Caving Knowledge Graphs in Latent Space with the Similarity Group

Kossi Amouzouvi[a, g, *, 1], **Yashrajsinh Chudasama**[d,e,1], **Jens Lehmann**[b], **Disha Purohit**[d,e,1], **Ariam Rivas**[d,1], **Bowen Song**[f], **Sahar Vahdati**[c] **and Maria-Esther Vidal**[d,e]

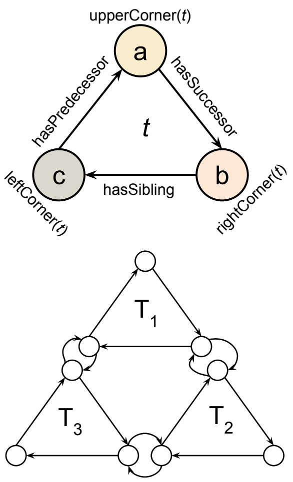
[a]Technische Universität Dresden, Dresden, Germany
[b]Amazon (work done outside of Amazon), Dresden, Germany
[c]Institute for Applied Informatics (InfAI), Leipzig, Germany
[d]Leibniz University Hannover, Hannover, Germany
[e]TIB Leibniz Information Centre for Science and Technology, Hannover, Germany
[f]Chinese University of Geosciences, Wuhan, China
[g]Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
ORCID (Kossi Amouzouvi): https://orcid.org/0000-0002-3614-0449, ORCID (Yashrajsinh Chudasama): https://orcid.org/0000-0003-3422-366X, ORCID (Jens Lehmann): https://orcid.org/0000-0001-9108-4278, ORCID (Disha Purohit): https://orcid.org/0000-0002-1442-335X, ORCID (Ariam Rivas): https://orcid.org/0000-0003-4916-5220, ORCID (Bowen Song): https://orcid.org/0000-0001-9287-2874, ORCID (Sahar Vahdati): https://orcid.org/0000-0002-7171-169X, ORCID (Maria-Esther Vidal): https://orcid.org/0000-0003-1160-8727

Abstract. Knowledge Graphs (KGs), with their intricate hierarchies and semantic relationships, present unique challenges for graph representation learning, necessitating tailored approaches to effectively capture and encode their complex structures into useful numerical representations. The fractal-like nature of these graphs, where patterns repeat at various scales and complexities, requires specialized algorithms that can adapt and learn from the multi-level structures inherent in the data. This similarity to fractals requires methods that preserve the recursive detail of knowledge graphs while facilitating efficient learning and extraction of relational patterns. In this study, we explore the integration of similarity group with attention mechanisms to represent knowledge graphs in complex spaces. In our approach, *SimE*, we make use of the algebraic (bijection) and geometric (similarity) properties of the similarity transformations to enhance the representation of self-similar fractals in KGs. We empirically validate the capability of providing representations of bijections and similarities in benchmark KGs. We also conducted controlled experiments that captured one-to-one, one-to-many, and many-to-many relational patterns and studied the behavior of state-of-the-art models including the proposed SimE model. Moreover, we created a set of fractal-like testbeds to assess the subgraph similarity learning ability of models. The observed results suggest that SimE captures the complex geometric structures of KGs whose statements satisfy these algebraic and geometric properties. In particular, SimE is competitive with state-of-the-art KG embedding models and is able to achieve high values of Hits@1. As a result, SimE is capable of effectively predicting correct links and ranking them with the highest ranks.

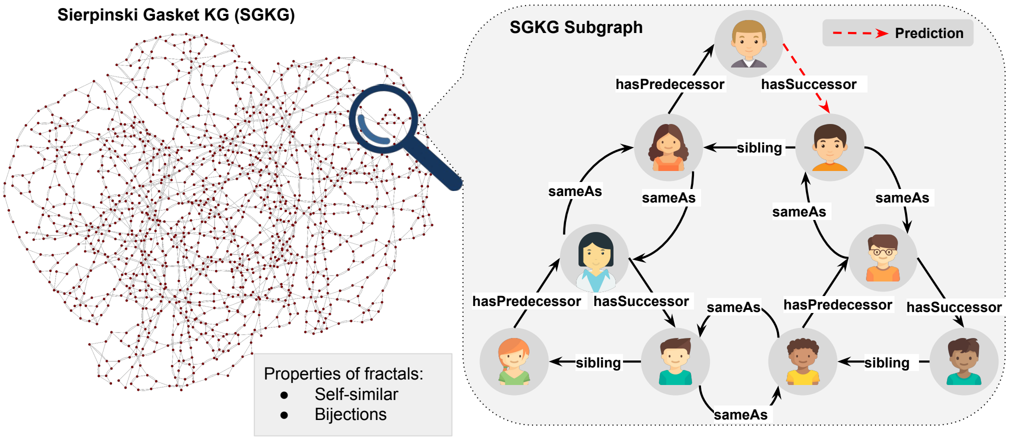
* Corresponding Author. Email: kossi.amouzouvi@tu-dresden.de.
[1] Equal contribution.

1 Introduction

Knowledge Graphs (KGs) continue to stand as an important pillar of AI-based approaches by providing a source of structured and verifiable knowledge. However, KGs in their symbolic form are not directly usable by most machine learning algorithms because these algorithms require numerical input to perform computations. Symbolic representations need to be converted into appropriate numerical forms (e.g., vectors) to enable the application of statistical techniques that can efficiently process and learn from the data [23, 24]. This necessitates the use of representation learning methods with an ability to preserve most of the symbolic structure of the KG, including relational and structural patterns. The objective of knowledge graph representation learning is to accurately transform these characteristics from the symbolic domain to a vector space. This is essential for enhancing both the functionality and reliability of systems across a variety of domains, including web search, recommender systems, and automated reasoning [13, 14, 26, 30, 32]. However, the effectiveness of these models is often limited due to the diverse structures present in KGs. A type of representation learning approaches namely Knowledge Graph embedding (KGE) models gained attention in their geometrical and algebraic design, where the preservation power lies on. KGEs perform link prediction by embedding entities and relations of a knowledge graph into a continuous vector space, where it learns to predict missing relations based on the observed geometric or semantic closeness of these embeddings. There has been a long list of KGE models [12], each of which can preserve a subset of structures including symmetry, composition, or hierarchies and tree-like structures. However, KGs encompass a variety of other structural types that have not been extensively explored. For instance, self-similar structures, which are analogous to *Fractals* in



(a) A triangle and one Sierpinski triangle with three Sierpinski triangles



Link Prediction: < head, relation, ? > Triple: < at_11663, hasSuccessor, bt_11663 >	Knowledge Graph Embedding (KGE) Models				
	TransE	ComplEx	RotE	AttE	SimE
Rank	14	3696	2	2	**1**

(b) Motivating Example

Figure 1: Figure 1a shows a triangle t with relationships *hasSibling*, *hasPredecessor*, *hasSuccessor* and with *upperCorner*(t), *leftrCorner*(t), and *rightCorner*(t). Furthermore, figure 1a shows triangles $T1$, $T2$, $T3$, and the Sierpinski triangle is composed of these three triangles of the same size. Figure 1b shows a Sierpinski KG (SGKG), of size 243. KGE models (e.g., TransE, ComplEx, RotE, AttE, and SimE) are utilized to perform a prediction task of the missing links between two entities. The illustration above demonstrates that the state-of-the-art KGE models were unable to capture the self-similar structures in fractals, while *SimE* outperforms and shows improved performance.

nature, frequently occur in domains such as biological, ecological, familial-social, and linguistic KGs. These complex patterns, often overlooked, offer significant potential for advancing our understanding and representation of interconnected data. In this work, we investigate self-similar structures on the example of familial-social as well as synthetic KGs comprising of fractals. We exclusively make use of the algebraic properties (bijection) of the similarity transformations to enhance the materialization of the properties represented in KGs.

