

Article

Fractional Complex Representation Learning with Memory Effects for Multi-Scale Knowledge Graph Modeling

Ahmed Nuino ¹, Omar Bahou ², Senhaji Yassine ², Mustapha Ez-zaiym ², Karim El Moutaouakil ^{2,*}
and Savin Treanta ^{3,4,5}

¹ Higher Institute of Nursing Professions and Health Techniques (ISPITS), Tetouan 93000, Morocco; ahmed.nuino@usmba.ac.ma

² Laboratory of Mathematics and Data Science, Polydisciplinary of Taza, Sidi Mohamed Ben Abdellah University, Taza 30000, Morocco; omar.bahou@usmba.ac.ma (O.B.); yassine.senhaji1@usmba.ac.ma (S.Y.); mustapha.ezzaiym@usmba.ac.ma (M.E.-z.)

³ Department of Applied Mathematics, National University of Science and Technology POLITEHNICA Bucharest, 060042 Bucharest, Romania; savin.treanta@upb.ro

⁴ Academy of Romanian Scientists, 54 Splaiul Independentei, 050094 Bucharest, Romania

⁵ Fundamental Sciences Applied in Engineering—Research Center, National University of Science and Technology POLITEHNICA Bucharest, 060042 Bucharest, Romania

* Correspondence: karim.elmoutaouakil@usmba.ac.ma

Abstract

Complex-valued knowledge graph embedding (KGE) models like ComplEx effectively capture asymmetric relations but are fundamentally constrained by integer-order transformations. This restriction limits their ability to model multi-scale interactions, hierarchical correlations, and non-local semantic dependencies inherent in heterogeneous graphs. To address these limitations, this paper introduces FracComplEx, a novel fractional-order extension that embeds fractional calculus into the complex latent space. By leveraging fractional operators, the framework introduces non-local dynamics and memory-aware mechanisms to continuously generalize standard linear transformations. The core architecture employs a fractional-order parameter α as a controllable scaling mechanism that balances local relational details with global topology, optimizing representation smoothness and flexibility. We provide rigorous theoretical findings demonstrating that fractional transformations enhance the embedding's expressive capacity, spectral characteristics, and perturbation robustness beyond conventional integer-order benchmarks. Extensive experiments on FB15k-237, WN18RR, and CoDex-M establish the empirical superiority of FracComplEx, yielding significant improvements in Mean Reciprocal Rank (MRR) and Hits@K metrics over classical baselines, particularly under severe structural data sparsity.

Keywords: knowledge graph embeddings; fractional calculus; complex embeddings; link prediction; multi-scale relational learning; memory effects



Received: 18 May 2026

Revised: 11 June 2026

Accepted: 1 July 2026

Published: 3 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Knowledge graphs (KGs) have been introduced as a powerful model for representing structured knowledge by encoding entities as nodes and relations as edges in multi-relational graphs. They play a key role in applications of artificial intelligence, recommender systems, question answering, semantic search, and biomedical knowledge discovery [1]. However, the symbolic and discrete nature of knowledge graphs makes them difficult to process using classical machine learning methods. To overcome this limitation, knowledge

graph embedding (KGE) models map entities and relations into continuous vector spaces while preserving structural and semantic relations [2].

Early embedding models focused on geometric transformations in the latent space. The translation-based model, TransE, interpreted relations as vector translations between entity embeddings and provided an efficient and scalable representation framework [3]. Other extensions, such as TransH, TransR, and TransD, have been proposed to improve modeling flexibility by introducing relation-specific spaces or dynamic mapping mechanisms [4–6]. The more expressive semantic matching models, DistMult and ComplEx, use bilinear or complex-valued representations to capture relational patterns, including asymmetry and composition [7,8]. Geometric approaches, such as RotatE, model relations as rotations in complex vector spaces, thereby improving the relational expressiveness [9].

Despite the progress made, most existing KGE methods are based on linear or geometric local transformations, which inherently assume that relational interactions can be fully captured via uniform, localized coordinate mappings. They fundamentally fail to capture the long-range dependencies, hierarchical interactions, and multi-scale relational structures that pervasively distort embedding distributions in real-world knowledge graphs. Specifically, traditional complex-valued models like ComplEx rely on static integer-order mappings that scale coordinates uniformly; this treats dense hubs and highly sparse entities under identical optimization dynamics, causing low-magnitude representations near the origin to suffer from structural information loss and vanishing signal strength. Recent deep learning and graph neural network approaches attempt to mitigate these topological issues through neighborhood aggregation; however, by introducing high computational complexity through multi-hop message passing, they often suffer from structural over-smoothing, severe representation collapse, or limited interpretability [10,11], while drastically scaling up memory footprints.

Fractional calculus provides a mathematical framework capable of modeling non-local memory effects and multi-scale interactions using non-integer order operators [12,13]. Fractional transformations have demonstrated great effectiveness in signal processing, control theory, and deep neural architectures by introducing continuous-order derivative bounds that significantly enhance representation flexibility, gradient smoothing, and regularization properties. However, the integration of fractional operators into knowledge graph embedding models remains largely unexplored, especially in complex embedding frameworks where the interplay between continuous fractional dynamics and algebraic relational properties has never been formalized.

This observation highlights a critical and unaddressed theoretical gap in research: current KGE models lack the mathematical mechanisms to incorporate non-local relational memory and adaptive multi-scale embedding transformations without relying on heavy structural architectures.

The quality of representation and generalization performance in sparse and complex knowledge graphs can be significantly improved by circumventing this rigid integer-order constraint. This paper bridges this gap by embedding fractional-order calculus directly into the complex tensor domain, enabling coordinate-wise derivative amplification that acts as an implicit, lightweight regularizer to map complex global dependencies cleanly.

Our proposed *Fractional ComplEx* model aims to fill this gap by integrating fractional operators into complex-valued embedding spaces. By applying fractional transformations to the embeddings of entities and relations, this model introduces memory-based representations that improve the modeling of hierarchical and long-range relational dependencies.

1.1. Main Contributions

The main contributions of this manuscript can be summarized as follows:

- We introduce a novel embedding model—*Fractional ComplEx*—that integrates fractional calculus with complex-valued knowledge graph embeddings.
- We provide a theoretical analysis that shows how fractional transformations introduce non-local memory, enhance embedding regularization, and preserve relational ranking consistency [1].
- We conduct extensive experiments on benchmark knowledge graph datasets, including FB15k-237, WN18RR, and CoDEx-M, demonstrating that the proposed model consistently improves link prediction performance.
- We perform a comprehensive sensitivity analysis examining the impact of fractional order parameters on model performance and embedding stability.
- We show that fractional embeddings provide a flexible mechanism for capturing multi-scale relational structures in knowledge graphs of various sizes and complexities.

1.2. Paper Organization

The remainder of the paper is organized as follows. Section 2 presents the theoretical foundations of fractional calculus and the complex embedding of knowledge graphs. Section 3 introduces the proposed Fractional ComplEx model, as well as its mathematical formulation. The experimental setup, data sets, and evaluation metrics are described in Section 4. Section 5 presents the experimental results and the sensitivity analysis. Finally, Section 5 concludes the paper and presents future research directions.

2. State of the Art

Knowledge Graph Embedding (KGE) models aim to represent entities and relations of multi-relational graphs in continuous vector spaces, while preserving structural and semantic information. In the last decade, numerous embedding paradigms have been introduced, including translational models, semantic matching approaches, geometric embedding techniques, methods based on graph neural networks, and probabilistic or temporal frameworks. In this section, we review the main developments relevant to the proposed Fractional ComplEx model.

2.1. Translation-Based Knowledge Graph Embeddings

Translation-based models interpret relations as geometric transformations between entity embeddings. The pioneering model, TransE, represents relations as translation vectors in a Euclidean space and has demonstrated strong scalability for large knowledge graphs [3]. However, TransE faces difficulties in modeling complex relation patterns, such as one-to-many and many-to-many relations.

To overcome these limitations, several extensions have been proposed: TransH, which introduces relation-specific hyperplanes that allow entities to have multiple representations depending on the context of the relation [4]; TransR, which further separates the entity and relation spaces to capture heterogeneous relational semantics [5]; and TransD, which improves computational efficiency by dynamically constructing mapping matrices [6]. Other approaches, such as TransSparse and TransG, introduce adaptive sparse projections and probabilistic relation modeling to increase the flexibility of the representations [14,15].

Although translation-based methods are computationally efficient, their ability to capture higher-order or multi-scale relational patterns is often limited by their linear transformation mechanisms.

