 2014. ISBN 9783319083322. URL https://books.google.co.kr/books?id=hN5cBAAAQBAJ.

Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. In *First Conference on Language Modeling*, 2024. URL https://openreview.net/forum?id=tEYskw1VY2.

Brian C Hall. Lie groups, lie algebras, and representations. In *Quantum Theory for Mathematicians*, pp. 333–366. Springer, 2013.

Katherine Henzler-Wildman and Dorothee Kern. Dynamic personalities of proteins. *Nature*, 450 (7172):964–972, 2007.

Elton P Hsu. *Stochastic analysis on manifolds*. Number 38. American Mathematical Soc., 2002. ISBN 9780821808023.

Bowen Jing, Bonnie Berger, and Tommi Jaakkola. AlphaFold meets flow matching for generating protein ensembles. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 22277–22303. PMLR, 21–27 Jul 2024a. URL https://proceedings.mlr.press/v235/jing24a.html.

Bowen Jing, Hannes Stärk, Tommi Jaakkola, and Bonnie Berger. Generative modeling of molecular dynamics trajectories. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 40534–40564. Curran Associates, Inc., 2024b. doi: 10.52202/079017-1282.

Kacper Kapusniak, Cristian Gabellini, Michael Bronstein, Prudencio Tossou, and Francesco Di Giovanni. Mars-fm: Generative modeling of molecular dynamics via markov state models. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 19073–19098, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/205f007604f14b702a7594b695de52bc-Paper-Conference.pdf.

Martin Karplus and J Andrew McCammon. Molecular dynamics simulations of biomolecules. *Nature structural biology*, 9(9):646–652, 2002.

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

B. Leimkuhler and C. Matthews. *Molecular Dynamics: With Deterministic and Stochastic Numerical Methods*. Interdisciplinary Applied Mathematics. Springer International Publishing, 2015. ISBN 9783319163758. URL https://books.google.co.kr/books?id=AGF1CQAAQBAJ.

Vittorio Limongelli, Luciana Marinelli, Sandro Cosconati, Concettina La Motta, Stefania Sartini, Laura Mugnaini, Federico Da Settimo, Ettore Novellino, and Michele Parrinello. Sampling protein motion and solvent effect during ligand binding. *Proceedings of the National Academy of Sciences*, 109(5):1467–1472, 2012. doi: 10.1073/pnas.1112181108.

Simon JA Malham and Anke Wiese. Stochastic lie group integrators. *SIAM Journal on Scientific Computing*, 30(2):597–617, 2008. doi: 10.1137/060666743.

Antonio Mirarchi, Toni Giorgino, and Gianni De Fabritiis. mdcath: A large-scale md dataset for data-driven computational biophysics. *Scientific Data*, 11(1):1299, 2024.

G.A. Pavliotis. *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. Texts in Applied Mathematics. Springer New York, 2014. ISBN 9781493913237. URL https://books.google.co.kr/books?id=8213BQAAQBAJ.

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

G.N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7(1):95–99, 1963. ISSN 0022-2836. doi: https://doi.org/10.1016/S0022-2836(63)80023-6. URL https://www.sciencedirect.com/science/article/pii/S0022283663800236.

Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller. Schnet–a deep learning architecture for molecules and materials. *The Journal of chemical physics*, 148(24):241722, 2018.

Aditya Sengar, Jiying Zhang, Pierre Vandergheynst, and PATRICK BARTH. Beyond ensembles: Simulating all-atom protein dynamics in a learned latent space. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=AwowReRWXI.

Till Siebenmorgen, Filipe Menezes, Sabrina Benassou, Erinc Merdivan, Kieran Didi, André Santos Dias Mourão, Radosław Kitel, Pietro Liò, Stefan Kesselheim, Marie Piraud, et al. Misato: machine learning dataset of protein–ligand complexes for structure-based drug discovery. *Nature computational science*, 4(5):367–378, 2024.

Chuan Tian, Koushik Kasavajhala, Kellon AA Belfon, Lauren Raguette, He Huang, Angela N Migues, John Bickel, Yuzhang Wang, Jorge Pincay, Qin Wu, et al. ff19sb: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *Journal of chemical theory and computation*, 16(1):528–552, 2019.

Oliver T Unke, Stefan Chmiela, Huziel E Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T Schutt, Alexandre Tkatchenko, and Klaus-Robert Muller. Machine learning force fields. *Chemical reviews*, 121(16):10142–10186, 2021.

Yann Vander Meersche, Gabriel Cretin, Aria Gheeraert, Jean-Christophe Gelly, and Tatiana Galochkina. Atlas: protein flexibility description from atomistic molecular dynamics simulations. *Nucleic acids research*, 52(D1):D384–D392, 2024.

Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9): 1157–1174, 2004. doi: https://doi.org/10.1002/jcc.20035.

Lee-Ping Wang, Todd J Martinez, and Vijay S Pande. Building force fields: An automatic, systematic, and reproducible approach. *The journal of physical chemistry letters*, 5(11):1885–1891, 2014.

Maciej P Wojtkowski and Carlangelo Liverani. Conformally symplectic dynamics and symmetry of the lyapunov spectrum. *Communications in mathematical physics*, 194(1):47–60, 1998.

Ziyang Yu, Wenbing Huang, and Yang Liu. Unisim: A unified simulator for time-coarsened dynamics of biomolecules. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=W598uW37zt.

Ziyang Yu, Wenbing Huang, and Yang Liu. Unified biomolecular trajectory generation via pretrained variational bridge. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=8HH9dBOxwu.

A Appendix

A.1 Notation Table

Symbol	Space	Role
$\mathcal{M} := T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$	-	Phase space.
$\mathbf{X}_t = (\mathbf{Q}_t, \mathbf{P}_t)$	$\mathbb{R}^{2n}$	Reference Langevin state.
$\mathbf{M}$	$\mathbb{R}^{n\times n}$	Mass matrix.
$\mathrm{U} : \mathbb{R}^n \to \mathbb{R}$	-	Potential energy.
γ, β	$\mathbb{R}$	Friction and inverse-temperature parameters.
G	Lie group	Inhomogeneous conformal symplectic group.
$\mathfrak{g}$	Lie algebra	Lie algebra of G.
$\rho : \mathsf{G} \times \mathcal{M} \to \mathcal{M}$	action	Group action on phase space.
$\mathbf{G}_t$	G-valued process	Group-valued stochastic flow.
$\circ d\mathbf{g}_t$	$\mathfrak{g}$-valued increment	Local Lie-algebra driver.
$\mathbf{Y}_t = (\mathbf{q}_t, \mathbf{p}_t)$	$\mathcal{M}$	Learned phase-space trajectory.
$\xi(\mathbf{X}_t, t; \theta)$	$\mathfrak{g}$-valued map	Drift.
$\Sigma_i(t)$	$\mathfrak{g}$-valued map	Diffusion directions.
V_ξ	vector field on $\mathcal{M}$	Fundamental vector field.
J	$\mathbb{R}^{2n\times 2n}$	Canonical symplectic matrix.
ω_J	2-form on $\mathcal{M}$	Canonical symplectic form.
$\mathrm{Sig}_t^{(m)}(\mathbf{g})$	tensor-algebra element	Segment signature.
H_t	local Lie coordinate	Logarithmic coordinate.
$\Omega_m, \Omega_{1:m}$	Lie-algebra element	Local / cumulative logarithmic increment.
$\boxplus$	binary operation	BCH scan product.
$\mathrm{dLog}_{\mathrm{H}_t}$	linear map	Inverse exponential differential.
B_n	scalar	Bernoulli number.
$T(\mathfrak{g}), \widehat{T}(\mathfrak{g})$	associative algebra	Tensor algebra and its completion.
$\mathcal{L}(\mathfrak{g}), \widehat{\mathcal{L}}(\mathfrak{g})$	Lie algebra	Free Lie algebra and its completion.
$\mathcal{L}^{[r]}(\mathfrak{g})$	homogeneous component	Degree-r part of the free Lie algebra.
$\mathcal{L}_{\le K} := \pi_{\le K}\widehat{\mathcal{L}}(\mathfrak{g})$	Lie algebra	Truncated free Lie algebra.
e_a	basis element of $\mathfrak{g}$	Degree-one basis.
e^c	basis element of $\mathfrak{g}^*$	Dual basis of $\{e_a\}$.
E_α	basis element of $\mathcal{L}_{\le K}$	Homogeneous Lie basis.
ε^α	basis element of $\mathcal{L}^*_{\le K}$	Dual basis of $\{\mathrm{E}_\alpha\}$.
$\mathrm{ev}_{\mathfrak{g}}$	Lie algebra homomorphism	Evaluation map into $\mathfrak{g}$.
$\mathsf{C}_{\alpha\beta}{}^{\gamma}$	scalar	Structure constant of $\{\mathrm{E}_\alpha\}$.
$\mathsf{A}_{\gamma}{}^{c}$	scalar	Evaluation coordinate of E_γ.
h_t^α	scalar process	Coordinate of $\pi_{\le K}\mathrm{H}_t$.
$\circ d\zeta_t^\beta$	scalar increment	Coordinate of $\pi_{\le K}\big(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})(\mathrm{Sig}_t^{(m)}(\mathbf{g}))^{-1}\big)$.
$\mathsf{D}^{\gamma}{}_{\beta}(h_t)$	scalar	Bernoulli–adjoint coefficient.
$\Omega_m^{\mathfrak{g},(K)}$	$\mathfrak{g}$	Evaluated local logarithmic increment.
$\Omega_{m,c}^{(K)}$	scalar	Coordinate of $\Omega_m^{\mathfrak{g},(K)}$.
F_i	basis element of $\mathfrak{csp}(2n)$	Linear conformal-symplectic basis.
$\mathbf{e}_\ell$	basis element of $\mathbb{R}^{2n}$	Translational basis.
$\circ dR_t, \circ ds_t$	increments	Evaluated linear / translational increments.
$\circ d\Psi_{n,t}^k$	scalar increment	Linear adjoint-recursion coordinate.
$\circ d\Theta_{n,t}^q$	scalar increment	Translational adjoint-recursion coordinate.
$\Lambda_m^{(K)}$	$\mathfrak{csp}(2n)$	Linear part of $\Omega_m^{\mathfrak{g},(K)}$.
$b_m^{(K)}$	$\mathbb{R}^{2n}$	Translational part of $\Omega_m^{\mathfrak{g},(K)}$.

A.2 EXPERIMENTAL DETAILS

Table 5: Hyperparameters for TSL-FLOW.

Parameter	Value / Description
Architecture	
Input representation	Heavy-atom positions q
Hidden dimension	256
Graph encoder	4 layers, 128 hidden dimension
Graph neighborhood	24 spatial neighbors, chain span 2
Temporal mixer	U-Net, 2 levels, kernel size 5
Control dimension	8 symmetric, 8 translation controls
Parameter count	2.87M
Training	
Loss function	Mean Squared Error on positions q
Optimizer	AdamW
Learning rate	1×10^{-3}
Weight decay	1×10^{-4}
Prediction horizon	10,240 steps

Data generation. We construct the benchmark trajectories by running explicit solvent molecular dynamics simulations in OpenMM (Eastman et al., 2017). We minimally repair initial protein structures from ATLAS (Vander Meersche et al., 2024) and mdCATH (Mirarchi et al., 2024) using PDBFixer, solvated them with TIP3P-FB water (Wang et al., 2014), and neutralize them with Na^+ or Cl^- counterions when needed. We run all simulations with the AMBER19 force field (Tian et al., 2019), a 400 K temperature, a 2 fs timestep, and a Langevin-middle style integrator. To form training and evaluation sets, we extract heavy-atom trajectory segments of 10,240 steps from the generated source trajectories, while retaining solvent and ions only in the reference simulations. We evaluate five systems, 1BA5, 2KCN, 2KSL, 2LMK, and 2LSM, from mdCATH, and five systems, 1H02, 2ERL, 3U5V, 5B1N, and 5L87, from ATLAS. For protein–ligand evaluation, we additionally use the MISATO 1F40 complex with ligand GPI (Siebenmorgen et al., 2024), generated with the same protein simulation setup and GAFF ligand parameters (Wang et al., 2004).

Model architecture. We illustrate the overall architecture of TSL-FLOW in Fig. 2. For the mdCATH and ATLAS experiments, we use the same architecture and hyperparameters, while training a separate model for each benchmark system. Starting from the initial conformation, an atom graph encoder builds local molecular context using chain connectivity, spatial neighbors, interatomic distances, directional features, and radial basis features. We then combine the encoded atom features with time embeddings and pass them through a temporal mixer to produce time-dependent hidden states. From these hidden states, we predict low-dimensional control coefficients, which are expanded through learned symmetric and translational bases to construct affine conformal-symplectic generators. We use the resulting generator to parameterize the local SDE update in phase space, where the linear component follows the conformal-symplectic form and the translational component represents the learned drift. Over trajectory segments, we compute local Lie-algebraic increments and compose them with the BCH product, enabling the cumulative update over the prediction horizon to be evaluated through an order-2 prefix-scan computation. Across benchmark systems, we keep the architecture fixed, and only the number of atoms changes.