Motivating Example. A fractal is a geometric structure that comprises self-similar geometric structures; Figure 1a illustrates *Sierpinski triangles*, a particular type of fractals (a formal definition is in Section 3). A fractal knowledge graph (KG) comprises nodes and edges representing *fractals*. Specifically, Figure 1b illustrates a portion of a Sierpinski Gasket KG (SGKG) with the size of 243 triangles. The extracted sub-graph describes an entity in the SGKG with relations, i.e., *hasSibling*, *hasPredecessor*, and *hasSuccessor*. Additionally, the sub-graph represents two similar nodes connected with *sameAs* relationship, resulting in Sierpinski triangles. In the real world, KGs operate on Open World Assumption (OWA) for the representation of knowledge, thus Figure 1b demonstrates missing information about a particular entity. For instance, completing the missing relations between entities would be relating two entities, i.e., at_11663 and bt_11663 using relation *hasSuccessor*, which represents the correct factual statement ⟨at_11663, hasSuccessor, bt_116633⟩ that is the missing link in the exemplar KG, further, utilized in the completion of KG.

Predicting the correct factual statement at ranking one is essential because it directly assesses the accuracy of the model's predictions, which is crucial for tasks such as link prediction and KG completion. Figure 1b illustrates the state-of-art approaches that worked on the mechanism of numerical learning methods (a.k.a., KGE models). KGE models predict the missing links via the representation of KGs in low-dimensional numerical vectors. Further, such latent vector is utilized to perform the link prediction task, i.e., analyzing the already known factual statements to deduce the missing ones. For instance, the translation-distance model, i.e., TransE predicts the missing tail at rank 14, whereas models working on complex hyperplane, i.e., ComplEx, and rotation mechanisms, i.e., RotE predicts the tail entity at rank 3969 and 2 respectively. Consequently, these techniques may not cover relevant knowledge in the neighborhood of an entity and fail to consider the properties of fractals such as self-similarity and bijection by randomly choosing facts.

In this paper, we present *SimE*, a KGE model that transforms the embedding vectors of entities and their relationships using Euclidean space to measure the plausibility of the triple. *SimE* is motivated by the lack of exploitation of the numerical learning methods to capture Sierpinski triangles. *SimE* utilizes the self-similar fractals to create valid candidate triples for a given Link Prediction (LP) task. As illustrated in the Figure 1b, *SimE* predicts the tail entity at rank 1, outperforming the state-of-the-art KGE models that are unable to detect the self-similar structures and bijections in the KGs. The experimental study reveals the impact and key role of *SimE* in capturing the self-similar structures of fractals, therefore increasing the performance of LP tasks. Thus, we utilize Hits@1 to quantify the percentage of times the predicted top-ranked plausible entity aligned with the actual relationship in the KG, as well as to report whenever the predicted correct links are ranked with the highest ranking scores.

2 Related Work

KGE models can be classified into tensor decomposition models, deep learning-based models, and geometric models based on the interaction dynamics of their embeddings. Among those, the geometric models can be further categorized based on the type of geometric space they utilize, such as Euclidean, complex, or hyperbolic spaces. In any of these embedding spaces, these models employ geometric methods such as rotation, reflection, scaling, and translation. In this work, we focus on geometric models, their spaces, and methods for the preservation of self-similar structures. Therefore, the rest of this section presents related work in terms of pattern preservation with regard to the KGE model designs so far.

Relational Pattern Preservation. Geometric models regard relations as geometric transformations between subject and object entities. TransE [6], a pioneer KGE model encoded relations by translations from subject entities to object entities in Euclidean space. Although it remains one of the simple and better-performing models, it has limitations in preserving certain multi-relational structures. These limitations gave rise to new translation-based models such as TransH [31], TransA [33], TransR [19]. In contrary to normal straight line translation models, RotatE [28] uses the complex products to rotate subject entities. Different from translation models, rotation can preserve symmetric and antisymmetric patterns well. It therefore, populates the Euclidean space by concentric circles and achieves translation of subject entities on circles. Subsequently, QuatE [35] extends the algebraic structure of RotatE to a hypercomplex structure; relation-entity interaction is represented by a quaternion product. DualE [8] uses dual-quaternion product to simultaneously achieve rotation and translation of subject entities by relations which yields in broader relational pattern preservation. DualE and QuatE subsume RotatE, and moreover, they introduce a non-commutative relation representation learning. These works fail to preserve hierarchical and tree structures.

Structural Pattern Preservation. Recently, it has been shown that hyperbolic geometry has the potential to facilitate learning over hierarchical relational patterns. These models define and project relation and entity embeddings onto relation-specific manifolds in a Poincaré ball. A model named MuRP [1] encodes relations by Möbius addition and Möbius matrix-vector multiplication to transform entity embeddings. The importance of using hyperbolic space is also assessed in MuRP by comparing it against its reduced version, named MuRE [1], which uses real addition and real matrix multiplication in Euclidean space. Another KGE model named AttH [11] brings rotation and reflection to the hyperbolic space in order to conjointly learn relational and hierarchical patterns. A variant of AttH in Euclidean space is also proposed dubbed AttE [11]. A recent work in geometry interactions, GIE [9], extends AttH by considering extra patterns in Euclidean and spherical space versions in the hyperbolic space. By interactively learning the spatial structures in the three spaces, GIE is able to simultaneously learn multiple types of structural patterns including hierarchies. Another type of structural pattern, however complex, is self-similar structure or fractal gasket subgraphs. Larsson et al., and Zhang et al. introduce the concept of relational fractal embedding. Several studies [3, 22, 25] have applied the concept of fractals in deep learning across networks. However, the preservation of these structures in KGs is an unexplored problem, which is addressed in this proposed approach.

3 Background and Preliminaries

In this work, we perform knowledge graph representation learning by exploiting graph-relations' embedding with similarity transformations of complex numbers. In this section, we introduce the key mathematical concepts that are used to define our model.

3.1 Similitude

In the complex plane, a similitude (also called similarity) is a bijection defined from the complex plane onto itself, that scales distances between two points by a constant positive real number. That is, if o and o' are similar to s, and s' by a similarity sim, then $\|o' - o\| = \alpha\|s' - s\|$ where $\alpha > 0$ is the scaling factor and $s, s', o, o' \in \mathbb{C}$. The couple (s, o) are analytical related by the complex equation

$$o = a \cdot s + b = \alpha e^{i\theta_a} \cdot \hat{s} + b \tag{1}$$

where $\alpha = |a|$ and θ_a are the modulus and the angle of a, and $\hat{s}$ is either s or its complex conjugate $\bar{s}$. $a, b \in \mathbb{C}$ define the similarity sim. Their values confer different properties to the similarity. sim is an isometry when $\alpha = 1$, otherwise, the product by $|a|$ allows the similarity to achieve scaling transformation. With the term b, the similarity is able to achieve translation transformation. The product of $e^{i\theta_a}$ and s results into a rotation transformation, whereas the product of $e^{i\theta_a}$ and $\bar{s}$ yields a reflection transformation. In the former case, sim is called direct similarity, and it is called indirect similitude in the later case. Direct similitude preserves orientation, while indirect similitude reverses the order. Thus, sim performs rotation or reflection combined with scaling and translation simultaneously.
We denote by $\mathbb{S}_d \subset \mathbb{C}^2$ the set of all direct similarities defined on the complex plane. When endowed with the operation of composition, the set $\mathbb{S}_d$ is a non-Abelian group since $(a, b) \circ (c, d) = (ac, ad + b)$ whereas $(c, d) \circ (a, b) = (ac, bc + d)$. We note that the complex couple (u, v) represents the transformation $u \cdot s + v$.

