FlorE: Integrating Full Lorentz Group and Directional Offsets for Effective Knowledge Graph Embedding

Zehua Duo[1,2,3], **Jiang Li**[1,2,3], **Xiangdong Su**[*1,2,3], **Guanglai Gao**[1,2,3],
[1]College of Computer Science, Inner Mongolia University
[2]National & Local Joint Engineering Research Center of Intelligent Information Processing Technology for Mongolian
[3]Inner Mongolia Key Laboratory of Multilingual Artificial Intelligence Technology
32309021@mail.imu.edu.cn

Abstract

Knowledge Graph Embedding (KGE) aims to map entities and relationships into a continuous vector space to facilitate reasoning and downstream tasks. Although previous KGE methods based on Euclidean, complex spaces, or hyperbolic spaces have performed well, they still struggle to effectively model Z-Paradox relation patterns, which account for a large proportion in each knowledge graph. To address this issue, we propose a novel KGE method **FlorE** which integrates the full Lorentz Group and directional offset operation in hyperbolic space for the KGE task. Specifically, we incorporate the full Lorentz Group to enable the same relation in the knowledge graph (KG) to perform indefinite isometry, thus avoiding the overlapping of entities. Meanwhile, we implement a directional offset operation via exponential mapping to transform the relations to the same Lorentz manifold of the entities, thus maintaining geometric consistency for the relations and entities in KG. By integrating these two techniques, FlorE can effectively model the Z-Paradox relation patterns and improve the representation learning ability for KGs. Experiments on the five benchmark datasets demonstrate that our method achieves state-of-the-art performance. For the Z-Paradox relation patterns, the improvement achieves **26.7%**, **15.6%**, **35.4%**, **33.7%**, and **31.5%** on FB15k-237, WN18RR, CoDEx-S, CoDEx-M and CoDEx-L, respectively.

Code — https://github.com/dzh597/FlorE

Introduction

The knowledge graph (KG) serves as a structured representation of entities and semantic relations among them and plays an important role in the fields of natural language processing (Ji et al. 2021), recommender systems (Zhang et al. 2025), and intelligent question answering (Ma et al. 2025). Knowledge Graph Embedding (KGE) aims to map the entities and relations in a KG into a continuous vector space to provide computation support for downstream tasks (Ge et al. 2024). Early KGE methods relied primarily on Euclidean space, such as the TransE model (Bordes et al. 2013), which models the semantic structure between triples based on distance. However, as graph structures become more complex, Euclidean space exhibits limitations in capturing hierarchical patterns (Nickel and Kiela 2017). To address this, later research introduced hyperbolic space due to its natural advantage in modeling hierarchical knowledge graphs.

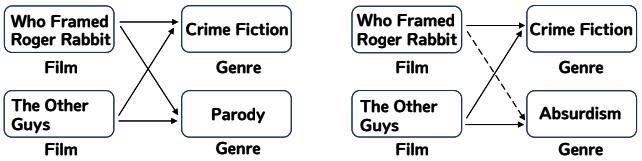


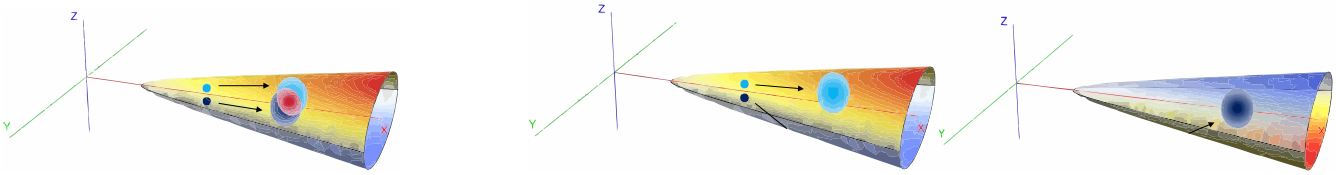
Figure 1: Illustration of Z-paradox in FB15k-237. The arrow indicates the relation /film/film/genre, , which means *the genre of the film is*. Solid lines represent actual facts, while dotted lines highlight erroneous predictions caused by Z-paradox.

Despite their strengths in handling complex structures and long-range dependencies, hyperbolic space-based KGE models still struggle with the Z-Paradox relation pattern, a structural ambiguity first proposed in Liu et al. (2024). This pattern arises when multiple head entities share a common tail under the same relation, e.g., (h_1, r, t_1), (h_2, r, t_1), and (h_2, r, t_2), leading many models to falsely infer (h_1, r, t_2) due to overly simplistic geometric assumptions. Figure 1 provides an illustrative example of this relation pattern. This transitive misgeneralization arises from translation or rotation biases (Li and Yang 2022; Song, Luo, and Huang 2021) and affects Euclidean, complex(Trouillon et al. 2016; Sun et al. 2019), and hyperbolic models (Ganea, Bécigneul, and Hofmann 2018; Zheng et al. 2024; Fan et al. 2024; Park et al. 2025; Lin et al. 2025).

A typical manifestation occurs when identical relations applied to different head entities yield overlapping geometric mappings, causing confusion between tail entities. Figure 2a provides a visual illustration of this problem under the positive Lorentz Group. The small circles on the left represent head entities, the arrows indicate the relation, and the large circles on the right denote tail entities. This setup highlights how different head entities sharing the same relation may incorrectly point to overlapping tail entities. As exemplified in Figure 1, the entities *Parody* and *Absurdism* are both located in the red-shaded region.

To address this issue, we propose **FlorE**, a new embed-

*Corresponding author.



(a) Tail confusion due to overlapping mappings under the positive Lorentz Group.

(b) The full Lorentz Group separates the mappings. The right figure is the result of rotating the left figure 180 degrees along the x-axis.

Figure 2: Geometric illustration of the effect of extending from the positive Lorentz Group to the full Lorentz Group.

ding framework that extends the relational transformation to the **f**ull **Lore**ntz Group, thus expanding the expression range of the Lorentz embedding model. As shown in Figure 2b, the full Lorentz Group introduces the reflection transformation with negative determinant, which enables the mapping results of the same relation to be separated on the manifold, and significantly alleviates the tail entity confusion. In particular, it allows tail entities associated with different head entities under the same relation to be properly distinguished rather than incorrectly overlapped. As illustrated in Figure 1, *Parody* and *Absurdism* are positioned in distinct regions, specifically the light blue and dark blue areas. However, this degrades the modeling of other relation patterns. Although the relations are represented in the full Lorentz Group, the entities are still embedded in the Lorentz model, which is defined only by the single-leaf hyperboloid. This mismatch causes a structural dissonance between the structure on which the relation transformations depend and the geometric basis in which the entities are embedded.