Training setup. We train TSL-FLOW for 3000 epochs using AdamW with learning rate 10^{-3} and weight decay 10^{-4}. We optimize the model with coordinate MSE on phase states and evaluate long-horizon generation over 10,240 steps. We implement all experiments in Python 3.12.7 with PyTorch 2.7.0+cu126 and CUDA runtime 12.6. We perform training and evaluation on four NVIDIA RTX 5000 Ada Generation GPUs. To measure runtime, we apply CUDA synchronization before and after model execution. For protein-only systems, we compute coordinate metrics on heavy-atom coordinates. For MISATO, we compute ligand interaction metrics after protein alignment using ligand heavy atoms and protein pocket heavy atoms.

Evaluation metrics. Let $x_{t,i}$ and $\hat{x}_{t,i}$ denote the reference and predicted coordinates of atom i at frame t, respectively. We first evaluate coordinate fidelity using mean squared error and aligned RMSD. Specifically, we compute

$$\mathrm{MSE} = \frac{1}{TN}\sum_{t=1}^{T}\sum_{i=1}^{N}\|\hat{x}_{t,i} - x_{t,i}\|_2^2, \qquad \mathrm{RMSD} = \frac{1}{T}\sum_{t=1}^{T}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\|\mathcal{A}_t(\hat{x}_{t,i}) - x_{t,i}\|_2^2}, \tag{7}$$

where $\mathcal{A}_t$ denotes the frame-wise rigid alignment. MSE is reported in nm^2, and RMSD is reported in nm.

To evaluate structural statistics, we compare pairwise distance distributions between reference and generated trajectories. Let $p_{\mathrm{ref}}(d)$ and $p_{\mathrm{pred}}(d)$ denote the empirical distributions of heavy-atom pairwise distances. We report

$$\text{Pair-JSD} = \mathrm{JSD}\left(p_{\mathrm{ref}}(d), p_{\mathrm{pred}}(d)\right), \qquad \mathrm{CSP} = \left\|A^\top J + JA + \gamma J\right\|_F^2, \tag{8}$$

where A is the linear generator of the learned phase-space update, J is the canonical symplectic matrix, and γ is the friction coefficient. The CSP residual measures whether the learned finite-map update remains consistent with the conformal symplectic structure.

For the protein–ligand experiments on MISATO, we follow the trajectory evaluation setting of PVB (Yu et al., 2026). Let ℓ_t and $\hat{\ell}_t$ denote aligned ligand heavy-atom coordinates, and let c_t and $\hat{c}_t$ denote ligand–pocket center-of-mass distances. We compute

$$\text{EMD-ligand} = W_1\left(\{\ell_t\}_{t=1}^T, \{\hat{\ell}_t\}_{t=1}^T\right), \qquad \text{EMD-CoM} = W_1\left(\{c_t\}_{t=1}^T, \{\hat{c}_t\}_{t=1}^T\right), \tag{9}$$

where W_1 denotes the Wasserstein-1 distance. Contact recovery is measured by

$$\text{RMSE-CONTACT} = \sqrt{\frac{1}{|\mathcal{C}|}\sum_{(i,j)\in\mathcal{C}}(\hat{o}_{ij} - o_{ij})^2}, \tag{10}$$

where o_{ij} and $\hat{o}_{ij}$ are reference and predicted ligand–pocket contact occupancies. We additionally report ligand RMSD in Å, the fraction of frames with ligand RMSD at most 2 Å, and ligand–protein heavy-atom clash counts below 2 Å.

Baseline implementations. We compare against representative learned biomolecular dynamics baselines, using the official public implementations of the respective authors whenever available. For fair comparison, all methods are evaluated on the same benchmark systems and reported on the common heavy-atom coordinate format. When a baseline uses a different internal representation or generation procedure, we preserve its original architecture and inference pipeline and only adapt the input and output processing needed for our datasets and metrics.

- ALPHAFLOW:`https://github.com/bjing2016/alphaflow`
- MDGEN: `https://github.com/bjing2016/mdgen`
- MARS-FM: `https://github.com/valence-labs/mars-fm`
- UNISIM: `https://github.com/yaledeus/UniSim`
- PVB: `https://github.com/yaledeus/PVB`
- GLDP: `https://github.com/adityasengar/GLDP`

Table 6: **Quantitative results on MISATO 1F40 protein–ligand heavy-atom trajectories.** We report coordinate fidelity, physical fidelity, and computational efficiency over the same short-horizon evaluation window. The best and second-best results are highlighted in **bold** and **blue**, respectively.

Method	Coordinate Fidelity		Physical Fidelity		Computational Efficiency	
	Mean MSE ↓	Mean RMSD ↓	Pair JSD ↓	CSP ↓	Inference Time ↓	Train Time ↓
Protein–Ligand Heavy-Atom Trajectory Generation						
UniSim (Yu et al., 2025)	6.97×10^{-1}	1.05×10^{0}	1.34×10^{-2}	4.66×10^{-1}	$\mathbf{2.86 \times 10^{2}}$	4.69×10^{2}
PVB (Yu et al., 2026)	$\mathbf{2.82 \times 10^{-2}}$	$\mathbf{2.13 \times 10^{-1}}$	$\mathbf{1.39 \times 10^{-4}}$	$\mathbf{7.66 \times 10^{-2}}$	3.52×10^{2}	$\mathbf{1.59 \times 10^{2}}$
Ours (TSL-Flow)	$\mathbf{7.68 \times 10^{-5}}$	$\mathbf{1.50 \times 10^{-2}}$	$\mathbf{7.45 \times 10^{-6}}$	$\mathbf{1.29 \times 10^{-3}}$	$\mathbf{6.47 \times 10^{0}}$	$\mathbf{2.03 \times 10^{1}}$

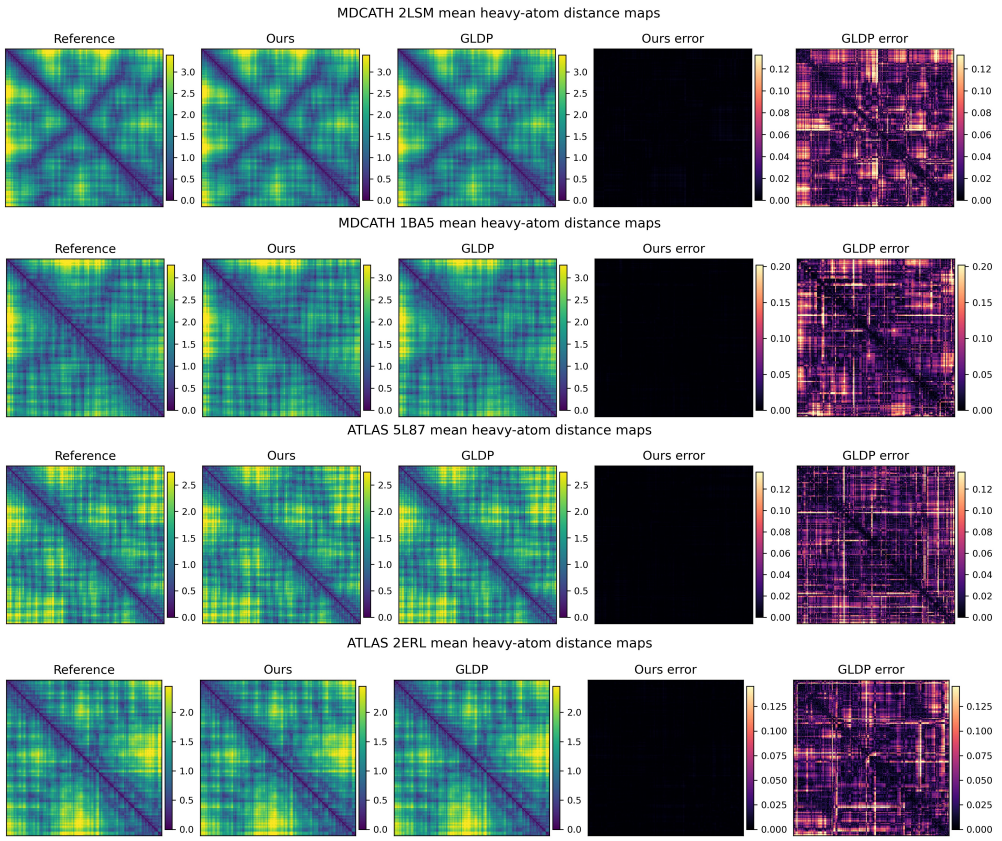


Figure 7: **Heavy-atom distance-map comparison.** We compare mean heavy-atom distance maps from the reference, TSL-Flow, and GLDP, along with absolute error maps relative to the reference. Compared with GLDP, TSL-Flow more closely preserves the reference distance structure across representative mdCATH and ATLAS systems.

A.3 Additional Experiments

Additional MISATO trajectory generation results. We report protein–ligand heavy-atom trajectory generation results on the MISATO 1F40 complex in Table 6. We use this evaluation to complement the ligand interaction metrics in the main text by measuring coordinate fidelity, phase-space structural consistency, and computational efficiency on the same protein–ligand system. Compared with both UniSim and PVB, TSL-Flow achieves the lowest error across all fidelity metrics and reduces mean MSE and mean RMSD by large margins. We observe the same trend in the physical fidelity metrics, where TSL-Flow obtains substantially lower Pair JSD and CSP residuals. These results indicate that the proposed phase-space formulation preserves both coordinate-level trajectory accuracy and heavy-atom distance statistics in the protein–ligand setting.

We observe a similar trend in computational efficiency. TSL-Flow requires 6.47 s for inference and 20.3 s for training time under the recorded evaluation setup, whereas the prior baselines require sub-

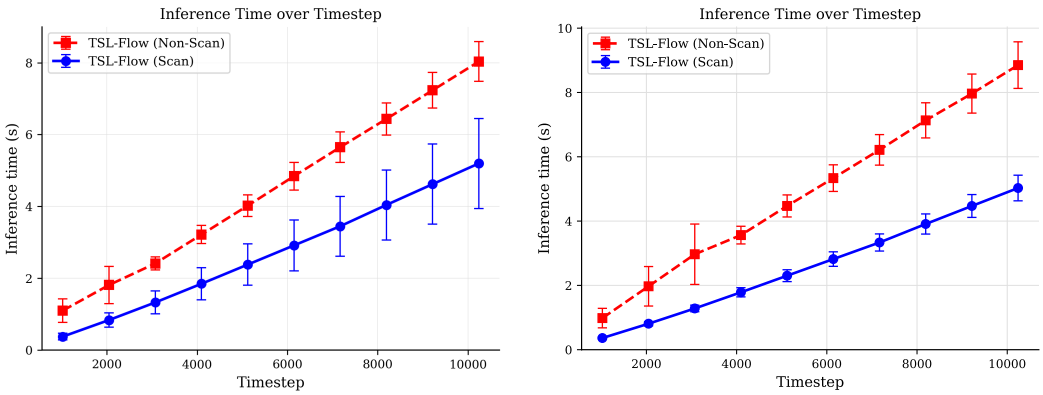


Figure 8: **Inference scaling of scan and non-scan variants.** We compare inference time across increasing trajectory timesteps with ATLAS and mdCATH. *(Left)* On ATLAS, the scan version of TSL-FLOW consistently reduces inference time compared with the non-scan variant. *(Right)* On mdCATH, the same trend is observed across the full prediction horizon, showing that scan-compatible composition improves long-horizon trajectory generation efficiency.

stantially larger wall-clock time. This result supports the main observation that the scan-compatible formulation provides an efficiency advantage beyond protein-only systems. Taken together, the MISATO 1F40 results suggest that the proposed model remains effective when ligand motion and pocket interactions are included in the generated trajectory.

Structural geometry analysis. We additionally examine whether the generated trajectories preserve heavy-atom distance geometry and local structural validity. For representative systems from mdCATH and ATLAS, we compute mean pairwise distance maps over sampled frames and compare them with the reference trajectories. Compared with GLDP, TSL-FLOW yields consistently lower distance-map errors, with mean absolute errors around 1.4–1.8×10^{-3} nm, while GLDP shows errors around 3.6–6.0×10^{-2} nm. This result supports the Pair-JSD comparison in the main tables and indicates that the proposed model better preserves heavy-atom distance relationships.

We further evaluate local geometry using bond length error, Cα neighbor distance error, local pair distance error, and close-contact counts. Across all evaluated systems, TSL-FLOW shows lower errors and produces substantially fewer close contacts than GLDP. Together with the protein–ligand contact visualization in Fig. 5, these diagnostics suggest that the generated trajectories better preserve both global distance structure and local molecular geometry.

Effect of Prefix-Scan Composition. To isolate the computational contribution of the proposed prefix-scan formulation, we compare scan and non-scan variants of TSL-FLOW in Figure 8. The non-scan variant applies the same learned phase-space update sequentially over trajectory timesteps, whereas the scan variant composes local Lie-algebraic updates through the scan-compatible product described in Section 3. Across both ATLAS and mdCATH, the scan variant achieves consistently lower inference time as the prediction horizon increases. This result supports that the efficiency gain comes from the compositional structure of the proposed formulation rather than from a difference in the learned local update itself.