3.2 Fractal Knowledge Graphs

Although the definition of a fractal is not unanimously accepted among mathematicians, a fractal can be seen as a geometrical shape formed by an identical pattern repeating on an ever decreasing scale. The pattern of a fractal persists at arbitrarily small scales. They are therefore called self-similar geometric shapes. The concept of fractals is present in nature, geometry, and algebra. Examples of fractals in the real world include trees, seashells, and river networks. Algebraic and geometric fractals can be formed by recursively solving analytical equations or repeatedly performing a simple geometric process. The Mandelbrot set and Sierpinski triangle are examples of algebraic and geometric fractals respectively. Although fractals find their applications in graph theory as Sierpinski gasket graphs, and in social networks as Sierpinski graphs, it has never been used for (multi-relational) KGs to the best of our knowledge. We introduce here the first example of a fractal KG, the Sierpinski Gasket KG (SGKG) while discussing the concepts related to its definition.

Knowledge Graphs. Given a set C of countable infinite constants. A *knowledge graph (KG)* is a directed edge-labeled graph $G = (V, E, L)$ where (1) $V \subseteq C$ is a set of nodes, (2) $L \subseteq C$ is a set of edge labels, and (3) $E \subseteq V \times L \times V$ is a set of edges.

Sierpinski triangles are inductively defined as follows:
Base case: A Sierpinski triangle is a triangle t, where upperCorner(t)=a, leftCorner(t)=c, rightCorner(t)=b and size(t)=1.
Inductive step: Given Sierpinski triangles T_1, T_2, and T_3 of the same size, i.e., size(T_1)=size(T_2)=size(T_3). A Sierpinski triangle T_4 is defined as the combination of T_1, T_2, and T_3 (denoted by $T_1 \bigoplus T_2 \bigoplus T_3$), where upperCorner($T_4$)=upperCorner($T_1$), leftCorner($T_4$)=leftCorner($T_3$), rightCorner($T_4$)= rightCorner($T_2$) and size($T_4$)=3*size($T_1$). Figure 1a illustrates two Sierpinski triangles; the one at the top corresponds to a triangle (as defined by

the base case), while the Sierpinski triangle at the bottom is created from Sierpinski triangles T_1, T_2, and T_3 (by the inductive step).

A Sierpinski Gasket KG (a.k.a. SGKG) is a directed edge-labeled graph $SGKG = (V, E, L)$ inductively defined as follows:
Base case: Given a triangle t with upperCorner(t) $= a$, leftCorner(t) $= c$, rightCorner(t) $= b$. The SGKG respective for t, $SGKG_t = (V, E, L)$, where *V = {a,b,c}, L = {hasSuccessor, sibling, hasPredecessor}, and E = {(c, hasSuccessor, a), (a, sibling, b), (b, hasPredecessor, c))}*.
Inductive step: Given three Sierpinski triangles T_1, T_2, and T_3, and their respective three Sierpinski Gasket KGs $SGKG_{T1} = (V_{T1}, E_{T1}, L_{T1})$, $SGKG_{T2} = (V_{T2}, E_{T2}, L_{T2})$, and $SGKG_{T3} = (V_{T3}, E_{T3}, L_{T3})$. Suppose T_4 is a Sierpinski triangle resulting of combining T_1, T_2, and T_3 (i.e., T_4=$T_1 \bigoplus T_2 \bigoplus T_3$). The Sierpinski Gasket KG for T_4, $SGKG_{T4}$= (V_{T4}, E_{T4}, L_{T4}) is defined as follows: $V_{T4} = V_{T1} \cup V_{T2} \cup V_{T3}$, $L_{T4} = L_{T1} \cup L_{T2} \cup L_{T3} \cup \{sameAs\}$, and $E_{T4} = E_{T1} \cup E_{T2} \cup E_{T3} \cup$ {(leftCorner(T_1), sameAs, upperCorner(T_3)), (upperCorner(T_3), sameAs leftCorner(T_1)), (rightCorner(T_1), sameAs, upperCorner(T_2)), (upperCorner(T_2), sameAs, rightCorner(T_1)), (rightCorner(T_3), sameAs, leftCorner(T_2)), (leftCorner(T_2), sameAs, rightCorner(T_3))}.

Knowledge Graph Embedding (KGE). Knowledge Graph Embedding models are capable of transforming entities and their relationships in a KG into a low-dimensional continuous latent vector space that preserves the KG's structure. Given a directed edge-labeled graph *G = (V, E, L)* and set of vectors ζ. A *KGE* of *G* is a pair of mappings (ϵ, σ) such that (1) ϵ: V $\rightarrow \zeta$, i.e., ϵ maps an entity e in V to a vector $\epsilon(e)$ in ζ, (2) σ: L $\rightarrow \zeta$, i.e., σ maps a directed edge labeled l to a vector $\sigma(l)$ in ζ. A score function ϕ: V$\times L \times$V $\rightarrow \mathbb{R}$ is used to measure the plausibility of candidate triples represented into latent vector space, triples $\langle s, r, o\rangle$ with the higher score $\phi(s, r, o)$ values conveys better plausibility of the prediction of $\langle s, r, o\rangle$.

4 The SimE Approach

The task of knowledge graph embedding completion consists of predicting the missing subject or object entity of incomplete facts on the form $(s, r, ?)$ or $(?, r, o)$. This tests the ability of the embedding model to distinguish the true entity among a batch of entities. To emphasize this, we extend the work of Sun et al.–in the RotatE model– by designing a model which uses attention mechanics to efficiently discriminate between true and false facts. Furthermore, the model uses similarity transformation which allows it to simultaneously learn rotation, translation, and scaling transformations important for structural and relational pattern preservation.
Let $\mathbf{s}, \mathbf{o} \in \mathbb{C}^n$ represent the n-dimensional complex embeddings of the subject and object entities. Thus, entities have real and imaginary parts which belong to $\mathbb{R}^n$. Relations between entities are represented by similarity transformations. That is, relations are embedded by $\mathbb{S}_d^n$ vectors that is $\mathbf{a}_r, \mathbf{b}_r \in \mathbb{C}^n$ so that a subject entity s is transformed to $\mathbf{a} \odot \mathbf{s} \oplus \mathbf{b}$ by the relation r. By $\odot$ and $\oplus$, we meant element-wise complex product and addition. In the polar form, the complex vector $\mathbf{a}$ becomes $|\mathbf{a}|e^{i\theta_a}$. Thus, $\mathbf{a} \odot \mathbf{s}$ rotates (by an angle θ_a) and then scales (enlarging or shrinking by a uniform scale factor $|\mathbf{a}|$) the subject embedding. It is worth noting that the similarity transformation allows relations to preserve patterns such as symmetry and antisymmetry, composition, hierarchy, and 1-to-1. To accurately predict the missing entities, KGE models assign a score to each test triple. The higher the score, the more plausible the triple. As most KGE models, our model *SimE* uses the Euclidean distance to define triple plausibility. The score of triple (s, r, o) is then given by

