Learning Continuous Temporal Dynamics on Symplectic Manifolds for Temporal Knowledge Graph Embedding

Jiang Li, Zehua Duo, Tian Lan, Feilong Bao, Guanglai Gao, Xiangdong Su*
[1] College of Computer Science, Inner Mongolia University, China
[2] National & Local Joint Engineering Research Center of Intelligent Information Processing Technology for Mongolian, China
[3] Inner Mongolia Key Laboratory of Multilingual Artificial Intelligence Technology, China
lijiangimu@gmail.com,cssxd@imu.edu.cn

Abstract

Temporal knowledge graph embedding (TKGE) aims to model the temporal evolution of relational facts. However, existing approaches predominantly rely on discrete timestamp lookup tables and high-dimensional embedding spaces, which lack explicit structural constraints for continuous-time dynamics. As a result, temporal patterns are often captured through capacity scaling rather than principled dynamic modeling, leading to limited parameter efficiency and scalability. To address these limitations, we propose *TDSym*, a physics-inspired framework that embeds temporal dynamics into a symplectic phase space. Our model introduces a structure-preserving Hamiltonian evolution mechanism based on a pairwise-decoupled Hamiltonian generator and its Cayley transform, ensuring that temporal transformations adhere to the symplectic group $\mathrm{Sp}(2d)$ and preserve phase-space volume with linear computational complexity. In addition, we design a Time-Aware Parameter Modulation mechanism that integrates continuous Rotary Time Embeddings via Feature-wise Linear Modulation, enabling smooth temporal evolution while capturing event-driven variations. Theoretical analysis establishes the geometric validity of the proposed framework. Extensive experiments on standard TKGE benchmarks demonstrate that *TDSym* achieves competitive performance with substantially lower embedding dimensions. Furthermore, empirical results show that the proposed continuous Hamiltonian evolution facilitates generalization to unseen timestamps by learning transferable temporal dynamics from the underlying geometric structure[1].

1 Introduction

Temporal Knowledge Graphs (TKGs) extend static knowledge graphs by associating facts with temporal information, enabling structured reasoning over time-evolving knowledge. By representing facts as quadruples (h, r, t, τ), where τ denotes a timestamp or a temporal interval, TKGs explicitly incorporate temporal semantics into relational representations, allowing models to capture the evolution of fact interactions across time. Such temporal representations are fundamental to a wide range of real-world applications, including event prediction (Xu et al., 2023; Tang et al., 2024; Tian et al., 2025), decision support (Chen et al., 2024; Huang and Zaslavsky, 2024; Wang et al., 2024), and recommendation systems (Mezni, 2021; Zhao et al., 2022; Hu et al., 2024; Chen et al., 2025), as relational semantics naturally evolve in dynamic environments. However, real-world TKGs are typically incomplete, with missing facts and sparse connections that limit their utility. To address this issue, Temporal Knowledge Graph Embedding (TKGE) has been studied as a means to learn low-dimensional representations of entities and relations for inferring missing links in TKGs.

To address these challenges, the research community has developed diverse TKGE architectures, primarily leveraging mechanisms such as high-order tensor decomposition and dynamic geometric transformations (Lacroix et al., 2020; Xu et al., 2021; Li et al., 2023; Ying et al., 2024; Liu and Wang, 2025). These mechanisms explicitly model the temporal evolution of relational facts and events. Despite these advances, existing TKGE methods encounter two fundamental limitations. **First**, many state-of-the-art models are heavily parameterized, often requiring embedding ranks in the thousands to reach competitive performance. While increasing the embedding rank improves expressiveness, it also substantially raises computational cost. This indicates that model performance still depends largely on capacity scaling. At the same time, it reveals that existing TKGE models lack sufficiently strong structural constraints to effectively

*Corresponding Author
[1]Code is available at https://github.com/dellixx/TDSym

Findings of the Association for Computational Linguistics: ACL 2026, pages 16341–16357

control model complexity, thereby limiting scalability. **Second**, time is predominantly treated as a set of discrete and independent symbols via lookup tables. While effective for capturing timestamp-specific patterns, this design overlooks the continuity of temporal dynamics. As a result, such discrete approaches tend to memorize time-specific states rather than learning evolution rules. Consequently, they struggle to generalize temporal dynamics beyond observed timestamps, which limits their ability to predict facts at unseen time points.

These observations naturally raise a pivotal question:

> *Can we construct an effective, structure-aware dynamical framework that captures continuous temporal evolution while maintaining strong expressiveness at low embedding ranks?*

To address this challenge, we propose *TDSym*, a novel temporal knowledge graph embedding framework in which relations and timestamps jointly parameterize structure-preserving Hamiltonian flows acting on entity representations in a symplectic phase space. Entities are embedded as low-dimensional phase-space vectors, and temporal relational transformations are realized as symplectic operators derived from Hamiltonian generators. By constraining these operators to the symplectic group, our approach introduces an explicit geometric inductive bias that preserves phase-space volume and enables stable, expressive modeling with compact representations.

Furthermore, *TDSym* bridges discrete supervision and continuous dynamics through a Time-Aware Parameter Modulation mechanism. Specifically, we employ differentiable Rotary Time Embeddings (RoTE) to encode continuous temporal signals and utilize Feature-wise Linear Modulation (FiLM) to condition the Hamiltonian vector field. This design enables the model to capture both timestamp-specific semantics and smoothly evolving temporal patterns within a unified dynamical system. To ensure strict geometric consistency, the learned Hamiltonian generators are mapped to symplectic transformations via the Cayley transform, enforcing symplectic constraints *by construction*. A distinctive advantage of the proposed continuous Hamiltonian evolution is that it learns an explicit vector field governing temporal dynamics, rather than memorizing timestamp-specific states. This enables TKG representations to be propagated to unobserved timestamps through the learned dynamics, allowing the model to infer facts at unseen time points.

Our main contributions are summarized as follows:

- We propose *TDSym*, a physics-informed framework that models relational evolution through structure-preserving Hamiltonian dynamics in a symplectic phase space, where relations and timestamps jointly parameterize symplectic transformations acting on entity representations. This introduces a structure-aware geometric inductive bias for TKGE.

- We design a unified temporal modeling mechanism that bridges discrete supervision with continuous dynamics. Specifically, we employ Rotary Time Embeddings (RoTE) and feature-wise linear modulation to condition Hamiltonian vector fields, while enforcing strict symplectic constraints via the Cayley transform.

- We show that *TDSym* achieves competitive performance with substantially lower embedding ranks, demonstrating improved parameter efficiency under the proposed geometric constraints. Moreover, experiments indicate that the continuous Hamiltonian evolution enables the model to infer facts at unseen timestamps.

Due to space constraints, the related work is presented in **Appendix A**.

2 Method

In this section, we present *TDSym*, a physics-informed framework for TKGE. As illustrated in Figure 1, the architecture models temporal relational reasoning as Hamiltonian dynamics in a symplectic phase space. The framework contains three key components: (1) *Time-Aware Parameter Modulation*, which fuses discrete and continuous temporal signals to generate dynamic Hamiltonian coefficients; (2) *Symplectic Phase Space Evolution*, which evolves head entities via a geometric flow on the manifold; and (3) *Structure-Preserving Symplectic Scoring*, which utilizes a twist map to compute the score function.

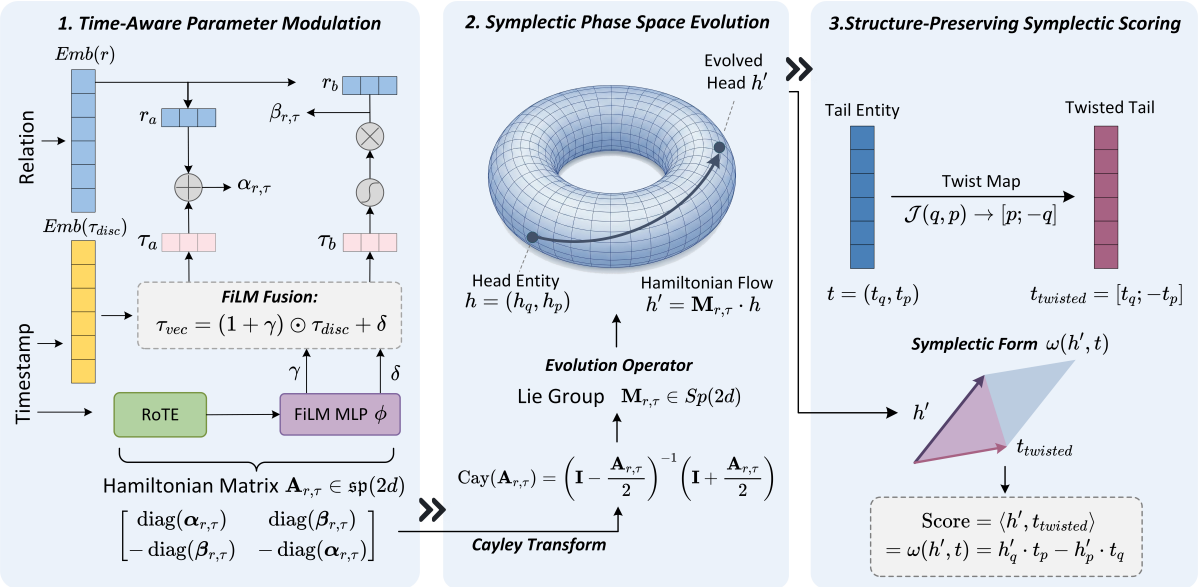


Figure 1: Illustration of *TDSym*. (1) **Time-Aware Parameter Modulation:** FiLM fuses continuous RoTE with discrete temporal features to generate dynamic Hamiltonian coefficients $\boldsymbol{\alpha}_{r,\tau}$ and $\boldsymbol{\beta}_{r,\tau}$. (2) **Symplectic Phase Space Evolution:** These coefficients construct a Hamiltonian generator that is mapped to the symplectic group $\mathrm{Sp}(2d)$ via the Cayley transform, ensuring structure-preserving evolution. (3) **Structure-Preserving Symplectic Scoring:** Scores are computed as the symplectic form between the evolved head h' and tail t using a geometric twist map $\mathcal{J}$.