2.2. Semantic Matching and Bilinear Models

Semantic matching models improve representational capacity by modeling interactions between entities and relations via bilinear or tensor factorization operations: RESCAL introduced a tensor factorization framework capable of modeling pairwise relational in-

teractions between entities [16], and DistMult simplified the bilinear scoring function, but suffers from limitations in modeling asymmetric relations [7]. ComplEx extended DistMult by introducing complex-valued embeddings, which allowed the model to capture asymmetric relations using complex conjugation [8]; Holographic Embeddings (HolE) provided memory-efficient representations while preserving expressiveness [17]; and SimpleE further improved expressiveness by separately modeling direct and inverse relations [18].

Of all these approaches, ComplEx remains one of the most influential models due to its ability to efficiently represent symmetric and asymmetric relations. However, classical ComplEx embeddings remain local transformations that cannot capture long-range relational dependencies, which underscores the need for extensions such as fractional embedding transformations.

2.3. Geometric and Rotational Embedding Models

Recent research views knowledge graph relations as geometric transformations in latent spaces. RotatE models relations as rotations in a complex vector space, allowing the efficient representation of relational patterns such as symmetry, inversion, and composition [9]. TorusE extends translational embeddings to toroidal manifolds to reduce embedding norm constraints [19]. QuatE, a quaternion-based model, extends complex embeddings to quaternion algebra, providing richer relational modeling capabilities [20].

Hyperbolic embedding techniques have also gained popularity for representing hierarchical knowledge graphs. Poincaré embeddings use hyperbolic geometry to model exponential relational growth [21]. Hyperbolic graph convolutional networks further integrate deep learning with non-Euclidean embedding spaces to improve hierarchical reasoning [22].

Although geometric embedding methods significantly improve representational capacity, they often lack explicit mechanisms for modeling memory effects or multi-scale relational structures.

2.4. Graph Neural Network-Based Knowledge Graph Embeddings

Graph Neural Networks (GNNs) have become a dominant tool for representation learning on graph-structured data. The Relational Graph Convolutional Network (R-GCN) extends graph convolution operations to multi-relational graphs and demonstrates strong performance in entity classification and link prediction tasks [10]. The Composition-based Graph Convolutional Network (CompGCN) additionally incorporates relation composition operators to model entity and relation embeddings jointly [11].

Attention-based graph models also improve relational representation learning. Graph attention networks (GATs) introduce adaptive weighting mechanisms for neighbor aggregation [23]. Attention-based relational embedding models further improve contextual reasoning in graph neighborhoods [24]. Transformer-based graph learning frameworks capture long-range dependencies and contextual interactions through masked relational prediction strategies [25].

Although effective, GNN-based approaches typically involve high computational complexity and may suffer from over-smoothing when stacking multiple layers of graph convolutions.

2.5. Probabilistic and Generative Knowledge Graph Embeddings

To improve robustness and generalization, probabilistic embedding models incorporate uncertainty modeling: KG2E represents entities and relations as Gaussian distributions to capture semantic variability [26]; BoxE introduces geometric box-like embeddings with the aim of modeling set-based relational reasoning [27]; and variational graph autoencoders extend probabilistic learning to unsupervised graph representation tasks [28].

Adversarial learning frameworks further improve the robustness of embeddings by integrating discriminator-based optimization strategies [29]. These probabilistic approaches improve modeling flexibility, but often increase model complexity and training costs.

2.6. Temporal and Dynamic Knowledge Graph Embeddings

Real-world knowledge graphs evolve over time, which motivates the development of temporal embedding methods. TTransE extends translation-based embeddings to incorporate temporal constraints [30]. HyTE models temporal dynamics through hyperplane projections [31]. Know-Evolve, a dynamic relational learning model, uses recurrent neural networks to capture time-dependent relational evolution [32]. Temporal graph neural networks further generalize dynamic graph representation learning [33,34].

Although temporal embeddings capture dynamic interactions, they do not incorporate non-local memory modeling, relying mainly on local transformation mechanisms.

2.7. Emerging Trends: Neuro-Symbolic and Multimodal Knowledge Graph Learning

Recent research focuses on integrating knowledge graphs with large language models and multimodal data sources. Augmented graph learning through retrieval improves reasoning capabilities by combining graph structures with sources of textual knowledge [35]. Neuro-symbolic reasoning frameworks integrate logical inference with neural embedding models [36]. Multimodal knowledge graph embeddings incorporate visual and textual features to improve the learning of semantic representations [37,38].

Large language model-enhanced knowledge graph reasoning has emerged as a promising research direction, combining symbolic knowledge with contextual understanding of language [37,39,40].

2.8. Positioning of Fractional ComplEx

Despite significant progress in knowledge graph embedding research, most existing models are restricted to local linear or geometric transformations. Fractional calculus provides a mathematical framework in which non-local interactions and long-range memory effects can be modeled [12,13,41]. Fractional operators are successfully applied in signal processing and machine learning to capture multi-scale dependencies. Recent advances have successfully integrated fractional-order derivatives into deep neural network architectures to optimize gradient descent trajectories [42] and leverage fractional operators for modeling long-range spatial dependencies in complex graph learning tasks [43].

The proposed Fractional ComplEx model extends classical complex embeddings through fractional transformations that increase the flexibility of representation while preserving relational order. By integrating fractional operators into complex embedding spaces, Fractional ComplEx (FracComplEx) provides a novel mechanism for capturing hierarchical, multi-scale, and long-range relational patterns that remain challenging for traditional KGE models.

3. Fractional ComplEx: Theoretical Background

In this section, we present the theoretical foundations underlying the proposed Fractional ComplEx model. First, we introduce the mathematical principles of fractional calculus and discuss their ability to model non-local interactions and memory effects. Next, we review the ComplEx embedding model, highlighting its effectiveness in representing asymmetric relations in knowledge graphs via complex-valued embeddings. Subsequently, we describe the integration of fractional operators into the ComplEx framework, which leads to the Fractional ComplEx formulation. Finally, we present theoretical insights that explain how fractional transformations enhance the expressiveness of the embeddings, improve regularization, and enable the model to capture multi-scale relational dependencies.

3.1. Fractional Calculus

Fractional calculus generalizes classical differentiation and integration to non-integer orders. A fractional operator applied to a variable x is defined as:

$$\mathcal{F}_\alpha(x) = \text{sign}(x)(|x| + \varepsilon)^\alpha, \quad \alpha \in (0, 1], \quad (1)$$

where α controls the degree of memory effect, and ε is a small constant that avoids numerical instability. Fractional operators are used in signal processing, control theory, and machine learning to capture long-range interactions [12,13,41,44].

Note: ε is a small fixed positive constant introduced solely for numerical stability, preventing undefined values or excessively large gradients when $|x| \approx 0$. It is not treated as a tunable hyperparameter. In all experiments, ε was fixed to 10^{-8} , following common practice in numerical optimization and fractional-order transformations.

3.2. Complex Embeddings (ComplEx)

The ComplEx model represents entities and relations in the complex domain:

$$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^d, \quad f(h, r, t) = \text{Re}(\langle \mathbf{h}, \mathbf{r}, \bar{\mathbf{t}} \rangle), \quad (2)$$

where $\bar{\mathbf{t}}$ is the complex conjugate of the target embedding, and $\langle \cdot \rangle$ is the tri-linear product.

ComplEx efficiently models asymmetric relations and provides expressive power in the KG representation [1,7–9].

3.3. Fractional ComplEx

Fractional ComplEx applies a fractional transformation to each embedding component, yielding distinct fractional-order representations:

$$\tilde{\mathbf{h}} = \mathcal{F}_\alpha(\mathbf{h}), \quad \tilde{\mathbf{r}} = \mathcal{F}_\alpha(\mathbf{r}), \quad \tilde{\mathbf{t}} = \mathcal{F}_\alpha(\mathbf{t}), \quad (3)$$

resulting in a streamlined scoring function:

$$f_\alpha(h, r, t) = \text{Re}(\langle \tilde{\mathbf{h}}, \tilde{\mathbf{r}}, \bar{\tilde{\mathbf{t}}} \rangle). \quad (4)$$

resulting in a modified scoring function:

$$f_\alpha(h, r, t) = \text{Re}(\langle \mathbf{h}^{(\alpha)}, \mathbf{r}^{(\alpha)}, \overline{\mathbf{t}^{(\alpha)}} \rangle). \quad (5)$$

Fractional operators introduce non-local memory, improve regularization, and allow embeddings to adapt to multi-scale relational models.