To avoid interference with other relation patterns, this paper proposes a directional offset operation that models each relation as an offset vector with direction (Zhuo and Rossignac 2012; Fang et al. 2025) and implements directed geometric offsets from entities in the Lorentz manifold. This operation not only ensures the legitimacy of the offset operation within the manifold but also effectively encodes the directional character of the relations, thus better preserving the model's ability to represent and reason about other relation patterns while maintaining the geometric consistency of the entities and relations.

Related Work

Euclidean space is the first geometric structure applied to knowledge graph embedding. TransE (Bordes et al. 2013) models relations as vector translations, which is intuitive and computationally efficient, but struggles with many-to-many, antisymmetric, and non-commutative structures (Xiao et al. 2015). DistMult (Yang et al. 2014) uses a bilinear scoring function and has seen expressive improvements, though its inherent symmetry limitation remains prominent (Vashishth et al. 2020). TuckER (Balažević, Allen, and Hospedales 2019b) introduces a tensor kernel mechanism to capture richer entity-relation interactions. Although these methods improve expressive capacity, they largely focus on triple-level modeling and overlook complex indirect structural patterns. Moreover, even when extending European embedding directly to quaternion representations (Li et al. 2022a, 2024b), it remains challenging to discern the Z-paradox relational patterns. MQuinE (Liu et al. 2024) is the only Euclidean method explicitly designed to mitigate the Z-Paradox through head-tail interaction modeling, effectively addressing reasoning errors caused by structural overlap.

ComplEx (Trouillon et al. 2016) first employed complex-valued embeddings to represent symmetric and antisymmetric relations via real and imaginary parts (Chao et al. 2020), while RotatE (Sun et al. 2019) modeled relations as rotations on the complex plane to enhance compositionality. Concurrently, DuASE (Li et al. 2024a)similarly employs a complex space and characterises relational patterns with reduced modelling complexity, serving as a comparable baseline for complex domain methods .Later extensions, such as QuatE (Zhang et al. 2019) and HAKE (Zhang et al. 2020b), introduced quaternion and modulus-phase spaces to improve isomorphism representation, and RulE (Tang et al. 2024) incorporated rule-based reasoning. QuatE-D (Fazael-Ardakani and Soltanian-Zadeh 2025) further enhances quaternion models with a distance-based scoring mechanism (Zhang et al. 2020a). Despite these advances, most of these models lack the ability to perceive indirect structural patterns (Abboud et al. 2020; Wang et al. 2019), making them prone to tail confusion in Z-Paradox scenarios where different head entities share a relation to overlapping tails.

Hyperbolic geometry has gained traction for its suitability in modeling hierarchical and long-range dependencies due to its negative curvature (Pan and Wang 2021). MuRP (Balazevic, Allen, and Hospedales 2019) introduces relational translation in the Poincaré model, while LorentzKG (Fan et al. 2024) improves numerical stability and expressiveness via the Lorentz model. Recent work incorporates 3D rotation and translation into hyperbolic space (Zhu and Shimodaira 2024), and tensor factorization has been extended with hyperbolic interaction components to better model hierarchical link structures (Yusupov, Rakhuba, and Frolov 2025; Balažević, Allen, and Hospedales 2019a). However, most approaches emphasize local embedding consistency and lack structural discrimination mechanisms, resulting in persistent inference errors under Z-Paradox patterns. This paper introduces a full Lorentz Group transformation to enhance the model's ability to distinguish structurally confusable cases and reduce such geometric misclassifications.

Methodology

Full Lorentz Group For Relation Modeling

Lorentz embedding spaces are constructed based on the Lorentz model, which provides a natural geometric framework for representing hyperbolic geometry via the upper sheet of a two-sheeted hyperboloid.

The Lorentz model $\mathbb{L}^n$ is defined as a submanifold in Minkowski space $\mathbb{R}^{1,n}$:

$$\mathbb{L}^n := \left\{ \mathbf{x} \in \mathbb{R}^{n+1} : \langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{L}} = -1,\ x_0 > 0 \right\}, \quad (1)$$

where the Lorentz inner product is defined as:

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}} = -x_0 y_0 + \sum_{i=1}^{n} x_i y_i, \quad (2)$$

this ensures that points lie on the upper sheet of the n-dimensional hyperboloid. The full Lorentz Group $\mathrm{O}(1, n)$ is the group of all linear transformations that preserve the Lorentz metric. It is defined as:

$$\mathrm{O}(1, n) = \left\{ A \in \mathrm{GL}(n+1, \mathbb{R}), A^\top J_n A = J_n \right\}, \quad (3)$$

where $J_n = \mathrm{diag}(-1, 1, \ldots, 1)$ is the Minkowski signature matrix. This group includes continuous and discrete transformations, such as rotations, boosts, and reflections.

For any two points $\mathbf{x}, \mathbf{y} \in \mathbb{L}^n$, the squared Lorentz distance is given by:

$$d^2_{\mathcal{L}}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle_{\mathcal{L}} = -2 - 2\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}}. \quad (4)$$

With these definitions established, we now describe how relation transformations are implemented under the full Lorentz Group $\mathrm{O}(1, k)$.

To implement relation transformations under the full Lorentz Group $\mathrm{O}(1, k)$, we apply orthogonal transformations only to the spatial dimensions of the embeddings of the entities while keeping the time dimension unchanged. Specifically, the transformation matrix Λ_r for each relation r is defined as:

$$\Lambda_r = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R_r \cdot D_r \end{bmatrix}, \quad (5)$$

where $R_r \in \mathrm{O}(k)$ is an orthogonal matrix that captures spatial rotations and reflections. The matrix R_r is constructed through a sequence of learnable reflections (Li et al. 2022b). Let $\{v_i\}_{i=1}^m$ be a set of learnable vectors, then:

$$R_r = \prod_{i=1}^{m} \left(I - \frac{2 v_i v_i^\top}{\|v_i\|^2} \right), \quad (6)$$

this construction guarantees that R_r is within the orthogonal group and satisfies $R_r^\top R_r = I$.

Next, D_r is a diagonal matrix that is used to introduce direction-sensitive reflection. It is defined as $D_r = \mathrm{diag}(1, \ldots, \sigma_r)$, where σ_r is a scalar that can be learned in the interval $(-1, 1)$. This scalar controls whether the final spatial dimension undergoes a directional flip. This is implemented by multiplying the last column of R_r by σ_r, allowing the model to reflect across arbitrary directions and angles.