2.1 Problem Formulation

A TKG is formalized as a collection of quadruples $(h, r, t, \tau) \in \mathcal{E} \times \mathcal{R} \times \mathcal{E} \times \mathcal{T}$, where $\mathcal{E}, \mathcal{R}, \mathcal{T}$ denote the sets of entities, relations, and timestamps, respectively. Given a query $(h, r, ?, \tau)$ or $(?, r, t, \tau)$, the objective is to infer the missing entity.

2.2 Time-Aware Parameter Modulation

Symplectic Phase-Space Representation Inspired by Hamiltonian dynamical systems (Arnol'd, 2013; Greydanus et al., 2019), we embed each entity into a symplectic phase space, where entities are represented as vectors in $\mathbb{R}^{2d}$ composed of generalized position ($\mathbf{q}$) and momentum ($\mathbf{p}$) components:

$$\mathbf{h} = [\mathbf{h}_q; \mathbf{h}_p], \quad \mathbf{t} = [\mathbf{t}_q; \mathbf{t}_p]. \tag{1}$$

This phase space is equipped with a symplectic structure governed by the canonical matrix $\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix}$.

To capture the temporal dynamics of relations, we employ a modulation mechanism that integrates both discrete semantic embeddings and continuous periodicity.

Hybrid Time Encoding (FiLM & RoTE). For a given timestamp τ, we encode it using a learnable discrete embedding $\boldsymbol{\tau}_{disc} \in \mathbb{R}^{2d}$ to capture timestamp-specific semantics. However, such discrete encodings lack the capacity to generalize beyond observed timestamps. To address this, we additionally employ a Rotary Time Embedding (RoTE) generator to capture continuous temporal evolution.

Inspired by positional encoding techniques in Transformers (Vaswani et al., 2017), RoTE generates a continuous representation by rotating a learnable base vector $\mathbf{b} \in \mathbb{R}^{2d}$ in the complex plane, rather than rotating input features directly. Specifically, we define a spectrum of frequencies $\Theta = \{\omega_j\}_{j=1}^{d}$, where $\omega_j = 10000^{-2j/d}$. Given a continuous timestamp value τ, the rotation angle for the j-th subspace is computed as $\theta_j = \tau \cdot \omega_j$. The continuous time embedding $\boldsymbol{\tau}_{rot}(\tau)$ is then obtained by applying a rotation matrix to pairs of elements in the base vector $\mathbf{b}$:

$$\begin{pmatrix} \boldsymbol{\tau}_{rot,2j} \\ \boldsymbol{\tau}_{rot,2j+1} \end{pmatrix} = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix} \begin{pmatrix} \mathbf{b}_{2j} \\ \mathbf{b}_{2j+1} \end{pmatrix}. \tag{2}$$

This mechanism ensures that temporal differences correspond to constant rotations in the embedding space, allowing the time representation to be smoothly extended to unseen timestamps by continuing the rotation angle.

To synthesize these signals, we employ Feature-wise Linear Modulation (FiLM). A lightweight

MLP ϕ projects the continuous RoTE output $\boldsymbol{\tau}_{rot}$ into scaling ($\boldsymbol{\gamma}$) and shifting ($\boldsymbol{\delta}$) parameters. The final fused temporal vector $\boldsymbol{\tau}_{vec}$ is computed via affine modulation:

$$\boldsymbol{\tau}_{vec} = (1 + \boldsymbol{\gamma}) \odot \boldsymbol{\tau}_{disc} + \boldsymbol{\delta}, \tag{3}$$

where $\odot$ denotes the Hadamard product.

Dynamic Coefficient Generation. The relation embedding is split into two components $\mathbf{r}_a$ and $\mathbf{r}_b$. Similarly, the modulated time vector $\boldsymbol{\tau}_{vec}$ is split into $\boldsymbol{\tau}_a$ and $\boldsymbol{\tau}_b$. These components interact to generate the time-specific Hamiltonian coefficients $\boldsymbol{\alpha}_{r,\tau}$ and $\boldsymbol{\beta}_{r,\tau}$. Consistent with the architecture in Figure 1, $\boldsymbol{\alpha}_{r,\tau}$ is derived via an additive shift, while $\boldsymbol{\beta}_{r,\tau}$ acts as a multiplicative gate:

$$\boldsymbol{\alpha}_{r,\tau} = \mathbf{r}_a + \boldsymbol{\tau}_a, \quad \boldsymbol{\beta}_{r,\tau} = \mathbf{r}_b \odot \sigma(\boldsymbol{\tau}_b), \tag{4}$$

where $\sigma(\cdot)$ is the sigmoid function. These coefficients define the effective scaling and modulation of the dynamical system for *TDSym*.

2.3 Symplectic Phase Space Evolution

Next, we construct a time-dependent Hamiltonian flow parameterized by the relation r and timestamp τ to govern the evolution of entity representations.

Hamiltonian Generator in the Symplectic Lie Algebra. Using the generated coefficients, we construct a pairwise-decoupled Hamiltonian matrix $\mathbf{A}_{r,\tau} \in \mathbb{R}^{2d \times 2d}$:

$$\mathbf{A}_{r,\tau} = \begin{bmatrix} \operatorname{diag}(\boldsymbol{\alpha}_{r,\tau}) & \operatorname{diag}(\boldsymbol{\beta}_{r,\tau}) \\ -\operatorname{diag}(\boldsymbol{\beta}_{r,\tau}) & -\operatorname{diag}(\boldsymbol{\alpha}_{r,\tau}) \end{bmatrix}, \tag{5}$$

where $\mathbf{A}_{r,\tau}$ satisfies $\mathbf{A}_{r,\tau}^\top \mathbf{J} + \mathbf{J}\mathbf{A}_{r,\tau} = \mathbf{0}$, ensuring that it lies in the symplectic Lie algebra $\mathfrak{sp}(2d)$.

Cayley Transform and Evolution. To obtain a symplectic operator on the Lie group $Sp(2d)$ while avoiding expensive matrix exponentiation, we apply the Cayley transform to $\mathbf{A}_{r,\tau}$. The evolution operator $\mathbf{M}_{r,\tau}$ is defined as:

$$\begin{aligned} \mathbf{M}_{r,\tau} &= \operatorname{Cay}(\mathbf{A}_{r,\tau}) \\ &= \left(\mathbf{I} - \frac{\mathbf{A}_{r,\tau}}{2}\right)^{-1} \left(\mathbf{I} + \frac{\mathbf{A}_{r,\tau}}{2}\right). \end{aligned} \tag{6}$$

The head entity is then evolved through the phase space:

$$\mathbf{h}' = \mathbf{M}_{r,\tau}\mathbf{h}. \tag{7}$$

This transformation guarantees that the symplectic structure and phase-space volume are preserved during evolution, as $\mathbf{M}_{r,\tau} \in Sp(2d)$ by construction.

2.4 Structure-Preserving Symplectic Scoring

Finally, to measure the plausibility of a triple, we compute the symplectic form between the evolved head $\mathbf{h}'$ and the tail $\mathbf{t}$.

Twist Map and Scoring. Direct computation of the symplectic bilinear form is efficiently implemented using a geometric "twist". We define a linear twist map $\mathcal{J}$ that rotates the tail entity coordinates:

$$\mathbf{t}_{twisted} = \mathcal{J}(\mathbf{t}_q, \mathbf{t}_p) = [\mathbf{t}_p; -\mathbf{t}_q], \tag{8}$$

where $\mathcal{J}$ denotes the linear operator induced by the canonical symplectic matrix $\mathbf{J}$, i.e., $\mathcal{J}(\mathbf{t}) = \mathbf{J}\mathbf{t}$. The scoring function is defined as the inner product between the evolved head and the twisted tail. Crucially, this operation is mathematically equivalent to computing the symplectic form $\omega(\mathbf{h}', \mathbf{t})$, which evaluates the symplectic form between the two entity states:

$$\begin{aligned} \text{Score} &= \langle \mathbf{h}', \mathbf{t}_{twisted} \rangle \\ &= \omega(\mathbf{h}', \mathbf{t}) = \mathbf{h}'_q \cdot \mathbf{t}_p - \mathbf{h}'_p \cdot \mathbf{t}_q. \end{aligned} \tag{9}$$

This formulation naturally encodes the directional and area-preserving properties of the symplectic manifold directly into the scoring mechanism.

2.5 Theoretical Analysis

In this subsection, we provide theoretical guarantees that our low-rank construction strictly preserves the geometric validity of the model. We formally state that *TDSym* adheres to the symplectic structure required for Hamiltonian dynamics, with detailed proofs provided in **Appendix B**.

Proposition 1 (Validity of the Hamiltonian Generator)**.** *The pairwise-decoupled Hamiltonian matrix $\mathbf{A}_{r,\tau}$ constructed in Eq. 5 lies in the symplectic Lie algebra $\mathfrak{sp}(2d)$, satisfying the defining condition $\mathbf{A}_{r,\tau}^\top \mathbf{J} + \mathbf{J}\mathbf{A}_{r,\tau} = \mathbf{0}$.*

Proposition 2 (Structure-Preserving Evolution)**.** *The evolution operator $\mathbf{M}_{r,\tau}$ derived via the Cayley transform is a valid symplectic matrix, i.e., $\mathbf{M}_{r,\tau} \in \mathrm{Sp}(2d)$. Consequently, the transformation strictly preserves the symplectic 2-form ω and phase space volume.*

Complexity In general, applying the Cayley transform to a dense symplectic generator incurs an $\mathcal{O}(d^3)$ computational cost. In contrast, our

pairwise-decoupled construction decomposes the computation into d independent 2×2 subsystems, so that the inverse $\left(\mathbf{I} - \frac{1}{2}\mathbf{A}_{r,\tau}\right)^{-1}$ can be computed via d closed-form inversions of independent 2×2 blocks, reducing the overall computational complexity to $\mathcal{O}(d)$. This structural constraint introduces a deliberate *inductive bias*: by enforcing independent dynamics for each conjugate pair (q_i, p_i), the model effectively represents the system as a collection of decoupled Hamiltonian subsystems. Such a design can help reduce the risk of overfitting in high-dimensional temporal settings, while preserving the global symplectic structure required for stable and coherent temporal evolution.