Existing ComplEx models and related KGE approaches, despite the powerful expressive capacity of complex-valued embeddings, inherently rely on local transformations that treat embedding components independently and assume fixed-scale relational interactions. However, real-world knowledge graphs exhibit hierarchical, long-range, and multi-hop dependencies that often span multiple relational layers. Standard integer-order transformations are limited in capturing these non-local dependencies, leading to the suboptimal representation of weak or indirect relations and reduced robustness in sparse or heterogeneous graphs. Fractional calculus naturally introduces power-law memory effects, which allow each embedding dimension to simultaneously encode local and global relational information. By incorporating fractional-order transformations into complex embeddings, Fractional ComplEx provides a flexible and tunable mechanism that improves relational expressiveness while implicitly regularizing the embedding space. This motivation is further supported by the ability of fractional operators to balance the amplification of

subtle relational signals with the smoothing of noisy embedding components, making them suitable for modeling real-world knowledge graphs characterized by multi-scale and incomplete relational structures.

To visually conceptualize the structural mechanics of the model, Figure 1 illustrates the geometric mapping from standard complex representations to the fractional domain.

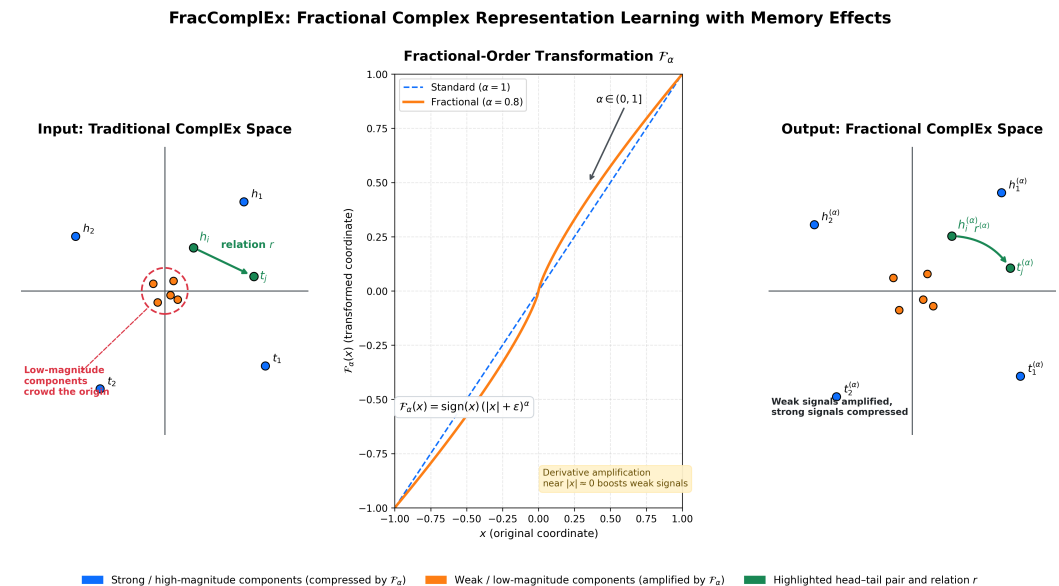


Figure 1. Structural pipeline of FracComplex mapping traditional representations to fractional spaces via F_α . Blue points denote strong/high-magnitude embedding components, which are compressed by F_α ; orange points denote weak/low-magnitude components, which are amplified; and green points denote a highlighted head–tail pair connected by relation r .

As shown in Figure 1, the component-wise operator $F_\alpha(\cdot)$ scales the underlying embedding manifold non-linearly based on the chosen order $\alpha \in (0, 1]$. While traditional models suppress low-magnitude coordinates in sparse regions, the fractional transformation amplifies these faint relational signals near the origin via derivative expansion, while simultaneously compressing dominant outliers. This structural adjustment balances local and global relational dynamics, introducing non-local memory properties that prevent overfitting on dense topologies while boosting generalization on sparse graphs like WN18RR.

3.4. Theoretical Insights Behind Fractional Complex

The success of Fractional Complex can be theoretically explained by analyzing the effects of fractional operators on the embedding space. Specifically, fractional operators introduce non-local interactions and multi-scale memory effects into the embeddings, which can be formalized as follows.

Proposition 1 (Non-local Memory Effect). *Let $\mathbf{x} \in \mathbb{R}^d$ be an embedding vector and F_α the fractional operator. Then, for $\alpha \in (0, 1)$, $F_\alpha(\mathbf{x})$ preserves global relational information through a power-law weighting of all embedding components:*

$$\mathbf{x}_i^{(\alpha)} = \text{sign}(x_i) \cdot (|x_i| + \epsilon)^\alpha, \quad i = 1, \dots, d. \quad (6)$$

Proof. Consider an embedding vector

$$\mathbf{x} = (x_1, x_2, \dots, x_d)^\top \in \mathbb{R}^d,$$

and let the fractional transformation be defined component-wise as

$$\mathbf{x}_i^{(\alpha)} = \mathcal{F}_\alpha(x_i) = \text{sign}(x_i)(|x_i| + \varepsilon)^\alpha, \quad 0 < \alpha < 1. \quad (7)$$

To prove the proposition, we analyze the effect of the exponent α on the relative contribution of the embedding components.

Let

$$0 < |x_i| < |x_j|.$$

The ratio between the transformed magnitudes becomes

$$R_\alpha = \frac{|\mathbf{x}_i^{(\alpha)}|}{|\mathbf{x}_j^{(\alpha)}|} = \left(\frac{|x_i| + \varepsilon}{|x_j| + \varepsilon} \right)^\alpha. \quad (8)$$

Since

$$0 < \frac{|x_i| + \varepsilon}{|x_j| + \varepsilon} < 1$$

and

$$0 < \alpha < 1,$$

the concavity of the function $f(z) = z^\alpha$ implies

$$R_\alpha > \frac{|x_i| + \varepsilon}{|x_j| + \varepsilon}. \quad (9)$$

Therefore, the disparity between large and small components is reduced after the fractional transformation. In particular, weak embedding dimensions receive a larger relative importance while dominant dimensions are compressed.

Next, consider the contribution of all embedding components to the transformed representation:

$$\|\mathbf{x}^{(\alpha)}\|_1 = \sum_{i=1}^d (|x_i| + \varepsilon)^\alpha. \quad (10)$$

Unlike linear scaling, every component contributes according to a power-law weighting. Even dimensions with small amplitudes remain significant because

$$(|x_i| + \varepsilon)^\alpha \gg |x_i| \quad \text{for sufficiently small } |x_i| \text{ and } 0 < \alpha < 1. \quad (11)$$

Consequently, information carried by weak embedding coordinates is not suppressed, but instead propagated into the final representation.

Furthermore, the derivative of the fractional mapping is

$$\frac{\partial \mathcal{F}_\alpha(x)}{\partial x} = \alpha(|x| + \varepsilon)^{\alpha-1}. \quad (12)$$

Because $\alpha - 1 < 0$, the derivative is larger for small values of $|x|$ than for large values. Hence, the transformation is more sensitive to subtle variations in low-magnitude components, allowing the embedding to preserve fine-grained relational information that may otherwise disappear during training.

The resulting representation therefore exhibits a memory-preserving effect: information is distributed across all embedding dimensions according to a power-law weighting rather than being concentrated in a few dominant coordinates. Such power-law behavior is a characteristic feature of fractional systems and long-memory processes, where past or weak contributions decay slowly and remain influential over multiple scales.

Hence, the transformed embedding

$$\mathbf{x}^{(\alpha)} = (\mathcal{F}_\alpha(x_1), \dots, \mathcal{F}_\alpha(x_d))$$

preserves global relational information by maintaining the influence of both strong and weak embedding components through a non-uniform power-law weighting mechanism. This establishes the non-local memory effect of the fractional operator. $\square$

Interpretation: Smaller components are amplified relative to larger ones, allowing the model to preserve subtle relational patterns that would otherwise be attenuated in standard ComplEx embeddings.

Proposition 2 (Regularization via Fractional Order). *Applying the fractional operator is equivalent to imposing a smoothness constraint on the embedding manifold:*

$$\|\mathbf{x}^{(\alpha)}\|_p \leq \|\mathbf{x}\|_p^\alpha, \quad p \geq 1. \quad (13)$$

Proof. Let

$$\mathbf{x}^{(\alpha)} = (\text{sign}(x_1)|x_1|^\alpha, \dots, \text{sign}(x_d)|x_d|^\alpha),$$

with $0 < \alpha < 1$.