A.4 Mathematical Preliminary

A.4.1 Lie-group actions and induced state-space vector fields

Let $\mathcal{M}$ be a smooth finite-dimensional manifold, and let $\mathfrak{X}(\mathcal{M})$ denote the Lie algebra of smooth vector fields on $\mathcal{M}$, equipped with the Lie–Jacobi bracket

$$[U, V] := U \cdot \nabla V - V \cdot \nabla U, \qquad U, V \in \mathfrak{X}(\mathcal{M}).$$

Let G be a finite-dimensional Lie group with Lie algebra $\mathfrak{g} = T_e\mathsf{G}$. For any elements $g_1, g_2 \in \mathfrak{g}$ and the identity e, we consider a smooth left action ρ satisfying

$$\rho : \mathsf{G} \times \mathcal{M} \to \mathcal{M}, \quad (g, y) \mapsto \rho_g(y), \quad \rho_e(y) = y, \quad \rho_{g_1 g_2}(y) = \rho_{g_1}\big(\rho_{g_2}(y)\big).$$

For a fixed reference point $\mathbf{Y}_0 \in \mathcal{M}$, we define the orbit map $\rho_{\mathbf{Y}_0} : \mathsf{G} \to \mathcal{M}$ where $\rho_{Y_0}(g) := \rho_g(\mathbf{Y}_0)$. When the action is transitive, every point in $\mathcal{M}$ can be written as $y = \rho_g(\mathbf{Y}_0)$ for some $g \in \mathsf{G}$. More generally, the same construction applies on the *orbit* $\mathcal{O}_{\mathbf{Y}_0} := \{\rho_g(\mathbf{Y}_0) : g \in \mathsf{G}\} \subseteq \mathcal{M}$. Then, the local and global isotropy subgroups at $\mathbf{Y}_0$ can be given as

$$\mathsf{G}_{Y_0} := \{g \in \mathsf{G} : \rho_g(Y_0) = Y_0\}, \quad \mathsf{G}_{\mathcal{M}} := \bigcap_{y \in \mathcal{M}} \{g \in \mathsf{G} : \rho_g(y) = y\}.$$

The action is effective if $\mathsf{G}_{\mathcal{M}} = \{e\}$, and it is free if $\mathsf{G}_y = \{e\}$ for every $y \in \mathcal{M}$. For a transitive action, the orbit is diffeomorphic to the homogeneous space $\mathcal{M} \simeq \mathsf{G}/\mathsf{G}_{\mathbf{Y}_0}$. Thus the action map $\rho_{\mathbf{Y}_0}$ may be viewed as a quotient projection from the transformation group to the state manifold. Let $\xi : \mathcal{M} \to \mathfrak{g}$ be a smooth Lie-algebra-valued map. Associated with ξ is a vector field $X_\xi \in \mathfrak{X}(\mathsf{G})$, called the generalized right-translation vector field, defined by

$$X_\xi(g) := \frac{d}{d\tau}\bigg|_{\tau=0} \exp\big(\tau\xi(\rho_{Y_0}(g))\big)g.$$

Equivalently, if $\mathrm{R}_g(h) := h \cdot g$ denotes right multiplication, then one can write $X_\xi(g) = T_e\mathrm{R}_g\big[\xi(\rho_{Y_0}(g))\big]$. When $\xi \in \mathfrak{g}$ is constant, X_ξ is the usual right-invariant vector field on G. Note that when ξ depends on the state $\rho_{Y_0}(g)$, X_ξ is generally not right-invariant, it is still obtained by right-translating the local Lie-algebra element $\xi(\rho_{Y_0}(g))$.

The vector field X_ξ on G induces a vector field on $\mathcal{M}$ through the orbit map $\rho_{Y_0} : \mathsf{G} \to \mathcal{M}$. More precisely, for $y = \rho_{Y_0}(g)$ with $V_\xi := (\rho_{Y_0})_* X_\xi$, we define

$$V_\xi(y) = D\rho_{Y_0}(g)\big[X_\xi(g)\big] = \frac{d}{d\tau}\bigg|_{\tau=0} \rho_{Y_0}\big(\gamma(\tau)\big) = \frac{d}{d\tau}\bigg|_{\tau=0} \rho_{\exp(\tau\xi(y))}(y), \tag{11}$$

where $\gamma(\tau) \in \mathsf{G}$ is the integral curve of X_ξ with $\gamma(0) = g$. In particular, for a fixed element $\eta \in \mathfrak{g}$, the corresponding fundamental vector field is $V_\eta(y) := \frac{d}{d\tau}\big|_{\tau=0} \rho_{\exp(\tau\eta)}(y)$. Hence, the Lie algebra acts infinitesimally on the state space by $\mathfrak{g} \ni \eta \longmapsto V_\eta \in \mathfrak{X}(\mathcal{M})$.

With the right-invariant convention above, the infinitesimal generator map $\eta \mapsto V_\eta$ is a Lie algebra homomorphism up to the standard sign convention determined by the choice of left- or right-invariant Lie algebra structure. Its image forms a finite-dimensional Lie algebra of vector fields on $\mathcal{M}$. The Lie-derivative form of the same push-forward relation is

$$V_\xi(y) = \mathcal{L}_{X_\xi}\rho_{Y_0}(g), \qquad y = \rho_{Y_0}(g),$$

where $\mathcal{L}_{X_\xi}\rho_{Y_0}(g)$ denotes the derivative of the orbit map along the vector field X_ξ at g.

A.4.2 Tensor algebra, Chen series, and $\mathfrak{g}$-valued log-signatures

In this subsection, we clarify the algebraic meaning of the Chen series introduced in the main manuscript. The key point is that, when forming tensor algebras, the Lie algebra $\mathfrak{g}$ is regarded first only as its underlying finite-dimensional vector space, whereas its Lie bracket is used afterward to interpret the resulting formal Lie series as concrete $\mathfrak{g}$-valued local coordinates.

Let $\mathfrak{g}$ be a d-dimensional real Lie algebra with basis $\{e_a\}_{a=1}^d$. We define

$$T(\mathfrak{g}) := \bigoplus_{k=0}^{\infty} \mathfrak{g}^{\otimes k}, \qquad \widehat{T}(\mathfrak{g}) := \prod_{k=0}^{\infty} \mathfrak{g}^{\otimes k}, \qquad \mathfrak{g}^{\otimes 0} := \mathbb{R}.$$

The algebra product on $T(\mathfrak{g})$ and $\widehat{T}(\mathfrak{g})$ is the concatenation product, written by juxtaposition. Explicitly, for $u \in \mathfrak{g}^{\otimes k}$ and $v \in \mathfrak{g}^{\otimes \ell}$, their product uv belongs to $\mathfrak{g}^{\otimes(k+\ell)}$. Inside $T(\mathfrak{g})$, we denote by $\mathcal{L}(\mathfrak{g}) \subset T(\mathfrak{g})$, the free Lie algebra generated by the degree-one space $\mathfrak{g}$, and by $\widehat{\mathcal{L}}(\mathfrak{g}) \subset \widehat{T}(\mathfrak{g})$ its completion.

Next, let $g = \{g_t\}_{t\in[s,t]}$ be a $\mathfrak{g}$-valued continuous semimartingale on a time interval $[s,t]$. Then, we write the local driver as $\circ dg_u = \sum_{a=1}^{d} e_a \circ dg_u^a$, where the symbol $\circ dg_u$ denotes the Stratonovich increment of the time-indexed $\mathfrak{g}$-valued driver. It is therefore not a differential form on the state space $\mathcal{M}$, but rather a degree-one increment taking values in $\mathfrak{g}$. Consequently, one may define a tensor-algebraic element as $\circ dg_{u_1} \otimes \cdots \otimes \circ dg_{u_k} \in \mathfrak{g}^{\otimes k}$ without loss of generality.

For each $k \geq 1$, define the k-th level iterated Stratonovich integral by

$$S_{s,t}^{(k)}(g) := \sum_{a_1,\ldots,a_k=1}^{d} \left(\int_{s<u_1<\cdots<u_k\leq t} \circ dg_{u_1}^{a_1} \cdots \circ dg_{u_k}^{a_k} \right) e_{a_1} \otimes \cdots \otimes e_{a_k} \in \mathfrak{g}^{\otimes k}.$$

The associated Chen series, or signature is defined as an infinite sum of iterated integrals:

$$\mathrm{Sig}_m(\mathfrak{g}) := 1 + \sum_{k=1}^{\infty} S_{s,t}^{(k)}(g) \in \widehat{T}(\mathfrak{g}).$$

For signatures of Stratonovich drivers, Chen's identity holds: for any $s < u < t$, $S_{s,t}(g) = S_{s,u}(g)S_{u,t}(g)$, where the product is taken in the concatenation algebra $\widehat{T}(\mathfrak{g})$. As a consequence, one may show that $\log S_{s,t}(g) \in \widehat{\mathcal{L}}(\mathfrak{g})$. Next, we define the formal log-signature on $[s,t]$ by

$$\Omega_{s,t}^{\mathrm{form}} := \log S_{s,t}(g) \in \widehat{\mathcal{L}}(\mathfrak{g}) \subset \widehat{T}(\mathfrak{g}).$$

In particular, for the m-th segment, $\Omega_m^{\mathrm{form}} := \log S_m$. At this stage, $\Omega_{s,t}^{\mathrm{form}}$ is still a Lie series in the completed free Lie algebra generated by the underlying vector space of $\mathfrak{g}$. To obtain a concrete $\mathfrak{g}$-valued object, we evaluate a finite truncation using the actual Lie bracket of $\mathfrak{g}$. Let

$$\pi_{\leq K} : \widehat{T}(\mathfrak{g}) \to \bigoplus_{k=0}^{K} \mathfrak{g}^{\otimes k}$$

denote the truncation map. By the universal property of the free Lie algebra, the identity map on the generating space extends uniquely to a Lie algebra homomorphism from the truncated free Lie algebra into the given Lie algebra $\mathfrak{g}$. We therefore define the K-th order $\mathfrak{g}$-valued log-signature by

$$\Omega_{s,t}^{\mathfrak{g},(K)} := \mathrm{ev}_{\mathfrak{g}}\big(\pi_{\leq K}\Omega_{s,t}^{\mathrm{form}}\big) \in \mathfrak{g}.$$

When no confusion arises, we simply write $\Omega_{s,t}^{(K)}$ or $\Omega_m^{(K)}$.

Here, we define $\mathrm{ev}_{\mathfrak{g}} : \mathcal{L}_{\leq K} \to \mathfrak{g}$ to be the unique Lie algebra homomorphism extending the identity on the degree-one generators. More precisely, viewing $\mathfrak{g}$ first as an underlying vector space, the inclusion $\mathfrak{g} = \mathcal{L}^{[1]}(\mathfrak{g}) \hookrightarrow \mathcal{L}_{\leq K}$ induces a unique map $\mathrm{ev}_{\mathfrak{g}}$ such that

$$\mathrm{ev}_{\mathfrak{g}}(v) = v, \quad v \in \mathfrak{g}, \quad \mathrm{ev}_{\mathfrak{g}}([U,V]_{\mathrm{free}}) = [\mathrm{ev}_{\mathfrak{g}}(U), \mathrm{ev}_{\mathfrak{g}}(V)]_{\mathfrak{g}}, \quad U, V \in \mathcal{L}_{\leq K}$$

Equivalently, $\mathrm{ev}_{\mathfrak{g}}$ is defined recursively by sending each formal Lie monomial in the free Lie algebra to the corresponding iterated bracket computed using the actual Lie bracket of $\mathfrak{g}$. Hence, if

$$\pi_{\leq K}\Omega_{s,t}^{\mathrm{form}} = \sum_{\alpha} \omega_{s,t}^{\alpha} E_{\alpha},$$

then

$$\Omega_{s,t}^{\mathfrak{g},(K)} = \mathrm{ev}_{\mathfrak{g}}\big(\pi_{\leq K}\Omega_{s,t}^{\mathrm{form}}\big) = \sum_{\alpha} \omega_{s,t}^{\alpha}\ \mathrm{ev}_{\mathfrak{g}}(E_{\alpha}) \in \mathfrak{g}.$$

In the following section, we will investigate local coordinate expression of Ω, when $(\mathsf{G}, \mathfrak{g})$ is specified.

A.5 PROOF

Proposition A.1. *For the fixed state representation (i.e., orbit map) $\mathbf{Y}_0 \in \mathcal{M}$, let us define the state-space process $\mathbf{Y}_t := \rho_{\mathbf{G}_t}(\mathbf{Y}_0)$. Then, the stochastic process $\mathbf{Y}_t$ is a solution to the following Stratonovich stochastic differential equation:*

$$d\mathbf{Y}_t = V_{\xi(t)}(\mathbf{Y}_t)dt + \sum_{i=1}^{r_w} V_{\Sigma_i(t)}(\mathbf{Y}_t) \circ d\mathbf{W}_t^i. \tag{12}$$

Equivalently, the physical state-space dynamics is the one-point motion of the Lie-group-valued stochastic flow $\mathbf{G}_t$ under the action ρ.