Together, the spatial transformation $R_r \cdot D_r$ allows the full matrix Λ_r to belong to $\mathrm{O}(1, k)$ while avoiding the use of Lorentz boosts (Fan et al. 2024). In this way, we achieve an implementation of indefinite isometry that accommodates richer relational transformations.

Finally, the transformed head entity under relation r is computed as:

$$\mathbf{h}_r = \Lambda_r \cdot \mathbf{h}, \quad (7)$$

where $\mathbf{h}$ is the original Lorentz embedding of the head entity. Since Λ_r preserves the Lorentz metric, the transformed entity $\mathbf{h}_r$ remains on the Lorentz manifold and satisfies $\langle \mathbf{h}_r, \mathbf{h}_r \rangle_{\mathcal{L}} = -1$.

Compared to models based solely on the positive Lorentz Group, our use of the full Lorentz Group offers broader geometric expressiveness and allows relation directions to be redefined through reflective operations. This provides a more stable and flexible foundation for modeling structural phenomena such as the Z-Paradox.

Directional Offset

Building on the transformation matrix of relation Λ_r from the full Lorentz Group $\mathrm{O}(1, k)$, we further introduce a directional offset operation to derive representations of the relation.

Each relation $r \in \mathcal{R}$ is represented by two learnable vectors: a center point $\mathbf{c}_r \in \mathbb{H}^n$, specifying the position of the relation on the manifold; and a direction vector $\mathbf{d}_r \in \mathbb{R}^{n+1}$, specifying its geometric direction. To map $\mathbf{d}_r$ to a valid tangent direction, we first project it orthogonally to the tangent space $T_{\mathbf{c}_r}\mathbb{H}^n$:

$$\alpha_r = \langle \mathbf{c}_r, \mathbf{d}_r \rangle_{\mathcal{L}}, \quad \boldsymbol{\xi}_r = \mathbf{d}_r + \alpha_r \cdot \mathbf{c}_r, \quad \hat{\boldsymbol{\xi}}_r = \frac{\boldsymbol{\xi}_r}{\|\boldsymbol{\xi}_r\|_{\mathcal{L}} + \varepsilon}, \quad (8)$$

where $\|\cdot\|_{\mathcal{L}}$ denotes the Lorentz norm, and ε is a small constant to avoid division by zero. The unit tangent vector $\hat{\boldsymbol{\xi}}_r$ captures the directional tendency of the relation.

To control the scale of geometric displacement, we introduce a scalar $\lambda_r > 0$ that is learnable for each relation and apply exponential mapping to generate the displacement point $\mathbf{q}_r$ on the manifold:

$$\begin{aligned} \mathbf{q}_r &= \exp_{\mathbf{c}_r}(\lambda_r \cdot \hat{\boldsymbol{\xi}}_r) \\ &= \cosh\left(\frac{\lambda_r}{\sqrt{k}}\right) \mathbf{c}_r + \sqrt{k} \sinh\left(\frac{\lambda_r}{\sqrt{k}}\right) \cdot \hat{\boldsymbol{\xi}}_r. \end{aligned} \quad (9)$$

where $k > 0$ is the absolute value of the negative curvature of the hyperbolic space. The resulting $\mathbf{q}_r$ encodes the geometric position and direction of relation.

To combine this displacement with entity embeddings while preserving manifold consistency, we do not use Euclidean vector addition. Instead, we use a log–transport–exp process akin to Möbius addition to apply $\mathbf{q}_r$ to an entity point $\mathbf{x} \in \mathbb{H}^n$.

We first log-map $\mathbf{q}_r$ to the tangent space at the origin $\mathbf{O} = (1, 0, \ldots, 0)$:

$$\boldsymbol{\zeta} = \log_{\mathbf{O}}(\mathbf{q}_r) = \frac{\mathrm{arcosh}(-\langle \mathbf{q}_r, \mathbf{O} \rangle_{\mathcal{L}})}{\sqrt{\langle \mathbf{q}_r, \mathbf{q}_r \rangle_{\mathcal{L}} + 1}} \cdot (\mathbf{q}_r - \langle \mathbf{q}_r, \mathbf{O} \rangle_{\mathcal{L}} \cdot \mathbf{O}). \quad (10)$$

Then, we parallel transport $\boldsymbol{\zeta}$ to the tangent space at $\mathbf{x}$:

$$\boldsymbol{\zeta}' = \boldsymbol{\zeta} + \frac{\langle \mathbf{x}, \boldsymbol{\zeta} \rangle_{\mathcal{L}}}{k - \langle \mathbf{x}, \mathbf{O} \rangle_{\mathcal{L}}} \cdot (\mathbf{x} + \mathbf{O}). \tag{11}$$

Finally, we map the result back to the manifold via the exponential mapping at $\mathbf{x}$:

$$\mathbf{x}_{\text{out}} = \exp_{\mathbf{x}} \left(\log_{\mathbf{O}}(\mathbf{q}_r) + \frac{\langle \mathbf{x}, \log_{\mathbf{O}}(\mathbf{q}_r) \rangle_{\mathcal{L}}}{k - \langle \mathbf{x}, \mathbf{O} \rangle_{\mathcal{L}}} \cdot (\mathbf{x} + \mathbf{O}) \right) \tag{12}$$

This process ensures geometrically consistent, curvature-aware displacement on the Lorentz manifold.

Score Function

To accurately model the semantic connectivity between entity pairs in a knowledge graph, we propose a scoring function based on the full Lorentz Group. This function integrates relation-guided geometric transformations, spatial proximity alignment, and directional consistency. It jointly evaluates whether a tail entity lies in the correct spatial location and along the correct directional path defined by the relation, thus improving the model's capacity to represent complex relational structures, especially in challenging scenarios such as the Z-Paradox.

Given any triple (h, r, t), the scoring function is defined as:

$$f(h,r,t) = -d^2_{\mathcal{L}}(\Lambda_{r,1}h,\ \Lambda_{r,2}t) + \gamma \cdot \langle \hat{\boldsymbol{\xi}}_r,\ \log_{\Lambda_{r,1}h}(t) \rangle_{\mathcal{L}} + b_h + b_t + \delta. \tag{13}$$

Here, $\Lambda_{r,1}h$ denotes the transformed embedding of the head entity h under relation r, and $\Lambda_{r,2}t$ represents the geometrically transformed embedding of the tail entity t. The term $d^2_{\mathcal{L}}(\cdot,\cdot)$ denotes the squared Lorentz distance, capturing spatial closeness between transformed head and tail entities. The vector $\hat{\boldsymbol{\xi}}_r$ is a unit directional offset associated with the relation r at its geometric center, while $\log_{\Lambda_{r,1}h}(t)$ denotes the tangent direction from the transformed head to the tail entity. Their Lorentz inner product $\langle \hat{\boldsymbol{\xi}}_r,\ \log_{\Lambda_{r,1}h}(t) \rangle_{\mathcal{L}}$ measures the directional alignment between the actual position of the tail entity and the expected direction of the relation r.