$$f(s, r, o) = -\|\tilde{\mathbf{o}} \ominus \mathbf{o}\|^2 \tag{2}$$

where $\tilde{\mathbf{o}}$ is the final-transformed subject and $\ominus$ is the elementwise complex subtraction. Existing KGE models use only the score of triples as a hard constraint to learn entity and relation embeddings in a way that corrupted entities are positioned far away from the target entity. To further facilitate this positional separation, *SimE* uses an attention mechanism in addition to the score function. First, *SimE* defines relation specific attention vector embeddings $\mathbf{r} \in \mathbb{C}^n$ and two intermediate transformed subjects $\mathbf{s}_u = \mathbf{a} \odot \mathbf{s}$ and $\mathbf{s}_d = \bar{\mathbf{a}} \odot \mathbf{s}$. We thought of $\mathbf{s}_u$ and $\mathbf{s}_d$ as resulting from an upward- and a downward-rotation respectively, since the former uses the angle-vector θ_a and the later uses $-\theta_a$. The reason behind the use of $\bar{a}$ for the second transformation is to empower the model with the ability to emphasize more on separating true from corrupted subjects.
These three embeddings $\mathbf{r}, \mathbf{s}_u$, and $\mathbf{s}_d$ are then combined to compute two attention coefficients,

$$\alpha_k^r(s) = \frac{e^{\mathbf{r}^T \mathbf{s}_k}}{e^{\mathbf{r}^T \mathbf{s}_u} + e^{\mathbf{r}^T \mathbf{s}_d}} \tag{3}$$

where $k \in \{u, d\}$ and $\mathbf{r}^T \mathbf{s}_k$ is the inner product of the two vectors. The final-transformed subject is given by

$$\tilde{\mathbf{o}} = \mathbf{b} \oplus \alpha_u^r(s)\mathbf{s}_u \oplus \alpha_d^r(s)\mathbf{s}_d. \tag{4}$$

The two attention coefficients sum to 1. The range of their values induces three scenarios: $\alpha_u \gg \alpha_d, \alpha_u \ll \alpha_d, \alpha_u \simeq \alpha_d$;, where $\gg$ ($\ll$) is the 'much greater (less) than' symbol. For a given relation and subject entity, these three scenarios mean the relation pays more attention to the upward-, downward-, and identity-rotation in transforming the subject entity, respectively. By identity rotation, we meant the angle of rotation is close to zero. We then define the overall adherence of relations to the type of rotation transformation by

$$\begin{aligned} \mathrm{adh}_u &= \frac{100}{N_E} \sum_{(s,r,o) \in \mathcal{KG}} \delta_{\{\alpha_u^r(s) \geq \alpha_d^r(s)\}}, \\ \mathrm{adh}_d &= 100 - \mathrm{adh}_u, \end{aligned} \tag{5}$$

where N_E and δ denote the cardinality of the set E, and the Kronecker delta function. A relation whose adherence adh_u is almost equal to 100 (or 0) is a relation that consistently uses the upward- (or downward-) rotation transformation. Otherwise, it is reduced to scaling and translation transformations.
We finally compare the transformed subject and the object. In order to train the model, the cross entropy loss is used with uniform negative sampling, where negative examples from (s, r, o) are sampled from all possible triples generated by replacing subject or object entities. True negative samples are filtered out from the negative sample batches. We used the cross-entropy loss function

$$loss = \sum_{(s,r,o) \in \mathcal{KG}} (-\log(\frac{f(s,r,o)}{\sum\limits_{o' \in \mathcal{E}} f(s,r,o')}) - \log(\frac{f(s,r,o)}{\sum\limits_{s' \in \mathcal{E}} f(s',r,o)})) \tag{6}$$

to evaluate the accuracy of the model. We denoted by s' or o' the corrupted subject or object entities obtain from the triple (s, r, o).
Theoretical Analysis of *SimE* on Subsumption. Our model *SimE* subsumes some existing KGE models. This means a difference in the performance of these models compared to *SimE* is a consequence

Table 1: **Statistics of Benchmarks**. The table on the left represents the triple counts of four real-world benchmarks and four synthetic benchmarks. Conversely, the Table on the right illustrates the triple counts for the four real-world benchmarks with their relationship types.

Benchmarks	Triples	Entities	Relations
French Royalty	12,554	2,656	12
Family	28,358	3,007	12
YAGO3-10	1,179,040	123,182	37
WN18RR	93,003	40,943	11
SGKG 1	72,720	36,450	4
SGKG 2	54,585	36,450	4
SGKG 3	54,585	36,450	6
SGKG 4	72,720	36,450	6

Relation Type	Benchmarks			
	French Royalty $n_V \mid n_L \mid n_E$	Family $n_V \mid n_L \mid n_E$	YAGO3-10 $n_V \mid n_L \mid n_E$	WN18RR $n_V \mid n_L \mid n_E$
1:1	2651, 12, 10693	2717, 12, 6887	122406, 38, 340992	40835, 11, 71015
1:N	1325, 6, 1696	1592, 12, 3462	114444, 34, 915478	26506, 11, 38938
N:1	2271, 11, 7393	1463, 10, 3397	121840, 37, 1042791	38623, 11, 65379
N:M	1675, 5, 3375	2578, 10, 5963	121681, 34, 1002950	40943, 11, 93003

Table 2: Experimental results on four real-world KGs. Results in bold and underlined represent the best and second-best, respectively.

KG	TransE		SimE		RotE		ComplEx		RotatE		RefE		MurE		CP		AttE	
	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1
French Royalty	0.441	0.304	**0.723**	**0.661**	0.713	0.638	0.570	0.516	0.540	0.484	0.610	0.504	0.627	0.520	0.479	0.358	0.627	0.522
Family	0.497	0.263	0.764	0.612	0.586	0.357	**0.906**	**0.862**	0.853	0.789	0.561	0.340	0.624	0.438	0.658	0.519	0.659	0.484
YAGO3-10	0.099	0.046	0.116	0.061	0.105	0.051	**0.381**	**0.306**	0.346	0.266	0.105	0.051	0.115	0.060	**0.381**	0.304	0.113	0.059
WN18RR	0.156	0.031	0.368	0.304	**0.374**	**0.316**	0.292	0.280	0.287	0.279	0.358	0.304	0.367	0.305	0.297	0.283	0.370	0.314

of the difference in their designs. We use the logical consequence symbol to formulate the statement, $m \models M$ which means that the model *M* subsumes the model *m*.

TransE $\models$ ***SimE***. TransE [5] is a KGE model which proposes a geometric interpretation of the latent space and interprets relation vectors as translations in vector space. Analytically, TransE attempts to achieve $\mathbf{s} \oplus \mathbf{r} \approx \mathbf{o}$. TransE can preserve antisymmetric and composition properties of relations, however, it can not naturally model 1-n, n-1, and n-m relationships. Observing that the transformation $\mathbf{s} \mapsto \mathbf{s} \oplus \mathbf{b} \in \mathbb{C}$ is the complex version of TransE, the set of translation transformations in the complex plane is a subspace of the similarity group. This means TransE corresponds to a similarity whose $\mathbf{a} = 1$. Thus, SimE subsumes TransE.