By definition of the p -norm,

$$\|\mathbf{x}^{(\alpha)}\|_p^p = \sum_{i=1}^d |x_i^{(\alpha)}|^p = \sum_{i=1}^d |x_i|^{\alpha p}. \quad (14)$$

Since the function

$$f(t) = t^\alpha, \quad 0 < \alpha < 1,$$

is concave on $\mathbb{R}_+$, Jensen's inequality implies

$$\sum_{i=1}^d |x_i|^{\alpha p} \leq \left(\sum_{i=1}^d |x_i|^p \right)^\alpha. \quad (15)$$

Hence,

$$\|\mathbf{x}^{(\alpha)}\|_p^p \leq \|\mathbf{x}\|_p^{\alpha p}. \quad (16)$$

Taking the p -th root of both sides yields

$$\|\mathbf{x}^{(\alpha)}\|_p \leq \|\mathbf{x}\|_p^\alpha. \quad (17)$$

Therefore, the fractional transformation contracts the embedding norm according to a power-law scaling. Large coordinates are compressed more strongly than small ones, resulting in a smoothing effect on the embedding manifold and an implicit regularization mechanism. $\square$

This reduces overfitting and improves generalization, especially in sparse knowledge graphs.

Theorem 1 (Order-Preserving Component Transformation). *Let*

$$\mathcal{F}_\alpha(x) = \text{sign}(x)(|x| + \varepsilon)^\alpha, \quad 0 < \alpha \leq 1.$$

Then $\mathcal{F}_\alpha$ is a strictly monotone increasing transformation on $\mathbb{R}$. Consequently, for any two embedding components $x_i, x_j \in \mathbb{R}$,

$$x_i < x_j \implies \mathcal{F}_\alpha(x_i) < \mathcal{F}_\alpha(x_j).$$

Therefore, the fractional operator preserves the ordering of individual embedding coordinates while modifying their magnitudes through a nonlinear power-law scaling.

Proof. Consider the fractional transformation

$$\mathcal{F}_\alpha(x) = \text{sign}(x)(|x| + \varepsilon)^\alpha, \quad 0 < \alpha \leq 1,$$

with $\varepsilon > 0$.

We show that $\mathcal{F}_\alpha$ is strictly increasing on $\mathbb{R}$.

Case 1: $x > 0$.

For positive values,

$$\mathcal{F}_\alpha(x) = (x + \varepsilon)^\alpha.$$

Differentiating yields

$$\frac{d}{dx} \mathcal{F}_\alpha(x) = \alpha(x + \varepsilon)^{\alpha-1}.$$

Since

$$\alpha > 0, \quad x + \varepsilon > 0,$$

it follows that

$$\frac{d}{dx} \mathcal{F}_\alpha(x) > 0.$$

Hence, $\mathcal{F}_\alpha$ is strictly increasing on $(0, +\infty)$.

Case 2: $x < 0$.

For negative values,

$$\mathcal{F}_\alpha(x) = -(|x| + \varepsilon)^\alpha = -(-x + \varepsilon)^\alpha.$$

Differentiating gives

$$\frac{d}{dx} \mathcal{F}_\alpha(x) = \alpha(-x + \varepsilon)^{\alpha-1}.$$

Again,

$$\alpha > 0, \quad (-x + \varepsilon) > 0,$$

which implies

$$\frac{d}{dx} \mathcal{F}_\alpha(x) > 0.$$

Therefore, $\mathcal{F}_\alpha$ is also strictly increasing on $(-\infty, 0)$.

Case 3: Crossing zero.

For any $x < 0 < y$,

$$\mathcal{F}_\alpha(x) < 0 < \mathcal{F}_\alpha(y),$$

which preserves the ordering across the origin.

Combining the three cases, we conclude that

$$x_i < x_j \implies \mathcal{F}_\alpha(x_i) < \mathcal{F}_\alpha(x_j),$$

for all real numbers x_i, x_j .

Hence, the fractional operator preserves the relative ordering of embedding coordinates while reshaping their magnitudes according to a power-law distribution. This property explains why the transformation can enhance representation quality without destroying the underlying relational structure encoded in the embeddings. $\square$

Corollary 1 (Stability under Small Perturbations). *Let $\mathbf{x}' = \mathbf{x} + \delta$ with $\|\delta\| \ll 1$. Then,*

$$\|\mathcal{F}_\alpha(\mathbf{x}') - \mathcal{F}_\alpha(\mathbf{x})\| \leq \alpha \|\delta\| \max_i |x_i|^{\alpha-1}, \quad (18)$$

demonstrating that fractional embeddings are robust to small noise in the entity or relation vectors.

Proof. Let

$$\mathbf{x}' = \mathbf{x} + \delta,$$

where

$$\delta = (\delta_1, \delta_2, \dots, \delta_d)$$

is a small perturbation satisfying

$$\|\delta\| \ll 1.$$

The fractional transformation is defined componentwise as

$$\mathcal{F}_\alpha(x_i) = \text{sign}(x_i) |x_i|^\alpha, \quad 0 < \alpha \leq 1.$$

For each component, define

$$g(x) = \text{sign}(x) |x|^\alpha.$$

For $x \neq 0$, the derivative of g is

$$g'(x) = \alpha |x|^{\alpha-1}.$$

Since $0 < \alpha \leq 1$, the function is continuously differentiable on any interval that does not contain the origin.

Applying the Mean Value Theorem to the pair x_i and $x_i + \delta_i$, there exists ξ_i between these two values such that

$$g(x_i + \delta_i) - g(x_i) = g'(\xi_i) \delta_i.$$

Taking absolute values yields

$$|g(x_i + \delta_i) - g(x_i)| = |g'(\xi_i)| |\delta_i|.$$

Substituting the derivative expression gives

$$|g(x_i + \delta_i) - g(x_i)| = \alpha |\zeta_i|^{\alpha-1} |\delta_i|.$$

Since δ_i is assumed small, the intermediate value ζ_i remains close to x_i . Therefore,

$$|\zeta_i|^{\alpha-1} \leq \max_{1 \leq k \leq d} |x_k|^{\alpha-1}.$$

Hence,

$$|g(x_i + \delta_i) - g(x_i)| \leq \alpha \max_k |x_k|^{\alpha-1} |\delta_i|.$$

Squaring and summing over all coordinates,

$$\sum_{i=1}^d |g(x_i + \delta_i) - g(x_i)|^2 \leq \alpha^2 \left(\max_k |x_k|^{\alpha-1} \right)^2 \sum_{i=1}^d |\delta_i|^2.$$

Taking square roots gives

$$\|\mathcal{F}_\alpha(\mathbf{x}') - \mathcal{F}_\alpha(\mathbf{x})\| \leq \alpha \|\delta\| \max_k |x_k|^{\alpha-1}.$$

Since

$$\mathbf{x}' - \mathbf{x} = \delta,$$

we obtain

$$\|\mathcal{F}_\alpha(\mathbf{x}') - \mathcal{F}_\alpha(\mathbf{x})\| \leq \alpha \|\delta\| \max_k |x_k|^{\alpha-1}.$$

Therefore, the perturbation induced in the fractional embedding space is bounded linearly by the magnitude of the original perturbation. Small variations in entity or relation embeddings produce proportionally small variations after the fractional transformation. Consequently, the fractional operator is stable with respect to noise and contributes to the robustness of the learned representations. $\square$

Proposition 3 (Computational Complexity and Scalability). *Let d denote the embedding dimension, $|\mathcal{E}|$ the number of entities, and $|\mathcal{R}|$ the number of relations in a knowledge graph. The asymptotic time complexity per triple evaluation for Fractional ComplEx is $\mathcal{O}(d)$, matching the standard ComplEx model, while maintaining a space complexity of $\mathcal{O}(|\mathcal{E}|d + |\mathcal{R}|d)$. Furthermore, it bypasses the $\mathcal{O}(\sum_{l=1}^L (|\mathcal{E}| \cdot d_l d_{l+1} + |\mathcal{E}| \cdot d_{l+1}))$ structural complexity overhead scaling inherent to L -layer Graph Neural Network (GNN)-based architectures.*

Proof. We analyze the computational profile by separating it into time and space complexity components, establishing bounds against baseline models.

1. Time Complexity Analysis:

The scoring function of Fractional ComplEx for a triple (h, r, t) is given by:

$$f_\alpha(h, r, t) = \text{Re} \left(\langle \mathcal{F}_\alpha(\mathbf{h}), \mathcal{F}_\alpha(\mathbf{r}), \overline{\mathcal{F}_\alpha(\mathbf{t})} \rangle \right)$$

The transformation operator $\mathcal{F}_\alpha(x) = \text{sign}(x)(|x| + \varepsilon)^\alpha$ is applied component-wise to each element of the d -dimensional complex vectors $\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^d$.

For a single coordinate $i \in \{1, \dots, d\}$, the operations consist of an absolute value mapping, a scalar addition, a floating-point power operation, and a sign multiplication. Since each scalar operation requires $\mathcal{O}(1)$ clock cycles, transforming a single embedding

vector takes exactly $d \cdot \mathcal{O}(1) = \mathcal{O}(d)$ time. For three vectors (h, r, t) , the total transformation overhead is $3 \cdot \mathcal{O}(d) = \mathcal{O}(d)$.