This formulation combines proximity and directional criteria. While the distance term encourages proximity, the direction term ensures that the tail entity appears in the correct semantic direction. To further illustrate how the scoring function geometrically captures factuality, especially in resolving structural ambiguities such as the Z-Paradox, we present a detailed example below.

To further illustrate how the scoring function distinguishes the factuality of structural relations in a geometric manner, we construct a simplified example in two-dimensional Lorentz space, demonstrating the necessity of the directional term in addressing the Z-Paradox.

We consider the space $\mathbb{R}^{1,1}$, where the Lorentz inner product is defined as:

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}} = -x_0 y_0 + x_1 y_1$$

We set the head entity as $h = (3, 0)$, and the relation direction vector as $\hat{\boldsymbol{\xi}}_r = (0, 1)$, i.e., pointing in the positive spatial direction. Two candidate tail entities are $t_1 = (4, 1)$ and $t_2 = (5, 3)$, both lying on the Lorentz manifold and satisfying $\langle t_1, t_1 \rangle_{\mathcal{L}} = -16$ and $\langle t_2, t_2 \rangle_{\mathcal{L}} = -16$.

We first compute the Lorentz squared distance between the head and each tail using:

$$d^2_{\mathcal{L}}(h,t) = -2 - 2\langle h, t \rangle_{\mathcal{L}}$$

Thus, we obtain:

$$\langle h, t_1 \rangle_{\mathcal{L}} = -3 \cdot 4 + 0 \cdot 1 = -12 \quad \Rightarrow \quad d^2(h, t_1) = 22$$
$$\langle h, t_2 \rangle_{\mathcal{L}} = -3 \cdot 5 + 0 \cdot 3 = -15 \quad \Rightarrow \quad d^2(h, t_2) = 28$$

From the perspective of distance, t_1 is clearly closer to h and would be preferred by models that rely solely on geometric proximity.

However, when incorporating the directional term, the outcome changes. We approximate the direction vector as $\log_h(t) \approx t - h$ and compute the Lorentz inner product with the relation direction:

$$\log_h(t_1) = (1, 1), \quad \langle \log_h(t_1), \hat{\boldsymbol{\xi}}_r \rangle_{\mathcal{L}} = 1$$
$$\log_h(t_2) = (2, 3), \quad \langle \log_h(t_2), \hat{\boldsymbol{\xi}}_r \rangle_{\mathcal{L}} = 3$$

In this illustrative example, we omit the bias terms b_h, b_t, and the δ to focus on the geometric and directional components.

By setting $\gamma = 4$, the final scores are:

$$f(h, r, t_1) = -22 + 4 \cdot 1 = -18,$$
$$f(h, r, t_2) = -28 + 4 \cdot 3 = -16.$$

Although t_2 is farther from h, its direction better aligns with the relation, yielding a higher score. The model thus correctly infers $h \to t_2$, avoiding the misjudgment caused by the Z-Paradox. This demonstrates that geometric distance alone is insufficient and that directional consistency is crucial for robust relational inference, especially in handling Z-Paradox relation patterns.

Experiment

Experimental Setup

Baseline. To evaluate the effectiveness of our proposed approach, we compare it against a set of representative baseline models. These baselines can be categorized based on their embedding space into the following three types:

- **Euclidean space**: including Tucker (Balažević, Allen, and Hospedales 2019b), HAKE (Zhang et al. 2020b), SpeedE (Pavlović and Sallinger 2024), and SMART(Amouzouvi et al. 2025), which model relations using classical translation and rotation mechanisms in flat geometry.
- **Complex space**: including ComplEx (Trouillon et al. 2016), RotatE (Sun et al. 2019), DCNE (Dong et al. 2024), QuatE-MI (Li et al. 2025), and ComplEX-DR (Kim, Park, and Kim 2025), which extend embedding representations using complex or quaternion-valued vectors to better capture asymmetry and compositionality.

Datasets	#Entities	#Relations	#Train	#Valid	#Test						
					#Total	#Sym.	#Inv.	#Com.	#Inj.	#Non-Inj.	#Z-Par.
FB15k-237	14541	237	272115	17535	20446	385	411	4491	1891	18575	6076
WN18RR	40943	11	86835	3034	3134	1112	11	19	928	2206	601
CoDEx-S	2034	42	32888	1827	1828	274	4	0	23	1805	583
CoDEx-M	17050	51	185584	10310	10310	366	11	23	203	10108	3193
CoDEx-L	77951	69	551193	30622	30622	874	114	62142	776	29846	9062

Table 1: Statistics of the FB15k-237, WN18RR, CoDEx-S, CoDEx-M and CoDEx-L.

Space	Model	FB15k-237				WN18RR			
		MRR ↑	H@1 ↑	H@3 ↑	H@10 ↑	MRR ↑	H@1 ↑	H@3 ↑	H@10 ↑
$\mathbb{E}$	Tucker(2019)	35.8	26.6	39.4	54.4	47.0	44.3	48.2	52.6
	HAKE(2020)	34.6	25.0	38.1	54.2	41.6	38.9	42.7	46.7
	SpeedE(2024)	32.0	22.7	35.6	50.4	49.3	44.6	51.2	58.4
	SMART(2025)	29.8	20.5	33.0	48.7	43.8	39.7	45.1	52.2
$\mathbb{C}$	ComplEx(2016)	25.7	16.5	29.3	44.3	43.2	39.6	45.2	50.0
	RotatE(2019)	29.0	20.8	31.6	54.8	37.8	33.0	41.7	49.1
	DCNE(2024)	35.4	25.7	29.3	54.7	49.2	44.8	51.0	58.1
	QuatE-MI(2025)	34.8	24.8	38.2	55.0	48.8	43.8	50.8	58.2
	ComplEX-DR(2025)	<u>38.8</u>	<u>29.5</u>	-	57.3	<u>50.3</u>	<u>46.0</u>	-	58.8
$\mathbb{H}$	RotH(2019)	31.4	22.3	34.6	49.7	47.2	42.8	49.0	55.3
	HyboNet(2022)	33.4	24.4	36.5	51.6	48.9	45.5	50.3	55.3
	LorentzKG(2024)	38.4	28.7	<u>42.2</u>	<u>57.9</u>	50.2	45.6	<u>52.3</u>	<u>58.9</u>
	MIG-TF(2025)	36.6	27.7	40.2	55.3	49.9	45.2	51.4	57.4
	FlorE(ours)	**40.2**	**30.9**	**44.3**	**60.7**	**50.9**	**46.5**	**53.7**	**61.0**