Proof. For the fixed initial state Y_0, introduce the orbit map

$$\Phi_{\mathbf{Y}_0} : \mathsf{G} \to \mathcal{M}, \qquad \Phi_{Y_0}(g) = \rho_g(\mathbf{Y}_0).$$

The key observation is that the differential of the orbit map sends a right-translated Lie-algebra element to the corresponding fundamental vector field. Indeed, for any $g \in \mathsf{G}$ and $\eta \in \mathfrak{g}$,

$$\begin{aligned} D\Phi_{\mathbf{Y}_0}(g)\big[T_e\mathrm{R}_g\eta\big] &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \Phi_{\mathbf{Y}_0}\big(\exp(\varepsilon\eta)g\big) && (13)\\ &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \rho_{\exp(\varepsilon\eta)g}(\mathbf{Y}_0) && (14)\\ &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \rho_{\exp(\varepsilon\eta)}\big(\rho_g(\mathbf{Y}_0)\big) && (15)\\ &= V_\eta\big(\rho_g(\mathbf{Y}_0)\big). && (16) \end{aligned}$$

The third equality uses the left-action property $\rho_{g_1 g_2} = \rho_{g_1} \circ \rho_{g_2}$. Since the stochastic differential equation for $\mathbf{G}_t$ is written in Stratonovich form, applying the ordinary chain rule proves the claim:

$$\begin{aligned} d\mathbf{Y}_t &= D\Phi_{\mathbf{Y}_0}(\mathbf{G}_t) \circ d\mathbf{G}_t && (17)\\ &= D\Phi_{\mathbf{Y}_0}(\mathbf{G}_t)\left[T_e\mathrm{R}_{\mathbf{G}_t}\xi(t)dt + \sum_{i=1}^{r_w} T_e\mathrm{R}_{\mathbf{G}_t}\Sigma_i(t) \circ d\mathbf{W}_t^i\right] && (18)\\ &= V_{\xi(t)}(\rho_{\mathbf{G}_t}(\mathbf{Y}_0))dt + \sum_{i=1}^{r_w} V_{\Sigma_i(t)}(\rho_{\mathbf{G}_t}(\mathbf{Y}_0)) \circ d\mathbf{W}_t^i && (19)\\ &= V_{\xi(t)}(\mathbf{Y}_t)dt + \sum_{i=1}^{r_w} V_{\Sigma_i(t)}(\mathbf{Y}_t) \circ d\mathbf{W}_t^i. && (20) \end{aligned}$$

□

Proposition A.2 (Thermo-symplectic Preservation in Stratonovich SDEs)**.** *Let $\Phi_{s,t}$ be the deterministic flow generated by $V_{\xi(t)}$, where $\xi(t)$ is a drift function of Stratonovich SDEs. If the dynamics satisfies thermo-symplectic symmetry, then the following is true:*

$$\Phi_{s,t}^* \omega_J = e^{-\gamma(t-s)}\omega_J.$$

Proof. Since $\mathcal{M} = T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$, we regard the canonical symplectic form as the constant 2-form

$$(\omega_J)_y(u, v) = u^\top J v, \qquad y \in \mathcal{M}, \quad u, v \in T_y\mathcal{M} \simeq \mathbb{R}^{2n}.$$

We first extract the differential constraint on the state-space vector field $V_{\xi(t)}$ from the thermo-symplectic condition $\mathcal{L}_{V_{\xi(t)}}\omega_J = -\gamma\omega_J$. Let us fix $y \in \mathcal{M}$ and $u, v \in T_y\mathcal{M}$. We extend u, v to constant vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ by

$$U(z) \equiv u, \qquad V(z) \equiv v, \qquad z \in \mathcal{M}.$$

Then, the differentials of two vector fields vanish, *i.e.*, $DU(z) = 0$ and $DV(z) = 0$, therefore the Lie brackets satisfy

$$[V_{\xi(t)}, U](y) = DU(y)V_{\xi(t)}(y) - DV_{\xi(t)}(y)U(y) = -DV_{\xi(t)}(y)u, \tag{21}$$

with the sign convention $[U, V] = U \cdot \nabla V - V \cdot \nabla U$ used in the appendix giving

$$[V_{\xi(t)}, U](y) = -DV_{\xi(t)}(y)u, \qquad [V_{\xi(t)}, V](y) = -DV_{\xi(t)}(y)v.$$

Moreover, since U and V are constant and ω_J is constant on $\mathbb{R}^{2n}$, one can observe that $\omega_J(U, V)(z) = u^\top Jv$ is independent of z, hence $V_{\xi(t)}\left[\omega_J(U, V)\right](y) = 0$. Now, we apply the definition of Lie derivative:

$$\begin{aligned}
(\mathcal{L}_{V_{\xi(t)}}\omega_J)_y(u, v) &= V_{\xi(t)}\left[\omega_J(U, V)\right](y) - \omega_J([V_{\xi(t)}, U], V)(y) - \omega_J(U, [V_{\xi(t)}, V])(y) \\
&= 0 - \omega_J(-DV_{\xi(t)}(y)u, v) - \omega_J(u, -DV_{\xi(t)}(y)v) \\
&= (DV_{\xi(t)}(y)u)^\top Jv + u^\top J(DV_{\xi(t)}(y)v) \\
&= u^\top DV_{\xi(t)}(y)^\top Jv + u^\top JDV_{\xi(t)}(y)v \\
&= u^\top \big(DV_{\xi(t)}(y)^\top J + JDV_{\xi(t)}(y)\big)v.
\end{aligned}$$

On the other hand, the thermo-symplectic condition gives

$$(\mathcal{L}_{V_{\xi(t)}}\omega_J)_y(u, v) = -\gamma(\omega_J)_y(u, v) = -\gamma u^\top Jv.$$

Comparing this with above estimation, we conclude that

$$u^\top \big(DV_{\xi(t)}(y)^\top J + JDV_{\xi(t)}(y)\big)v = -\gamma u^\top Jv \qquad \text{for all } u, v \in \mathbb{R}^{2n}.$$

Hence

$$DV_{\xi(t)}(y)^\top J + JDV_{\xi(t)}(y) = -\gamma J \qquad \text{for all } y \in \mathcal{M}. \tag{22}$$

We now turn to the flow $\Phi_{s,t}$ generated by $V_{\xi(t)}$, namely

$$\frac{d}{dt}\Phi_{s,t}(y) = V_{\xi(t)}(\Phi_{s,t}(y)), \qquad \Phi_{s,s}(y) = y.$$

Define its Jacobian by $F_{s,t}(y) := D\Phi_{s,t}(y) \in \mathbb{R}^{2n\times 2n}$. Differentiating the flow equation with respect to the initial state y gives

$$\frac{d}{dt}F_{s,t}(y) = DV_{\xi(t)}(\Phi_{s,t}(y))F_{s,t}(y). \tag{23}$$

Next, by definition of pullback,

$$\begin{aligned}
(\Phi_{s,t}^*\omega_J)_y(u, v) &= (\omega_J)_{\Phi_{s,t}(y)}\big(D\Phi_{s,t}(y)u, D\Phi_{s,t}(y)v\big) &(24)\\
&= \omega_J\big(F_{s,t}(y)u, F_{s,t}(y)v\big) &(25)\\
&= u^\top F_{s,t}(y)^\top JF_{s,t}(y)v. &(26)
\end{aligned}$$

Differentiate this with respect to t:

$$\begin{aligned}
\frac{d}{dt}(\Phi_{s,t}^*\omega_J)_y(u, v) &= \frac{d}{dt}\Big(u^\top F_{s,t}(y)^\top JF_{s,t}(y)v\Big) \\
&= u^\top\Big(\dot{F}_{s,t}(y)^\top JF_{s,t}(y) + F_{s,t}(y)^\top J\dot{F}_{s,t}(y)\Big)v \\
&= u^\top\Big(F_{s,t}(y)^\top DV_{\xi(t)}(\Phi_{s,t}(y))^\top JF_{s,t}(y) + F_{s,t}(y)^\top JDV_{\xi(t)}(\Phi_{s,t}(y))F_{s,t}(y)\Big)v \\
&= u^\top F_{s,t}(y)^\top\Big(DV_{\xi(t)}(\Phi_{s,t}(y))^\top J + JDV_{\xi(t)}(\Phi_{s,t}(y))\Big)F_{s,t}(y)v.
\end{aligned}$$

By using Eq. 22 at the point $\Phi_{s,t}(y)$, we obtain

$$\begin{aligned}
\frac{d}{dt}(\Phi_{s,t}^*\omega_J)_y(u, v) &= u^\top F_{s,t}(y)^\top(-\gamma J)F_{s,t}(y)v &(27)\\
&= -\gamma u^\top F_{s,t}(y)^\top JF_{s,t}(y)v &(28)\\
&= -\gamma(\Phi_{s,t}^*\omega_J)_y(u, v). &(29)
\end{aligned}$$

Since this holds for every $y \in \mathcal{M}$ and every $u, v \in T_y\mathcal{M}$, it follows that

$$\frac{d}{dt}\Phi^*_{s,t}\omega_J = -\gamma\Phi^*_{s,t}\omega_J.$$

Finally, at $t = s$, one has $\Phi_{s,s} = \mathrm{Id}_{\mathcal{M}}$, thus we have $F_{s,s}(y) = I_{2n}$. This leads to the following calculation:

$$(\Phi^*_{s,s}\omega_J)_y(u, v) = u^\top I_{2n}^\top J I_{2n} v = u^\top J v = (\omega_J)_y(u, v). \tag{30}$$

Hence, we have shown $\Phi^*_{s,s}\omega_J = \omega_J$ where it solves the linear ODE and conclude the relation $\Phi^*_{s,t}\omega_J = e^{-\gamma(t-s)}\omega_J$

$$\frac{d}{dt}\Phi^*_{s,t}\omega_J = -\gamma\Phi^*_{s,t}\omega_J.$$

□

Proposition A.3 (Underdamped Langevin flow is conformally symplectic)**.** *Assume that* $\mathbf{M} \in \mathbb{R}^{n\times n}$ *is symmetric positive definite and that* $\mathrm{U} \in C^2(\mathbb{R}^n)$*. Then, one can rewrite Langevin dynamics in Eq. 1 in Stratonovich form as*

$$d\mathbf{Q}^i_t = \sum_{a=1}^n (\mathbf{M}^{-1})_{ia}\mathbf{P}^a_t dt, \quad d\mathbf{P}^i_t = -\partial_{q_i}\mathrm{U}(\mathbf{Q}_t)dt - \gamma\mathbf{P}^i_t dt + \sqrt{2\gamma\beta^{-1}} \circ d\mathbf{B}^i_t. \tag{31}$$

Let us define the drift vector field of the form given as follows:

$$\mathsf{X}(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^n \left(\sum_{a=1}^n (\mathbf{M}^{-1})_{ia}\mathbf{p}_a\right)\partial_{q_i} + \sum_{i=1}^n \big(-\partial_{q_i}\mathrm{U}(\mathbf{q}) - \gamma\mathbf{p}_i\big)\partial_{p_i}.$$

Then, the flow on $\mathcal{M}$ *generated by* X*,* $\Phi_{s,t}(\mathbf{q}, \mathbf{p}) = \big(\mathbf{Q}_{s,t}(\mathbf{q}, \mathbf{p}), \mathbf{P}_{s,t}(\mathbf{q}, \mathbf{p})\big)$ *satisfies temporal thermo-symplectic preservation conditions:*

$$\Phi^*_{s,t}\omega_J = e^{-\gamma(t-s)}\omega_J.$$

Proof. Since the diffusion coefficient of Langevin dynamics is constant, the Itô and Stratonovich forms coincide. Hence Eq. 1 may indeed be written as Eq. 31. We now compute the pullback of ω_J under the flow $\Phi_{s,t}$ in coordinates. Let us write the flow in component-wise manner:

$$\Phi_{s,t}(\mathbf{q}, \mathbf{p}) = \big(\mathbf{Q}^1_{s,t}(\mathbf{q}, \mathbf{p}), \ldots, \mathbf{Q}^n_{s,t}(\mathbf{q}, \mathbf{p}), \mathbf{P}^1_{s,t}(\mathbf{q}, \mathbf{p}), \ldots, \mathbf{P}^n_{s,t}(\mathbf{q}, \mathbf{p})\big).$$

where,

$$\frac{d}{dt}\begin{pmatrix}\mathbf{Q}_{s,t}\\ \mathbf{P}_{s,t}\end{pmatrix} = \begin{pmatrix}\mathbf{M}^{-1}\mathbf{P}_{s,t}\\ -\nabla\mathrm{U}(\mathbf{Q}_{s,t}) - \gamma\mathbf{P}_{s,t}\end{pmatrix}, \quad (\mathbf{Q}_{s,s}, \mathbf{P}_{s,s}) = (\mathbf{q}, \mathbf{p}). \tag{32}$$

By definition of pullback, the coordinate differential can be obtained as

$$\Phi^*_{s,t}dq_i = d\big(\mathbf{Q}^i_{s,t}\big), \quad \Phi^*_{s,t}dp_i = d\big(\mathbf{P}^i_{s,t}\big), \tag{33}$$