RotatE $\models$ ***SimE***. The KGE model RotatE [27] represents each relation as a rotation from the subject entity to the object entity in the complex latent space. This is expressed as $e^{i\theta} \odot \mathbf{s}$ which is a similarity that satisfies the conditions, $|\mathbf{a}| = 1$, and $\mathbf{b} = 0$. In other words, RotatE uses a subspace of the similarity group to embed relations. This is to say, SimE subsumes RotatE as well. RotatE can efficiently model symmetric relations.

Model Complexity. As mentioned in Cao et al. [7], *SimE* has the same time complexity as other Euclidean models, i.e., $O(n)$ where n is the space dimensionality of the KG entities. However, *SimE* does not require RotE's rotation-translation transformation twice; contrary, *SimE* resorts to an attention mechanism to select rotation or translation. Thus, *SimE* space complexity is $O(N_V n + 2N_L n)$, where N_V: entities' number and N_L relations' number.

5 Experimental Study

KGE models are commonly evaluated using LP tasks. Here, an LP task involves identifying missing entities to complete a triple *(s, r, ?)*. For instance, (*Toni Kroos, playsFor, ?*), where *? is Real Madrid*. To identify plausible entities for an incomplete triple, the score function is obtained to estimate the likelihood of the entities. To evaluate LP, we remove the tail entities from the testing set, calculate the scoring function for each triple in the training data, and rank them from higher to lower. Finally, the average values for *MRR* and *Hits@1* are presented. Overall, our empirical study includes 432 testbeds; and 9 KGE models on 24 benchmark KGs assessed with 2 metrics. Among these, 126 testbeds demonstrate the effectiveness of *SimE* model. We assess the effectiveness of *SimE* to capture knowledge and enrich the KG completion process. In particular, this work explores the following research questions: **RQ1)** What is the effectiveness of *SimE* for each type of relationship in the KG completion task? **RQ2)** How does SimE use the attention vector to embed the entities? **RQ3)** Could SimE take advantage of the geometric properties of similarity transformations to preserve self-similar sub-graphs?

5.1 Benchmarks

To assess the efficacy of SimE in representing bijections in KGs, controlled experiments were conducted utilizing real-world KGs; namely the French Royalty KG [15], the Family KG [16], YAGO3-10 [20], WN18RR [21] and four synthetic benchmarks. Table 1 shows the statistics of the benchmark KGs. Each benchmark KG is divided into four categories: one-to-one (1:1) , one-to-many (1:N) , many-to-one (N:1) , and many-to-many (N:M) relationships, demonstrate how entities are connected in KGs to any other entity. For example, in a 1:1 relationship, each entity in the relationship pair can be associated with at most one entity in the KG, representing a bijection. Similarly, 1:N relationships state that an entity in a KG is associated with multiple other entities in a KG, and so on. In summary, the cardinalities in a KG help to accurately model the relationships between different entities, ensuring that the data structure represents real-world scenarios effectively. These subgraphs of the state-of-the-art benchmarks were created to evaluate the performance of *SimE* in all scenarios. The Family KG represents the biological links (e.g., father, sibling, etc.) that connect several family members. The French Royalty KG, collected from DBpedia, depicts French royal families with various relationships, such as successors, child, and siblings. Moreover, YAGO3-10 and WN18RR use a hierarchical taxonomy based on Wikipedia to represent many entities and their relationships. *SimE* can identify patterns like fractals and bijections, as illustrated in Figure 1b. To demonstrate the distinctive behavior of *SimE*, we generated the synthetic benchmarks (*SGKG 1*, *SGKG 2*, *SGKG 3*, and *SGKG 4*) comprising fractals of Sierpinski triangles (*a.k.a. triangles*), where each fractal is connected by isomorphic relations (e.g., sameAs), and the benchmarks comprises of fractals of size 243. SimE detects patterns, e.g., bijections and isomorphisms, observed in fractal-like KGs. The synthetic benchmarks comprise two types of connections: symmetric and isomorphic. When Sierpinski triangles are connected isomorphically, it indicates that there is a structural similarity between the triangles such that their vertices and edges can be mapped onto each other without altering the connections. Symmetry in connections implies that the relationship is bidirectional. If node A is connected to node B, then node B is also connected to node A. *SGKG1* includes isomorphic connections between the Sierpinski triangles, as seen in Figure 1a using symmetric *sameAs* relationship. The triangles T_1, T_2, and T_3 are connected symmetrically using *sameAs*. The *leftCorner* of T_1 is symmetrically connected to the *upperCorner* of T_3, similarly *rightCorner* of T_1 is connected to the *upperCorner* of T_2, and *rightCorner* of T_3 is connected to the *leftCorner* of T_2. However, in *SGKG2*, the triangles were not connected symmetrically. That is, the *sameAs* relationships were connected in an isomorphic manner, but in a unidirectional fashion. In other words, the relationships between the triangles T_1, T_2, and T_3 are not symmetric. In *SGKG3*, different relationships, in-

Table 3: Experimental results on the synthetic benchmarks. Results in bold and underlined represent the best and second-best, respectively.

KG	TransE		SimE		RotE		ComplEx		RotatE		RefE		MurE		CP		AttE	
	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1
SGKG 1	0.232	0.000	0.906	**0.900**	**0.944**	<u>0.892</u>	0.393	0.393	0.393	0.393	0.578	0.438	0.848	0.789	0.394	0.393	<u>0.909</u>	<u>0.892</u>
SGKG 2	0.404	0.233	0.679	<u>0.676</u>	**0.864**	**0.753**	0.000	0.000	0.000	0.000	0.170	0.084	<u>0.693</u>	0.673	0.000	0.000	0.687	0.674
SGKG 3	0.395	0.225	0.668	0.668	**0.707**	**0.697**	0.000	0.000	0.000	0.000	0.170	0.084	<u>0.693</u>	0.673	0.000	0.000	0.687	<u>0.674</u>
SGKG 4	0.233	0.000	0.902	<u>0.902</u>	**0.948**	**0.923**	0.400	0.400	0.400	0.400	0.591	0.477	0.899	0.876	0.400	0.400	<u>0.914</u>	0.899

Table 4: Experimental results on the four benchmarks and their respective sub-graphs based on the relation type. Results in bold and underlined represent the best and second-best, respectively. Among all baseline models, SimE-TransE exhibits the highest statistical significant p-value of 1.54×10^{-120}.