Table 2: Performances of various KGE models on WN18RR and FB15k-237 datasets, with most results obtained from (Fan et al. 2024) and original papers (Balazevic, Allen, and Hospedales 2019; Pavlović and Sallinger 2024; Dong et al. 2024). $\mathbb{E}$, $\mathbb{C}$ and $\mathbb{H}$ are denoted as the presentation space of Euclidean, complex and hyperbolic space, respectively. Our model uses an embedding dimension of 32. The best result in terms of each metric is shown in **bold**, and the second best one is <u>underlined</u>. ↑: higher is better; ↓: lower is better. "–" denotes unavailable entries. The same settings are applied to Table 3 and Table 4.

Model	CoDEx-S			CoDEx-M			CoDEx-L		
	MRR ↑	H@1 ↑	H@10 ↑	MRR ↑	H@1 ↑	H@10 ↑	MRR ↑	H@1 ↑	H@10 ↑
RESCAL♣	40.4	29.3	62.3	31.7	24.4	45.6	30.4	24.2	41.9
TransE♣	35.4	21.9	63.4	30.3	22.3	45.4	18.7	11.6	31.7
ComplEx♣	<u>46.5</u>	<u>37.2</u>	64.6	33.7	26.2	47.6	29.4	23.7	40.0
ConvE♣	44.4	34.3	63.5	31.8	23.9	46.4	30.3	24.0	42.0
TuckER♣	44.4	33.9	63.8	32.8	25.9	45.8	30.9	24.4	43.0
MuRP♠	42.0	31.1	63.2	30.6	22.6	45.6	-	-	-
ATTH♠	40.2	28.6	63.2	31.5	23.7	46.4	-	-	-
CoPE♠	44.6	35.0	63.1	32.6	25.1	46.6	-	-	-
LorentzKG	45.5	36.7	<u>65.4</u>	<u>37.4</u>	<u>27.5</u>	<u>51.3</u>	<u>33.7</u>	<u>25.6</u>	<u>44.9</u>
FlorE(ours)	**48.7**	**39.3**	**74.9**	**40.2**	**30.5**	**55.7**	**35.9**	**28.0**	**47.1**

Table 3: Performances of various KGE models on CoDEx-S, CoDEx-M and CoDEx-L datasets. ♣ indicates that the results are from (Liu et al. 2024), and ♠ indicates that the results are from (Liang et al. 2024). We adopt different baseline models because the selected baselines are commonly used in prior works, ensuring fair and consistent comparison with existing studies.

- **Hyperbolic space**: including RotH (Chami et al. 2020), HyboNet (Chen et al. 2021), and LorentzKG (Fan et al. 2024), and MIG-TF (Yusupov, Rakhuba, and Frolov 2025), which leverages negatively curved geometry to encode hierarchical and structural patterns in knowledge graphs.

Dataset. We evaluated the proposed model on five benchmark knowledge graph datasets: WN18RR (Dettmers et al. 2018), FB15k-237 (Toutanova et al. 2015), CoDEx-S, CoDEx-M, and CoDEx-L (Safavi and Koutra 2020), with dataset statistics summarized in Table 1.

Evaluation. To evaluate the performance of the link prediction task under the filtering setting (Bordes et al. 2013), we ranked the test triples by ranking them against all other candidate triples.

Implementation Details. Our model is implemented in PyTorch and trained on a single 80 GB NVIDIA A100 GPU.

Space	Dataset	FB15k-237						WB18RR					
	Relation Partten	Sym.	Inv.	com.	Inj.	Non-Inj.	**Z-Par.**	Sym.	Inv.	com.	Inj.	Non-Inj.	**Z-Par.**
	Proportion %	1.8	2.0	22.0	9.2	90.1	29.7	35.5	0.3	0.6	30.0	70.3	19.1
$\mathbb{E}$	Tucker(2019)	36.3	35.7	29.4	30.8	37.5	19.4	84.6	43.4	88.2	37.5	46.9	21.0
	HAKE(2020)	35.9	36.5	27.1	29.6	36.8	21.4	83.7	44.0	87.1	36.3	45.2	22.5
	SpeedE(2024)	33.9	34.2	27.5	29.0	36.4	18.4	81.1	42.3	85.6	36.7	44.8	23.4
	SMART(2025)	32.5	31.8	26.1	27.3	35.5	19.4	81.2	42.0	82.4	35.2	43.6	24.7
$\mathbb{C}$	ComplEx(2016)	33.6	29.2	22.4	26.5	33.8	20.4	82.4	38.0	80.1	32.7	40.9	22.3
	RotatE(2019)	31.4	32.7	25.6	28.4	35.2	19.2	80.1	41.8	83.1	35.3	43.5	25.4
	DCNE(2024)	33.7	35.9	28.4	31.2	38.1	<u>25.6</u>	82.3	43.5	87.7	37.8	47.1	<u>30.2</u>
	QuatE-MI(2025)	35.0	36.3	26.7	29.8	36.5	22.9	83.8	42.1	86.3	35.4	45.6	27.7
	ComplEX-DR(2025)	37.5	**38.6**	29.7	<u>33.4</u>	<u>41.2</u>	22.7	86.0	**50.3**	<u>92.2</u>	36.4	**50.2**	26.3
$\mathbb{H}$	RotH(2019)	34.6	36.0	28.6	30.1	34.5	23.5	81.6	44.5	84.9	38.5	44.7	26.4
	HyboNet(2022)	29.8	32.3	26.6	32.1	33.5	22.6	78.4	38.7	72.5	<u>40.3</u>	45.3	29.6
	LorentzKG(2024)	**39.1**	38.2	**31.7**	32.9	**42.1**	24.0	**88.5**	44.5	**92.7**	39.5	<u>49.8</u>	29.4
	MIG-TF(2025)	35.4	35.8	28.9	31.5	38.3	23.2	82.7	38.4	85.9	37.9	47.3	27.6
	FlorE(ours)	<u>38.5</u>	<u>38.4</u>	<u>30.3</u>	**33.8**	39.6	**52.3**	<u>86.4</u>	<u>45.3</u>	86.7	**41.9**	48.9	**45.8**

Table 4: Relation Pattern-wise MRR Analysis on FB15k-237 and WN18RR.