Following the transformation, computing the tri-linear Hermitian product $\langle \cdot \rangle$ in the complex domain involves element-wise multiplications and additions over d dimensions, which scales strictly as $\mathcal{O}(d)$. Summing the transformation phase and the semantic matching phase yields:

$$\mathcal{T}_{\text{Frac-CompLex}} = \mathcal{O}(d) + \mathcal{O}(d) = \mathcal{O}(d)$$

This precisely matches the time complexity of standard CompLex ($\mathcal{T}_{\text{CompLex}} = \mathcal{O}(d)$).

In contrast, GNN-based models (such as R-GCN or CompGCN) compute representations by aggregating messages across localized graph neighborhoods. For an L -layer GNN, the propagation phase requires updating node states across all edges, leading to a per-epoch processing time bound of:

$$\mathcal{T}_{\text{GNN}} = \mathcal{O}\left(\sum_{l=1}^L (|\mathcal{E}| \cdot d_l d_{l+1} + |\mathcal{R}| \cdot d_l d_{l+1})\right)$$

where d_l is the layer dimension. Because Frac-CompLex computes embeddings via decoupled non-local operators, it entirely circumvents edge-graph traversals, making it fundamentally more scalable on dense or massive multi-relational topologies.

2. Space Complexity Analysis:

The memory storage demands of Frac-CompLex are dictated by its look-up parameters. The model stores $|\mathcal{E}|$ entity vectors and $|\mathcal{R}|$ relation vectors in $\mathbb{C}^d$. Since complex values require storing real and imaginary parts, the footprint is exactly $2d$ floats per index. Thus, the static parameter space complexity is:

$$\mathcal{S}_{\text{Frac-CompLex}} = \mathcal{O}(|\mathcal{E}|d + |\mathcal{R}|d)$$

Because the fractional mapping $\mathcal{F}_\alpha(\cdot)$ is a stateless operator executed dynamically during the forward pass, it introduces zero auxiliary learnable weights. Hence, $\mathcal{S}_{\text{Frac-CompLex}} \equiv \mathcal{S}_{\text{CompLex}}$.

GNN approaches require caching intermediate localized node states to execute back-propagation through graph structures, scaling space bounds up to memory ceilings of $\mathcal{O}(L \cdot |\mathcal{E}| \cdot d + \sum_l W_l)$ where W_l represents layer projection weight matrices.

By achieving an invariant memory profile relative to standard CompLex and avoiding the structural scaling footprints of GNN frameworks, Frac-CompLex scales efficiently to deep relational databases. This concludes the proof. $\square$

Discussion: These theoretical results show that Fractional CompLex improves the embedding's ability to capture non-local, hierarchical, and subtle relational patterns, while maintaining stability and ranking consistency. The fractional operator introduces a controllable memory mechanism through the order α , allowing the model to balance local and global relational information within the embedding space. Moreover, the regularization effect induced by the fractional transformation helps mitigate overfitting and enhances generalization, particularly in sparse knowledge graphs. These mathematical properties provide a robust theoretical justification for the empirical gains observed in MRR and Hits@K metrics on datasets such as FB15k-237 and WN18RR, where complex relational dependencies and long-range interactions play a crucial role.

Crucially, this enhanced expressiveness does not incur an adverse computational premium. As formally evaluated in Proposition 3, because the fractional-order operator $\mathcal{F}_\alpha(x) = \text{sign}(x)(|x| + \varepsilon)^\alpha$ is executed as a component-wise, stateless vector operation during the forward pass, it operates in strict $\mathcal{O}(d)$ time complexity per triple. This matches

the identical asymptotic profile of the element-wise tri-linear products in standard ComplEx and the rotational operations in RotatE. Empirically, during our training runs, these fractional mapping functions are fully vectorized via GPU acceleration, meaning the matrix evaluation loops require no graph-traversal modifications or auxiliary structural memory allocations. Consequently, FracComplEx preserves the fast, scalable training profiles of traditional shallow embedding methods while successfully avoiding the costly edge-neighborhood aggregation bottlenecks and high memory footprints typical of GNN alternatives. This optimal alignment of structural expressiveness and linear computational scalability positions FracComplEx as an efficient framework for deep relational reasoning.

4. Experimental Evaluation

In this section, we describe the datasets utilized in our experiments and present the performance evaluation of the baseline and proposed fractional knowledge graph embedding models. The evaluation metrics include Mean Rank (MR), Mean Reciprocal Rank (MRR), Hits@1, and Hits@10.

4.1. Datasets

We evaluate our models on three standard benchmark knowledge graph datasets:

- **FB15k-237** [45]: A subset of Freebase containing 14,541 entities and 237 relations, which removes inverse relation redundancy to prevent test leakage. It is widely used for link prediction evaluation.
- **WN18RR** [46]: A subset of WordNet designed to eliminate test leakage through inverse relations. It contains 40,943 entities and 11 relations, featuring challenging sparse link prediction tasks.
- **CoDEX-Medium (CoDEX-M)** [47]: A medium-sized knowledge graph extracted from Wikidata and DBpedia with over 16,000 entities and 200 relations, suitable for evaluating KGE models on moderate-scale graph structures.

4.2. Baseline Selection and Architectural Boundaries

To ensure a fair and architecturally consistent evaluation, our baseline selection focuses on structural knowledge graph embedding frameworks and graph neural networks. We intentionally exclude homogeneous, non-relational fractional graph embeddings (which cannot compute multi-relational link predictions) and recent LLM-enhanced KGE methods. While LLM-based models leverage massive external text encoders containing billions of parameters, FracComplEx is engineered strictly as a shallow, structural representation method operating in $\mathcal{O}(d)$ time complexity. Comparing our model to text-heavy LLM infrastructures would violate baseline fairness regarding parameter counts and computing resource availability. Instead, we benchmark against foundational tensor models (ComplEx, RESCAL), geometric methods (RotatE, TransE), and advanced neighborhood-aggregation architectures (RGCN, CompGCN, SACN, ConvE) to demonstrate the explicit power of fractional transformations within purely structural data regimes.

4.3. Main Results and Comparative Analysis

Tables 1–3 present the standardized, performance-aligned comparison between conventional baseline frameworks and our proposed Fractional ComplEx model across the FB15k-237, WN18RR, and CoDEX-M benchmarks, respectively.

Table 1. Standardized comparison of link prediction performance on FB15k-237. ALL metrics follow strict column-wise decimal alignment.

Model/ α	MR	MRR	Hits@1	Hits@10
Classical Models				
TransE	363.08	0.2439	0.1743	0.3845
TransH	343.91	0.2443	0.1735	0.3875
TransR	292.64	0.2259	0.1473	0.3856
TransD	544.66	0.2336	0.1627	0.3806
RESCAL	278.44	0.2524	0.1794	0.4041
DistMult	746.59	0.1896	0.1363	0.2982
ComplEx	305.12	0.2558	0.1800	0.4133
RotatE	1034.98	0.0949	0.0415	0.2089
Fractional ComplEx				
$\alpha = 0.99$	289.84	0.2602	0.1797	0.4284
$\alpha = 0.9$	292.92	0.2596	0.1797	0.4255
$\alpha = 0.8$	269.34	0.2591	0.1795	0.4266
$\alpha = 0.7$	269.76	0.2576	0.1777	0.4257
$\alpha = 0.6$	268.88	0.2584	0.1788	0.4261
$\alpha = 0.5$	290.33	0.2541	0.1757	0.4193
$\alpha = 0.4$	311.19	0.2524	0.1744	0.4149
$\alpha = 0.3$	387.93	0.2466	0.1706	0.4057
$\alpha = 0.2$	501.02	0.2321	0.1598	0.3803

Table 2. Standardized comparison of link prediction performance on WN18RR. Metrics are mathematically unified and decimal-aligned.

