Dihedron Algebraic Embeddings for Spatio-Temporal Knowledge Graph Completion

Mojtaba Nayyeri[1,5], Sahar Vahdati[2], Md Tansen Khan[1(✉)], Mirza Mohtashim Alam[2], Lisa Wenige[2], Andreas Behrend[3], and Jens Lehmann[1,4]

[1] University of Bonn, Bonn, Germany
nayyeri@cs.uni-bonn.de, s6mmkhan@uni-bonn.de
[2] Institute for Applied Informatics (InfAI), Dresden, Germany
{vahdati,alam,wenige}@infai.org
[3] Institute for Telecommunications (INT), TH Köln, Cologne, Germany
andreas.behrend@th-koeln.de
[4] Fraunhofer IAIS, Dresden, Germany
jens.lehmann@iais.fraunhofer.de
[5] University of Stuttgart, Stuttgart, Germany

Abstract. Many knowledge graphs (KG) contain spatial and temporal information. Most KG embedding models follow triple-based representation and often neglect the simultaneous consideration of the spatial and temporal aspects. Encoding such higher dimensional knowledge necessitates the consideration of true algebraic and geometric aspects. Hypercomplex algebra provides the foundation of a well defined mathematical system among which the Dihedron algebra with its rich framework is suitable to handle multidimensional knowledge. In this paper, we propose an embedding model that uses Dihedron algebra for learning such spatial and temporal aspects. The evaluation results show that our model performs significantly better than other adapted models.

Keywords: Knowledge graph · Embedding · Spatio-temporal

1 Introduction

Large cross-domain Knowledge Graphs (KGs), such as DBpedia [15] and Wikidata [27], leverage a triple representation of facts in the form of (h, r, t) where h, t and r refer to head and tail entities and the relation respectively. Despite the availability of huge amounts of such data, one of the major challenges is the impossibility of capturing all (true) facts of the target domain. Thus, sparsity and incompleteness are major problems of KGs. The objective of KG completion is to generate true triples that are not explicitly given in the KG. For example, the query $(PrinceWilliam, met, ?)$, with an unknown tail as "?", means *"With whom did Prince William meet?"* for which $h = VolodymyrZelensky$ is a possible answer. This leads to predict the $(PrinceWilliam, met, VolodymyrZelensky)$ triple.

P. Groth et al. (Eds.): ESWC 2022, LNCS 13261, pp. 253–269, 2022.
https://doi.org/10.1007/978-3-031-06981-9_15

Knowledge Graph Embedding (KGE) models have shown high performance for KG completion. A KGE model usually maps the entities and relations into a d dimensional vector space (e.g. $\mathbb{R}^d$). The plausibility of a triple (h, r, t) is measured via a score function $f(h, r, t)$. In this way, a KGE model performs the KG completion task by replacing the potential entities or relations in incomplete triple patterns. After measuring the plausibility, the triples with high scores are regarded to be likely true and can be added to the KG for completing it or undergoing further human inspection. This is a widely used approach for major KGE models that are only designed for triple-based KGs. However, important semantic aspects of some facts are neglected when not considering Spatio-temporal aspects. For example, for the triple $(PrinceWilliam, met, VolodymyrZelensky)$ it would be highly relevant to know the location and time of the meeting. We found that at least 13% of the resources in DBpedia fall into this category where triples are connected to additional information via time, location, or both.

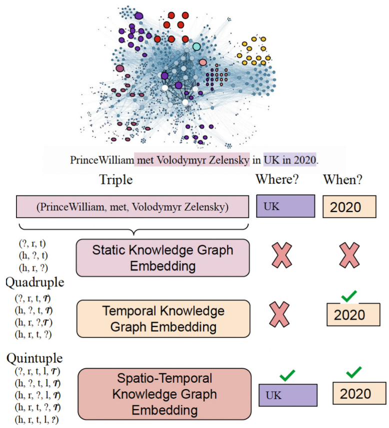

Fig. 1. Spatio-temporal treated by KGEs.

However, the existing KGE models often only consume triples and are not capable of exploiting the additional spatial and temporal dimension of facts. Recent attempts in temporal knowledge graph embeddings advance consideration of temporal aspects [31], but do not consider spatial information. In those models, facts are then represented as quadruples (h, r, t, τ) where τ is the temporal information. Therefore, the respective models are capable to complete quadruples of the form $(?, r, t, \tau)$, $(h, ?, t, \tau)$ or $(h, r, ?, \tau), (h, r, t, ?)$. However, none of the current models can directly consider spatial information. Spatio-temporal facts can be represented as a quintuple (h, r, t, l, τ) where l reflects the location information (spatial). Such quintuples, e.g. $(PrinceWilliam, met, VolodymyrZelensky, 2020, UK)$, can simultaneously associate time and space information to a given fact. As shown in Fig. 1 time and location are excluded from most current KGE models and thus potentially valuable Spatio-temporal information remains unused for completion tasks. This might in turn reduce the performance as well as the meaningfulness and interpretability of machine learning results. Therefore, we propose a family of KGE models to fully exploit the spatial and temporal information for KG completion. We specifically take advantage of the 4D algebra of hypercomplex vectors to complete a quintuple representation in the form $(?, r, t, l, \tau), (h, ?, t, l, \tau)$, $(h, r, ?, l, \tau)$, $(h, r, t, ?, \tau)$ or $(h, r, t, l, ?)$. This is especially suitable because for each incomplete quintuple, four of the five elements are always present. All of these four elements (entity (h or t), relation r, location l, time τ) are assumed to be mutually independent which will be represented by 4 orthogonal bases. In this work, we employ the Dihedron algebra, which is a rich 4D algebra of hypercomplex spaces and provides a well-suited theoretical foundation for the embedding of quintuples considering time and location aspects. Our contributions are as follows: We present a) a family of novel KGE models using Dihedron algebra to capture Spatio-temporal information, b) a technique that allows existing KGE models to be adapted to include Spatio-temporal information, c) a geometric interpretation of the used algebra for the problem of encoding Spatio-temporal information, d) several Spatio-temporal KGs (ST-KGs) derived from YAGO3K, DBpedia34K, and WikiData53K.