Embeddings are initialized on the manifold using the Geoopt library (Kochurov, Karimov, and Kozlukov 2020). We perform a grid search over the learning rate, negative sample size, batch size, and margin to determine optimal hyperparameters. The best settings are WN18RR: 0.0343, 200, 512, 1.01 and FB15k-237: 0.0072, 100, 512, 1.07.

Main Result

Table 2 presents the link prediction performance of our proposed FlorE model compared to mainstream KGE methods on five benchmark datasets, spanning Euclidean, complex, and hyperbolic embedding spaces. FlorE consistently outperforms baselines across all evaluation metrics (MRR, Hits@1, Hits@10), validating its robustness under diverse geometries and data distributions.

On FB15k-237, FlorE achieves an MRR of 0.402, surpassing LorentzKG by 1.8%. Hits@1 improves by 2.2%, and Hits@3 and Hits@10 by 2–3 points, reflecting its superior capacity for modeling complex relational structures. On the simpler WN18RR, FlorE maintains its lead with an MRR of 0.509 versus 0.502 for LorentzKG, and achieves 0.9 and 2.1 point gains on Hits@1 and Hits@10, respectively. Even with strong baselines, FlorE shows clear advantages.

To assess generalization, we evaluated FlorE on the CoDEx dataset series: CoDEx-S, CoDEx-M, and CoDEx-L. As shown in Table 3, FlorE achieves the best results across all metrics. On CoDEx-S, it reaches 48.7% MRR, with 39.3% Hits@1 and 74.9% Hits@10. On CoDEx-M, it scores 40.2% MRR, 30.5% Hits@1, and 55.7% Hits@10, outperforming LorentzKG and CoPE. Even on the large-scale CoDEx-L, FlorE maintains its lead with 35.9% MRR, 28.0% Hits@1, and 47.1% Hits@10, demonstrating strong generalization on complex graphs.

To evaluate FlorE's structural modeling capacity, we analyze its performance across six relation patterns: Symmetric, Inverse, Composition, Injective, Non-Injective, and Z-Paradox. Table 4 shows FlorE's clear advantage, particularly in addressing the challenging Z-Paradox. On FB15k-237 and WN18RR, FlorE achieves 52.3% and 45.8% accuracy on Z-Paradox relations, while the best-performing baselines reach only 25.6% and 30.2%, yielding absolute gains of 26.7% and 15.6%, respectively. The effectiveness of our method is positively correlated with the prevalence of Z-Paradox relations in the dataset. Detailed MRR results on Z-Paradox for the CoDEx-S, CoDEx-M, and CoDEx-L datasets are provided in **Appendix A**.

Although FlorE is slightly outperformed in some patterns (e.g., Sym. or Com.), it consistently matches or exceeds state-of-the-art results on Inverse and Injective relations, which are key for structural reasoning. This balance highlights FlorE's enhanced expressiveness, especially under structurally ambiguous scenarios such as the Z-Paradox.

Ablation Study

We conducted ablation experiments on two benchmark knowledge graph datasets, FB15k-237 and WN18RR, to assess the individual and combined effects of the full Lorentz Group and Directional Offset modules. We analyzed both link prediction performance and relation structure patterns. The results are presented in Tables 5 and 6.

The results show that FlorE w/o DO is the primary source of performance gains, while FlorE w/o FLG enhances precision and supports fine-grained relational reasoning. Notably, the complete model achieves the best results across all metrics, confirming the complementary nature of the two components. Significant improvements in Hits@1 and Hits@3 further suggest that the directional offset enhances the precise localization of target entities. **Appendix B** provides further analysis.

Case Study

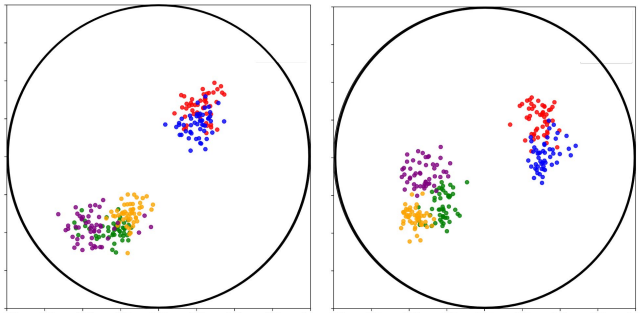
To validate the effectiveness of the full Lorentz Group in mitigating inference ambiguity, we conducted a case study on FB15k-237 to compare the distribution of tail entity embeddings between models using only the positive Lorentz Group and those using the full Lorentz Group. The goal is to examine whether the full Lorentz Group better distinguishes overlapping tail entities under shared relations, thereby alleviating the Z-Paradox. Figure 3 illustrates the tail entity distributions learned by two models using 2D Poincaré disk projections. Relation group definitions are provided in **Appendix C**.

Module	FB15k-237				WB18RR			
	MRR ↑	H@1 ↑	H@3 ↑	H@10 ↑	MRR ↑	H@1 ↑	H@3↑	H@10 ↑
FLorE w/o DO	37.3	27.6	41.5	55.8	48.4	43.8	50.1	59.3
FLorE w/o FLG	24.1	14.2	26.4	44.9	32.8	21.2	39.0	45.8
FLorE	40.2	30.9	44.3	60.7	50.9	46.5	53.7	61.0

Table 5: Ablation results on FB15k-237 and WN18RR.

Dataset	FB15k-237						WN18RR					
Relation Partten	Sym	Inv	com	Inj	Non-Inj	**Z-Par**	Sym	Inv	com	Inj	Non-Inj	**Z-Par**
FLorE w/o DO	28.4	32.3	22.6	32.4	34.8	45.1	78.4	38.7	72.5	40.3	52.7	53.0
FLorE w/o FLG	27.9	27.5	21.8	24.3	25.7	21.6	62.9	31.8	69.0	29.4	31.2	27.3
FLorE(ours)	38.5	38.4	30.3	33.8	39.6	42.3	86.4	45.8	86.7	41.9	47.0	52.8

Table 6: Relation Pattern-wise MRR in the Ablation Study on FB15k-237 and WN18RR.



(a) Positive Lorentz Group (b) Full Lorentz Group

Figure 3: Embedding of the tail entities on five relations by positive Lorentz Group model (a) and full Lorentz Group model (b). Red, blue, green, yellow, and purple represent groups 1 to 5 respectively.