We further provide formal proofs demonstrating that our proposed symplectic evolution operator possesses the capacity to model diverse relation patterns in **Appendix I**.

3 Experiments

3.1 Datasets

To comprehensively evaluate the effectiveness of our approach under diverse temporal dynamics, we conduct experiments on four widely used temporal knowledge graph benchmarks: **ICEWS14**, **ICEWS05-15** (García-Durán et al., 2018), **YAGO11k**, and **Wikidata12k** (Dasgupta et al., 2018). These datasets encompass two fundamentally different paradigms: *event-centric* graphs (typically characterized by discrete timestamps) and *fact-centric* graphs (often involving continuous validity intervals).

Additionally, to investigate whether the proposed continuous Hamiltonian formulation yields the expected emergent benefit of temporal extrapolation, we introduce two time-sensitive variants: **ICEWS14-Ext** and **ICEWS05-15-Ext**. Unlike the standard random partitioning used in interpolation tasks, we adopt a strict chronological split strategy, where the datasets are divided into training (80%), validation (10%), and testing (10%) sets based on timestamp order. This setting ensures that the model is evaluated exclusively on future timestamps unseen during training, serving as a rigorous testbed to verify if *TDSym* has successfully learned the intrinsic laws of evolution rather than merely memorizing static snapshots. Detailed statistics of all datasets are reported in **Appendix C**.

3.2 Baselines

We compare our method against a diverse set of state-of-the-art temporal knowledge graph embedding (TKGE), including TComplEx (Lacroix et al., 2020) (ICLR 2020), TeLM (Xu et al., 2021) (NAACL 2021), TeAST (Li et al., 2023) (*ACL 2023*), TCompoundE (Ying et al., 2024) (*ACL 2024*), and TeRDy (Liu and Wang, 2025) (*ACL 2025*). These baselines cover a broad spectrum of modeling paradigms, including tensor factorization, geometry-aware transformations, and structured temporal dynamics, thereby providing a comprehensive comparison across representative TKGE approaches. In particular, TCompoundE and TeRDy represent strong recent baselines that demonstrate competitive performance on multiple temporal knowledge graph benchmarks. All selected baselines provide publicly available implementations, enabling reproducible evaluation and fair comparison under consistent experimental settings.

For additional experimental settings, please refer to **Appendix D**.

4 Results and Analysis

4.1 Link Prediction Results

Results on ICEWS datasets. As shown in Table 1, *TDSym* consistently achieves competitive performance on the highly dynamic ICEWS datasets. For instance, on ICEWS14 with rank 64, our model attains an MRR of **0.492**, outperforming the strongest baseline, TeAST (0.467), by a significant margin. Similarly, on the longer-term ICEWS05-15 dataset, *TDSym* maintains robust performance, effectively capturing both short-term temporal variations and long-range historical dependencies. These results validate that the proposed symplectic evolution mechanism provides a stable and expressive framework for modeling intense temporal dynamics.

Performance on Temporal Extrapolation Table 2 presents results on the extrapolation benchmarks (ICEWS14-Ext, ICEWS05-15-Ext), where the test set strictly comprises future timestamps ($T\_test > T\_train$). A striking observation is the catastrophic performance collapse of state-of-the-art baselines (e.g., TComplEx MRR drops to 0.009 at Rank=32). This failure is structurally inevitable for discrete approaches: they treat time as independent symbols via lookup tables, rendering unseen

Rank	Model	ICEWS14					ICEWS05-15				
		MRR	H@1	H@3	H@10	Params (MB)	MRR	H@1	H@3	H@10	Params (MB)
32	TComplEx	0.421	0.327	0.463	0.608	1.94	0.489	0.381	0.545	0.700	3.66
	TeLM	0.398	0.309	0.437	0.569	3.88	0.419	0.324	0.461	0.606	7.33
	TeAST	0.456	0.353	0.508	0.650	2.03	0.492	**0.384**	0.547	0.699	4.64
	TCompoundE	0.411	0.306	0.459	0.617	1.94	0.456	0.345	0.509	0.674	3.66
	TeRDy	0.427	0.322	0.475	0.634	1.94	0.457	0.345	0.511	0.677	3.66
	***TDSym* (ours)**	**0.467**	**0.357**	**0.524**	**0.678**	1.97	**0.493**	0.378	**0.552**	**0.713**	3.70
64	TComplEx	0.438	0.350	0.478	0.607	3.88	0.518	0.426	0.564	0.695	7.33
	TeLM	0.398	0.326	0.427	0.539	7.77	0.439	0.361	0.471	0.592	14.66
	TeAST	0.467	0.375	0.511	0.645	4.06	**0.524**	**0.434**	0.571	0.698	9.29
	TCompoundE	0.420	0.312	0.473	0.627	3.88	0.477	0.366	0.535	0.692	7.33
	TeRDy	0.435	0.329	0.485	0.641	3.88	0.476	0.365	0.533	0.691	7.33
	***TDSym* (ours)**	**0.492**	**0.384**	**0.550**	**0.698**	4.01	0.520	0.407	**0.584**	**0.738**	7.45

Table 1: Link prediction results on ICEWS14 and ICEWS05-15 grouped by embedding rank.

Rank	Model	ICEWS14-Ext					ICEWS05-15-Ext				
		MRR	H@1	H@3	H@10	Params (MB)	MRR	H@1	H@3	H@10	Params (MB)
32	TComplEx	0.009	0.003	0.006	0.017	1.94	0.012	0.004	0.011	0.025	3.66
	TeLM	0.005	0.002	0.004	0.009	3.88	0.012	0.005	0.011	0.024	7.33
	TeAST	0.029	0.009	0.026	0.065	2.03	0.050	0.019	0.048	0.107	4.64
	TCompoundE	0.012	0.005	0.010	0.022	1.94	0.021	0.007	0.019	0.045	3.66
	TeRDy	0.013	0.004	0.011	0.027	1.94	0.018	0.007	0.017	0.038	3.66
	***TDSym* (ours)**	**0.362**	**0.269**	**0.404**	**0.537**	1.97	**0.361**	**0.259**	**0.405**	**0.559**	3.70
64	TComplEx	0.008	0.003	0.007	0.016	3.88	0.012	0.004	0.010	0.024	7.33
	TeLM	0.009	0.003	0.007	0.017	7.77	0.011	0.004	0.010	0.021	14.66
	TeAST	0.031	0.009	0.025	0.064	4.06	0.057	0.023	0.056	0.120	9.29
	TCompoundE	0.011	0.004	0.009	0.021	3.88	0.010	0.003	0.008	0.019	7.33
	TeRDy	0.007	0.002	0.005	0.014	3.88	0.012	0.004	0.010	0.025	7.33
	***TDSym* (ours)**	**0.365**	**0.270**	**0.409**	**0.547**	4.01	**0.382**	**0.279**	**0.427**	**0.584**	7.45

Table 2: Link prediction results on ICEWS14-Ext and ICEWS05-15-Ext grouped by embedding rank.

future timestamps as random noise that corrupts inference. In contrast, *TDSym* maintains robust performance. This capability is not a targeted design for forecasting, but an emergent property that verifies the validity of our continuous symplectic formulation. Unlike discrete embeddings that structurally fail at unobserved indices, our model encodes time as continuous phase rotations via RoTE, where accessing a future timestamp corresponds to a mathematically well-defined rotation by a larger angle. Coupled with a Hamiltonian objective that models the laws of evolution rather than static snapshots, the system naturally propagates entity states along the symplectic trajectory. This confirms that *TDSym* has internalized the underlying dynamics, rendering extrapolation a seamless consequence of its geometric integrity rather than a learned heuristic.

Model Efficiency *TDSym* achieves a highly favorable trade-off between expressiveness and complexity, driven by its $\mathcal{O}(d)$ pairwise-decoupled evolution mechanism. As detailed in Table 1, this design allows *TDSym* to outperform high-capacity baselines like TeLM on ICEWS14 (Rank 64) while using only 4.01 MB of parameters compared to TeLM's 7.77 MB. As observed in both Rank 32 and Rank 64 settings, *TDSym* consistently maintains its performance lead over tensor-based baselines, confirming that our symplectic formulation provides a robust and parameter-efficient representation regardless of the embedding dimension.

Supplementary results on YAGO11k and Wikidata12k are provided in **Appendix E** to further demonstrate this robustness.

4.2 Ablation Study

To verify the effectiveness of the proposed components, we conduct both quantitative and qualitative analyses on the ICEWS14 dataset.

4.2.1 Quantitative Analysis

To assess the contribution of the proposed temporal components, we conduct an ablation study on

Model Variant	Components		ICEWS14				ICEWS05-15			
	RoTE	FiLM	Rank=32		Rank=64		Rank=32		Rank=64	
			MRR	H@10	MRR	H@10	MRR	H@10	MRR	H@10
w/o Continuous Time	✗	✓	0.445	0.642	0.468	0.655	0.472	0.683	0.505	0.701
w/o FiLM Modulation	✓	✗	0.452	0.655	0.479	0.673	0.483	0.696	0.511	0.709
***TDSym* (Full)**	✓	✓	**0.467**	**0.678**	**0.492**	**0.698**	**0.493**	**0.713**	**0.520**	**0.738**

Table 3: Ablation study of *TDSym* on ICEWS14 and ICEWS05-15 datasets with embedding ranks $d \in \{32, 64\}$.

ICEWS14 and ICEWS05-15 (ranks $d \in \{32, 64\}$), focusing on Continuous Time Encoding (RoTE) and Time-Aware Parameter Modulation (FiLM). Table 3 reports the results.