2 Related Work

Static Knowledge Graph Embedding Most existing KGE models learn over KGs containing triples of the form (h, r, t). One of the primary KGE models is TransE [2]. For a given positive triple, TransE represents a relation r as translation from head to tail i.e. $\boldsymbol{h} + \boldsymbol{r} = \boldsymbol{t}$ where $\boldsymbol{h}, \boldsymbol{r}, \boldsymbol{t}$ are embedding vectors of head, relation and tail. In order to address the limitations of TransE in encoding various relational patterns such as symmetry, transitivity etc., several variants of TransE have been proposed such as TransH [28], TransR [18], TransD [8] etc. In RotatE, each relation is represented as a rotation in the complex space. Due to algebraic characteristics of rotation, e.g. a) every rotation matrix has a unique inverse, b) composition of two rotation matrices is a rotation matrix, this model can encode various relational patterns such as inversion and composition [24]. The ComplEx model embeds KGs into a complex vector space which together with N3 regularization as in ComplEx-N3, has become one of the state of the art KGE models [13,26]. ComplEx and ComplEx-N3 efficiently model symmetric and anti-symmetric patterns. QuatE extends ComplEx to Quaternion vector representation and obtains state-of-the-art results on link prediction over static KGs when it is combined with N3 regularization i.e. QuatE-N3 [33]. The matrix representation from Dihedral groups [30] has been used for modeling each relation of a KG to represent various patterns such as skew-symmetric, inversion, and composition in static KGs. There are several other static KGE models that can be found in [3,9,17,21].

Temporal Knowledge Graph Embedding Temporal Knowledge Graph Embedding (TKGE) models focus on dynamic KGs with an additional temporal part. In this way, triple-based representation is formed as quadruple. Most of the early TKGE models have been built on top of the already existing KGEs. The HyTE [4] model is one of the early TKGEs that first projects the head, relation, and tail embeddings to the space of the timestamp. Furthermore, for the final scoring of the newly predicted facts, it employs TransE on the projected embeddings. There are several other TKGEs which have been proposed as extensions of TransE such as TTransE [14] and TA-TransE [6]. HyTE and other extensions do not consider any hypercomplex algebraic aspects that could let the model cover the spatial information beside the temporal ones. The other state of the art model among TKGEs is the ConT model that is an extension of the Tucker [1] KGE. There are also several extensions of DisMult [32] that were proposed for encoding of temporal KGs such as TDistMult [19] and TA-DistMult [6]. These models are based on recurrent neural network (RNNs) that captures the entity embeddings for head and tail parts. RE-NET is another RNN-based TKGE that captures pair-wise knowledge in the form of (head, relation) or (tail, relation) by using specific patterns from the historical information between entities [10]. However, none of these extensions exploits the algebraic aspects of the embedding models or attempts to consider the spatial aspects.

Another issue of these models is that all of them inherit the problems of the underlying base models on top of which they are extended. For example, the TKGEs that were built on top of TransE suffer from encoding of certain relational patterns. The TeRo model [31] has been recently proposed to overcome such problems of the already existing TKGEs on inference of relational patterns. Although TeRo solves the limitation of other models to some extend, it does not leverage the characteristics of hypercomplex spaces, also does not target ST-KGs. The TComplEx model [12] is temporal version of ComplEx-N3, that obtained state-of-the art performance on link prediction over temporal KGs.

Other Works related to Spatio-Temporal In our work, a spatio-temporal context is denoted as a combination of a location as cities or countries and a time slot with the granularity of year, e.g., Spain, 1982. It shall be noted that in other domains, the term "spatial" also refers to the detailed geographical aspects of locations either with satellite information or geo-coordinates information. There are different tracks of interdisciplinary research related to the later meaning of spatial data such as geospatial artificial intelligence (GeoAI) combining geography, earth science and artificial intelligence which are not directly in the scope of our work [7,16,22]. In [34], a general framework for analysing multi-source spatio-temporal data is given. Although, this approach is based on KGEs, the meaning of spatial data lies in the urban data, land maps and satellite data.