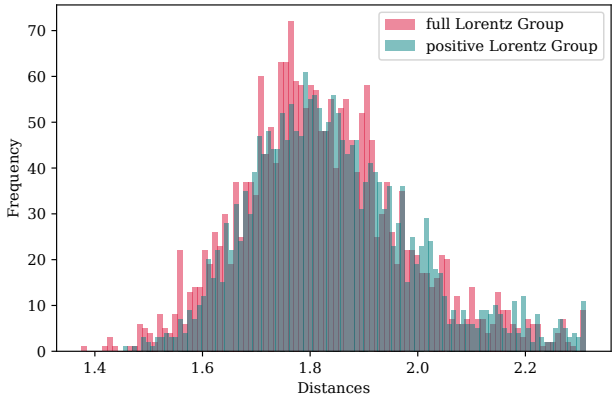


Figure 4: Distribution histogram for a hierarchical relation.

As shown in Figure 3a, the positive Lorentz model yields highly overlapped clusters, especially when multiple relations share tail entities, making it difficult to separate their semantics. In contrast, Figure 3b shows that the full Lorentz Group leads to clearer separation, with distinct relational clusters and reduced interference even in overlapping regions. This stems from the additional transformation degrees of freedom in the full Lorentz Group, which improves geometric expressiveness.

In summary, the full Lorentz Group enables more robust structural modeling by reducing semantic overlap and enhancing relational distinction, ultimately supporting more accurate reasoning in link prediction tasks. Comparisons of model capacity and embedding dimensions are provided in **Appendix D** and **Appendix E**, respectively.

Analysis of Hierarchical Pattern

By analyzing the distribution of entity embedding distances under the _derivationally_related_form relationship in WN18RR, we verified that our proposed full Lorentz Group model enhances the modeling capability of Z-paradox-type relationships without weakening its ability to express hierarchical relationships.

As shown in Figure 4, _derivationally_related_form is a relation that connects entities at the same level in the semantic hierarchy, so its embedding distances should be concentrated in the middle interval. The results indicate that both the full Lorentz Group model and the positive Lorentz Group model exhibit similar distance distribution patterns for this relationship, suggesting that the full Lorentz Group can also effectively capture hierarchical structures. Additionally, further analysis of hierarchical patterns can be found in **Appendix F**.

Conclusion

In this work, we present a knowledge graph embedding framework that leverages the full Lorentz Group and a directional offset operation to address Z-Paradox. By extending the relational transformation space beyond the positive Lorentz Group, our model captures a broader set of geometric operations, enabling better discrimination between structurally similar relations. Furthermore, the directional offset operation provides a flexible and curvature-aware way to incorporate relation directionality while preserving manifold consistency. Experiments on five benchmarks show our method outperforms prior models.

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

Abboud, R.; Ceylan, I.; Lukasiewicz, T.; and Salvatori, T. 2020. Boxe: A box embedding model for knowledge base completion. *Advances in Neural Information Processing Systems*, 33: 9649–9661.

Amouzouvi, K.; Song, B.; Coletta, A.; Bellomarini, L.; Lehmann, J.; and Vahdati, S. 2025. SMART: Relation-Aware Learning of Geometric Representations for Knowledge Graphs. *arXiv preprint arXiv:2507.13001*.

Balazevic, I.; Allen, C.; and Hospedales, T. 2019. Multi-relational poincaré graph embeddings. *Advances in neural information processing systems*, 32.

Balažević, I.; Allen, C.; and Hospedales, T. M. 2019a. Hypernetwork knowledge graph embeddings. In *International conference on artificial neural networks*, 553–565. Springer.

Balažević, I.; Allen, C.; and Hospedales, T. M. 2019b. Tucker: Tensor factorization for knowledge graph completion. *arXiv preprint arXiv:1901.09590*.

Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; and Yakhnenko, O. 2013. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.

Chami, I.; Wolf, A.; Juan, D.-C.; Sala, F.; Ravi, S.; and Ré, C. 2020. Low-dimensional hyperbolic knowledge graph embeddings. *arXiv preprint arXiv:2005.00545*.

Chao, L.; He, J.; Wang, T.; and Chu, W. 2020. Pairre: Knowledge graph embeddings via paired relation vectors. *arXiv preprint arXiv:2011.03798*.

Chen, W.; Han, X.; Lin, Y.; Zhao, H.; Liu, Z.; Li, P.; Sun, M.; and Zhou, J. 2021. Fully hyperbolic neural networks. *arXiv preprint arXiv:2105.14686*.

Dettmers, T.; Minervini, P.; Stenetorp, P.; and Riedel, S. 2018. Convolutional 2d knowledge graph embeddings. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.

Dong, Y.; Kong, Q.; Wang, L.; and Luo, Y. 2024. Dual complex number knowledge graph embeddings. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 5391–5400.

Fan, X.; Xu, M.; Chen, H.; Chen, Y.; Das, M.; and Yang, H. 2024. Enhancing hyperbolic knowledge graph embeddings via lorentz transformations. In *Findings of the Association for Computational Linguistics ACL 2024*, 4575–4589.

Fang, Y.; Liu, X.; Lu, W.; Pedrycz, W.; Lang, Q.; and Yang, J. 2025. Knowledge graph completion with low-dimensional gated hierarchical hyperbolic embedding. *Knowledge-Based Systems*, 309: 112804.

Fazael-Ardakani, H.-S.; and Soltanian-Zadeh, H. 2025. QuatE-D: A Distance-Based Quaternion Model for Knowledge Graph Embedding. *arXiv preprint arXiv:2504.13983*.

Ganea, O.; Bécigneul, G.; and Hofmann, T. 2018. Hyperbolic neural networks. *Advances in neural information processing systems*, 31.

Ge, X.; Wang, Y. C.; Wang, B.; Kuo, C.-C. J.; et al. 2024. Knowledge graph embedding: An overview. *APSIPA Transactions on Signal and Information Processing*, 13(1).

Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; and Yu, P. S. 2021. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE transactions on neural networks and learning systems*, 33(2): 494–514.

Kim, J.; Park, J.; and Kim, K. 2025. RSCF: Relation-Semantics Consistent Filter for Entity Embedding of Knowledge Graph. *arXiv preprint arXiv:2505.20813*.

Kochurov, M.; Karimov, R.; and Kozlukov, S. 2020. Geoopt: Riemannian optimization in pytorch. *arXiv preprint arXiv:2005.02819*.

Li, J.; Su, X.; Duo, Z.; Lan, T.; Guo, X.; and Gao, G. 2025. A Mutual Information Perspective on Knowledge Graph Embedding. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 22152–22166.