KG	TransE		SimE		RotE		ComplEx		RotatE		RefE		MurE		CP		AttE	
	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1	MRR	Hits@1
French 1:1	0.500	0.376	<u>0.743</u>	<u>0.688</u>	**0.747**	**0.696**	0.525	0.460	0.487	0.415	0.645	0.562	0.657	0.550	0.481	0.385	0.663	0.576
Family 1:1	0.421	0.270	**0.642**	**0.530**	<u>0.624</u>	<u>0.497</u>	0.385	0.350	0.351	0.317	0.507	0.367	0.530	0.396	0.372	0.322	0.556	0.428
YAGO 1:1	0.122	0.074	**0.142**	**0.102**	0.130	0.087	0.090	0.072	0.073	0.046	0.129	0.087	<u>0.141</u>	<u>0.101</u>	0.093	0.073	0.140	0.100
WN18RR 1:1	0.082	0.013	**0.198**	**0.142**	<u>0.193</u>	**0.142**	0.114	0.101	0.108	0.095	0.178	0.134	0.191	0.138	0.114	0.102	0.186	<u>0.139</u>
French 1:N	0.189	0.061	0.046	0.023	0.254	0.155	**0.261**	**0.211**	<u>0.247</u>	<u>0.193</u>	0.187	0.096	0.183	0.079	0.174	0.082	0.221	0.114
Family 1:N	0.269	0.099	0.306	0.153	0.341	0.151	0.080	0.053	0.066	0.045	0.306	0.134	<u>0.322</u>	<u>0.154</u>	0.086	0.045	**0.336**	**0.161**
YAGO 1:N	0.087	0.037	0.101	0.046	0.093	0.045	**0.406**	**0.323**	0.396	<u>0.315</u>	0.094	0.047	0.101	0.047	<u>0.399</u>	0.313	0.097	0.044
WN18RR 1:N	0.221	0.011	0.531	0.443	0.546	0.474	0.480	0.449	0.454	0.426	0.541	0.468	<u>0.560</u>	<u>0.492</u>	0.495	0.457	**0.562**	**0.501**
French N:1	0.615	0.555	0.672	0.628	0.664	0.624	0.423	0.347	0.392	0.313	0.653	0.614	**0.691**	**0.651**	0.421	0.346	<u>0.685</u>	<u>0.641</u>
Family N:1	0.307	0.150	0.382	0.199	**0.399**	**0.221**	0.082	0.048	0.052	0.032	0.376	0.211	0.376	0.202	0.090	0.045	<u>0.387</u>	<u>0.216</u>
YAGO N:1	0.096	0.044	0.111	0.057	0.100	0.047	<u>0.386</u>	<u>0.309</u>	0.361	0.283	0.100	0.047	0.110	0.056	**0.392**	**0.314**	0.109	0.056
WN18RR N:1	0.147	0.022	**0.376**	0.302	0.371	0.301	0.280	0.258	0.271	0.251	0.357	0.295	<u>0.372</u>	<u>0.303</u>	0.290	0.266	<u>0.372</u>	**0.306**
Family N:M	0.291	0.129	0.360	0.175	**0.376**	<u>0.186</u>	0.121	0.087	0.102	0.073	0.347	0.183	0.337	0.157	0.137	0.097	<u>0.363</u>	**0.188**
FrenchN-M	0.289	0.059	0.480	0.416	**0.607**	<u>0.507</u>	<u>0.582</u>	**0.544**	0.551	0.506	0.513	0.410	0.521	0.396	0.475	0.368	0.526	0.391
YAGO N:M	0.087	0.037	0.106	0.051	0.094	0.043	<u>0.380</u>	**0.303**	0.370	0.294	0.093	0.047	0.106	0.051	**0.381**	<u>0.301</u>	0.104	0.050
WN18RR N:M	0.157	0.031	0.369	0.302	**0.373**	**0.318**	0.293	0.281	0.286	0.278	0.359	<u>0.310</u>	0.365	0.306	0.297	0.286	<u>0.370</u>	**0.318**

cluding *sameAs*, *equivalent*, and *equal* relationships, were employed to connect triangles T_1, T_2, and T_3, isomorphically in a unidirectional manner. While the nature of these relationships is inherently symmetric, if not explicitly stated, KGE models are unable to consider them as symmetric relationships. In contrast, the *SGKG4* triangles T_1, T_2, and T_3 were connected isomorphically with different relationships, *sameAs*, *equivalent*, and *equal*. This was explicitly indicated by establishing bidirectional connections, thereby signifying the presence of symmetric connections.

5.2 *Evaluations Setup*

In our evaluation, we compare *SimE* with eight baselines KGE models, which includes MurE [2], TransE [4], AttE [10], CP [17], RotatE [28], ComplEx [29], RotE [10], RefE [10]. The translation distance model, such as TransE, transforms the head entity's vector closer to the tail entity with a given relation. Canonical polyadic (*CP*) is a tensor factorization approach where entities and relations are represented by n-dimensional real vectors. MurE is a hyperbolic interaction model capable of effectively modeling hierarchies in KG. ComplEx uses complex-valued representations for the entities and relations. Entities and relations are represented as vectors, and the plausibility score is calculated using the Hadamard product. The Hadamard product is non-commutative in complex space, allowing ComplEx to successfully simulate asymmetric relations.

Metrics. We use metrics to evaluate the quality of the KG completion process. *Hits@K* evaluates the ratio of correct entity predictions at the top *K* predicted links by counting how frequently the true (ground truth) links occur in the top *K* ranked predictions provided by the model. Mean reciprocal rank (MRR) is the average of the reciprocal ranks of all the triples; it is an indicator of mean rank after removing the effect of outliers. The evaluation metrics range from 0 to 1, with higher values indicating better performance of the models.

Implementation. *SimE* is implemented using PyTorch in Python 3.9, and all the experiments are performed in a virtual machine on Google Colab with 40 GiB VRAM and 1 GPU NVIDIA A100-SXM4, with CUDA version 12.2 (Driver 535.104.05). The benchmark KGs are divided into an 80-10-10 train-test-valid ratio. To prevent overfitting, the default settings for the training KGE models include a learning rate of $1e^{-1}$, a batch size of 1000, and *Adagrad* as the regularization optimizer with a negative sample size of 50. For reproducibility, the benchmarks and code are available on GitHub[2].

[2] https://github.com/NIMI-research/SimE

5.3 *Effectiveness of* SimE *on Link Prediction*

We report in Table 2, 3, and 4 the experimental evaluation of the baseline KGE models and our model *SimE* on the French Royalty KG, the Family KG, YAGO3-10, WN18RR and their sub-graphs suffixed with 1:1, 1:N, N:1, N:M as well as four synthetic KGs. These evaluations aim to assess the ability of the model to handle self-similar and topological patterns in the KGs. This section aims to answer the research questions **RQ1**), **RQ2**), and **RQ3**).

Whole KG. Our findings outline the detailed analysis of several KGE models demonstrating the impact of each model in predicting the tail entities more accurately. We evaluated the end-to-end performance of *SimE* analogous to the execution of baseline KGE models in 72 testbeds (Table 2). Across all testbeds, *SimE* outperforms baselines in the French Royalty KG, indicating the effectiveness of the proposed model in capturing the bijection properties (e.g., *hasPredecessor*) and self-similar entities. Our baselines, and in particular ComplEx and RotatE, still achieve competitive performance, while the gap from *SimE* widens in the French Royalty KG and WN18RR. *SimE* showcased the results in capturing similar patterns based on Euclidean distance with *Hits@1* values ranging from 0.061 to 0.661. However, TransE demonstrates comparatively lower performance, depicting the model's ability to handle symmetric and complex relationships. ComplEx and RotatE show robust performance on YAGO3-10 and the Family KG benchmarks, with *Hits@1* values ranging from 0.306 to 0.892, indicating the effectiveness in capturing complex relationships. Models operating on Euclidean distance, such as RefE, MurE, and AttE illustrate the static performance in comparison to *SimE*, which aligned with the aforementioned fact in Section 4. Despite promising results, existing KGE models struggle to capture bijection and self-similar fractals in benchmarks.