The removal of RoTE ("w/o Continuous Time") results in the most significant performance degradation across all settings. On ICEWS14 (Rank 32), MRR drops from 0.467 to 0.445. This confirms that discrete embeddings alone suffer from sparsity and fail to capture the periodicity of historical events. The continuous manifold provided by RoTE is essential for generalizing to unseen timestamps and modeling smooth temporal evolution.

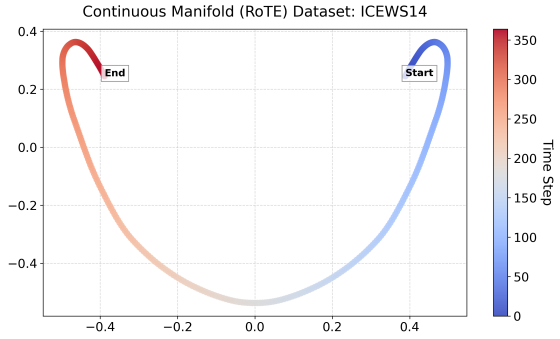
Replacing FiLM with simple additive fusion ("w/o FiLM Modulation") leads to consistent performance drops. This indicates that temporal information functions as a dynamic gate rather than a static bias. The multiplicative mechanism in FiLM allows time features to scale the relational Hamiltonian matrix, effectively modeling the varying intensity of relations over time.

The full *TDSym* model consistently outperforms both variants. The experimental hierarchy (Full $>$ w/o FiLM $>$ w/o RoTE) demonstrates that RoTE provides the necessary geometric granularity, while FiLM effectively injects this signal into the symplectic dynamics, yielding a robust time-aware Hamiltonian system.

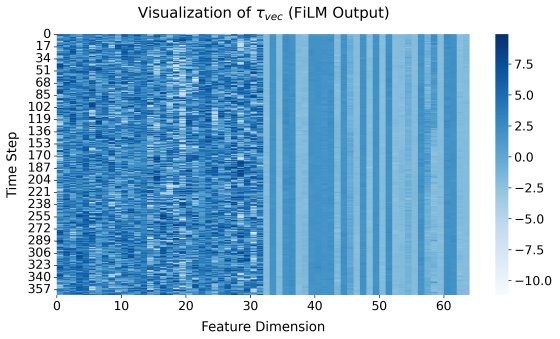
4.2.2 Qualitative Analysis

To intuitively understand *why* these components are critical, we visualize the learned latent representations in Figure 2.

Impact of RoTE. Figure 2(a) visualizes the 2D PCA projection of the Rotary Time Embeddings. The trajectory forms a smooth, closed manifold, geometrically confirming that RoTE effectively encodes temporal periodicity and continuity. This explains the performance drop in the "w/o Continuous Time" variant: without this structured manifold, the model loses the topological basis required for temporal extrapolation and handling unseen timestamps.



(a) RoTE Trajectory



(b) FiLM Heatmap

Figure 2: Visualization of temporal representations on ICEWS14. **(a)** The 2D PCA projection of RoTE reveals a smooth, cyclic manifold, confirming the preservation of temporal periodicity. **(b)** The heatmap of $\boldsymbol{\tau}_{vec}$ exhibits distinct striations, indicating the effective fusion of discrete event semantics into the continuous temporal flow.

Impact of FiLM. Figure 2(b) displays the heatmap of the final modulated vector $\boldsymbol{\tau}_{vec}$ in Eq. 3. Interestingly, we observe a distinct structural separation along the feature dimensions: the earlier dimensions exhibit dense, irregular patterns characteristic of discrete event semantics, while the latter dimensions display clear vertical striations indicative of continuous temporal consistency. This

suggests that *TDSym* effectively disentangles temporal information into heterogeneous subspaces. The "discrete-dominant" subspace captures high-frequency event triggers (e.g., sudden shocks), while the "continuous-dominant" subspace maintains the smooth evolutionary flow. This multi-dimensional representation provides a rich, complementary basis for the subsequent Hamiltonian construction, allowing the model to simultaneously respect historical facts and varying temporal trends.

We present additional experimental results on Dynamics and Liouville's Theorem in **Appendix F**.

4.3 Geometric Verification

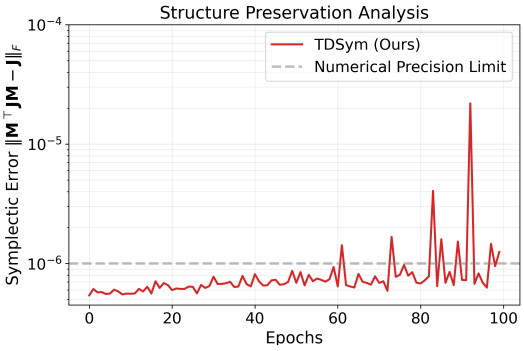
A core theoretical contribution of *TDSym* is the guarantee that the learned time evolution strictly adheres to symplectic geometry, thereby preserving the phase space volume. While Proposition 2 provides a theoretical proof based on the Cayley transform, it is crucial to verify that this property holds numerically during the optimization process.

Metric Definition. We quantify the structural deviation using the Frobenius norm of the symplectic constraint error:

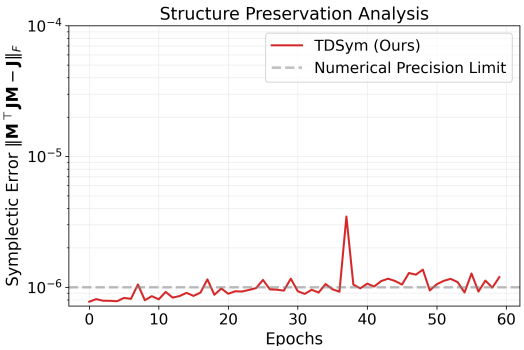
$$\mathcal{E}_{sym} = \|\mathbf{M}_{r,\tau}^{\top}\mathbf{J}\mathbf{M}_{r,\tau} - \mathbf{J}\|_F, \qquad (10)$$

where $\mathbf{J}$ is the standard symplectic form and $\mathbf{M}_{r,\tau}$ is the evolution operator generated by the model. A value of zero indicates perfect structure preservation.

Symplecticity Verification. Figure 3 illustrates the evolution of the symplectic constraint error $\mathcal{E}_{\text{sym}}$ on ICEWS14 across different model capacities (Rank=32 and Rank=64). As observed, the error magnitude remains consistently negligible throughout training, fluctuating around 10^{-6}. This level is close to the numerical precision limit of single-precision floating-point arithmetic, indicating that the residual error arises from floating-point round-off rather than violations of the symplectic constraint. Importantly, the error does not accumulate over training epochs nor increase with higher embedding dimensions (Rank=64). These results empirically confirm that *TDSym* preserves the symplectic structure and phase-space volume up to numerical precision in practice. Such structure preservation supports numerically stable entity evolution and helps prevent trajectory divergence or embedding collapse commonly observed in unconstrained models during long-term temporal reasoning.



(a) Symplectic Error (Rank=32)



(b) Symplectic Error (Rank=64)

Figure 3: Evolution of the symplectic constraint error (defined in Eq. 10) on ICEWS14.

5 Conclusion

In this paper, we presented *TDSym*, a novel TKGE framework that grounds temporal knowledge graph evolution in the mathematical rigor of symplectic geometry. By modeling entity dynamics as a Hamiltonian system, we introduce a continuous and structure-preserving architecture that adheres to Liouville's theorem, ensuring long-term numerical stability with errors bounded at the level of machine precision ($\approx 10^{-6}$). *TDSym* learns the underlying continuous laws of temporal evolution. Our empirical results further indicate that this geometric consistency leads to robust generalization to unseen future timestamps. By driving the Hamiltonian flow with continuous Rotary Time Embeddings (RoTE), the model propagates entity states along well-defined symplectic trajectories, capturing intrinsic temporal dynamics rather than overfitting to observed temporal patterns. Finally, we demonstrate that this theoretical rigor is achieved with high computational efficiency. The proposed low-rank pairwise-decoupled formulation reduces the complexity of symplectic evolution to $\mathcal{O}(d)$, enabling *TDSym* to scale effectively while often

outperforming strong baselines on standard TKGE benchmark datasets.

Limitations

While *TDSym* achieves competitive performance with high parameter efficiency, two limitations remain. First, the reliance on a pairwise-decoupled Hamiltonian structure, while crucial for ensuring $\mathcal{O}(d)$ complexity and preventing overfitting, theoretically limits the modeling of dense cross-dimensional interactions compared to full-rank matrices. Second, like many temporal knowledge graph embedding models, *TDSym* cannot generalize to unseen entities. Nevertheless, benefiting from our continuous symplectic evolution design, *TDSym* excels in temporal extrapolation, effectively predicting facts at future timestamps unseen during training.

Acknowledgments

This work was funded by National Natural Science Foundation of China (Grant No. 62366036), Outstanding Youth Fund Project of Inner Mongolia Autonomous Region (Grant No. 2025JQ010), Program for Young Talents of Science and Technology in Universities of Inner Mongolia Autonomous Region (Grant No. NJYT24033), Major Science and Technology Projects of Inner Mongolia Autonomous Region (Grant No. 2025ZDSF0029), Key R&D and Achievement Transformation Program of Inner Mongolia Autonomous Region (Grant No. 2025YFDZ0011, 2025YFDZ0026, 2025YFSH0021, 2025YFHH0073), Hohhot Science and Technology Project (Grant No. 2023-Zhan-Zhong-1).

References

Vladimir Igorevich Arnol'd. 2013. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media.

Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, pages 2787–2795.

Li Cai, Xin Mao, Yuhao Zhou, Zhaoguang Long, Changxu Wu, and Man Lan. 2024. A survey on temporal knowledge graph: Representation learning and applications. *arXiv preprint arXiv:2403.04782*.

Ines Chami, Adva Wolf, Da-Cheng Juan, Frederic Sala, Sujith Ravi, and Christopher Ré. 2020. Low-dimensional hyperbolic knowledge graph embeddings. *arXiv preprint arXiv:2005.00545*.