3 Dihedron Algebra

Dihedron is a hypercomplex number system, extending complex number to 4D space. The space is denoted by $\mathbb{D}$. A Dihedron number $d = x_r\mathbf{1} + x_i\boldsymbol{i} + x_j\boldsymbol{j} + x_k\boldsymbol{k} \in \mathbb{D}$ includes one real part x_r and three imaginary parts $x_i\boldsymbol{i}$, $x_j\boldsymbol{j}, x_k\boldsymbol{k}$, where the bases are $\mathbf{1} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \boldsymbol{i} = \begin{bmatrix}0 & 1\\ \bar{1} & 0\end{bmatrix}, \boldsymbol{j} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \boldsymbol{k} = \begin{bmatrix}1 & 0\\ 0 & \bar{1}\end{bmatrix}$, $\bar{a} = -a$, and the following equations hold $\boldsymbol{i^2} = \bar{\mathbf{1}}$, $\boldsymbol{j^2} = \mathbf{1}$, $\boldsymbol{k^2} = \mathbf{1}, \boldsymbol{ij} = \boldsymbol{k}$, $\boldsymbol{ij} = \boldsymbol{k}$, $\boldsymbol{jk} = \bar{\boldsymbol{i}}$, $\boldsymbol{ki} = \boldsymbol{j}$, $\boldsymbol{ji} = \bar{\boldsymbol{k}}$, $\boldsymbol{kj} = \boldsymbol{i}$, and $\boldsymbol{ik} = \bar{\boldsymbol{j}}$. Some other mathematical representations and operations are as follows:

Dihedron Matrix Representation A Dihedron number $d \in \mathbb{D}$ can be represented as the following matrix form

$$q = x_r\mathbf{1} + x_i\boldsymbol{i} + x_j\boldsymbol{j} + x_k\boldsymbol{k} = \begin{bmatrix}x_r + x_k & x_i + x_j\\ \bar{x}_i + x_j & x_r + \bar{x_k}\end{bmatrix} \equiv (x_r, x_i, x_j, x_k) \equiv$$
$$x_r + v_q, \;\; v_q = (x_i, x_j, x_k) \in \mathbb{R}^{3d}\, q \in \mathbb{D}.$$

Dihedron Product The Dihedron product (denoted by $\otimes$) between two Dihedron numbers $q_x, q_y \in \mathbb{D}$ is defined as follows

$$\begin{aligned} q_z = z_r\mathbf{1} + v_{q_z} &= q_x \otimes q_y = (x_r\mathbf{1} + v_{q_x}) \otimes (y_r\mathbf{1} + v_{q_y}) \\ &= (x_r y_r - \langle v_{q_x}, v_{q_y}\rangle)\mathbf{1} + (x_r v_{q_y} + y_r v_{q_x}) + v_{q_x} \times v_{q_y}, \end{aligned} \tag{1}$$

where $\langle v_{q_x}, v_{q_y} \rangle = x_i y_i - x_j y_j - x_k y_k, \quad v_{q_x} \times v_{q_y} = \begin{bmatrix} -x_j y_k + x_k y_j \\ x_k y_i - x_i y_k \\ x_i y_j - x_y y_i \end{bmatrix}.$

The product incorporates all element factors into computation. We will explain the advantage of this product in modeling spatio-temporal data in the next section.

4 Proposed Approach

Here, we introduce the proposed approach dubbed ST-NewDE based on Dihedron Algebra. Let us assume the following fact *"Prince William met Volodymyr Zelensky in UK in 2020."* that contains five parts namely subject (*Prince William*), relation (*meet*), object (*Volodymyr Zelensky*), adverb of place (*UK*), and adverb of time (*2020*) . Incompleteness occurs when one of the parts is missing at a time. In a ST-KG, those could be seen as a quintuple in the form $(?, r, t, l, \tau)$, $(h, ?, t, l, \tau)$, $(h, r, ?, l, \tau)$, $(h, r, t, ?, \tau)$ or $(h, r, t, l, ?)$. One efficient way towards completing such a KG is to represent such spatio-temporal queries in the vector space where four elements are embedded separately in a real vector space. Then, each of those four parts are combined to build up the query representation as 4D hypercomplex vectors. To complete the missing part $?$, we employ a rich 4D algebra of hypercomplex space named Dihedron that is used to measure the similarity of the query and the plausible answers (i.e. possible entities for object), while capturing the mutual correlation between each pair elements.

Spatio-Temporal KG Let us have a spatio-temporal KG $\mathcal{K} = \{(h, r, t, l, \tau)|\ h, t \in \mathcal{E}, r \in \mathcal{R}, l \in \mathcal{L}, \tau \in \mathcal{T}\}$, where $\mathcal{E}, \mathcal{R}, \mathcal{L}, \mathcal{T}$ are entity, relation, location and time dictionaries respectively.

Embedding Space Each entity ($e = h, t$), relation (r), location (l) and time (τ) are embedded into d dimensional real vector space, shown in bold i.e. $\boldsymbol{e}, \boldsymbol{r}, \boldsymbol{l}, \boldsymbol{\tau} \in \mathbb{R}^d$. Note that an entity ($e \in \mathcal{E}$) plays both roles of subject (head h) or object (tail t) in a quintuple. If an entity is in the subject role, the embedding is shown as $\boldsymbol{h}$ ($\boldsymbol{t}$ for subject role).
Incomplete Quintuples and Answers Given a quintuple $\{(h, r, t, l, \tau)$, we split it into two parts: incomplete quintuples (IQ) and answer (A) as shown in Table 1. In this way,

Table 1. Incomplete quintuples and their answers in spatio-temporal knowledge graphs.