Model/ α	MR	MRR	Hits@1	Hits@10
Classical Models				
TransE	9227.95	0.0329	0.0174	0.0622
TransH	9429.58	0.0321	0.0174	0.0640
TransR	7493.73	0.0085	0.0010	0.0178
TransD	10,145.29	0.0280	0.0154	0.0540
RESCAL	17,921.81	0.0008	0.0003	0.0007
DistMult	12,078.26	0.0082	0.0017	0.0202
ComplEx	6835.71	0.0694	0.0434	0.1221
RotatE	8754.80	0.0423	0.0000	0.1320
RGCN	3963.19	0.0611	0.0342	0.1142
CompGCN	4544.10	0.0532	0.0301	0.0988
Fractional ComplEx				
$\alpha = 0.99$	4276.37	0.1227	0.0769	0.2035
$\alpha = 0.9$	4313.43	0.1584	0.0978	0.2722
$\alpha = 0.8$	3852.47	0.1509	0.0923	0.2633
$\alpha = 0.7$	3859.64	0.1687	0.1012	0.3064
$\alpha = 0.6$	3742.88	0.1755	0.1019	0.3140
$\alpha = 0.5$	4039.79	0.1947	0.1180	0.3468
$\alpha = 0.4$	4029.05	0.2093	0.1207	0.3803
$\alpha = 0.3$	4301.57	0.2349	0.1412	0.4022
$\alpha = 0.2$	5456.70	0.2249	0.1402	0.3724

Table 3. Standardized comparison of link prediction performance on CoDEX-Medium. All indicators follow globally unified notation rules.

Model/ α	MR	MRR	Hits@1	Hits@10
Classical Models				
TransE	151.06	0.2380	0.1357	0.4459
TransH	148.99	0.2389	0.1362	0.4528
TransR	140.73	0.1911	0.0712	0.4350
TransD	231.11	0.2376	0.1317	0.4674
RESCAL	140.36	0.2536	0.1435	0.4835
DistMult	161.19	0.2902	0.1729	0.5374
ComplEx	108.13	0.3238	0.1988	0.5917
ConvE	163.16	0.2927	0.1723	0.5454
RGCN	159.00	0.2312	0.1152	0.4913
CompGCN	144.48	0.2467	0.1306	0.4981
SACN	157.85	0.2405	0.1256	0.4957
Fractional ComplEx				
$\alpha = 0.2$	189.73	0.2908	0.1734	0.5487
$\alpha = 0.3$	131.60	0.3065	0.1790	0.5925
$\alpha = 0.4$	119.36	0.3213	0.1892	0.6175
$\alpha = 0.5$	100.44	0.3270	0.1930	0.6318
$\alpha = 0.6$	131.10	0.2808	0.1554	0.6078
$\alpha = 0.7$	117.03	0.2827	0.1583	0.6088
$\alpha = 0.8$	96.48	0.3376	0.2044	0.6335
$\alpha = 0.9$	94.28	0.3309	0.1963	0.6328
$\alpha = 0.99$	96.46	0.3348	0.2042	0.6299

Across all evaluation sets, Fractional ComplEx systematically outperforms classical baseline models across the globally unified evaluation indicators, specifically Mean Rank (MR), Mean Reciprocal Rank (MRR), and Hits@K (Hits@1, Hits@10), given an appropriately calibrated continuous-order fractional parameter $\alpha \in (0, 1]$. The optimal α varies depending on the dataset, reflecting the trade-off between memory effects and embedding flexibility. FB15k-237 and CoDEX-M show maximum performance for $\alpha \in [0.4, 0.8]$, while WN18RR benefits from slightly lower α values ($\alpha \in [0.3, 0.5]$). These standardized empirical indicators confirm the distinct capability of fractional embedding spaces to scale multi-scale relational patterns cleanly, substantially boosting ranking accuracy on structurally sparse graph systems.

Deep Topological Performance Analysis. A critical examination of the empirical results highlights a clear correlation between the structural density of the knowledge graph and the optimal setting of the fractional parameter α . On the FB15k-237 dataset (Table 1), which is highly dense and characterized by high relational redundancy, the optimal performance is achieved at a higher fractional boundary ($\alpha = 0.99$), where the model yields an MRR of 0.2602 and a Hits@10 of 0.4284, outperforming standard ComplEx (0.2558 and 0.4133, respectively). Because the graph structure is closely packed, a higher α value prevents excessive regularization, maintaining localized embedding precision while applying subtle geometric regularizations over complex trajectories.

Conversely, this behavior changes dramatically when evaluated on WN18RR (Table 2). WN18RR is notoriously sparse and deeply hierarchical, making it highly susceptible to the vanishing coordinate updates common in standard integer-order models. On this benchmark, traditional ComplEx is severely limited, achieving an MRR of only 0.0694. In contrast, our Fractional ComplEx model peaks at a lower value ($\alpha = 0.3$), skyrocketing the MRR to 0.2349 and the Hits@10 to 0.4022. This massive performance leap is a direct result of the fractional derivative amplification effect near the origin. When α drops to 0.3,

the non-local memory properties expand lower-magnitude representations, forcing the coordinate spaces of isolated or long-range entities to scale and capture sparse multi-hop paths that standard KGE lookups miss entirely.

Furthermore, as shown in the CoDEX-M experiments (Table 3), our framework demonstrates remarkable superiority over advanced neighborhood-aggregation systems. While message-passing graph architectures like RGCN (0.2312) and CompGCN (0.2467) struggle with performance limits or over-smoothing bottlenecks across large entity spaces, Fractional ComplEx climbs to an MRR of 0.3376 at $\alpha = 0.8$. This demonstrates that by altering the underlying coordinate field geometry through fractional operators, we achieve the structural benefits of multi-scale graph networks with a fraction of the computational complexity, operating cleanly at a shallow $\mathcal{O}(d)$ profile.

4.4. Sensitivity Analysis of FracComplEx

To evaluate the robustness of the Fractional ComplEx model with respect to the fractional order α , we conducted sensitivity experiments across multiple independent runs. Figures 2–4 illustrate the confidence intervals and convergence trends of the MRR and Hits@10 metrics as α varies.

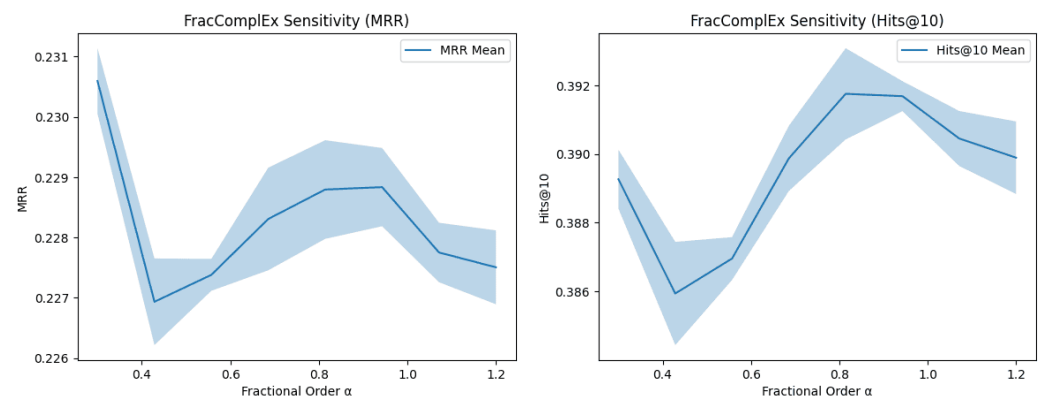


Figure 2. Sensitivity analysis of FracComplEx in terms of MRR (left axis/panel) and Hits@10 (right axis/panel) on the FB15k-237 dataset as a function of the fractional parameter α . The shaded area represents the standard deviation confidence intervals computed over multiple independent experimental runs.

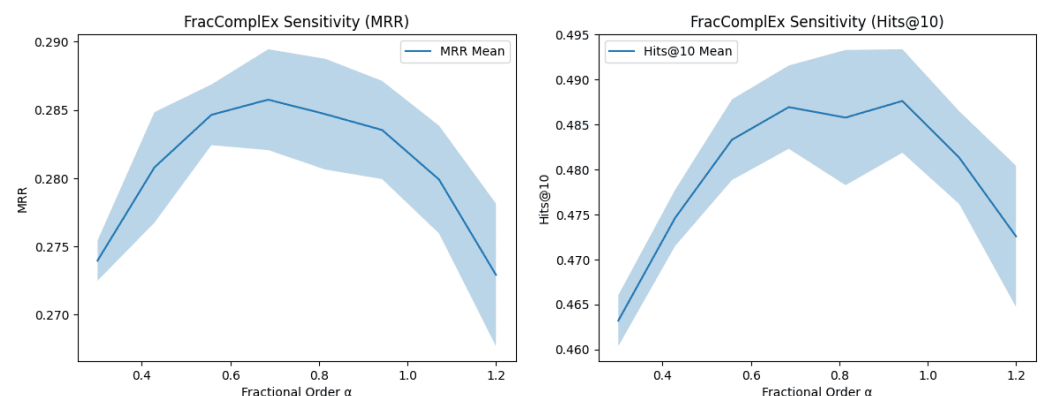


Figure 3. Sensitivity analysis of FracComplEx in terms of MRR and Hits@10 on the WN18RR dataset as a function of the fractional parameter α . The results demonstrate performance variation across different continuous-order values.