Kai Chen, Ye Wang, Yitong Li, and Aiping Li. 2022. Rotateqvs: Representing temporal information as rotations in quaternion vector space for temporal knowledge graph completion. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 5843–5857. Association for Computational Linguistics.

Minze Chen, Zhenxiang Tao, Weitong Tang, Tingxin Qin, Rui Yang, and Chunli Zhu. 2024. Enhancing emergency decision-making with knowledge graphs and large language models. *International Journal of Disaster Risk Reduction*, 113:104804.

Wei Chen, Haoyu Huang, Zhiyu Zhang, Tianyi Wang, Youfang Lin, Liang Chang, and Huaiyu Wan. 2025. Next-poi recommendation via spatial-temporal knowledge graph contrastive learning and trajectory prompt. *IEEE Transactions on Knowledge and Data Engineering*.

Yuhan Chen, Takashi Matsubara, and Takaharu Yaguchi. 2021. Neural symplectic form: Learning hamiltonian equations on general coordinate systems. *Advances in Neural Information Processing Systems*, 34:16659–16670.

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

Shib Sankar Dasgupta, Swayambhu Nath Ray, and Partha P. Talukdar. 2018. Hyte: Hyperplane-based temporally aware knowledge graph embedding. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 2001–2011. Association for Computational Linguistics.

Alberto García-Durán, Sebastijan Dumancic, and Mathias Niepert. 2018. Learning sequence encoders for temporal knowledge graph completion. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 4816–4821. Association for Computational Linguistics.

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

Ernst Hairer, Christian Lubich, and Gerhard Wanner. 2006. Structure-preserving algorithms for ordinary differential equations. *Geometric numerical integration*, 31.

Ainaz Hajimoradlou and Mehran Kazemi. 2022. Stay positive: knowledge graph embedding without negative sampling. *arXiv preprint arXiv:2201.02661*.

Hengchang Hu, Wei Guo, Xu Liu, Yong Liu, Ruiming Tang, Rui Zhang, and Min-Yen Kan. 2024. User behavior enriched temporal knowledge graphs for sequential recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, pages 266–275.

Guang-Li Huang and Arkady Zaslavsky. 2024. Contextual knowledge graph approach to bias-reduced decision support systems. *Journal of Decision Systems*, 33(sup1):29–46.

Pengzhan Jin, Zhen Zhang, Aiqing Zhu, Yifa Tang, and George Em Karniadakis. 2020. Sympnets: Intrinsic structure-preserving symplectic networks for identifying hamiltonian systems. *Neural Networks*, 132:166–179.

Timothée Lacroix, Guillaume Obozinski, and Nicolas Usunier. 2020. Tensor decompositions for temporal knowledge base completion. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Jennifer Lautenschlager, Steve Shellman, and Michael Ward. 2015. Icews event aggregations.

Julien Leblay and Melisachew Wudage Chekol. 2018. Deriving validity time in knowledge graph. In *Companion of the The Web Conference 2018 on The Web Conference 2018, WWW 2018, Lyon , France, April 23-27, 2018*, pages 1771–1776. ACM.

Jiang Li, Xiangdong Su, and Guanglai Gao. 2023. Teast: Temporal knowledge graph embedding via archimedean spiral timeline. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 15460–15474. Association for Computational Linguistics.

Ziyang Liu and Chaokun Wang. 2025. Terdy: Temporal relation dynamics through frequency decomposition for temporal knowledge graph completion. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9611–9622.

Haithem Mezni. 2021. Temporal knowledge graph embedding for effective service recommendation. *IEEE Transactions on Services Computing*, 15(5):3077–3088.

Sebastien Montella, Lina M Rojas Barahona, and Johannes Heinecke. 2021. Hyperbolic temporal knowledge graph embeddings with relational and time curvatures. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3296–3308.

Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. 2019. Rotate: Knowledge graph embedding by relational rotation in complex space. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.

Xing Tang, Ling Chen, Hongyu Shi, and Dandan Lyu. 2024. Dhyper: A recurrent dual hypergraph neural network for event prediction in temporal knowledge graphs. *ACM Transactions on Information Systems*, 42(5):1–23.

Xiaowei Tian, Xiaoyan Zhang, Xiaofeng Du, and Tianbo Lu. 2025. Cohn: Context-aware hawkes graph network for temporal knowledge graph reasoning. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, pages 2895–2904.

Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. 2016. Complex embeddings for simple link prediction. In *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*, volume 48 of *JMLR Workshop and Conference Proceedings*, pages 2071–2080. JMLR.org.

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

Weiguang Wang, Yingying Feng, Haiyan Zhao, Xin Wang, Ruikai Cai, Wei Cai, and Xia Zhang. 2024. Mdpg: a novel multi-disease diagnosis prediction method based on patient knowledge graphs. *Health Information Science and Systems*, 12(1):15.

Chengjin Xu, Yung-Yu Chen, Mojtaba Nayyeri, and Jens Lehmann. 2021. Temporal knowledge graph completion using a linear temporal regularizer and multivector embeddings. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 2569–2578. Association for Computational Linguistics.

Chengjin Xu, Mojtaba Nayyeri, Fouad Alkhoury, Hamed Shariat Yazdi, and Jens Lehmann. 2020. Tero: A time-aware knowledge graph embedding via temporal rotation. In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, pages 1583–1593. International Committee on Computational Linguistics.

Yi Xu, Junjie Ou, Hui Xu, and Luoyi Fu. 2023. Temporal knowledge graph reasoning with historical contrastive learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pages 4765–4773.

Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2015. Embedding entities and relations for learning and inference in knowledge bases. In

3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings.

Rui Ying, Mengting Hu, Jianfeng Wu, Yalan Xie, Xiaoyi Liu, Zhunheng Wang, Ming Jiang, Hang Gao, Linlin Zhang, and Renhong Cheng. 2024. Simple but effective compound geometric operations for temporal knowledge graph completion. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11074–11086, Bangkok, Thailand. Association for Computational Linguistics.

Yuyue Zhao, Xiang Wang, Jiawei Chen, Yashen Wang, Wei Tang, Xiangnan He, and Haiyong Xie. 2022. Time-aware path reasoning on knowledge graph for recommendation. *ACM Transactions on Information Systems*, 41(2):1–26.

Yaofeng Desmond Zhong, Biswadip Dey, and Amit Chakraborty. 2019. Symplectic ode-net: Learning hamiltonian dynamics with control. *arXiv preprint arXiv:1909.12077*.

A Related Work

A.1 Temporal Knowledge Graph Embedding

Static knowledge graph embedding (KGE) methods aim to represent entities and relations in continuous vector spaces under the assumption that relational facts remain time-invariant. Representative approaches include translational models such as TransE (Bordes et al., 2013), which interpret relations as translation operations in the embedding space, as well as semantic matching models such as DistMult (Yang et al., 2015), ComplEx (Trouillon et al., 2016), and RotatE (Sun et al., 2019), which capture relational patterns through bilinear interactions or geometric transformations. While effective for static relational reasoning, these models are inherently limited in their ability to represent temporal dynamics, as they treat all facts as timeless and fail to account for the evolution of entities and relations over time.

To overcome the limitations of static knowledge graph embeddings, a growing body of work has focused on incorporating temporal information into representation learning, giving rise to Temporal Knowledge Graph Embedding (TKGE) models. Temporal KGE (TKGE) incorporates temporal information to model the validity and evolution of facts over time. Early approaches treat time as discrete indices and learn time-specific transformations, e.g., TTransE (Leblay and Chekol, 2018) and HyTE (Dasgupta et al., 2018), which project or translate embeddings conditioned on timestamps. Recent factorization-based methods extend tensor decompositions to temporal settings, such as TComplEx (Lacroix et al., 2020), and typically require relatively large embedding ranks to achieve strong performance. Rotation- or modulation-based families (e.g., TeRo (Xu et al., 2020) and TeLM (Xu et al., 2021)) further model time-conditioned relational dynamics through multiplicative interactions over embeddings. More recently, TeAST (Li et al., 2023) maps each relation to an Archimedean spiral timeline and introduces a temporal spiral regularizer to encode temporal order and evolution patterns. TCompoundE (Ying et al., 2024) models temporal facts via compound geometric operations by combining relation-specific and time-specific transformations, improving expressiveness for complex temporal dynamics. TeRDy (Liu and Wang, 2025) characterizes relation evolution through frequency decomposition, leveraging Fourier transform to represent temporal relational dynamics and introducing dedicated temporal regularization to handle both long-term trends and short-term variations.

Beyond Euclidean embedding spaces, hyperbolic geometry has been explored for modeling structured and potentially hierarchical relational patterns. Montella et al. (2021) propose HERCULES, a time-aware extension of AttH (Chami et al., 2020), in which manifold curvature is parameterized jointly by relations and time, enabling time-conditioned representations in hyperbolic space. While hyperbolic geometry provides a powerful inductive bias, existing hyperbolic TKGE approaches typically rely on discrete-time supervision and negative sampling, and their non-Euclidean operations together with curvature parameterization can introduce additional optimization and scalability challenges (Montella et al., 2021; Chami et al., 2020; Hajimoradlou and Kazemi, 2022). Moreover, these methods generally treat time as discrete indices (e.g., timestamp embeddings or discretized observations), and thus lack an explicit mechanism to model continuous temporal evolution across neighboring time points.