Thus,

$$\Phi^*_{s,t}\omega_J = \Phi^*_{s,t}\left(\sum_{i=1}^n dq_i \wedge dp_i\right) \tag{34}$$

$$= \sum_{i=1}^n \Phi^*_{s,t}dq_i \wedge \Phi^*_{s,t}dp_i \tag{35}$$

$$= \sum_{i=1}^n d\big(\mathbf{Q}^i_{s,t}\big) \wedge d\big(\mathbf{P}^i_{s,t}\big). \tag{36}$$

Next, we differentiate each pulled-back coordinate 1-form with respect to t. From Eq. 32,

$$\frac{d}{dt}\mathbf{Q}^i_{s,t} = \sum_{a=1}^n (\mathbf{M}^{-1})_{ia}\mathbf{P}^a_{s,t}, \qquad \frac{d}{dt}\mathbf{P}^i_{s,t} = -\partial_{q_i}\mathrm{U}(\mathbf{Q}_{s,t}) - \gamma\mathbf{P}^i_{s,t}.$$

Hence,

$$\frac{d}{dt}\big(\Phi^*_{s,t}dq_i\big) = \frac{d}{dt}d\big(\mathbf{Q}^i_{s,t}\big) \tag{37}$$

$$= d\left(\frac{d}{dt}\mathbf{Q}^i_{s,t}\right) \tag{38}$$

$$= d\left(\sum_{a=1}^{n}(\mathbf{M}^{-1})_{ia}\mathbf{P}^a_{s,t}\right) \tag{39}$$

$$= \sum_{a=1}^{n}(\mathbf{M}^{-1})_{ia}d\big(\mathbf{P}^a_{s,t}\big) \tag{40}$$

$$= \sum_{a=1}^{n}(\mathbf{M}^{-1})_{ia}\Phi^*_{s,t}dp_a, \tag{41}$$

and similarly

$$\frac{d}{dt}\big(\Phi^*_{s,t}dp_i\big) = \frac{d}{dt}d\big(\mathbf{P}^i_{s,t}\big) \tag{42}$$

$$= d\left(\frac{d}{dt}\mathbf{P}^i_{s,t}\right) \tag{43}$$

$$= d\left(-\partial_{q_i}\mathrm{U}(\mathbf{Q}_{s,t}) - \gamma\mathbf{P}^i_{s,t}\right) \tag{44}$$

$$= -\sum_{a=1}^{n}\partial^2_{q_aq_i}\mathrm{U}(\mathbf{Q}_{s,t})d\big(\mathbf{Q}^a_{s,t}\big) - \gamma d\big(\mathbf{P}^i_{s,t}\big) \tag{45}$$

$$= -\sum_{a=1}^{n}\partial^2_{q_aq_i}\mathrm{U}(\mathbf{Q}_{s,t})\Phi^*_{s,t}dq_a - \gamma\Phi^*_{s,t}dp_i. \tag{46}$$

Now differentiate Eq. 36:

$$\frac{d}{dt}\big(\Phi^*_{s,t}\omega_J\big) = \frac{d}{dt}\left(\sum_{i=1}^{n}\Phi^*_{s,t}dq_i \wedge \Phi^*_{s,t}dp_i\right) \tag{47}$$

$$= \sum_{i=1}^{n}\frac{d}{dt}\big(\Phi^*_{s,t}dq_i\big) \wedge \Phi^*_{s,t}dp_i + \sum_{i=1}^{n}\Phi^*_{s,t}dq_i \wedge \frac{d}{dt}\big(\Phi^*_{s,t}dp_i\big). \tag{48}$$

Substituting 41 and 46 into 48,

$$\begin{aligned}\frac{d}{dt}\big(\Phi^*_{s,t}\omega_J\big) = &\sum_{i=1}^{n}\left(\sum_{a=1}^{n}(\mathbf{M}^{-1})_{ia}\Phi^*_{s,t}dp_a\right) \wedge \Phi^*_{s,t}dp_i \\ &+ \sum_{i=1}^{n}\Phi^*_{s,t}dq_i \wedge \left(-\sum_{a=1}^{n}\partial^2_{q_aq_i}\mathrm{U}(\mathbf{Q}_{s,t})\Phi^*_{s,t}dq_a - \gamma\Phi^*_{s,t}dp_i\right)\end{aligned} \tag{49}$$

$$\begin{aligned}= &\sum_{i=1}^{n}\sum_{a=1}^{n}(\mathbf{M}^{-1})_{ia}\Phi^*_{s,t}dp_a \wedge \Phi^*_{s,t}dp_i \\ &- \sum_{i=1}^{n}\sum_{a=1}^{n}\partial^2_{q_aq_i}\mathrm{U}(\mathbf{Q}_{s,t})\Phi^*_{s,t}dq_i \wedge \Phi^*_{s,t}dq_a \\ &- \gamma\sum_{i=1}^{n}\Phi^*_{s,t}dq_i \wedge \Phi^*_{s,t}dp_i.\end{aligned} \tag{50}$$

We now simplify the first two sums. For the $dp \wedge dp$ term, since $\mathbf{M}^{-1}$ is symmetric,

$$\sum_{i,a}^{n}(\mathbf{M}^{-1})_{ia}\Phi^*_{s,t}dp_a \wedge \Phi^*_{s,t}dp_i = \frac{1}{2}\sum_{i,a}^{n}\Big((\mathbf{M}^{-1})_{ia} - (\mathbf{M}^{-1})_{ai}\Big)\Phi^*_{s,t}dp_a \wedge \Phi^*_{s,t}dp_i \tag{51}$$

$$= 0. \tag{52}$$

For the $dq \wedge dq$ term, since $\mathrm{U} \in C^2$, its Hessian is symmetric, hence

$$\sum_{i,a}^{n} \partial^2_{q_a q_i} \mathrm{U}(\mathbf{Q}_{s,t}) \Phi^*_{s,t} dq_i \wedge \Phi^*_{s,t} dq_a = \frac{1}{2} \sum_{i,a} \Big(\partial^2_{q_a q_i} \mathrm{U}(\mathbf{Q}_{s,t}) - \partial^2_{q_i q_a} \mathrm{U}(\mathbf{Q}_{s,t}) \Big) \Phi^*_{s,t} dq_i \wedge \Phi^*_{s,t} dq_a = 0. \quad (53)$$

Using 52, 53, and Eq. 36, Eq. 50 reduces to

$$\frac{d}{dt} \big(\Phi^*_{s,t} \omega_J \big) = -\gamma \sum_{i=1}^{n} \Phi^*_{s,t} dq_i \wedge \Phi^*_{s,t} dp_i = -\gamma \Phi^*_{s,t} \omega_J. \quad (54)$$

Finally, at $t = s$, we have $\Phi_{s,s} = \mathrm{Id}$, and the following is true: $\Phi^*_{s,s} \omega_J = \omega_J$. Therefore, $\Phi^*_{s,t} \omega_J$ is the solution to the linear ODE

$$\frac{d}{dt} \big(\Phi^*_{s,t} \omega_J \big) = -\gamma \Phi^*_{s,t} \omega_J, \quad \Phi^*_{s,s} \omega_J = \omega_J, \quad \Phi^*_{s,t} \omega_J = e^{-\gamma(t-s)} \omega_J.$$

This proves the claim. □

Proposition A.4. *Assume that $V_{\Sigma_i(t)}$ are C^1 in the state variable and that, for every $(t, y) \in [0, T] \times \mathbb{R}^{2n}$ with $y = (q, p)$, let us define parametrized function*

$$u_\theta(t, y) := \begin{pmatrix} 0 & \sqrt{\frac{\beta}{2\gamma}} I_n \end{pmatrix} \left(\begin{pmatrix} \mathbf{M}^{-1} p \\ -\nabla \mathrm{U}(q) - \gamma p \end{pmatrix} - V_{\xi(t)}(y) - \frac{1}{2} \sum_{i=1}^{n} D\big(V_{\Sigma_i(t)}\big)(y) V_{\Sigma_i(t)}(y) \right).$$

If Novikov's condition holds:

$$\mathbb{E}_{\mathbb{P}_\mathbf{Y}} \left[\exp \left(\frac{1}{2} \int_0^T \|u_\theta(t, \mathbf{Y}_t)\|^2 dt \right) \right] < \infty \quad (55)$$

Then,

$$\mathrm{KL}(\mathbb{P}_\mathbf{Y} \| \mathbb{P}_\mathbf{X}) = \frac{1}{2} \mathbb{E}_{\mathbb{P}_\mathbf{Y}} \int_0^T \|u_\theta(t, \mathbf{Y}_t)\|^2 dt. \quad (56)$$

Proof. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P})$ be a filtered probability space supporting an n-dimensional Brownian motion $\mathbf{W}$ and a weak solution $\mathbf{Y} = (\mathbf{q}, \mathbf{p})$ of the learned Stratonovich SDE

$$d\mathbf{Y}_t = V_{\xi(t)}(\mathbf{Y}_t) dt + \sum_{i=1}^{n} V_{\Sigma_i(t)}(\mathbf{Y}_t) \circ d\mathbf{W}^i_t,$$

such that $\mathrm{Law}_\mathbb{P}(\mathbf{Y}) = \mathbb{P}_\mathbf{Y}$. We first rewrite this equation in Itô form. Since the state space is Euclidean and each $V_{\Sigma_i(t)}$ is C^1 in the state variable, the standard Stratonovich-to-Itô conversion gives

$$d\mathbf{Y}_t = \left(V_{\xi(t)}(\mathbf{Y}_t) + \frac{1}{2} \sum_{i=1}^{n} D\big(V_{\Sigma_i(t)}\big)(\mathbf{Y}_t) V_{\Sigma_i(t)}(\mathbf{Y}_t) \right) dt + \begin{pmatrix} V_{\Sigma_1(t)}(\mathbf{Y}_t) & \cdots & V_{\Sigma_n(t)}(\mathbf{Y}_t) \end{pmatrix} d\mathbf{W}_t,$$

By assumption,

$$\begin{pmatrix} V_{\Sigma_1(t)}(y) & \cdots & V_{\Sigma_n(t)}(y) \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2\gamma\beta^{-1}} I_n \end{pmatrix} =: \sigma,$$

with

$$b_Y(t, y) := V_{\xi(t)}(y) + \frac{1}{2} \sum_{i=1}^{n} D\big(V_{\Sigma_i(t)}\big)(y) V_{\Sigma_i(t)}(y).$$

On the other hand, the reference process $\mathbf{X}_t = (\mathbf{Q}_t, \mathbf{P}_t)$ satisfies $d\mathbf{X}_t = b_X(\mathbf{X}_t) dt + \sigma d\mathbf{B}_t$, where $b_X(y) := \big(\mathbf{M}^{-1} p, -\nabla \mathrm{U}(q) - \gamma p\big)^T$. Hence the learned and reference processes have the same

diffusion matrix σ. We now compute the difference of the two Itô drifts. By the assumption in the proposition,

$$\begin{pmatrix} I_n & 0 \end{pmatrix} b_Y(t,y) = \mathbf{M}^{-1}p = \begin{pmatrix} I_n & 0 \end{pmatrix} b_X(y), \quad \begin{pmatrix} I_n & 0 \end{pmatrix} \big(b_X(y) - b_Y(t,y)\big) = 0.$$

Therefore, the drift discrepancy lies entirely in the momentum coordinates. Equivalently,

$$b_X(y) - b_Y(t,y) \in \mathrm{Im}(\sigma) = \left\{ \begin{pmatrix} 0 \\ z \end{pmatrix} : \ z \in \mathbb{R}^n \right\}.$$

Similarly, one can define the unique vector $u_\theta(t,y) = \sigma^\dagger \big(b_X(y) - b_Y(t,y)\big)$. Substituting the definitions of b_X and b_Y yields

$$u_\theta(t,y) = \begin{pmatrix} 0 & \sqrt{\dfrac{\beta}{2\gamma}} I_n \end{pmatrix} \left(\begin{pmatrix} \mathbf{M}^{-1}p \\ -\nabla \mathrm{U}(q) - \gamma p \end{pmatrix} - V_{\xi(t)}(y) - \frac{1}{2} \sum_{i=1}^{n} D\big(V_{\Sigma_i(t)}\big)(y) V_{\Sigma_i(t)}(y) \right),$$

which is precisely the expression stated in the proposition. Next, we define the exponential process

$$Z_t := \exp\left(\int_0^t u_\theta(s, \mathbf{Y}_s)^\top d\mathbf{W}_s - \frac{1}{2} \int_0^t \|u_\theta(s, \mathbf{Y}_s)\|^2 ds \right), \qquad 0 \le t \le T.$$

By Novikov's condition,

$$\mathbb{E}_{\mathbb{P}} \left[\exp\left(\frac{1}{2} \int_0^T \|u_\theta(t, \mathbf{Y}_t)\|^2 dt \right) \right] < \infty,$$

so Z_t is a martingale on $[0,T]$. In particular, $\mathbb{E}_{\mathbb{P}}[Z_T] = 1$ where a new probability measure $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ can be defined as $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = Z_T$. By Girsanov's theorem, the following process is an n-dimensional Brownian motion under $\widetilde{\mathbb{P}}$:

$$\widetilde{\mathbf{W}}_t := \mathbf{W}_t - \int_0^t u_\theta(s, \mathbf{Y}_s) ds$$