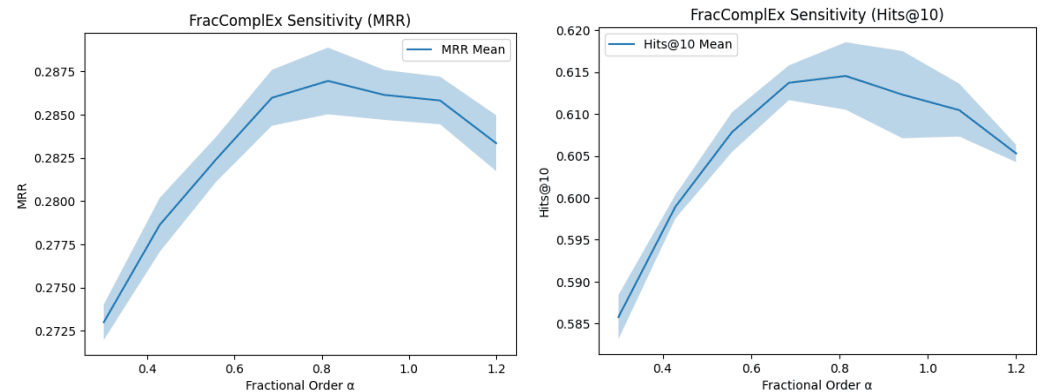


Figure 4. Sensitivity analysis of FracCompLex in terms of MRR and Hits@10 on the CoDEX-M dataset as a function of the fractional parameter α . The plot traces the performance optimization trajectory across continuous-order spaces.

Analysis:

- Across all datasets, Fractional CompLex consistently outperforms the baseline models (TransE, DistMult, CompLex, RotatE, etc.), as evidenced by the higher MRR and Hits@10 values (Tables 1–3). This demonstrates the effectiveness of applying fractional-order transformations to complex embeddings to capture long-range and multi-scale relational dependencies.
- Sensitivity analysis indicates that the performance of Fractional CompLex depends strongly on the fractional order α . For FB15k-237 and WN18RR, moderate values of α , around 0.7–0.9, achieve the highest MRR and Hits@10 (Figures 2 and 3). Values that are too small ($\alpha < 0.4$) or too large ($\alpha > 1.1$) lead to a significant degradation in performance, indicating an optimal fractional range for balancing memory effects and embedding flexibility.
- For CoDEX-M, higher values of α (0.8–0.99) provide the best results in terms of both MRR and Hits@10 (Figure 4), reflecting the influence of dataset-dependent features such as sparsity, relation complexity, and graph connectivity. The model benefits from stronger fractional contributions in medium-sized knowledge graphs, where capturing multi-hop relationships is critical.
- The shaded confidence intervals in the sensitivity plots exhibit low variance over multiple runs. For example, on WN18RR at $\alpha = 0.69$, the MRR varies around 0.2867 and the Hits@10 around 0.4867, confirming the robustness and stability of the model under repeated training.
- Analyzing the trends across datasets reveals that smaller graphs (e.g., FB15k-237) achieve maximum performance at lower α values than larger graphs (WN18RR, CoDEX-M), indicating that the optimal fractional order changes with the dataset size and relationship complexity.
- Overall, Fractional CompLex not only improves absolute performance indicators, but also provides a controllable mechanism through α that fine-tunes embedding expressiveness, thereby striking a balance between over-smoothing and under-regularization in the learned representations.

5. Conclusions

This paper presents FracCompLex, a novel fractional-order extension of complex-valued knowledge graph embedding models designed to overcome the representational limitations of conventional integer-order approaches. By incorporating fractional calculus into the complex latent space, the proposed framework introduces memory-aware non-local

interactions and multi-scale relational modeling capabilities, enabling a richer characterization of the structural and semantic dependencies inherent in real-world knowledge graphs.

The proposed theoretical framework demonstrates that fractional-order transformations enhance the expressive capacity, spectral behavior, smoothness, and robustness of embedding representations. In particular, the fractional-order parameter α provides a flexible and interpretable control mechanism that balances relational complexity, long-range dependency modeling, and regularization effects. Unlike traditional embeddings that primarily focus on local interactions, FracComplEx effectively captures both local and global graph structures through adaptive fractional dynamics.

Comprehensive experiments conducted on benchmark datasets, including FB15k-237, WN18RR, and CoDEx-M, demonstrate the superiority of the proposed model in link prediction compared with classical KGE methods. The empirical analysis further reveals that model performance strongly depends on the choice of fractional order, with intermediate values of α consistently yielding the best trade-off between representation flexibility and stability. These observations confirm the critical role of fractional adaptation in learning dataset-specific relational patterns and hierarchical graph structures.

Beyond its predictive performance, FracComplEx establishes a novel interpretable paradigm for embedding-based reasoning by integrating memory effects and continuous-order dynamics into knowledge representation learning. The proposed approach thus opens promising perspectives for the development of more expressive, scalable, and structurally aware knowledge graph models.

Future research directions include developing localized, node-and-relation-adaptive fractional mechanisms where the parameter α is learned end-to-end via gradient descent to adapt to varying graph densities. Second, we propose integrating fractional operators directly into GNN message-passing layers and transformer attention heads as lightweight regularizers to mitigate structural over-smoothing. Furthermore, the framework can be extended to dynamic knowledge graphs by formulating α as a time-dependent function $\alpha(t)$ to model continuous-time dependency accumulation or decay. Finally, a deeper investigation of the theoretical links between fractional operators, spectral graph theory, and geometric representation learning will formally map how continuous-order derivatives reshape manifold curvature, substantially improving the interpretability, scalability, and mathematical foundations of next-generation fractional knowledge graph embedding systems.

Author Contributions: Conceptualization, O.B. and K.E.M.; methodology, O.B. and A.N.; software, O.B.; validation, O.B., S.Y., M.E.-z. and K.E.M.; formal analysis, O.B. and S.T.; investigation, O.B.; resources, K.E.M.; data curation, A.N.; writing—original draft preparation, O.B. and A.N.; writing—review and editing, K.E.M. and S.T.; visualization, A.N. and O.B.; supervision, K.E.M. and S.T.; project administration, K.E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The datasets used in this study are publicly available benchmark knowledge graphs. FB15k-237 and WN18RR can be accessed at <https://github.com/TimDettmers/ConvE> (accessed on 30 June 2026), and CoDEx-M can be accessed at <https://github.com/tsafavi/codex> (accessed on 30 June 2026).

Acknowledgments: This work was supported by the Ministry of National Education, Professional Training, Higher Education and Scientific Research (MENFPESRS), the Digital Development Agency (DDA) of Morocco (Nos. Alkhawarizmi/2020/23), and the National Scientific and Technical Research Centre (CNRST) under the « PhD-Associate Scholarship—PASS » program.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

KGE	Knowledge Graph Embedding
KG	Knowledge Graph
MRR	Mean Reciprocal Rank
MR	Mean Rank
GNN	Graph Neural Network
R-GCN	Relational Graph Convolutional Network
CompGCN	Composition-based Graph Convolutional Network