A.2 Physics-Informed and Symplectic Representation Learning

Physics-informed representation learning incorporates inductive biases from dynamical systems, such as conservation laws and geometric invariants. Hamiltonian Neural Networks (HNNs) learn a Hamiltonian function from data and generate dy-

namics through Hamilton's equations, improving long-horizon stability by respecting energy-like conserved quantities (Greydanus et al., 2019). To further preserve geometric structure, symplectic learning methods explicitly constrain the learned evolution to be symplectic. For example, Symplectic Recurrent Neural Networks (SRNNs) combine a neural Hamiltonian with symplectic integration to mitigate numerical drift in long-term rollouts (Chen et al., 2019). Related directions include continuous-time formulations such as Symplectic ODE-Net (SymODEN) (Zhong et al., 2019), structure-preserving map approximators such as SympNets (Jin et al., 2020), and approaches that learn the symplectic form itself from data (Chen et al., 2021). While symplectic and Hamiltonian inductive biases are well studied for physical simulation and time-series modeling, they remain largely unexplored in temporal multi-relational settings. Recent surveys on temporal knowledge graph representation learning focus primarily on time-indexed embeddings and discrete-time modeling paradigms, and do not cover symplectic structure as a core design principle (Cai et al., 2024). Motivated by this gap, *TDSym* introduces symplectic structure into TKGE by parameterizing relation-induced evolution as symplectic transformations via the Cayley transform, enabling stable and parameter-efficient modeling of temporal dynamics.

B Detailed Proofs

B.1 Proof of Proposition 1

Proof. Let us denote the diagonal blocks as $\mathbf{D}_\alpha = \mathrm{diag}(\boldsymbol{\alpha}_{r,\tau})$ and $\mathbf{D}_\beta = \mathrm{diag}(\boldsymbol{\beta}_{r,\tau})$. The Hamiltonian generator is defined as:

$$\mathbf{A}_{r,\tau} = \begin{bmatrix} \mathbf{D}_\alpha & \mathbf{D}_\beta \\ -\mathbf{D}_\beta & -\mathbf{D}_\alpha \end{bmatrix}. \tag{11}$$

Substituting this into the Lie algebra condition with the standard symplectic form $\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix}$, we verify the symplectic product:

$$\begin{aligned} &\mathbf{A}_{r,\tau}^\top \mathbf{J} + \mathbf{J}\mathbf{A}_{r,\tau} = \\ &\begin{bmatrix} \mathbf{D}_\alpha & -\mathbf{D}_\beta \\ \mathbf{D}_\beta & -\mathbf{D}_\alpha \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{D}_\alpha & \mathbf{D}_\beta \\ -\mathbf{D}_\beta & -\mathbf{D}_\alpha \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{D}_\beta & \mathbf{D}_\alpha \\ \mathbf{D}_\alpha & \mathbf{D}_\beta \end{bmatrix} + \begin{bmatrix} -\mathbf{D}_\beta & -\mathbf{D}_\alpha \\ -\mathbf{D}_\alpha & -\mathbf{D}_\beta \end{bmatrix} = \mathbf{0}. \end{aligned} \tag{12}$$

Since the sum vanishes identically for any $\mathbf{D}_\alpha, \mathbf{D}_\beta$, the matrix $\mathbf{A}_{r,\tau}$ is a valid generator in $\mathfrak{sp}(2d)$. □

B.2 Proof of Proposition 2

Proof. The Cayley transform maps elements from the Lie algebra to the Lie group. Let $\mathbf{S} = \frac{1}{2}\mathbf{A}_{r,\tau}$. Since $\mathbf{A}_{r,\tau} \in \mathfrak{sp}(2d)$, it follows that $\mathbf{S}^\top \mathbf{J} = -\mathbf{J}\mathbf{S}$. Consider the operator $\mathbf{M}_{r,\tau} = (\mathbf{I} - \mathbf{S})^{-1}(\mathbf{I} + \mathbf{S})$. We examine the condition $\mathbf{M}_{r,\tau}^\top \mathbf{J} \mathbf{M}_{r,\tau} = \mathbf{J}$:

$$\begin{aligned} &\mathbf{M}_{r,\tau}^\top \mathbf{J} \mathbf{M}_{r,\tau} = \\ &(\mathbf{I} + \mathbf{S})^\top (\mathbf{I} - \mathbf{S})^{-\top} \mathbf{J} (\mathbf{I} - \mathbf{S})^{-1} (\mathbf{I} + \mathbf{S}). \end{aligned} \tag{13}$$

Using $\mathbf{J}\mathbf{S} = -\mathbf{S}^\top \mathbf{J}$, we derive the identity $(\mathbf{I} - \mathbf{S})^{-\top} \mathbf{J} = \mathbf{J}(\mathbf{I} + \mathbf{S})^{-1}$ and $(\mathbf{I} + \mathbf{S}^\top)\mathbf{J} = \mathbf{J}(\mathbf{I} - \mathbf{S})$. Substituting these:

$$\begin{aligned} \mathbf{M}_{r,\tau}^\top \mathbf{J} \mathbf{M}_{r,\tau} &= (\mathbf{I} + \mathbf{S}^\top) \left[\mathbf{J}(\mathbf{I} + \mathbf{S})^{-1}\right] (\mathbf{I} - \mathbf{S})^{-1}(\mathbf{I} + \mathbf{S}) \\ &= \left[\mathbf{J}(\mathbf{I} - \mathbf{S})\right] (\mathbf{I} + \mathbf{S})^{-1}(\mathbf{I} - \mathbf{S})^{-1}(\mathbf{I} + \mathbf{S}). \end{aligned} \tag{14}$$

Since rational functions of the same matrix $\mathbf{S}$ commute, $(\mathbf{I}+\mathbf{S})^{-1}$ and $(\mathbf{I}-\mathbf{S})^{-1}$ can be swapped. This leads to cancellation:

$$\begin{aligned} \ldots &= \mathbf{J}(\mathbf{I} - \mathbf{S})(\mathbf{I} - \mathbf{S})^{-1}(\mathbf{I} + \mathbf{S})^{-1}(\mathbf{I} + \mathbf{S}) \\ &= \mathbf{J} \cdot \mathbf{I} \cdot \mathbf{I} = \mathbf{J}. \end{aligned} \tag{15}$$

Thus, $\mathbf{M}_{r,\tau} \in \mathrm{Sp}(2d)$, ensuring the evolution is strictly symplectic. □

C Detailed Dataset Statistics

Table 4 reports detailed statistics for all datasets, including the number of entities and relations, temporal coverage, and the sizes of the training, validation, and test splits.

ICEWS14 and ICEWS05-15. ICEWS14 and ICEWS05-15 are event-centric temporal knowledge graphs derived from the Integrated Crisis Early Warning System (ICEWS) (Lautenschlager et al., 2015). Each fact is represented as a quadruple (h, r, t, τ), indicating that an event involving head entity h and tail entity t under relation r occurs at timestamp τ. ICEWS14 contains events from a single year (2014), whereas ICEWS05-15 spans a longer period from 2005 to 2015, exhibiting substantially richer temporal diversity and longer-range dependencies. As summarized in Table 4, ICEWS14 includes 7,128 entities and 230 relations, while ICEWS05-15 contains 10,488 entities and 251 relations, making it both larger in scale and more temporally complex. Following standard practice in TKGE, timestamps are treated as discrete indices, and we adopt the official train/validation/test splits. These datasets are particularly suitable for evaluating a model's ability to capture fast-evolving temporal dynamics.

Dataset	#Ent	#Rel	Period (year)	#Train	#Valid	#Test
ICEWS14	7,128	230	2014	72,826	8,941	8,963
ICEWS05-15	10,488	251	2005–2015	368,962	46,275	46,092
YAGO11k	10,623	10	-431–2844	16,406	2,050	2,051
Wikidata12k	12,554	24	19–2020	32,497	4,062	4,062

Table 4: Statistics of the datasets used in our experiments.

Dataset	# Ent.	# Rel.	# Time	Split	# Quadruples	Time Range
ICEWS14-Ext	7,128	230	365	Train	72,584	2014-01-01 → 2014-10-23
				Valid	8,545	2014-10-23 → 2014-11-24
				Test	8,549	2014-11-24 → 2014-12-31
ICEWS05-15-Ext	10,488	251	4,017	Train	369,063	2005-01-01 → 2013-11-19
				Valid	41,921	2013-11-19 → 2014-12-27
				Test	41,562	2014-12-27 → 2015-12-31

Table 5: Statistics of the constructed extrapolation datasets. Unlike random splits, these datasets are partitioned strictly by chronological order (80%/10%/10%) to evaluate the model's capability in predicting future events.

YAGO11k and Wikidata12k. YAGO11k and Wikidata12k are fact-centric temporal knowledge graphs, where relations persist over extended time intervals rather than occurring instantaneously. Each fact is typically represented as $(h, r, t, [\tau_s, \tau_e])$, indicating that the relation holds continuously from start time τ_s to end time τ_e. Following common practice in prior work (Xu et al., 2021), temporal intervals are discretized into dataset-specific time indices to enable unified temporal modeling. As shown in Table 4, YAGO11k and Wikidata12k contain 10,623 and 12,554 entities, respectively, with relatively fewer relations but substantially longer temporal spans. These datasets emphasize long-term relational consistency and temporal validity, making them well suited for evaluating a model's ability to preserve structured temporal semantics over extended horizons.

ICEWS14-Ext and ICEWS05-15-Ext. Table 5 presents the detailed statistics of the constructed extrapolation benchmarks. Distinct from standard random partitioning, we adopt a strict chronological splitting strategy to ensure *disjoint time intervals* across subsets. As shown in the table, the training, validation, and testing sets follow a sequential timeline (e.g., for ICEWS14-Ext, the training phase concludes on October 23, 2014, immediately followed by the validation phase). This setup guarantees that the model is evaluated exclusively on future events involving known entities, providing a rigorous assessment of temporal predicting capabilities.

D Additional Experimental Setups

D.1 Evaluation Protocol

We follow the standard evaluation protocol for TKGE to ensure fair and reproducible comparison across all models (Lacroix et al., 2020; Ying et al., 2024; Liu and Wang, 2025). Given a query quadruple $(h, r, ?, \tau)$ or $(?, r, t, \tau)$, the task is to rank the correct entity among all candidate entities at the corresponding timestamp. We adopt the widely used *filtered setting* introduced in prior work, where all valid triples appearing in the training, validation, or test sets are removed from the candidate list before ranking. This prevents penalizing the model for predicting alternative but valid facts and ensures a fair evaluation of predictive quality. We evaluate model performance using standard ranking-based metrics, including Mean Reciprocal Rank (MRR) and Hits@k (for $k \in \{1, 3, 10\}$). For each test quadruple, we compute scores for all possible entity substitutions and rank the correct entity among all candidates under the filtered setting. Higher MRR and Hits@k values indicate better predictive performance.