IQ	$(?, r, t, l, \tau)$	$(h, ?, t, l, \tau)$	$(h, r, ?, l, \tau)$	$(h, r, t, ?, \tau)$	$(h, r, t, l, ?)$
A	h	r	t	l	τ

each incomplete quintuple contains four fixed parts. Because 4D spaces contain four mutually orthogonal bases, they are the most suitable algebraic representation for such quintuples with four fixed parts. Among 4D spaces, Dihedron and Quaternion are two of the main and well-defined hypercomplex algebras [23, 29]. Dihedron representation covers a wider range of geometric shapes than the Quaternion in the 4D vector space. Consequently, it is more expressive and thus better suited for encoding the complex spatio-temporal incomplete quintuples in the 4D space. Figure 2 illustrates the flexibility of Dihedron over Quaternion in a 4D vector space. From the geometric view point, Dihedron algebra represents various hyperboloid models (two-sheet, one-sheet) as well as conic surfaces and spheres. In this regard, we consider Dihedron algebra as a *true (well-suited) algebra* for modeling complex spatio-temporal incomplete quintuples.

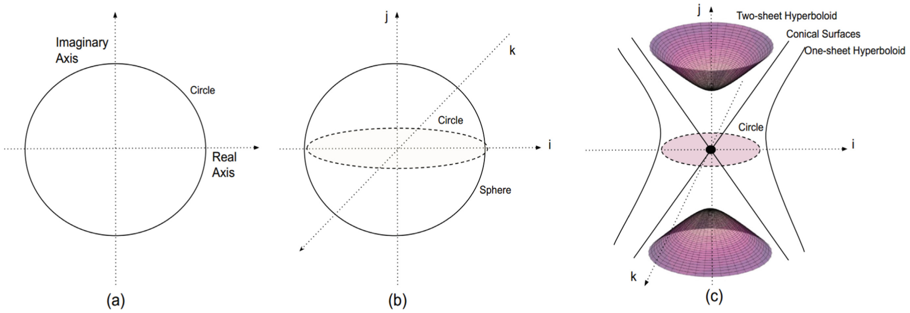

Fig. 2. Illustration of (a) Complex plane; (b) Quaternion space; (c) Dihedrons space.

Dihedron Representation of Incomplete Quintuples: We represent a 4D spatio-temporal query (incomplete quintuples) of the form $(h, r, ?, l, \tau)$ as a $d \times 4$D vector in a Dihedron space $\mathbb{D}^d$ i.e. $\boldsymbol{q} = \boldsymbol{h1} + \boldsymbol{ri} + \boldsymbol{lj} + \boldsymbol{\tau k}$, where $\boldsymbol{q}$ is the generalization of complex numbers with three imaginary units $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$, and also we have $\boldsymbol{h}, \boldsymbol{r}, \boldsymbol{l}, \boldsymbol{\tau} \in \mathbb{R}^d$. We specifically focus on Dihedron algebra where the detailed description was introduced in previous section. Therefore, a Dihedron query is represented as follows

$$\boldsymbol{q} = \boldsymbol{h1} + \boldsymbol{ri} + \boldsymbol{lj} + \boldsymbol{\tau k} = \begin{bmatrix} \boldsymbol{h} + \boldsymbol{\tau} & \boldsymbol{r} + \boldsymbol{l} \\ \bar{\boldsymbol{r}} + \boldsymbol{l} & \boldsymbol{h} + \bar{\boldsymbol{\tau}} \end{bmatrix} \equiv (\boldsymbol{h}, \boldsymbol{r}, \boldsymbol{l}, \boldsymbol{\tau}) \equiv$$
$$\boldsymbol{h} + \boldsymbol{v_q}, \;\; \boldsymbol{v_q} = (\boldsymbol{r}, \boldsymbol{l}, \boldsymbol{\tau}) \in \mathbb{R}^{3d}, \; \boldsymbol{q} \in \mathbb{D}^d.$$

Object Dihedron Completion of Quintuples: Here we show how the answers to the above incomplete quintuples are computed by a KGE using Dihedron algebra. Let t e.g. *Volodymyr Zelensky* be the answer of an incomplete quintuple $(h, r, ?, l.\tau)$ for a query e.g. *Whom Prince William (h) did meet (r) in 2020 (τ) in UK (l)?*. In order to complete the quintuple in the vector space, the similarity of the query ($\boldsymbol{q}$) and its given answer ($\boldsymbol{t}$) is maximized. Here we use the Dihedron product ($\otimes$) between the query q and its answer t representations. Note that while the query $\boldsymbol{q}$ lies on the Dihedron space, its answer $\boldsymbol{t} = \boldsymbol{e}$ lies on the real space. To match these two representations on the same space, we add three real auxiliary ($\boldsymbol{t_x}, \boldsymbol{t_y}, \boldsymbol{t_z}$) vectors to tail embedding i.e. $\boldsymbol{t} = \boldsymbol{e1} + \boldsymbol{v_t}$, $\boldsymbol{v_t} = \boldsymbol{t_x i} + \boldsymbol{t_y j} + \boldsymbol{t_z k}$. Using two representations for an entity e according to its role (head $\boldsymbol{h} = \boldsymbol{e}$ or tail $\boldsymbol{t} = \boldsymbol{e1} + \boldsymbol{v_t}$) enables our model to differentiate between each roles. Such representations facilitate efficient capturing of the natural role of entities per each triple for the underlying KG. After matching the spaces, now, we can apply the Dihedron product to measure the plausibility of tail t to be the answer of the query q. This leads us to define the score function as