By substituting the dynamics $d\mathbf{W}_t = d\widetilde{\mathbf{W}}_t + u_\theta(t, \mathbf{Y}_t)dt$ into the Itô dynamics of $\mathbf{Y}_t$, we obtain

$$\begin{aligned} d\mathbf{Y}_t &= b_Y(t, \mathbf{Y}_t)dt + \sigma\big(d\widetilde{\mathbf{W}}_t + u_\theta(t, \mathbf{Y}_t)dt\big) \\ &= \big(b_Y(t, \mathbf{Y}_t) + \sigma u_\theta(t, \mathbf{Y}_t)\big)dt + \sigma d\widetilde{\mathbf{W}}_t \\ &= b_X(\mathbf{Y}_t)dt + \sigma d\widetilde{\mathbf{W}}_t. \\ &= \begin{pmatrix} \mathbf{M}^{-1}\mathbf{p}_t \\ -\nabla \mathrm{U}(\mathbf{q}_t) - \gamma \mathbf{p}_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sqrt{2\gamma\beta^{-1}} I_n \end{pmatrix} d\widetilde{\mathbf{W}}_t \end{aligned} \tag{57}$$

Since $\mathbf{X}_0$ and $\mathbf{Y}_0$ have the same law, it follows that under $\widetilde{\mathbb{P}}$ the process $\mathbf{Y}$ has the same path law as the reference Langevin process $\mathbf{X}$, namely $\mathrm{Law}_{\widetilde{\mathbb{P}}}(\mathbf{Y}) = \mathbb{P}_{\mathbf{X}}$. To identify the Radon–Nikodým derivative on path space, observe that σ is constant and has full column rank, so from

$$d\mathbf{Y}_t = b_Y(t, \mathbf{Y}_t)dt + \sigma d\mathbf{W}_t$$

we recover $\mathbf{W}$ from $\mathbf{Y}$ by

$$\mathbf{W}_t = \sigma^\dagger \left(\mathbf{Y}_t - \mathbf{Y}_0 - \int_0^t b_Y(s, \mathbf{Y}_s) ds \right), \qquad \mathbb{P}\text{-a.s.}$$

Hence Z_T is measurable with respect to the path $\mathbf{Y}$ alone. Therefore there exists a measurable functional $\mathcal{Z}$ on $C([0,T], \mathbb{R}^{2n})$ such that

$$Z_T = \mathcal{Z}(\mathbf{Y}), \qquad \mathbb{P}\text{-a.s.}$$

Now, let F be any bounded measurable functional on path space. Then

$$\mathbb{E}_{\mathbb{P}_{\mathbf{X}}}[F] = \mathbb{E}_{\widetilde{\mathbb{P}}}[F(\mathbf{Y})] = \mathbb{E}_{\mathbb{P}}[Z_T F(\mathbf{Y})] = \mathbb{E}_{\mathbb{P}}[\mathcal{Z}(\mathbf{Y})F(\mathbf{Y})] = \mathbb{E}_{\mathbb{P}_{\mathbf{Y}}}[\mathcal{Z}F].$$

Since this holds for every bounded measurable F, we conclude that

$$\frac{d\mathbb{P}_{\mathbf{X}}}{d\mathbb{P}_{\mathbf{Y}}} = \mathcal{Z}, \quad \frac{d\mathbb{P}_{\mathbf{Y}}}{d\mathbb{P}_{\mathbf{X}}} = \mathcal{Z}^{-1}.$$

Therefore,

$$\mathrm{KL}(\mathbb{P}_{\mathbf{Y}}\|\mathbb{P}_{\mathbf{X}}) = \mathbb{E}_{\mathbb{P}_{\mathbf{Y}}}\left[\log\frac{d\mathbb{P}_{\mathbf{Y}}}{d\mathbb{P}_{\mathbf{X}}}\right] = -\mathbb{E}_{\mathbb{P}_{\mathbf{Y}}}[\log\mathcal{Z}] = -\mathbb{E}_{\mathbb{P}}[\log Z_T]$$
$$= -\mathbb{E}_{\mathbb{P}}\left[\int_0^T u_\theta(t,\mathbf{Y}_t)^\top d\mathbf{W}_t\right] + \frac{1}{2}\mathbb{E}_{\mathbb{P}}\int_0^T \|u_\theta(t,\mathbf{Y}_t)\|^2 dt$$

It remains to show that the stochastic integral has zero mean. From Novikov's condition we have

$$\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{1}{2}\int_0^T \|u_\theta(t,\mathbf{Y}_t)\|^2 dt\right)\right] < \infty.$$

Since $x \le 2e^{x/2}$ for every $x \ge 0$, this implies

$$\mathbb{E}_{\mathbb{P}}\int_0^T \|u_\theta(t,\mathbf{Y}_t)\|^2 dt < \infty.$$

Therefore the Itô integral $\int_0^T u_\theta(t,\mathbf{Y}_t)^\top d\mathbf{W}_t$ is square-integrable, hence a martingale with mean zero:

$$\mathbb{E}_{\mathbb{P}}\left[\int_0^T u_\theta(t,\mathbf{Y}_t)^\top d\mathbf{W}_t\right] = 0.$$

Consequently,

$$\mathrm{KL}(\mathbb{P}_{\mathbf{Y}}\|\mathbb{P}_{\mathbf{X}}) = \frac{1}{2}\mathbb{E}_{\mathbb{P}}\int_0^T \|u_\theta(t,\mathbf{Y}_t)\|^2 dt = \frac{1}{2}\mathbb{E}_{\mathbb{P}_{\mathbf{Y}}}\int_0^T \|u_\theta(t,\mathbf{Y}_t)\|^2 dt$$

Since $\mathrm{Law}_{\mathbb{P}}(\mathbf{Y}) = \mathbb{P}_{\mathbf{Y}}$. For the equivalent expression, substitute the explicit formula for u_θ:

$$\|u_\theta(t,\mathbf{Y}_t)\|^2 = \left\|\begin{pmatrix}0 & \sqrt{\frac{\beta}{2\gamma}}I_n\end{pmatrix}\left(\begin{pmatrix}\mathbf{M}^{-1}\mathbf{p}_t\\ -\nabla \mathrm{U}(\mathbf{q}_t) - \gamma\mathbf{p}_t\end{pmatrix} - V_{\xi(t)}(\mathbf{Y}_t) - \frac{1}{2}\sum_{i=1}^n D\big(V_{\Sigma_i(t)}\big)(\mathbf{Y}_t)V_{\Sigma_i(t)}(\mathbf{Y}_t)\right)\right\|^2$$
$$= \frac{\beta}{2\gamma}\left\|\begin{pmatrix}0 & I_n\end{pmatrix}\left(\begin{pmatrix}\mathbf{M}^{-1}\mathbf{p}_t\\ -\nabla \mathrm{U}(\mathbf{q}_t) - \gamma\mathbf{p}_t\end{pmatrix} - V_{\xi(t)}(\mathbf{Y}_t) - \frac{1}{2}\sum_{i=1}^n D\big(V_{\Sigma_i(t)}\big)(\mathbf{Y}_t)V_{\Sigma_i(t)}(\mathbf{Y}_t)\right)\right\|^2.$$

Multiplying by the prefactor $1/2$ yields

$$\mathrm{KL}(\mathbb{P}_{\mathbf{Y}}\|\mathbb{P}_{\mathbf{X}})$$
$$= \frac{\beta}{4\gamma}\mathbb{E}_{\mathbb{P}_{\mathbf{Y}}}\int_0^T \left\|\begin{pmatrix}0 & I_n\end{pmatrix}\left(\begin{pmatrix}\mathbf{M}^{-1}\mathbf{p}_t\\ -\nabla \mathrm{U}(\mathbf{q}_t) - \gamma\mathbf{p}_t\end{pmatrix} - V_{\xi(t)}(\mathbf{Y}_t) - \frac{1}{2}\sum_{i=1}^n D\big(V_{\Sigma_i(t)}\big)(\mathbf{Y}_t)V_{\Sigma_i(t)}(\mathbf{Y}_t)\right)\right\|^2 dt,$$

which proves the claim. □

Proposition A.5. *For a homogeneous basis $\{E_\alpha\}_{\alpha\in\mathcal{A}_{\le K}}$ of $\mathcal{L}_{\le K}$, and its dual basis $\{\varepsilon^\alpha\}_{\alpha\in\mathcal{A}_{\le K}}$, there exists a stochastic process h_t such that*

$$\Omega_m^{\mathfrak{g},(K)} = \sum_{c=1}^{d} \Omega_{m,c}^{(K)} e_c, \quad \Omega_{m,c}^{(K)} = \sum_{\gamma,\beta\in\mathcal{A}_{\le K}} \mathsf{A}_\gamma^{\ c} \int_{I_m} \mathsf{D}^\gamma{}_\beta(h_t) \circ d\zeta_t^\beta,$$

where B_n denotes Bernoulli numbers, and one can write

$$\mathsf{D}^\gamma{}_\beta(h_t) = \delta_\beta^\gamma + \sum_{n=1}^{K-1} \frac{B_n}{n!} \sum_{\alpha_1,\dots,\alpha_n\in\mathcal{A}_{\le K}} \sum_{\nu_1,\dots,\nu_{n-1}\in\mathcal{A}_{\le K}} h_t^{\alpha_1}\cdots h_t^{\alpha_n} C_{\alpha_n\beta}^{\ \ \nu_1} C_{\alpha_{n-1}\nu_1}^{\ \ \nu_2} \cdots C_{\alpha_1\nu_{n-1}}^{\ \ \gamma},$$

and

$$\circ d\zeta_t^\beta = \sum_{a=1}^{d} \left[\delta_a^\beta + \sum_{n=1}^{K-1} \frac{1}{n!} \sum_{\alpha_1,\dots,\alpha_n} \sum_{\nu_1,\dots,\nu_{n-1}} h_t^{\alpha_1}\cdots h_t^{\alpha_n} \mathsf{C}_{\alpha_n a}{}^{\nu_1} \mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2} \cdots \mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\beta} \right] \circ dg_t^a.$$

Proof. For each $r = 1,\dots,K$, let us define notations as $d_r := \dim \mathcal{L}^{[r]}(\mathfrak{g})$ and $\mathcal{A}^{[r]} := \{1,\dots,d_r\}$. Then, the finite index set for fixed K is defined as $\mathcal{A}_{\le K} := \bigsqcup_{r=1}^K \mathcal{A}^{[r]}$. We choose a basis $\{E_\alpha\}_{\alpha\in\mathcal{A}^{[r]}} \subset \mathcal{L}^{[r]}(\mathfrak{g})$ so that it is a homogeneous basis of $\mathcal{L}_{\le K} = \bigoplus_{r=1}^K \mathcal{L}^{[r]}(\mathfrak{g})$. Specifically, the base space of free Lie algebra can be defined as

$$\mathcal{L}_{\le K} := \pi_{\le K}\widehat{\mathcal{L}}(\mathfrak{g}) = \bigoplus_{r=1}^{K} \mathcal{L}^{[r]}(\mathfrak{g}) = \left\{ \sum_{r=1}^{K} X^{[r]} \;\middle|\; X^{[r]} \in \mathcal{L}^{[r]}(\mathfrak{g}) \right\} \subset \bigoplus_{r=1}^{K} \mathfrak{g}^{\otimes r}.$$

For instance, in first degree , *i.e.*, $K = 1$, one may identify $\mathcal{A}^{[1]} = \{1,\dots,d\}$ and denote $E_a = e_a$, where $\{\varepsilon^\alpha\}_{\alpha\in\mathcal{A}_{\le K}} \subset \mathcal{L}^*_{\le K}$ is the dual basis of $\{E_\alpha\}_{\alpha\in\mathcal{A}_{\le K}}$, so that $\varepsilon^\alpha(E_\beta) = \delta_\beta^\alpha$. Next, we introduce the scalar-valued coordinate coefficients of basis as

$$\mathsf{C}_{\alpha\beta}{}^{\gamma} := \varepsilon^\gamma([E_\alpha, E_\beta]) \in \mathbb{R}, \qquad \mathsf{A}_\gamma{}^{c} := e^c\big(\mathrm{ev}_{\mathfrak{g}}(E_\gamma)\big) \in \mathbb{R}.$$

Equivalently, one may rewrite above relation as follows:

$$[E_\alpha, E_\beta] = \sum_{\gamma\in\mathcal{A}_{\le K}} \mathsf{C}_{\alpha\beta}{}^{\gamma} E_\gamma, \qquad \mathrm{ev}_{\mathfrak{g}}(E_\gamma) = \sum_{c=1}^{d} \mathsf{A}_\gamma{}^{c} e_c.$$

As a next step, we aim to explicate the local-coordinate expressions of two processes, H_t and $\circ d\mathrm{Sig}(\cdot)(\mathrm{Sig}(\cdot))^{-1}$. To be more specific, with the coordinate functions (*i.e.*, basis) introduced above, we define the local-coordinate processes:

$$h_t^\alpha := \varepsilon^\alpha\big(\pi_{\le K}\mathrm{H}_t\big), \qquad \circ d\zeta_t^\beta := \varepsilon^\beta\left(\pi_{\le K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right).$$