D.2 Implementation Details

All baseline methods are implemented using their officially released codebases, following the recommended training protocols and hyperparameter configurations reported in the original papers. We do not modify the original model architectures or optimization strategies, ensuring a fair and re-

Rank	Model	YAGO11k					Wikidata12k				
		MRR	H@1	H@3	H@10	Params (MB)	MRR	H@1	H@3	H@10	Params (MB)
32	TComplEx	0.119	0.094	0.117	0.165	2.70	0.231	0.165	0.255	0.358	3.23
	TeLM	0.126	0.102	0.122	0.174	5.40	0.247	**0.183**	0.276	0.367	6.47
	TeAST	0.120	0.094	0.120	0.164	2.79	0.242	0.176	0.269	0.369	3.38
	TCompoundE	0.134	0.095	0.137	**0.204**	2.70	0.243	0.167	0.262	0.393	3.23
	TeRDy	0.132	0.081	0.146	0.216	2.70	0.241	0.162	0.261	0.389	3.23
	***TDSym* (ours)**	**0.138**	**0.104**	**0.137**	0.191	2.73	**0.252**	0.180	**0.281**	**0.397**	3.27
64	TComplEx	0.128	0.101	0.126	0.178	5.40	0.247	0.183	0.276	0.367	6.47
	TeLM	0.132	**0.107**	0.132	0.186	10.79	0.255	0.185	0.283	0.381	12.94
	TeAST	0.123	0.098	0.126	0.168	5.59	0.251	0.182	0.279	0.381	6.76
	TCompoundE	0.131	0.095	0.136	0.205	5.40	0.254	0.180	0.276	0.401	6.47
	TeRDy	0.127	0.086	0.131	0.200	5.40	0.253	0.178	0.274	0.399	6.47
	***TDSym* (ours)**	**0.140**	0.100	**0.139**	**0.210**	5.52	**0.262**	**0.190**	**0.293**	**0.414**	6.60

Table 6: Link prediction results on YAGO11k and wikidata12k grouped by embedding rank.

producible comparison across methods. For our method, we implement the proposed symplectic temporal modeling framework within the same training pipeline. All models are trained using identical data splits, negative sampling strategies, and evaluation protocols. The learning rate is fixed to 0.1, and the batch size is set to 1000 for all experiments. The embedding dimension is selected from {32, 64}, and we report results using the best-performing configuration on the validation set. To prevent overfitting, early stopping is applied with a maximum of 200 training epochs. All experiments are conducted on a single NVIDIA Tesla A100 GPU, and random seeds are fixed to ensure reproducibility.

E Results on YAGO11k and Wikidata12k

On YAGO11k and Wikidata12k, which are characterized by sparse timestamps and long-term factual validity, *TDSym* continues to demonstrate strong performance. In particular, on Wikidata12k with rank 64, *TDSym* achieves the highest MRR of 0.262 and Hits@1 of 0.190, surpassing high-capacity baselines like TeLM (0.255 MRR). In contrast, despite their larger parameter budgets, models such as TeLM often exhibit signs of overfitting on these sparse datasets. These results suggest that the proposed Time-Aware Parameter Modulation mechanism enables *TDSym* to adapt robustly to irregular and long-range temporal patterns without requiring excessive model depth.

F Dynamics and Liouville's Theorem

A fundamental property of Hamiltonian systems is the conservation of phase space volume, known as Liouville's Theorem. This property is crucial for learning long-term dependencies, as it prevents the vanishing or exploding of state information over time.

To visually demonstrate this capability, we conduct a particle cloud evolution experiment. We initialize a cluster of $N = 6000$ particles sampling from a Gaussian distribution $\mathcal{N}(\mathbf{0}, 0.3\mathbf{I})$ in the phase plane (q_0, p_0) at $\tau = 0$. We then evolve these particles forward to $\tau = 30$ using both *TDSym* and a non-symplectic baseline (diagonal scaling without the Cayley constraint).

Figure 4 presents the evolution results. Under the symplectic dynamics of *TDSym* (Figure 4, center), the particle cloud is stretched and rotated due to the learned Hamiltonian vector field. However, the total area occupied by the cloud remains invariant, with an area ratio of 1.000 relative to the initial state. This confirms that our model acts as a valid volume-preserving flow. Conversely, the non-symplectic baseline (Figure 4, right) exhibits severe volume contraction, with the area ratio dropping to 0.106. This "volume collapse" physically corresponds to the dissipation of information, which theoretically leads to the vanishing gradient problem in recurrent architectures. By preserving phase-space volume, *TDSym* helps mitigate information dissipation and supports effective gradient propagation over long temporal horizons.

G Inference Efficiency

Table 7 reports the inference time on the ICEWS14 test set with rank 64. Despite incorporating structured symplectic dynamics, *TDSym* incurs only a modest runtime overhead, requiring approximately $\times 1.16$ the inference time of the fastest baseline (TCompoundE). This demonstrates that the pro-

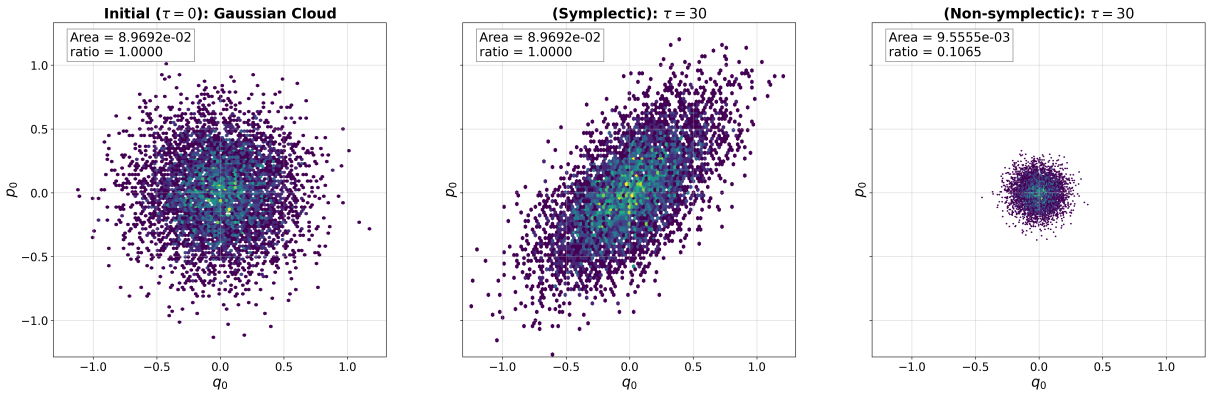


Figure 4: Phase space volume preservation analysis. We visualize the evolution of a particle cloud in the (q, p) phase plane. **(Left)** Initial Gaussian distribution at $\tau = 0$. **(Center)** Under *TDSym*, the cloud at $\tau = 30$ is deformed but strictly preserves its area (ratio ≈ 1.0), confirming adherence to Liouville's Theorem. **(Right)** A non-symplectic baseline exhibits severe volume collapse (ratio ≈ 0.1), indicating information dissipation.

Model	Time (ms)	Relative Time
TCompoundE	2201.53	×1.00
TeAST	2211.81	×1.00
TeRDy	2213.67	×1.01
TComplEx	2213.83	×1.01
TeLM	2245.17	×1.02
***TDSym* (Ours)**	**2547.61**	**×1.16**

Table 7: Inference time comparison on the ICEWS14 test set (rank = 64). Relative runtime is normalized by the fastest baseline, TCompoundE.

posed model achieves substantially improved predictive performance while maintaining competitive inference efficiency.

H Long-term Numerical Stability Analysis

To examine the long-term numerical behavior induced by the Cayley-based evolution, we conduct a recursive simulation experiment. Unconstrained recurrent models are known to exhibit numerical instability when the spectral radius of the transition operator deviates from unity. Accordingly, we compare *TDSym* against a synthetic Unconstrained Linear Baseline ($\mathbf{h}_{t+1} = \mathbf{W} \cdot \mathbf{h}_t$), where $\mathbf{W}$ is initialized with a spectral radius $\rho = 1.05$ to illustrate the typical instability of generic linear dynamics.

As shown in Figure 5, the unconstrained baseline exhibits characteristic exponential divergence under recursive application. In contrast, *TDSym* maintains bounded numerical behavior over 1,000 steps, avoiding both explosion and vanishing of entity representations. The resulting trajectories display quasi-periodic patterns, consistent with the structure-preserving nature of the learned dynamics.

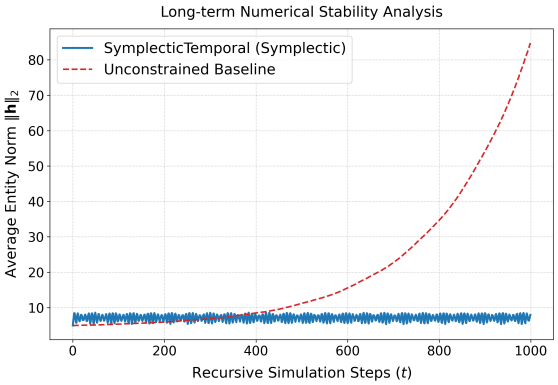


Figure 5: Long-term numerical stability analysis over 1,000 recursive steps. The *TDSym* model (Blue) exhibits bounded oscillatory behavior, characteristic of Hamiltonian dynamics where phase-space volume is strictly preserved.