$$S(q, t) = -\|\boldsymbol{s1} + \boldsymbol{v_s}\|, \tag{2}$$

where

$$\boldsymbol{s1} + \boldsymbol{v}_s = \boldsymbol{q} \otimes \boldsymbol{t} = (\boldsymbol{h1} + \boldsymbol{v}_q) \otimes (\boldsymbol{e1} + \boldsymbol{v}_t), \tag{3}$$

$$\boldsymbol{s} = \boldsymbol{h}.\boldsymbol{e} - \boldsymbol{v}_q.\boldsymbol{v}_t,\ \boldsymbol{v}_s = \boldsymbol{h}\boldsymbol{v}_t + \boldsymbol{e}\boldsymbol{v}_q + \boldsymbol{v}_q \times \boldsymbol{v}_t,\ \boldsymbol{v}_q.\boldsymbol{v}_t = \boldsymbol{r}\boldsymbol{t}_x - \boldsymbol{l}\boldsymbol{t}_y - \boldsymbol{\tau}\boldsymbol{t}_z,$$
$$\boldsymbol{v}_q \times \boldsymbol{v}_t = \begin{bmatrix} -\boldsymbol{l}\boldsymbol{t}_z + \boldsymbol{\tau}\boldsymbol{t}_y \\ \boldsymbol{\tau}\boldsymbol{t}_x - \boldsymbol{r}\boldsymbol{t}_z \\ \boldsymbol{r}\boldsymbol{t}_y - \boldsymbol{l}\boldsymbol{t}_x \end{bmatrix}. \tag{4}$$

The advantage of Eq. 2 is its efficiency in memory and time due to representing query and answer separately in Dihedron space while to compute the query, no mathematical operation (e.g., addition/subtraction/multiplication) are used. All operations are done in the query answering phase which reduces the complexity significantly.

Subject Dihedron Completion of Quintuples: The previous formulation for incomplete quintuple representation and their corresponding answers was used when tail (object) was queried i.e. $(h, r, ?, l, t)$. Here we present our formulation for an incomplete quintuple and its answer where the head (subject) is queried i.e. $(?, r, t, l, \tau)$. Let us assume that, the example question is *Who met with Volodymyr Zelensky in UK in 2020?* and the answer is *Barack Obama*. We first use the common approach from [13] of augmentation on the KG by adding reverse relations (here reverse quintuples) (t, r^{-1}, h, l, τ) for each quintuple (h, r, t, l, τ) present in the KG (train set). During testing, we use the reverse quintuple as following $\boldsymbol{q} = \boldsymbol{t}\,\boldsymbol{1} + \boldsymbol{r}^{-1}\boldsymbol{i} + \boldsymbol{l}\boldsymbol{j} + \boldsymbol{\tau}\boldsymbol{k}$, and the answer for this Dihedron representation is the following $\boldsymbol{h}$ where entity is attached with three auxiliary vectors i.e., $\boldsymbol{h} = \boldsymbol{e1} + \boldsymbol{h}_x\boldsymbol{i} + \boldsymbol{h}_y\boldsymbol{j} + \boldsymbol{h}_z\boldsymbol{k}$. In this way, we preserve the basis of entity ($\boldsymbol{1}$), relation (verb) ($\boldsymbol{i}$), adverb of location ($\boldsymbol{j}$) and adverb of time ($\boldsymbol{k}$) regardless of head (subject) or tail (object) Incomplete Quintuples.

Mathematical Interpretation We highlight the advantages of our formulation from various mathematical view points namely *Spatio-temporal coordinate representation*, *Capturing Spatio-temporal and Relational Correlations* and *Geometric Interpretation*. *Spatio-temporal Coordinate*. We represent a query with $\boldsymbol{q} = \boldsymbol{h1} + \boldsymbol{ri} + \boldsymbol{lj} + \boldsymbol{\tau k} = \boldsymbol{h1} + \boldsymbol{v}_q,\ \ \boldsymbol{v}_q = (\boldsymbol{r}, \boldsymbol{l}, \boldsymbol{\tau}) \in \mathbb{R}^{3d},\ \boldsymbol{q} \in \mathbb{D}^d$. In this representation, we assign distinct coordinate bases to each of the distinct quintuple elements i.e. h, r, l, τ (i, j, k in Fig. 2 part (c)). This is consistent with the nature of KGs where entity, relation, location and time are four distinct components. It is noteworthy that the orthogonality of bases in Dihedron space is related to the position of elements of a quintuple pattern which affects the order of the elements in 4D Dihedron space corresponding to **1, i, j, k**, and it is not related to the embedding vectors $\boldsymbol{h}, \boldsymbol{r}, \boldsymbol{\tau}, \boldsymbol{l}$. In addition, this representation enables the head entity ($\boldsymbol{h}$) to move towards the tail entity ($\boldsymbol{t}$), by using the vector $\boldsymbol{v}_q$. In order to determine the direction of movement, the $\boldsymbol{v}_q$ vector is not only dependent on the relation, but also on the location and time adverbs.