Relationship Types. As described above, the performance of *SimE* was evaluated over the subgraphs generated with different relationship types, i.e., 1:1, 1:N, N:1, and N:M. We report the empirical evaluation over these subgraphs in Table 4. The results suggest that SimE outperforms the baselines on the 1:1 downstream task with *Hits@1* values ranging from 0.102 to 0.688, except on the French Royalty KG, where it ranks second in both *MRR* and *Hits@1* to RotE. While the best-performing baselines based on the 1:N relation are ComplEx and AttE having values 0.323 and 0.161 respectively, in the N:1 relation, euclidean distance models such as RotE, MurE, and CP show robust performance, and in the N:M relations, they are RotE, ComplEx, and CP. Nonetheless, the performance of *SimE* is notably stagnant compared to other state-of-the-art models. However, given the attention mechanism for capturing the similarity between entities, our proposed *SimE* model is better suited to capture the characteristics of 1:1 relationships. That mechanism empowers the model to emphasize the separation of true and corrupted entities. Furthermore, the similarity transformations allow *SimE* to preserve the structural patterns and 1:1 relations as shown in Table 4.

Two important algebraic properties of similarity group elements are non-commutativity and invertibility. The relations in the French Royalty and Family KGs are not necessarily commutative. For example, consider the triples ⟨`Empress_Josephine, hasChild, Napoleon_II`⟩ and ⟨`Napoleon_II, hasPredecessor, Empress_Josephine`⟩; if these properties were commutative, the triple ⟨`Napoleon_II, hasChild, Empress_Josephine`⟩ would be part of the French Royalty KG which would mean that the triple ⟨`Empress_Josephine, hasPredecessor, Napoleon_II`⟩ could be predicted, contributing to the outstanding performance of *SimE* on these two family-related KGs.

Synthetic SGKGs. Table 3 displays the results for the synthetic benchmarks. The results represent a detailed study of numerous KGE models, revealing how each model affects the accuracy with which tail entities are predicted. We tested the end-to-end performance of *SimE* in 72 testbeds, comparing it to baseline KGE model execution. *SimE* outperforms baselines in *SGKG 1*, suggesting the effectiveness of the proposed model in capturing symmetric properties (e.g., *sameAs*) and self-similar structures with *Hits@1* of 0.900 in *SGKG 1*. Furthermore, baselines, particularly RotE and AttE show the second-best performance for *SGKG 1* with *Hits@1* of 0.892. When comparing *MRR* in *SGKG 1*, RotE surpasses *SimE* by 0.944. *SGKG 2* does not contain symmetrical relationships, *SimE* has the second-best performance with *Hits@1*. RotE exhibits competitive behavior in *SGKG 2* with the highest *Hits@1* of 0.753, followed by *SimE* and AttE with a minimal difference of *Hits@1* of 0.676 and 0.674. *SGKG 3* also lacks symmetric relationships but contains different types of isomorphic relations, e.g., *equal*, *equivalent*, and *sameAs*. RotE ranked first with *Hits@1* of 0.697, followed by AttE with *Hits@1* of 0.674. *SimE* is very competitive, with *Hits@1* of 0.668. Lastly, in *SGKG 4*, symmetric relationships are utilized to connect the *Sierpinski triangles* with different relationships. RotE performed the best and *SimE* performed at the second-best position with *Hits@1* of 0.923 and 0.902.

The findings suggest that existing KGE models produce promising results but fall short in capturing the bijection and self-similar fractals included in the *SGKG 1* and *SGKG 4* benchmarks, whereas *SimE* outperforms the state-of-the-art KGE models. Synthetic benchmarks of *Sierpinski triangles*, coupled with isomorphic relations, revealed new features of *SimE* and indicated the necessity for a KGE model capable of identifying complicated geometric structures. Furthermore, the empirical evaluation showed that increasing cycles or complex geometric structures with isomorphic connections following bijections would improve *SimE*'s performance. The *SimE* model employs attention mechanics to effectively distinguish between true and false facts, resulting in improved performance and KG completion with true facts demonstrating higher efficiency. Our experimental assessment shows that *SimE* can recognize complicated geometric structures like *fractals*, with effective performance. In Euclidean space, *SimE* outperformed KGE models, but translational models like TransE performed ineffectively in recognizing complicated symmetric structures in *SGKG 1*. However, KGE models such as ComplEx and RotatE struggled to identify complex structures and bijections, necessitating the use of *SimE*.

5.4 Discussion

Our results from different KGs demonstrate that the *SimE* model is ranked among the top-performing models. However, *SimE* has a few shortcomings that lie in the model design and the considered similitude. *SimE* is an attention-based model, and weights two intermediate relation embeddings learned from the dataset. These weights are the attention coefficients. The performance of the model comes with an effective attention coefficient learning. Failing to do this, *SimE* is not expected to surpass its subsumed models. On another level, the choice of the relational transformations to achieve the intermediate transformations can rend *SimE* to perform low. The results presented in Table 2, 3 and 4 are trained with default hyperparameters settings, indicating optimizing the hyperparameters will enhance *SimE* model's performance. We conducted two statistical tests (Wilcoxon test and Spearman's rho) to evaluate *SimE* model effectiveness among baselines. The Wilcoxon test evaluates the differences between the accurate tail predictions among models. For instance, in French Royalty KG, the values depict that among all the models, SimE is statistically significant with the p-value ranging from 5.37×10^{-34} to 1.43×10^{-4}. However, Spearman's rho reveals the strength of the association between ranks of predictions. In this case, the p-value ranges from 2.75×10^{-96} to 0.00×10^{-1}. In the synthetic benchmark, *SGKG 1*, the performance of SimE is significant with other models. The p-values for Wilcoxon and Spearman's rho range from 3.80×10^{-9} to $0.00 \times 10^{+00}$ and from 1.66×10^{-64} to 9.83×10^{-127} respectively. In French 1:1 KG, the Wilcoxon p-value is 5.37×10^{-34} to 1.43×10^{-4}. The Spearman's rho values are 8.01×10^{-5} and 0.737. Thus, the statistical tests shed light on SimE's significant performance among other models.

6 Conclusions and Future Work

In our study, we focus on leveraging the algebraic (bijection) and geometric (similarity) properties of similarity transformations to enhance the representation of self-similar fractals in KGs. Our method, named *SimE*, is validated through empirical analysis and demonstrates its proficiency in capturing bijections and similarities within benchmark KGs. We conducted comprehensive experiments that encompassed one-to-one, one-to-many, many-to-one, and many-to-many relational patterns, analyzing both the performance of state-of-the-art models and our proposed model, *SimE*. Additionally, we developed a fractal-like KG designed to evaluate the models' effectiveness in learning subgraph similarities. This approach allows us to assess how well each model handles the complex structures and relationships inherent in KGs. In future work, we will extend the testbeds to include more complex fractals. We also plan to investigate real-world KGs in their self-similar structures and further extend *SimE* to be capable of preserving those. Furthermore, KGE models require hyperparameter tuning to achieve improved efficiency in the performance of KG completion with true facts, allowing models to rank higher. In terms of the theoretical framework of *SimE*, it is planned to be generalized further to subsume further state-of-the-art models. This would allow *SimE* to outperform and effectively detect complex geometric structures in real-world KGs.