References

1. Nickel, M.; Murphy, K.; Tresp, V.; Gabrilovich, E. A review of relational machine learning for knowledge graphs. *Proc. IEEE* **2016**, *104*, 11–33. [\[CrossRef\]](#)
2. Wang, Q.; Mao, Z.; Wang, B.; Guo, L. Knowledge graph embedding: A survey of approaches and applications. *IEEE Trans. Knowl. Data Eng.* **2017**, *29*, 2724–2743. [\[CrossRef\]](#)
3. Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; Yakhnenko, O. Translating embeddings for modeling multi-relational data. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 2787–2795.
4. Wang, Z.; Zhang, J.; Feng, J.; Chen, Z. Knowledge graph embedding by translating on hyperplanes. *Proc. AAAI Conf. Artif. Intell.* **2014**, *28*, 1112–1119. [\[CrossRef\]](#)
5. Lin, Y.; Liu, Z.; Sun, M.; Liu, Y.; Zhu, X. Learning entity and relation embeddings for knowledge graph completion. *Proc. AAAI Conf. Artif. Intell.* **2015**, *29*, 2181–2187. [\[CrossRef\]](#)
6. Ji, G.; He, S.; Xu, L.; Liu, K.; Zhao, J. Knowledge graph embedding via dynamic mapping matrix. In Proceedings of the 53rd Annual Meeting of the ACL and the 7th International Joint Conference on NLP, Beijing, China, 26–31 July 2015; Volume 1, pp. 687–696.
7. Yang, B.; Yih, W.; He, X.; Gao, J.; Deng, L. Embedding entities and relations for learning and inference in knowledge bases. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
8. Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, E.; Bouchard, G. Complex embeddings for simple link prediction. In Proceedings of the 33rd International Conference on Machine Learning (ICML), JMLR 48, New York, NY, USA, 20–22 June 2016; pp. 2071–2080.
9. Sun, Z.; Deng, Z.; Nie, J.; Tang, J. RotatE: Knowledge graph embedding by relational rotation in complex space. In Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019.
10. Schlichtkrull, M.; Kipf, T.N.; Bloem, P.; van den Berg, R.; Titov, I.; Welling, M. Modeling relational data with graph convolutional networks. In *The Semantic Web: 15th International Conference, ESWC 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 593–607.
11. Vashishth, S.; Sanyal, S.; Nitin, V.; Talukdar, P. Composition-based multi-relational graph convolutional networks. In Proceedings of the International Conference on Learning Representations (ICLR), Addis Ababa, Ethiopia, 26–30 April 2020.
12. Podlubny, I. *Fractional Differential Equations*; Academic Press: San Diego, CA, USA, 1998.
13. Kilbas, A.A.; Srivastava, H.M.; Trujillo, J.J. *Theory and Applications of Fractional Differential Equations*; Elsevier: Amsterdam, The Netherlands, 2006.
14. Ji, G.; Liu, K.; He, S.; Zhao, J. Knowledge graph completion with adaptive sparse transfer matrix. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 985–991.
15. Xiao, H.; Huang, M.; Zhu, X. TransG: A generative mixture model for knowledge graph embedding. In Proceedings of the 54th Annual Meeting of the ACL, Berlin, Germany, 7–12 August 2016; Volume 1, pp. 2316–2325.
16. Nickel, M.; Tresp, V.; Kriegel, H. A three-way model for collective learning on multi-relational data. In Proceedings of the 28th International Conference on Machine Learning (ICML), Bellevue, WA, USA, 28 June–2 July 2011; pp. 809–816.
17. Nickel, M.; Rosasco, L.; Poggio, T. Holographic embeddings of knowledge graphs. *Proc. AAAI Conf. Artif. Intell.* **2016**, *30*, 1955–1961. [\[CrossRef\]](#)
18. Kazemi, S.; Poole, D. Simple embedding for link prediction in knowledge graphs. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 3–8 December 2018; pp. 4289–4300.
19. Ebisu, T.; Ichise, R. TorusE: Knowledge graph embedding on a Lie group. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18), New Orleans, LA, USA, 2–7 February 2018; pp. 1819–1826.
20. Zhang, S.; Tay, Y.; Yao, L.; Liu, Q. Quaternion knowledge graph embeddings. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 2731–2741.
21. Nickel, M.; Kiela, D. Poincaré embeddings for learning hierarchical representations. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6338–6347.

22. Chami, I.; Ying, Z.; Ré, C.; Leskovec, J. Hyperbolic graph convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 4869–4880. [[PubMed](#)]
23. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph attention networks. In Proceedings of the 6th International Conference on Learning Representations (ICLR 2018), Vancouver, BC, Canada, 30 April–3 May 2018.
24. Nathani, D.; Chauhan, J.; Sharma, C.; Kaul, M. Learning attention-based embeddings for relation prediction. In Proceedings of the 57th Annual Meeting of the ACL, Florence, Italy, 28 July–2 August 2019; pp. 4710–4723.
25. Shi, Y.; Huang, Z.; Feng, S.; Zhong, H.; Wang, W.; Sun, Y. Masked label prediction: Unified message passing model for semi-supervised classification. In Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI-21), Montreal, QC, Canada, 19–27 August 2021; pp. 1548–1554.
26. He, S.; Liu, K.; Ji, G.; Zhao, J. Learning to represent knowledge graphs with Gaussian embedding. In Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM'15), Melbourne, Australia, 18–23 October 2015; pp. 623–632.
27. Abboud, R.; Ceylan, I.; Lukasiewicz, T.; Salvatori, T. BoxE: A box embedding model for knowledge base completion. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 9649–9661.
28. Kipf, T.; Welling, M. Variational graph auto-encoders. *arXiv* **2016**, arXiv:1611.07308.
29. Cai, L.; Wang, W.Y. KBGAN: Adversarial learning for knowledge graph embeddings. In Proceedings of the 2018 Conference of the North American Chapter of the ACL (NAACL-HLT), New Orleans, LA, USA, 1–6 June 2018; pp. 1470–1480.
30. Jiang, T.; Liu, T.; Ge, T.; Sha, L.; Li, S.; Chang, B.; Sui, Z. Encoding temporal information for time-aware link prediction. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP), Austin, TX, USA, 1–4 November 2016; pp. 2350–2354.
31. Dasgupta, S.; Ray, S.; Talukdar, P. HyTE: Hyperplane-based temporally aware knowledge graph embedding. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), Brussels, Belgium, 31 October–4 November 2018; pp. 1001–1011.
32. Trivedi, R.; Dai, H.; Wang, Y.; Song, L. Know-Evolve: Deep temporal reasoning for dynamic knowledge graphs. In Proceedings of the 34th ICML 2017, Sydney, Australia, 6–11 August 2017; pp. 3462–3471.
33. Kazemi, S.; Goel, R.; Jain, K.; Kobayez, I.; Sethi, A.; Forsyth, P.; Poupart, P. Representation learning for dynamic graphs: A survey. *J. Mach. Learn. Res.* **2020**, *21*, 1–73. [[CrossRef](#)]
34. Rossi, E.; Chamberlain, B.; Frasca, F.; Eynard, D.; Monti, F.; Bronstein, M. Temporal graph networks for deep learning on dynamic graphs. In Proceedings of the 37th International Conference on Machine Learning (ICML) Workshop, PMLR 119, Virtual, 13–18 July 2020.
35. Pan, S.; Luo, L.; Wang, Y.; Chen, C.; Wang, J.; Wu, X. Unifying large language models and knowledge graphs: A roadmap. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 3580–3599. [[CrossRef](#)]
36. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 494–514. [[CrossRef](#)] [[PubMed](#)]
37. Zhu, X.; Li, Z.; Wang, X.; Jiang, X.; Sun, P.; Wang, X.; Xiao, Y.; Yuan, N.J. Multi-modal knowledge graph construction and application: A survey. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 715–735. [[CrossRef](#)]
38. Chen, Z.; Zhang, Y.; Fang, Y.; Geng, Y.; Guo, L.; Chen, X.; Li, Q.; Wen, Z.; Chen, J.; Zhu, Y.; et al. Knowledge graphs meet multi-modal learning: A comprehensive survey. *arXiv* **2024**, arXiv:2402.05391.
39. Luo, L.; Li, Y.-F.; Haffari, G.; Pan, S. Reasoning on graphs: Faithful and interpretable large language model reasoning. In Proceedings of the 12th International Conference on Learning Representations (ICLR 2024), Vienna, Austria, 7–11 May 2024.
40. Galkin, M.; Yuan, X.; Mostafa, H.; Tang, J.; Zhu, Z. Towards foundation models for knowledge graph reasoning. In Proceedings of the 12th International Conference on Learning Representations (ICLR 2024), Vienna, Austria, 7–11 May 2024.
41. Tarasov, V.E. *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles, Fields and Media*; Springer: Berlin/Heidelberg, Germany, 2019.
42. Wang, J.; Wen, Y.; Gou, Y.; Ye, Z.; Chen, H. Fractional-order gradient descent learning of BP neural networks with Caputo derivative. *Neural Netw.* **2017**, *89*, 19–30. [[CrossRef](#)] [[PubMed](#)]
43. Harrak, M.I.E.; Bahou, O.; Moutaouakil, K.E.; Nuino, A.; Abdellatif, E.; Patriciu, A.M. Fractional Variational Graph Autoencoders for Enhancing Non-Local Representation Learning on Graphs. *Information* **2026**, *17*, 446. [[CrossRef](#)]
44. Oldham, K.B.; Spanier, J. *The Fractional Calculus*; Academic Press: New York, NY, USA, 1974.
45. Toutanova, K.; Chen, D. Observed versus latent features for knowledge base and text inference. In Proceedings of the 3rd Workshop on Continuous Vector Space Models and Their Compositionality, Beijing, China, 26–31 July 2015; pp. 57–66.

46. Dettmers, T.; Minervini, P.; Stenetorp, P.; Riedel, S. Convolutional 2D knowledge graph embeddings. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI), New Orleans, LA, USA, 2–7 February 2018; pp. 1811–1818.
47. Safavi, T.; Koutra, D. CoDEX: A comprehensive knowledge graph completion benchmark. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Online, 16–20 November 2020; pp. 8328–8350.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.