Capturing Spatio-Temporal and Relational Correlations While entity, relation, location and time are considered as orthogonal bases in the Dihedron space, the correlation between i) head entity-(relation, location, time) ($\boldsymbol{h}\boldsymbol{v}_t$), ii) tail entity-(relation, location, time) ($\boldsymbol{e}\boldsymbol{v}_q$), iii) head-tail entity ($\boldsymbol{he}$), iv) relation-location $\boldsymbol{r}\boldsymbol{t}_y$, v) relation-time ($\boldsymbol{r}\boldsymbol{t}_z$), vi) location-time ($\boldsymbol{l}\boldsymbol{t}_z$), vii) location-relation ($\boldsymbol{l}\boldsymbol{t}_x$), viii) time-relation ($\boldsymbol{\tau}\boldsymbol{t}_x$), etc. are involved in final score calculation. Such comprehensive correlation capturing is enabled by properly formulating the query and the answer and efficiently using the Dihedron product in score calculation (see Eq. 2, 3, 4).

Geometric Interpretation (Two/Three Dimensional Subalgebras of $\mathbb{D}$) Without loss of generality, let $d = 1$ be the length of a Dihedron $\boldsymbol{h1} + \boldsymbol{ri} + \boldsymbol{lj} + \boldsymbol{\tau k}$ is $\boldsymbol{h^2} + \boldsymbol{r^2} - \boldsymbol{l^2} - \boldsymbol{\tau^2} = c$ where c is a constant value. This representation is rich in terms of geometry and covers the following geometric objects (see part (c) of Fig. 2):

i **Circle**: Let us fix $(-\boldsymbol{l^2} - \boldsymbol{\tau^2})$ in $\boldsymbol{h^2} + \boldsymbol{r^2} - \boldsymbol{l^2} - \boldsymbol{\tau^2} = c$ to a constant v alue as $-c_1^2$. We then have a circle of the form $\boldsymbol{h^2} + \boldsymbol{r^2} = c_2$, $c_2 = c + c_1^2$.
ii **Two Sheet Hyperboloid**: Let us then fix $\boldsymbol{h}^2$ to a constant value c_2^2. Therefore, $\boldsymbol{r^2} - \boldsymbol{l^2} - \boldsymbol{\tau^2} = c - c_2^2$. If $c - c_2^2 < 0$, then the object is one sheet Hyperboloid.
iii **One Sheet Hyperboloid**: In the previous representation, if $c - c_2^2 > 0$, the object is two sheet Hyperboloid.
iv **Conical Surface**: Following the above definition for $\boldsymbol{r^2} - \boldsymbol{l^2} - \boldsymbol{\tau^2} = c$, if $c - c_2^2 = 0$, then the object is then conical surfaces.

Regarding the mentioned advantages, our proposed model is capable of efficiently embedding the spatio-temporal KGs into a well-suited geometric space. For the other incomplete quintuples $(h, r, t, ?, \tau)$, $(h, r, t, l, ?)$ that were not discussed here, the same approach is followed. We add the potential location or time of the query part in the Eq. 4 and then compute the score using Eq. 2. If the resulted score is high, the selected location or time leads to a plausible quintuple.

5 Experiments

Here, we provide evaluation of our model against the adapted models for ST KGs. **Baseline Models.** We compare our model with several baselines and state-of-the-art KGE models. Overall, there are three categories of KGE models selected for comparison. *Static KGE Models.* Many KGE models learn over triples. The TransE [2] model has been widely used as baseline for KGEs. Although this model is simple, it obtained high performance in link prediction task on several benchmark KGs. RotatE [24] is another model that uses self-adversarial negative sampling that led to obtain state-of-the art performance on the distance based class of embeddings (i.e. the models that use distance function in their score functions). In order to have a fair comparison, we train this model without using self-adversarial negative sampling (in same setting as ours). The ComplEx model trained with N3 regularization [13] got state-of-the-art performance in most of the used static KGs. To have a fair comparison, we trained all the models (including our models and their competitors) using N3 regularization. QuatE [33] took the advantage of the Quaternion space and obtained state-of the-art performance with N3 regularization. *Temporal KGE Models.* There is less research on the topic of temporal KGEs in comparison to the work done on static KGEs. However, there are several strong baseline methods for temporal KGEs. Similar to ComplEx-N3, temporal version of this model obtained state-of-the art performance on link prediction over temporal KGs. Therefore, we chose TComplEx [12] that was trained with N3 regularization, as a competitor. For other models, we extended the three state of the art KGE models (TransE, RotatE, and QuatE) in terms of capability of encoding temporal aspects. Therefore, T-TransE, T-RotatE and T-QuatE are our extension over existing static KGE models for temporal embedding. The formulation of the models are specified in the Table 2.