Acknowledgements

This work has been partially supported by TrustKG- Transforming Data in Trustable Insights with grant P99/2020, and by the Lower Saxony Ministry of Science and Culture (MWK) with funds from the

Volkswagen Foundation's zukunft.niedersachsen program (CAIMed - Lower Saxony Center for AI and Causal Methods in Medicine; GA No. ZN4257). We also acknowledge the financial support by the Federal Ministry of Education and Research of Germany and by Sächsisches Staatsministerium für Wissenschaft, Kultur und Tourismus in the programme Center of Excellence for AI-research „Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig" (ScaDS.AI). The work is also supported by the SECAI project (grant 57616814) funded by DAAD (German Academic Exchange Service).

References

[1] I. Balazevic, C. Allen, and T. Hospedales. Multi-relational poincaré graph embeddings. *Advances in Neural Information Processing Systems*, 32, 2019.

[2] I. Balažević, C. Allen, and T. Hospedales. Multi-relational poincaré graph embeddings, 2019.

[3] E. Bieberich. Recurrent fractal neural networks: a strategy for the exchange of local and global information processing in the brain. *Biosystems*, 66(3):145–164, 2002.

[4] A. Bordes, N. Usunier, A. García-Durán, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In C. J. C. Burges, L. Bottou, Z. Ghahramani, and K. Q. Weinberger, editors, *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*. URL https://proceedings.neurips.cc/paper/2013/hash/1cecc7a77928ca8133fa24680a88d2f9-Abstract.html.

[5] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013. URL https://proceedings.neurips.cc/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf.

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

[7] J. Cao, J. Fang, Z. Meng, and S. Liang. Knowledge graph embedding: A survey from the perspective of representation spaces. *ACM Comput. Surv.*, 56(6), mar 2024. doi: 10.1145/3643806.

[8] Z. Cao, Q. Xu, Z. Yang, X. Cao, and Q. Huang. Dual quaternion knowledge graph embeddings. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 6894–6902, 2021.

[9] Z. Cao, Q. Xu, Z. Yang, X. Cao, and Q. Huang. Geometry interaction knowledge graph embeddings, 2022.

[10] I. Chami, A. Wolf, D.-C. Juan, F. Sala, S. Ravi, and C. Ré. Low-dimensional hyperbolic knowledge graph embeddings. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6901–6914. Association for Computational Linguistics, July 2020. doi: 10.18653/v1/2020.acl-main.617.

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

[12] S. Choudhary, T. Luthra, A. Mittal, and R. Singh. A survey of knowledge graph embedding and their applications. *arXiv preprint arXiv:2107.07842*, 2021.

[13] Z. Du, C. Zhou, J. Yao, T. Tu, L. Cheng, H. Yang, J. Zhou, and J. Tang. Cogkr: Cognitive graph for multi-hop knowledge reasoning. *IEEE Transactions on Knowledge and Data Engineering*, 35(2):1283–1295, 2021.

[14] X. Ge, Y.-C. Wang, B. Wang, and C. J. Kuo. Core: A knowledge graph entity type prediction method via complex space regression and embedding. *Pattern Recognition Letters*, 157:97–103, 2022.

[15] N. Halliwell, F. Gandon, and F. Lecue. User Scored Evaluation of Non-Unique Explanations for Relational Graph Convolutional Network Link Prediction on Knowledge Graphs. In *K-CAP*. ACM, 2021.

[16] S. Kok and P. M. Domingos. Statistical predicate invention. In *ICML*, 2007. URL https://doi.org/10.1145/1273496.1273551.

[17] T. Lacroix, N. Usunier, and G. Obozinski. Canonical tensor decomposition for knowledge base completion, 2018.

[18] G. Larsson, M. Maire, and G. Shakhnarovich. Fractalnet: Ultra-deep neural networks without residuals. *arXiv preprint arXiv:1605.07648*, 2016.

[19] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu. Learning entity and relation embeddings for knowledge graph completion. In *Proceedings of the AAAI conference on artificial intelligence*, volume 29, 2015.

[20] F. Mahdisoltani, J. Biega, and F. M. Suchanek. YAGO3: A knowledge base from multilingual wikipedias. In *CIDR*, 2015. URL http://cidrdb.org/cidr2015/Papers/CIDR15_Paper1.pdf.

[21] G. A. Miller. Wordnet: A lexical database for english. *Commun. ACM*, 1995. doi: 10.1145/219717.219748.

[22] G. Ning, Z. Zhang, and Z. He. Knowledge-guided deep fractal neural networks for human pose estimation. *IEEE Transactions on Multimedia*, 20(5):1246–1259, 2017.

[23] D. Purohit, Y. Chudasama, A. Rivas, and M.-E. Vidal. Sparkle: Symbolic capturing of knowledge for knowledge graph enrichment with learning. In *Proceedings of the 12th Knowledge Capture Conference 2023*, K-CAP '23, page 44–52, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701412. doi: 10.1145/3587259.3627547. URL https://doi.org/10.1145/3587259.3627547.

[24] A. Rivas, D. Collarana, M. Torrente, and M.-E. Vidal. A neuro-symbolic system over knowledge graphs for link prediction. *Semantic Web Journal. Special Issue on Neuro-Symbolic Artificial Intelligence and the Semantic Web*, pages 1–25, 2023. doi: 10.3233/SW-233324. URL https://doi.org/10.3233/SW-233324.

[25] G. F. Roberto, A. Lumini, L. A. Neves, and M. Z. do Nascimento. Fractal neural network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images. *Expert Systems with Applications*, 166:114103, 2021.

[26] Y. Shen, N. Ding, H.-T. Zheng, Y. Li, and M. Yang. Modeling relation paths for knowledge graph completion. *IEEE Transactions on Knowledge and Data Engineering*, 33(11):3607–3617, 2020.

[27] Z. Sun, Z.-H. Deng, J.-Y. Nie, and J. Tang. Rotate: Knowledge graph embedding by relational rotation in complex space, 2019. URL https://arxiv.org/abs/1902.10197.

[28] Z. Sun, Z.-H. Deng, J.-Y. Nie, and J. Tang. Rotate: Knowledge graph embedding by relational rotation in complex space, 2019.

[29] T. Trouillon, J. Welbl, S. Riedel, E. Gaussier, and G. Bouchard. Complex embeddings for simple link prediction. In M. F. Balcan and K. Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2071–2080, New York, New York, USA, 20–22 Jun 2016. PMLR. URL https://proceedings.mlr.press/v48/trouillon16.html.

[30] X. Wang, D. Wang, C. Xu, X. He, Y. Cao, and T.-S. Chua. Explainable reasoning over knowledge graphs for recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 5329–5336, 2019.

[31] Z. Wang, J. Zhang, J. Feng, and Z. Chen. Knowledge graph embedding by translating on hyperplanes. In *Proceedings of the AAAI conference on artificial intelligence*, volume 28, 2014.

[32] Y. Xian, Z. Fu, S. Muthukrishnan, G. De Melo, and Y. Zhang. Reinforcement knowledge graph reasoning for explainable recommendation. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*, pages 285–294, 2019.

[33] H. Xiao, M. Huang, Y. Hao, and X. Zhu. Transa: An adaptive approach for knowledge graph embedding. *arXiv preprint arXiv:1509.05490*, 2015.

[34] J. Zhang, L. Tan, X. Tao, D. Wang, J. J.-C. Ying, and X. Wang. Learning relational fractals for deep knowledge graph embedding in online social networks. In *Web Information Systems Engineering–WISE 2019: 20th International Conference, Hong Kong, China, January 19–22, 2020, Proceedings 20*, pages 660–674. Springer, 2019.

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