Equivalently, one can rewrite as follows:

$$\pi_{\le K}\mathrm{H}_t = \sum_{\alpha\in\mathcal{A}_{\le K}} h_t^\alpha \mathrm{E}_\alpha, \quad \pi_{\le K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right) = \sum_{\beta\in\mathcal{A}_{\le K}} (\circ d\zeta_t^\beta)\mathrm{E}_\beta$$

Differentiating the upper limit t of each iterated Stratonovich integral, we obtain

$$\begin{aligned}
\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g}) &= \sum_{k=1}^{\infty} \sum_{a_1,\dots,a_k=1}^{d} \left(\int_{t_{m-1}<u_1<\cdots<u_{k-1}\le t} \circ dg_{u_1}^{a_1}\cdots\circ dg_{u_{k-1}}^{a_{k-1}} \right) (\mathrm{E}_{a_1}\otimes\cdots\otimes\mathrm{E}_{a_k}) \circ dg_t^{a_k} \\
&= \left(1 + \sum_{r=1}^{\infty} \sum_{a_1,\dots,a_r=1}^{d} \left(\int_{t_{m-1}<u_1<\cdots<u_r\le t} \circ dg_{u_1}^{a_1}\cdots\circ dg_{u_r}^{a_r} \right) \mathrm{E}_{a_1}\otimes\cdots\otimes\mathrm{E}_{a_r} \right) \circ d\mathbf{g}_t \\
&= \mathrm{Sig}_t^{(m)}(\mathbf{g}) \circ d\mathbf{g}_t.
\end{aligned}$$

Therefore,

$$\circ d\zeta_t^{\beta} = \varepsilon^{\beta}\left(\pi_{\le K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right) = \varepsilon^{\beta}\left(\pi_{\le K}\left(\mathrm{Sig}_t^{(m)}(\mathbf{g}) \circ d\mathbf{g}_t\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right).$$

Now, we use the logarithmic coordinate $\mathrm{H}_t = \log(\mathrm{Sig}_t^{(m)}(\mathbf{g}))$, so that the following relation holds: $\pi_{\le K}\mathrm{H}_t = \sum_{\alpha\in\mathcal{A}_{\le K}} h_t^{\alpha}\mathrm{E}_{\alpha}$. Then,

$$\mathrm{Sig}_t^{(m)}(\mathbf{g}) \circ d\mathbf{g}_t\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1} = \mathrm{Ad}_{\exp(\mathrm{H}_t)}(\circ d\mathbf{g}_t) = e^{\mathrm{ad}_{\mathrm{H}_t}}(\circ d\mathbf{g}_t).$$

Hence, after projection to degree less than $\le K$, the inner term of coordinate expression in $\circ d\zeta_t$ can be simplified as follows:

$$\pi_{\le K}\left(\mathrm{Sig}_t^{(m)}(\mathbf{g}) \circ d\mathbf{g}_t\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right) = \pi_{\le K}\left(\sum_{n=0}^{\infty}\frac{1}{n!}\,\mathrm{ad}_{\mathrm{H}_t}^n(\circ d\mathbf{g}_t)\right) = \sum_{n=0}^{K-1}\frac{1}{n!}\,\mathrm{ad}_{\pi_{\le K}\mathrm{H}_t}^n(\circ d\mathbf{g}_t),$$

Since $\circ d\mathbf{g}_t$ has degree 1, and therefore $\mathrm{ad}^n_{\pi_{\le K}H_t}(\circ d\mathbf{g}_t)$ has degree at least $n+1$, so all terms with $n \ge K$ vanish under $\pi_{\le K}$. Following by the definition of adjoint operator, one can show that for every $n \ge 1$,

$$\mathrm{ad}^n_{\pi_{\le K}H_t}(\circ d\mathbf{g}_t) = \sum_{a=1}^{d}\sum_{\alpha_1,\dots,\alpha_n}\sum_{\nu_1,\dots,\nu_{n-1},\beta} h_t^{\alpha_1}\cdots h_t^{\alpha_n}\mathsf{C}_{\alpha_n a}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\beta}\mathrm{E}_{\beta}\circ dg_t^a.$$

Applying coordinate function ε^{β} to above expression gives the first argument:

$$\circ d\zeta_t^{\beta} = \sum_{a=1}^{d}\left[\delta_a^{\beta} + \sum_{n=1}^{K-1}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n}\sum_{\nu_1,\dots,\nu_{n-1}} h_t^{\alpha_1}\cdots h_t^{\alpha_n}\mathsf{C}_{\alpha_n a}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\beta}\right]\circ dg_t^a.$$

By definition, the proposed local logarithmic increment is given as

$$\Omega_m^{\mathfrak{g},(K)} = \mathrm{ev}_{\mathfrak{g}}\left(\pi_{\le K}\int_{I_m}\mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right).$$

Since $\pi_{\le K}$, $\mathrm{ev}_{\mathfrak{g}}$, and the time integral are linear on the finite-dimensional space $\mathcal{L}_{\le K}$, it suffices to compute the integrand in the basis $\{\mathrm{E}_{\alpha}\}_{\alpha\in\mathcal{A}_{\le K}}$. Next, we apply the Bernoulli-series formula for dexp^{-1} to obtain the degree-$\le K$ part of the logarithmic differential:

$$\begin{aligned}
&\pi_{\le K}\,\mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\\
&\quad= \pi_{\le K}\,\mathrm{dexp}^{-1}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\\
&\quad= \pi_{\le K}\left(\sum_{n=0}^{\infty}\frac{B_n}{n!}\,\mathrm{ad}_{\mathrm{H}_t}^n\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right)\\
&\quad= \sum_{n=0}^{\infty}\frac{B_n}{n!}\pi_{\le K}\left(\mathrm{ad}_{\mathrm{H}_t}^n\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right)\\
&\quad= \sum_{n=0}^{\infty}\frac{B_n}{n!}\pi_{\le K}\Big(\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t+\pi_{>K}\mathrm{H}_t}\Big(\pi_{\le K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\\
&\qquad\qquad+\pi_{>K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\Big)\Big)\\
&\quad= \sum_{n=0}^{\infty}\frac{B_n}{n!}\pi_{\le K}\left(\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t}\left(\pi_{\le K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right)\right)\\
&\quad= \sum_{n=0}^{K-1}\frac{B_n}{n!}\,\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t}\left(\pi_{\le K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right).
\end{aligned}$$

Here the sum truncates at $n = K-1$. Indeed, each E_β has degree at least 1, so $\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t}(\mathrm{E}_\beta)$ has degree at least $n+1$. Consequently, its projection onto $\mathcal{L}_{\le K}$ vanishes for every $n \ge K$. Following the result, for each $n \ge 1$,

$$\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t}\left(\sum_\beta (\circ d\zeta_t^\beta)\mathrm{E}_\beta\right) = \sum_{\alpha_1,\dots,\alpha_n,\beta}\ \sum_{\nu_1,\dots,\nu_{n-1},\gamma} h_t^{\alpha_1}\cdots h_t^{\alpha_n}(\circ d\zeta_t^\beta) \cdot \mathsf{C}_{\alpha_n\beta}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\gamma}\mathrm{E}_\gamma.$$

Substituting the above adjoint expansions into the Bernoulli series and collecting the coefficients of the basis vectors E_γ, we compute

$$\begin{aligned}
&\pi_{\le K}\,\mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\\
&\quad= \sum_{n=0}^{K-1}\frac{B_n}{n!}\,\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t}\left(\sum_{\beta\in\mathcal{A}_{\le K}}(\circ d\zeta_t^\beta)\mathrm{E}_\beta\right)\\
&\quad= \mathrm{ad}^0_{\pi_{\le K}\mathrm{H}_t}\left(\sum_{\beta\in\mathcal{A}_{\le K}}(\circ d\zeta_t^\beta)\mathrm{E}_\beta\right) + \sum_{n=1}^{K-1}\frac{B_n}{n!}\,\mathrm{ad}^n_{\pi_{\le K}\mathrm{H}_t}\left(\sum_{\beta\in\mathcal{A}_{\le K}}(\circ d\zeta_t^\beta)\mathrm{E}_\beta\right)\\
&\quad= \sum_{\beta\in\mathcal{A}_{\le K}}(\circ d\zeta_t^\beta)\mathrm{E}_\beta + \sum_{n=1}^{K-1}\frac{B_n}{n!}\sum_{\alpha_1,\dots,\alpha_n\in\mathcal{A}_{\le K}}\sum_{\beta\in\mathcal{A}_{\le K}}\sum_{\nu_1,\dots,\nu_{n-1}\in\mathcal{A}_{\le K}}\sum_{\gamma\in\mathcal{A}_{\le K}} h_t^{\alpha_1}\cdots h_t^{\alpha_n}(\circ d\zeta_t^\beta)\\
&\qquad\qquad\cdot \mathsf{C}_{\alpha_n\beta}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\gamma}\mathrm{E}_\gamma\\
&\quad= \sum_{\gamma\in\mathcal{A}_{\le K}}\left(\sum_{\beta\in\mathcal{A}_{\le K}}\delta_\beta^\gamma\circ d\zeta_t^\beta\right)\mathrm{E}_\gamma\\
&\qquad+\sum_{\gamma\in\mathcal{A}_{\le K}}\left(\sum_{\beta\in\mathcal{A}_{\le K}}\sum_{n=1}^{K-1}\frac{B_n}{n!}\sum_{\alpha_1,\dots,\alpha_n\in\mathcal{A}_{\le K}}\sum_{\nu_1,\dots,\nu_{n-1}\in\mathcal{A}_{\le K}} h_t^{\alpha_1}\cdots h_t^{\alpha_n}\right.\\
&\qquad\qquad\left.\cdot \mathsf{C}_{\alpha_n\beta}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\gamma}\circ d\zeta_t^\beta\right)\mathrm{E}_\gamma\\
&\quad= \sum_{\gamma\in\mathcal{A}_{\le K}}\left(\sum_{\beta\in\mathcal{A}_{\le K}}\left[\delta_\beta^\gamma + \sum_{n=1}^{K-1}\frac{B_n}{n!}\sum_{\alpha_1,\dots,\alpha_n\in\mathcal{A}_{\le K}}\sum_{\nu_1,\dots,\nu_{n-1}\in\mathcal{A}_{\le K}} h_t^{\alpha_1}\cdots h_t^{\alpha_n}\right.\right.\\
&\qquad\qquad\left.\left.\cdot \mathsf{C}_{\alpha_n\beta}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\gamma}\right]\circ d\zeta_t^\beta\right)\mathrm{E}_\gamma\\
&\quad= \sum_{\gamma\in\mathcal{A}_{\le K}}\left(\sum_{\beta\in\mathcal{A}_{\le K}}\mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)\mathrm{E}_\gamma,
\end{aligned}$$

where each term of coefficient function D can be computed as

$$\mathsf{D}^\gamma{}_\beta(h_t) = \delta_\beta^\gamma + \sum_{n=1}^{K-1}\frac{B_n}{n!}\sum_{\alpha_1,\dots,\alpha_n\in\mathcal{A}_{\le K}}\sum_{\nu_1,\dots,\nu_{n-1}\in\mathcal{A}_{\le K}} h_t^{\alpha_1}\cdots h_t^{\alpha_n}\mathsf{C}_{\alpha_n\beta}{}^{\nu_1}\mathsf{C}_{\alpha_{n-1}\nu_1}{}^{\nu_2}\cdots\mathsf{C}_{\alpha_1\nu_{n-1}}{}^{\gamma}.$$

Our next aim is to apply the evaluation map $\mathrm{ev}_{\mathfrak{g}}$, which was defined as $\mathrm{ev}_{\mathfrak{g}}(\mathrm{E}_\gamma) = \sum_{c=1}^{d} \mathsf{A}_\gamma{}^c e_c$. By using linearity, we obtain

$$\begin{aligned}
\mathrm{ev}_{\mathfrak{g}}\left(\pi_{\leq K}\,\mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right) &= \mathrm{ev}_{\mathfrak{g}}\left(\sum_\gamma\left(\sum_\beta \mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)\mathrm{E}_\gamma\right)\\
&= \sum_\gamma\left(\sum_\beta \mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)\mathrm{ev}_{\mathfrak{g}}(\mathrm{E}_\gamma)\\
&= \sum_\gamma\left(\sum_\beta \mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)\left(\sum_{c=1}^{d}\mathsf{A}_\gamma{}^c e_c\right)\\
&= \sum_{c=1}^{d}\left(\sum_{\gamma,\beta}\mathsf{A}_\gamma{}^c\mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)e_c.
\end{aligned}$$

Finally, integrating over I_m gives

$$\begin{aligned}
\Omega_m^{\mathfrak{g},(K)} &= \int_{I_m}\mathrm{ev}_{\mathfrak{g}}\left(\pi_{\leq K}\,\mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right)\\
&= \int_{I_m}\sum_{c=1}^{d}\left(\sum_{\gamma,\beta}\mathsf{A}_\gamma{}^c\mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)e_c\\
&= \sum_{c=1}^{d}\left(\sum_{\gamma,\beta}\mathsf{A}_\gamma{}^c\int_{I_m}\mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta\right)e_c.
\end{aligned}$$

Therefore,

$$\Omega_m^{\mathfrak{g},(K)} = \sum_{c=1}^{d}\Omega_{m,c}^{(K)}e_c, \qquad \Omega_{m,c}^{(K)} = \sum_{\gamma,\beta}\mathsf{A}_\gamma{}^c\int_{I_m}\mathsf{D}^\gamma{}_\beta(h_t)\circ d\zeta_t^\beta,$$

which is exactly the claimed coordinate formula. □

Proposition A.6. *Let us specify the Lie algebra as $\mathfrak{g} = \mathfrak{csp}(2n) \ltimes \mathbb{R}^{2n}$, where $r := \dim \mathfrak{csp}(2n)$. Then, there exist local coordinate processes $\{\circ d\Psi_{n,t}^k\}_{k=1}^r$ and $\{\circ d\Theta_{n,t}^q\}_{q=1}^{2n}$, determined by the semidirect-product adjoint recursion, such that*

$$\Omega_{m,k}^{(K)} = \int_{I_m}\sum_{n=0}^{K-1}\frac{B_n}{n!}\circ d\Psi_{n,t}^k, \qquad \Omega_{m,r+q}^{(K)} = \int_{I_m}\sum_{n=0}^{K-1}\frac{B_n}{n!}\circ d\Theta_{n,t}^q,$$

for $k = 1,\dots,r$ and $q = 1,\dots,2n$.