Li, J.; Su, X.; Ma, X.; and Gao, G. 2022a. Quatse: Spherical linear interpolation of quaternion for knowledge graph embeddings. In *CCF International Conference on Natural Language Processing and Chinese Computing*, 209–220. Springer.

Li, J.; Su, X.; Zhang, F.; and Gao, G. 2024a. Learning low-dimensional multi-domain knowledge graph embedding via dual archimedean spirals. In *Findings of the Association for Computational Linguistics ACL 2024*, 1982–1994.

Li, J.; Su, X.; Zhang, F.; and Gao, G. 2024b. TransERR: Translation-based knowledge graph embedding via efficient relation rotation. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 16727–16737.

Li, J.; and Yang, Y. 2022. Star: Knowledge graph embedding by scaling, translation and rotation. In *International Conference on AI and Mobile Services*, 31–45. Springer.

Li, R.; Zhao, J.; Li, C.; He, D.; Wang, Y.; Liu, Y.; Sun, H.; Wang, S.; Deng, W.; Shen, Y.; et al. 2022b. House: Knowledge graph embedding with householder parameterization. In *International conference on machine learning*, 13209–13224. PMLR.

Liang, Q.; Wang, W.; Bao, F.; and Gao, G. 2024. Fully hyperbolic rotation for knowledge graph embedding. In *ECAI 2024*, 1615–1622. IOS Press.

Lin, L.; Qin, H.; Qi, Q.; Gu, R.; Zuo, P.; and Cheng, Y. 2025. Multi-relation-pattern knowledge graph embeddings for link prediction in hyperbolic space. *International Journal of Intelligent Networks*.

Liu, Y.; Fang, H.; Cai, Y.; and Sun, M. 2024. MQuinE: a Cure for "Z-paradox" in Knowledge Graph Embedding. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, 9837–9850.

Ma, C.; Chen, Y.; Wu, T.; Khan, A.; and Wang, H. 2025. Large Language Models Meet Knowledge Graphs for Question Answering: Synthesis and Opportunities. *arXiv preprint arXiv:2505.20099*.

Nickel, M.; and Kiela, D. 2017. Poincaré embeddings for learning hierarchical representations. *Advances in neural information processing systems*, 30.

Pan, Z.; and Wang, P. 2021. Hyperbolic hierarchy-aware knowledge graph embedding for link prediction. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, 2941–2948.

Park, J.; Han, S.; Shin, W.-Y.; and Lim, S. 2025. Metapath-based Hyperbolic Contrastive Learning for Heterogeneous Graph Embedding. *arXiv preprint arXiv:2506.16754*.

Pavlović, A.; and Sallinger, E. 2024. Speede: Euclidean geometric knowledge graph embedding strikes back. In *Findings of the Association for Computational Linguistics: NAACL 2024*, 69–92.

Safavi, T.; and Koutra, D. 2020. Codex: A comprehensive knowledge graph completion benchmark. *arXiv preprint arXiv:2009.07810*.

Song, T.; Luo, J.; and Huang, L. 2021. Rot-pro: Modeling transitivity by projection in knowledge graph embedding. *Advances in Neural Information Processing Systems*, 34: 24695–24706.

Sun, Z.; Deng, Z.-H.; Nie, J.-Y.; and Tang, J. 2019. Rotate: Knowledge graph embedding by relational rotation in complex space. *arXiv preprint arXiv:1902.10197*.

Tang, X.; Zhu, S.-C.; Liang, Y.; and Zhang, M. 2024. RulE: Knowledge Graph Reasoning with Rule Embedding. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Findings of the Association for Computational Linguistics: ACL 2024*, 4316–4335. Bangkok, Thailand: Association for Computational Linguistics.

Toutanova, K.; Chen, D.; Pantel, P.; Poon, H.; Choudhury, P.; and Gamon, M. 2015. Representing text for joint embedding of text and knowledge bases. In *Proceedings of the 2015 conference on empirical methods in natural language processing*, 1499–1509.

Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, É.; and Bouchard, G. 2016. Complex embeddings for simple link prediction. In *International conference on machine learning*, 2071–2080. PMLR.

Vashishth, S.; Sanyal, S.; Nitin, V.; Agrawal, N.; and Talukdar, P. 2020. Interacte: Improving convolution-based knowledge graph embeddings by increasing feature interactions. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 3009–3016.

Wang, Q.; Huang, P.; Wang, H.; Dai, S.; Jiang, W.; Liu, J.; Lyu, Y.; Zhu, Y.; and Wu, H. 2019. Coke: Contextualized knowledge graph embedding. *arXiv preprint arXiv:1911.02168*.

Xiao, H.; Huang, M.; Hao, Y.; and Zhu, X. 2015. TransG: A generative mixture model for knowledge graph embedding. *arXiv preprint arXiv:1509.05488*.

Yang, B.; Yih, W.-t.; He, X.; Gao, J.; and Deng, L. 2014. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575*.

Yusupov, V.; Rakhuba, M.; and Frolov, E. 2025. Knowledge Graph Completion with Mixed Geometry Tensor Factorization. *arXiv preprint arXiv:2504.02589*.

Zhang, H.; Wang, D.; Sun, Z.; Li, Y.; Sun, Y.; Liang, H.; and Wang, W. 2025. KG4RecEval: Does Knowledge Graph Really Matter for Recommender Systems? *ACM Transactions on Information Systems*, 43(3): 1–36.

Zhang, S.; Tay, Y.; Yao, L.; and Liu, Q. 2019. Quaternion knowledge graph embeddings. *Advances in neural information processing systems*, 32.

Zhang, Y.; Yao, Q.; Dai, W.; and Chen, L. 2020a. AutoSF: Searching scoring functions for knowledge graph embedding. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, 433–444. IEEE.

Zhang, Z.; Cai, J.; Zhang, Y.; and Wang, J. 2020b. Learning hierarchy-aware knowledge graph embeddings for link prediction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 3065–3072.

Zheng, Z.; Zhou, B.; Yang, H.; Tan, Z.; Waaler, A.; Kharlamov, E.; and Soylu, A. 2024. Low-dimensional hyperbolic knowledge graph embedding for better extrapolation to under-represented data. In *European Semantic Web Conference*, 100–120. Springer.

Zhu, Y.; and Shimodaira, H. 2024. 3D Rotation and Translation for Hyperbolic Knowledge Graph Embedding. In Graham, Y.; and Purver, M., eds., *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1497–1515. St. Julian's, Malta: Association for Computational Linguistics.

Zhuo, W.; and Rossignac, J. 2012. Curvature-based offset distance: Implementations and applications. *Computers & Graphics*, 36(5): 445–454.