Spatio-Temporal KGE Models We are not aware of any knowledge graph embedding model over spatio-temporal KGs (ST KGs). Therefore, we extended the formulation of the above-mentioned static and temporal KGE models to learn over ST KGs. ST-TransE, ST-RotatE, ST-ComplEx, ST-QuatE are the resulting models. Their characteristics and formulations are included in Table 2.

Table 2. Specification of baseline and state of the art KGE models."Ours" denote our proposed models which are mostly extensions of already existing models for temporal and ST-KG embeddings. DyHE, T-DyHE and ST-DyHE are models we propose for the ablation study on the effect of translation and rotation in Dihedron space for static, temporal and spatio-temporal KGEs. Note that r_1, r_2 are Dihedron rotation and translation vectors. The symbols $\circ$, $\otimes_H$ and $\otimes_D$ denote the complex, Hamilton and Dihedron products, respectively. $Re(.)$ refers to the real part of complex numbers. (T-L) refer to time and location. $\boldsymbol{r_{1,2}} = \boldsymbol{r_1} + \boldsymbol{r_2}$

Model	(T-L)	Score function	Pattern	Embeddings
TransE [2]	(✗-✗)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}+\boldsymbol{r}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t}\in\mathbb{R}^d$
T-TransE (ours)	(✓-✗)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}+\boldsymbol{r}+\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{\tau}\in\mathbb{R}^d$
ST-TransE (ours)	(✓-✓)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}+\boldsymbol{r}+\boldsymbol{l}+\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{\tau},\boldsymbol{l}\in\mathbb{R}^d$
RotatE [24]	(✗-✗)	$\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}\circ\boldsymbol{r}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t}\in\mathbb{C}^d$
T-RotatE (ours)	(✓-✗)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}\circ\boldsymbol{r}\circ\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{\tau}\in\mathbb{C}^d$
ST-RotatE (ours)	(✓-✓)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}\circ\boldsymbol{r}\circ\boldsymbol{l}\circ\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{\tau},\boldsymbol{l}\in\mathbb{C}^d$
ComplEx [26]	(✗-✗)	$Re(\boldsymbol{q}\times\bar{\boldsymbol{t}})$	$\boldsymbol{q}=\boldsymbol{h}\times\boldsymbol{r}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t}\in\mathbb{C}^d$
T-ComplEx [12]	(✓-✗)	$Re(\boldsymbol{q}\times\bar{\boldsymbol{t}})$	$\boldsymbol{q}=\boldsymbol{h}\times\boldsymbol{r}\times\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{\tau}\in\mathbb{C}^d$
ST-ComplEx	(✓-✓)	$Re(\boldsymbol{q}\times\bar{\boldsymbol{t}})$	$\boldsymbol{q}=\boldsymbol{h}\times\boldsymbol{r}\times\boldsymbol{l}\times\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{l},\boldsymbol{\tau}\in\mathbb{C}^d$
QuatE [33]	(✗-✗)	$Re(\boldsymbol{q}\otimes_H\bar{\boldsymbol{t}})$	$\boldsymbol{q}=\boldsymbol{h}\otimes_H\boldsymbol{r}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t}\in\mathbb{Q}^d$
T-QuatE (ours)	(✓-✗)	$Re(\boldsymbol{q}\otimes_H\bar{\boldsymbol{t}})$	$\boldsymbol{q}=\boldsymbol{h}\otimes_H\boldsymbol{r}\otimes_H\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{\tau}\in\mathbb{Q}^d$
ST-QuatE (ours)	(✓-✓)	$Re(\boldsymbol{q}\otimes_H\bar{\boldsymbol{t}})$	$\boldsymbol{q}=\boldsymbol{h}\otimes_H\boldsymbol{r}\otimes_H\boldsymbol{l}\otimes_H\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r},\boldsymbol{t},\boldsymbol{l},\boldsymbol{\tau}\in\mathbb{Q}^d$
DyHE (ours)	(✗-✗)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}\otimes_D\boldsymbol{r_{1,2}}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r_{1,2}},\boldsymbol{t}\in\mathbb{D}^d$
T-DyHE (ours)	(✓-✗)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}\otimes_D\boldsymbol{r_{1,2}}+\boldsymbol{\tau}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r_{1,2}},\boldsymbol{t},\boldsymbol{\tau}\in\mathbb{D}^d$
ST-DyHE (ours)	(✓-✓)	$-\|\boldsymbol{q}-\boldsymbol{t}\|$	$\boldsymbol{q}=\boldsymbol{h}\otimes_D\boldsymbol{r_{1,2}}+\boldsymbol{\tau}+\boldsymbol{l}$	$\boldsymbol{q},\boldsymbol{h},\boldsymbol{r_{1,2}},\boldsymbol{t},\boldsymbol{l},\boldsymbol{\tau}\in\mathbb{D}^d$

Spatio-Temporal Knowledge Graphs We constructed three spatio-temporal KGs. The distribution shown in Fig. 3 depicts that temporal aspects mostly appear after year 1900. Former to this time, there is a steady number of temporal information.