Proof. Fix a basis $\{F_i\}_{i=1}^r$ of $\mathfrak{csp}(2n)$ and let $\{\mathbf{e}_\ell\}_{\ell=1}^{2n}$ be the standard basis of $\mathbb{R}^{2n}$. Adapted to the direct sum splitting $\mathfrak{g} = \mathfrak{csp}(2n)\oplus\mathbb{R}^{2n}$, one may rewrite the coordinate expression of the degree-one basis $e_i = (F_i, 0), e_{r+\ell} = (0, \mathbf{e}_\ell)$. Let us first define the local expression of Lie-bracket as follows:

$$[F_i, F_j] = \sum_{k=1}^{r}\mathsf{f}_{ij}{}^k F_k, \qquad F_i\mathbf{e}_\ell = \sum_{q=1}^{2n}\rho_{i\ell}{}^q\mathbf{e}_q.$$

Similarly, the coordinate expression of H_t can be further specified as

$$\mathrm{ev}_{\mathfrak{g}}(\pi_{\leq K}\mathrm{H}_t) = (L_t, a_t), \qquad L_t = \sum_{i=1}^{r}\mathfrak{l}_t^i F_i, \qquad a_t = \sum_{\ell=1}^{2n}\mathfrak{a}_t^\ell\mathbf{e}_\ell,$$

In this local coordinate, one may rewrite

$$\mathrm{ev}_{\mathfrak{g}}\left(\pi_{\leq K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right) = \left[\circ dR_t = \sum_{j=1}^{r}(\circ d\mathfrak{r}_t^j)F_j, \circ ds_t = \sum_{\ell=1}^{2n}(\circ d\mathfrak{s}_t^\ell)\mathbf{e}_\ell\right]$$

For simplicity, let us introduce the variables

$$\widetilde{H}_t := \mathrm{ev}_{\mathfrak{g}}(\pi_{\leq K}\mathrm{H}_t) = (L_t, a_t),$$
$$\widetilde{Z}_t := \mathrm{ev}_{\mathfrak{g}}\left(\pi_{\leq K}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right) = (\circ dR_t, \circ ds_t).$$

By the identical Bernoulli-adjoint expansion shown in Proposition A.5, we have

$$\mathrm{ev}_{\mathfrak{g}}\left(\pi_{\leq K}\,\mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}_t^{(m)}(\mathbf{g})\big(\mathrm{Sig}_t^{(m)}(\mathbf{g})\big)^{-1}\right)\right) = \sum_{n=0}^{K-1}\frac{B_n}{n!}\,\mathrm{ad}^n_{\widetilde{H}_t}(\widetilde{Z}_t).$$

Hence it is enough to compute the coordinate-based expression of $\mathrm{ad}^n_{(L_t,a_t)}(\circ dR_t, \circ ds_t)$ with respect to the basis adapted to $\mathfrak{csp}(2n) \ltimes \mathbb{R}^{2n}$. Next, we introduce the group law and the induced Lie bracket as follows:

$$(A, a)\cdot(B, b) = (AB, Ab + a), \quad [(X, u), (Y, v)] = ([X, Y], Xv - Yu)$$

Our next aim is to provide a local-coordinate expression of this Lie bracket. To this end, let $(X, u), (Y, v) \in \mathfrak{csp}(2n) \ltimes \mathbb{R}^{2n}$, with $X, Y \in \mathfrak{csp}(2n)$ and $u, v \in \mathbb{R}^{2n}$, and write

$$X = \sum_{i=1}^{r}\mathfrak{x}^i F_i, \qquad u = \sum_{\ell=1}^{2n}\mathfrak{u}^\ell \mathbf{e}_\ell, \qquad Y = \sum_{j=1}^{r}\mathfrak{y}^j F_j, \qquad v = \sum_{\ell=1}^{2n}\mathfrak{v}^\ell \mathbf{e}_\ell.$$

Then,

$$\begin{aligned}[X, Y] &= \left[\sum_{i=1}^{r}\mathfrak{x}^i F_i, \sum_{j=1}^{r}\mathfrak{y}^j F_j\right] = \sum_{i,j}\mathfrak{x}^i\mathfrak{y}^j[F_i, F_j] \\ &= \sum_{i,j}\mathfrak{x}^i\mathfrak{y}^j\left(\sum_{k=1}^{r} f_{ij}{}^k F_k\right) = \sum_{k=1}^{r}\left(\sum_{i,j}\mathfrak{x}^i\mathfrak{y}^j f_{ij}{}^k\right)F_k,\end{aligned}$$

while

$$Xv = \left(\sum_{i=1}^{r}\mathfrak{x}^i F_i\right)\left(\sum_{\ell=1}^{2n}\mathfrak{v}^\ell\mathbf{e}_\ell\right) = \sum_{i,\ell}\mathfrak{x}^i\mathfrak{v}^\ell F_i\mathbf{e}_\ell = \sum_{i,\ell}\mathfrak{x}^i\mathfrak{v}^\ell\left(\sum_{q=1}^{2n}\rho_{i\ell}{}^q\mathbf{e}_q\right) = \sum_{q=1}^{2n}\left(\sum_{i,\ell}\mathfrak{x}^i\mathfrak{v}^\ell\rho_{i\ell}{}^q\right)\mathbf{e}_q,$$

and similarly

$$Yu = \left(\sum_{j=1}^{r}\mathfrak{y}^j F_j\right)\left(\sum_{\ell=1}^{2n}\mathfrak{u}^\ell\mathbf{e}_\ell\right) = \sum_{j,\ell}\mathfrak{y}^j\mathfrak{u}^\ell F_j\mathbf{e}_\ell = \sum_{j,\ell}\mathfrak{y}^j\mathfrak{u}^\ell\left(\sum_{q=1}^{2n}\rho_{j\ell}{}^q\mathbf{e}_q\right) = \sum_{q=1}^{2n}\left(\sum_{j,\ell}\mathfrak{y}^j\mathfrak{u}^\ell\rho_{j\ell}{}^q\right)\mathbf{e}_q.$$

Therefore, the Lie bracket can be summarized as

$$[(X, u), (Y, v)] = \left(\sum_{k=1}^{r}\left(\sum_{i,j}\mathfrak{x}^i\mathfrak{y}^j f_{ij}{}^k\right)F_k,\ \sum_{q=1}^{2n}\left(\sum_{i,\ell}\mathfrak{x}^i\mathfrak{v}^\ell\rho_{i\ell}{}^q - \sum_{j,\ell}\mathfrak{y}^j\mathfrak{u}^\ell\rho_{j\ell}{}^q\right)\mathbf{e}_q\right).$$

Next, we define processes $\circ d\Psi^k_{0,t} := \circ d\mathfrak{r}^k_t$ and $\circ d\Theta^q_{0,t} := \circ d\mathfrak{s}^q_t$ such that

$$\mathrm{ad}^n_{\widetilde{H}_t}(\widetilde{Z}_t) = \left(\sum_{j=1}^{r}\circ d\Psi^j_{n,t}F_j,\ \sum_{\ell=1}^{2n}\circ d\Theta^\ell_{n,t}\mathbf{e}_\ell\right).$$

Then,

$$\begin{aligned}\mathrm{ad}^{n+1}_{\widetilde{H}_t}(\widetilde{Z}_t) &= [(L_t, a_t), \mathrm{ad}^{n}_{\widetilde{H}_t}(\widetilde{Z}_t)] \\ &= \left[\left(\sum_{i=1}^{r} \mathfrak{l}_t^i F_i, \sum_{\ell=1}^{2n} \mathfrak{a}_t^\ell \mathbf{e}_\ell\right), \left(\sum_{j=1}^{r} \circ d\Psi^j_{n,t} F_j, \sum_{\ell=1}^{2n} \circ d\Theta^\ell_{n,t} \mathbf{e}_\ell\right)\right] \\ &= \left(\sum_{k=1}^{r}\left(\sum_{i,j=1}^{r} \mathfrak{l}_t^i \circ d\Psi^j_{n,t} {f_{ij}}^k\right) F_k,\ \sum_{q=1}^{2n}\left(\sum_{i=1}^{r}\sum_{\ell=1}^{2n} \mathfrak{l}_t^i \circ d\Theta^\ell_{n,t} {\rho_{i\ell}}^q - \sum_{j=1}^{r}\sum_{\ell=1}^{2n} \mathfrak{a}_t^\ell \circ d\Psi^j_{n,t} {\rho_{j\ell}}^q\right)\mathbf{e}_q\right).\end{aligned}$$

Hence the recursion is

$$\circ d\Psi^k_{n+1,t} = \sum_{i,j=1}^{r} \mathfrak{l}_t^i \circ d\Psi^j_{n,t} {f_{ij}}^k, \qquad k = 1, \dots, r,$$

and

$$\circ d\Theta^q_{n+1,t} = \sum_{i=1}^{r}\sum_{\ell=1}^{2n} \mathfrak{l}_t^i \circ d\Theta^\ell_{n,t} {\rho_{i\ell}}^q - \sum_{j=1}^{r}\sum_{\ell=1}^{2n} \mathfrak{a}_t^\ell \circ d\Psi^j_{n,t} {\rho_{j\ell}}^q, \qquad q = 1, \dots, 2n.$$

Therefore, for every $n \geq 0$,

$$\mathrm{ad}^{n}_{\widetilde{H}_t}(\widetilde{Z}_t) = \left(\sum_{k=1}^{r} \circ d\Psi^k_{n,t} F_k,\ \sum_{q=1}^{2n} \circ d\Theta^q_{n,t} \mathbf{e}_q\right).$$

Substituting this into the Bernoulli expansion gives

$$\begin{aligned}\mathrm{ev}_{\mathfrak{g}}\left(\pi_{\leq K}\, \mathrm{dLog}_{\mathrm{H}_t}\left(\circ d\mathrm{Sig}^{(m)}_t(\mathbf{g})\big(\mathrm{Sig}^{(m)}_t(\mathbf{g})\big)^{-1}\right)\right) &= \sum_{n=0}^{K-1} \frac{B_n}{n!}\left(\sum_{k=1}^{r} \circ d\Psi^k_{n,t} F_k,\ \sum_{q=1}^{2n} \circ d\Theta^q_{n,t} \mathbf{e}_q\right) \\ &= \left(\sum_{k=1}^{r}\sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Psi^k_{n,t} F_k,\ \sum_{q=1}^{2n}\sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Theta^q_{n,t} \mathbf{e}_q\right).\end{aligned}$$

Finally, integrating over I_m yields

$$\Omega^{\mathfrak{g},(K)}_m = \left(\sum_{k=1}^{r}\left(\int_{I_m} \sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Psi^k_{n,t}\right) F_k,\ \sum_{q=1}^{2n}\left(\int_{I_m} \sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Theta^q_{n,t}\right)\mathbf{e}_q\right).$$

Thus, one may find the expression $\Omega^{\mathfrak{g},(K)}_m = (\Lambda^{(K)}_m, b^{(K)}_m)$ such that

$$\Lambda^{(K)}_m = \sum_{k=1}^{r}\left(\int_{I_m} \sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Psi^k_{n,t}\right) F_k, \qquad b^{(K)}_m = \sum_{q=1}^{2n}\left(\int_{I_m} \sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Theta^q_{n,t}\right)\mathbf{e}_q.$$

Reading coefficients in the degree-one basis $e_i = (F_i, 0)$ and $e_{r+q} = (0, \mathbf{e}_q)$, we prove the claim by showing that

$$\Omega^{(K)}_{m,k} = \int_{I_m} \sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Psi^k_{n,t}, \qquad \Omega^{(K)}_{m,r+q} = \int_{I_m} \sum_{n=0}^{K-1} \frac{B_n}{n!} \circ d\Theta^q_{n,t}.$$

□