This behavior can be attributed to the geometric constraints imposed by the symplectic formulation. The Cayley transform $\mathbf{M} = (\mathbf{I} - \frac{1}{2}\mathbf{A}_{r,\tau})^{-1}(\mathbf{I} + \frac{1}{2}\mathbf{A}_{r,\tau})$ enforces symplectic structure by construction, restricting the evolution to volume-preserving transformations. While the Cayley map is closely related to the implicit midpoint scheme used in geometric integration (Hairer et al., 2006), we emphasize that our approach does not aim to provide formal stability guarantees for arbitrary generators. Rather, the preserved symplectic structure empirically promotes numerically stable behavior under long-term recursive application of the learned evolution operators.

I Theoretical Analysis of Relation Pattern Modeling

In this section, we analyze the expressiveness of *TDSym* with respect to relation pattern modeling. Following established literature in temporal knowledge graph embedding (Chen et al., 2022; Li et al., 2023), we adopt the standard definitions of four fundamental relation patterns: Symmetry, Asymmetry, Inversion, and Temporal Evolution. We then provide formal proofs showing that the proposed symplectic evolution operator, $\mathbf{M}_{r,\tau} \in \mathrm{Sp}(2d)$, has sufficient algebraic capacity to model each of these patterns.

I.1 Definitions of Relation Patterns

Following previous studies, we define the key relation patterns as follows.

Definition 1 (Symmetry)**.** *A relation r is symmetric if, for all timestamps τ, the existence of a fact implies its reverse:*

$$r(h,t,\tau) \implies r(t,h,\tau). \tag{16}$$

Definition 2 (Asymmetry)**.** *A relation r is asymmetric if, for all timestamps τ, the existence of a fact implies the negation of its reverse:*

$$r(h,t,\tau) \implies \neg r(t,h,\tau). \tag{17}$$

Definition 3 (Inversion)**.** *A relation r_1 is the inverse of relation r_2 if, for all timestamps τ,*

$$r_1(h,t,\tau) \iff r_2(t,h,\tau). \tag{18}$$

Definition 4 (Temporal Evolution)**.** *A relation r evolves over time if its interaction logic changes between timestamps τ_1 and τ_2, such that*

$$r(h,t,\tau_1) \not\Leftrightarrow r(h,t,\tau_2). \tag{19}$$

I.2 Modeling Capabilities of *TDSym*

Recall that the scoring function of *TDSym* is defined as

$$f_r(h,t,\tau) = \langle \mathbf{M}_{r,\tau}\mathbf{h}, \mathbf{t} \rangle_\omega = (\mathbf{M}_{r,\tau}\mathbf{h})^\top \mathbf{J}\mathbf{t}, \tag{20}$$

where $\mathbf{h}, \mathbf{t} \in \mathbb{R}^{2d}$ are the embeddings of the head entity h and tail entity t, respectively, and $\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix}$ is the canonical symplectic matrix. The matrix $\mathbf{M}_{r,\tau}$ lies in the symplectic group $\mathrm{Sp}(2d)$. We now show that this formulation can model the four relation patterns defined above.

I.2.1 Modeling Symmetry

Proposition 3. *TDSym can model symmetric relations.*

Proof. To model symmetry, it suffices to show that there exists a valid symplectic operator such that

$$f_r(h,t,\tau) = f_r(t,h,\tau). \tag{21}$$

For all entity embeddings $\mathbf{h}, \mathbf{t}$. Consider the choice $\mathbf{M}_{r,\tau} = \mathbf{J}$. We first verify that $\mathbf{J} \in \mathrm{Sp}(2d)$. Since $\mathbf{J}^\top = -\mathbf{J}$ and $\mathbf{J}^2 = -\mathbf{I}$, we have

$$\mathbf{J}^\top \mathbf{J}\mathbf{J} = (-\mathbf{J})\mathbf{J}^2 = (-\mathbf{J})(-\mathbf{I}) = \mathbf{J}. \tag{22}$$

Hence, $\mathbf{J}$ is a valid symplectic matrix. Under this choice, the score becomes

$$f_r(h,t,\tau) = (\mathbf{J}\mathbf{h})^\top \mathbf{J}\mathbf{t} = \mathbf{h}^\top \mathbf{J}^\top \mathbf{J}\mathbf{t} = \mathbf{h}^\top \mathbf{t}. \tag{23}$$

Similarly,

$$f_r(t,h,\tau) = (\mathbf{J}\mathbf{t})^\top \mathbf{J}\mathbf{h} = \mathbf{t}^\top \mathbf{h} = \mathbf{h}^\top \mathbf{t}. \tag{24}$$

Therefore,

$$f_r(h,t,\tau) = f_r(t,h,\tau),$$

which shows that *TDSym* can model symmetric relation pattern. □

I.2.2 Modeling Asymmetry

Proposition 4. *TDSym can model asymmetric relations.*

Proof. Asymmetry follows naturally from the skew-symmetry of the symplectic form. Consider the identity evolution

$$\mathbf{M}_{r,\tau} = \mathbf{I}, \tag{25}$$

which is a valid symplectic matrix. Then the score of the forward fact is

$$f_r(h,t,\tau) = \mathbf{h}^\top \mathbf{J}\mathbf{t}. \tag{26}$$

For the reversed fact, we have

$$f_r(t,h,\tau) = \mathbf{t}^\top \mathbf{J}\mathbf{h}. \tag{27}$$

Using the skew-symmetry of $\mathbf{J}$, namely $\mathbf{J}^\top = -\mathbf{J}$, we obtain

$$\mathbf{t}^\top \mathbf{J}\mathbf{h} = -\mathbf{h}^\top \mathbf{J}\mathbf{t}. \tag{28}$$

Hence,

$$f_r(t,h,\tau) = -f_r(h,t,\tau). \tag{29}$$

Therefore, whenever $f_r(h,t,\tau) \neq 0$, the model assigns opposite scores to a fact and its reverse, which is sufficient to model asymmetric relation pattern. □

I.2.3 Modeling Inversion

Proposition 5. *TDSym can model inverse relation pairs* (r_1, r_2).

Proof. To model inversion, we need

$$f_{r_1}(h, t, \tau) = f_{r_2}(t, h, \tau). \tag{30}$$

Let $\mathbf{M}_{r_1,\tau} \in \mathrm{Sp}(2d)$ be the symplectic operator associated with relation r_1. We construct the operator of the inverse relation r_2 as

$$\mathbf{M}_{r_2,\tau} = -\mathbf{M}_{r_1,\tau}^{-1}. \tag{31}$$

Since $\mathbf{M}_{r_1,\tau}^{-1} \in \mathrm{Sp}(2d)$ and $-\mathbf{I} \in \mathrm{Sp}(2d)$, it follows that

$$\mathbf{M}_{r_2,\tau} \in \mathrm{Sp}(2d). \tag{32}$$

Now consider the score of the reversed fact under relation r_2:

$$\begin{aligned} f_{r_2}(t, h, \tau) &= (\mathbf{M}_{r_2,\tau}\mathbf{t})^\top \mathbf{J}\mathbf{h} \\ &= (-\mathbf{M}_{r_1,\tau}^{-1}\mathbf{t})^\top \mathbf{J}\mathbf{h} \\ &= -\mathbf{t}^\top \mathbf{M}_{r_1,\tau}^{-\top}\mathbf{J}\mathbf{h}. \end{aligned} \tag{33}$$

Because $\mathbf{M}_{r_1,\tau}$ is symplectic, it satisfies $\mathbf{M}_{r_1,\tau}^{-\top}\mathbf{J} = \mathbf{J}\mathbf{M}_{r_1,\tau}$. Thus,

$$f_{r_2}(t, h, \tau) = -\mathbf{t}^\top \mathbf{J}\mathbf{M}_{r_1,\tau}\mathbf{h}. \tag{34}$$

Using the skew-symmetry of the symplectic form,

$$\mathbf{t}^\top \mathbf{J}\mathbf{M}_{r_1,\tau}\mathbf{h} = -(\mathbf{M}_{r_1,\tau}\mathbf{h})^\top \mathbf{J}\mathbf{t}, \tag{35}$$

we obtain

$$\begin{aligned} f_{r_2}(t, h, \tau) &= (\mathbf{M}_{r_1,\tau}\mathbf{h})^\top \mathbf{J}\mathbf{t} \\ &= f_{r_1}(h, t, \tau). \end{aligned} \tag{36}$$

Therefore, *TDSym* can model inverse relation pattern. □

I.2.4 Modeling Temporal Evolution

Proposition 6. *TDSym can model relations that evolve dynamically over time.*

Proof. In *TDSym*, the relation operator explicitly depends on the timestamp τ:

$$\mathbf{M}_{r,\tau} = \mathrm{Cay}(\mathbf{A}_{r,\tau}), \tag{37}$$

where $\mathbf{A}_{r,\tau}$ is determined by the time-aware coefficients $\boldsymbol{\alpha}_{r,\tau}$ and $\boldsymbol{\beta}_{r,\tau}$. These coefficients are generated from the relation embedding together with time-dependent modulation signals, including RoTE-based continuous temporal encoding.

Therefore, for two timestamps $\tau_1 \neq \tau_2$, the model can realize

$$\boldsymbol{\alpha}_{r,\tau_1} \neq \boldsymbol{\alpha}_{r,\tau_2} \quad \text{or} \quad \boldsymbol{\beta}_{r,\tau_1} \neq \boldsymbol{\beta}_{r,\tau_2}, \tag{38}$$

which implies

$$\mathbf{A}_{r,\tau_1} \neq \mathbf{A}_{r,\tau_2} \quad \Longrightarrow \quad \mathbf{M}_{r,\tau_1} \neq \mathbf{M}_{r,\tau_2}. \tag{39}$$

Hence, the scoring function becomes time-dependent:

$$f_r(h, t, \tau_1) \neq f_r(h, t, \tau_2). \tag{40}$$

Moreover, because the continuous RoTE component varies smoothly with τ, the induced family of operators $\mathbf{M}_{r,\tau}$ can be viewed as a continuous time-dependent trajectory on the symplectic manifold whenever the Cayley transform is well-defined. Thus, *TDSym* is able to capture changing relation behavior across time rather than representing relations as static mappings. □