We formed quintuples from YAGO [25] dataset namely YAGO3K by extracting triples with time and location information. Our initial analysis suggested that spatio-temporal information is connected to specific relations (i.e., wroteMusicFor, created). We have worked with YAGO3-10 [20] used in ConvE [5]. The training set of YAGO3-10 has 107,9040 triples. After extraction of time and location, we ended up with a total number of 9734 quintuples with 3619 entities, 8 relations, 422 locations and 195 time. We also constructed quintuples from DBpedia and Wikidata using dedicated SPARQL queries on the respective public SPARQL endpoints. The datasets are named DBpedia34K & WikiData53K. We utilize the DBpedia release of January 2021 which contains more than 900 million triples[1] and the Wikidata release of April 2021 which contains more than 1.26 billion statements.[2]. In the case of DBpedia, it can be assumed that at least 13% of the resources have temporal and/or spatial references. For Wikidata a comparable figure could not be determined because the number of resources in the knowledge base is much higher. Despite the timeout errors, it is safe to say that spatio-temporal information is abundant in both of the resources since we were able to extract a high number of representative triples From DBpedia, we queried more than 82,000 quintuples comprising information on multimedia items (books, music and movies), space missions, battles or buildings. Similar to the DBpedia dataset, the Wikidata dataset also contains a high number of quintuples (approx. 103,000). Overall, we gained 53849 number of entities, and 8 relations from WikiData where 296 different locations is present in the data and 627 different time information. For the obtained dataset from DBpedia, we have 34604 entities and 7 relations with 5687 different locations and 896 time information.

Evaluation Metrics We use common metrics namely Mean Reciprocal Rank (MRR) and Hits@k(k=1,3,10) for the evaluations. Here we explain each of the metrics. Given a set of test quintuple of the form (h, r, t, l, τ), we first remove the head entity and generate a query in the form of $(?, r, t, l, \tau)$. We then replace $?$ with all entities e' in the KG to generate n_e number of quintuples (e', r, t, l, τ), where n_e is the number of entities in the KG. We compute the score of each triple, and sort them to return the rank of the original triple (h, r, t, l, τ). We denote the left rank by r_l. With a similar procedure, the right rank r_r is computed by completion of the right query $(h, r, ?, l, \tau)$. The average rank for the ith quintuple is computed by $r_{ai} = \frac{r_l + r_r}{2}$. **MRR** is computed by taking the average of the reciprocals from the ranks in the testing triples i.e. $\sum_{i=1}^{n_t} \frac{1}{r_{ai}}$ where n_t is the number of testing triples. **Hits@k** is the percentage of testing triples that are ranked lower than k. For static and temporal KGE models, we used the following queries $(?, r, t), (h, r, ?)$ and $(?, r, t, \tau), (h, r, ?, \tau)$ and present their ranking in the Table 3. We additionally report the evaluation results on time and location completion in the Table 4.

Evaluation Results We implemented[3] all models using the Pytorch library. We used full-cross entropy loss [26] with N3 regularization [13] for training each of the models.

Hyperparameter settings: A wide array of testing has been done based on searching hyperparameters. The hyperparameters for which we achieved the best results are provided in Table 5. Throughout the experiments the *Adam* [11] optimizer has performed well and is thus used majorly. The hyperparameters mentioned in Table 5 are provided as d-dimension, B-Batch size, LR-Learning rate and R-regularization parameters.

In Table 6, we provide a selected list of incomplete quintuples that were used for an ablation study between the models. As can be seen, the results on predicted answers for each incomplete quintuple are fully correct by our proposed model where ComplEx and ST-ComplEx were not capable of correct predictions. When the query is about incompleteness of head or tail, we compared it to the ComplEx model as otherwise it

1 https://www.dbpedia.org/resources/latest-core/.

2 https://grafana.wikimedia.org/d/000000175/wikidata-datamodel-statements?orgId=1&refresh=30m.

3 https://github.com/mojtabanayyeri/Spatio-temporal-KGEs.

Table 3. Head/Tail completion results. Models without prefix consume triples. Models with "T" as prefix consume quadruples (triple plus time). Models with "ST" as prefix consume quintuple (triple plus location and time). The best results are written bold.

	YAGO3K			DBpedia34K			WikiData53K		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
TransE	0.561	0.496	0.689	0.454	0.415	0.519	0.245	0.152	0.451
T-TransE	0.709	0.676	0.779	0.501	0.475	0.544	0.396	0.337	0.538
ST-TransE	0.705	0.670	0.775	0.500	0.451	0.577	0.565	0.546	0.599
RotatE	0.564	0.503	0.688	0.461	0.425	0.521	0.246	0.153	0.458
T-RotatE	0.700	0.679	0.736	0.505	0.487	0.534	0.356	0.283	0.525
ST-RotatE	0.682	0.668	0.702	0.428	0.412	0.455	0.523	0.486	0.588
ComplEx	0.562	0.501	0.686	0.462	0.427	0.523	0.250	0.154	0.464
T-ComplEx	0.702	0.674	0.753	0.500	0.482	0.529	0.376	0.306	0.539
ST-ComplEx	0.689	0.668	0.727	0.450	0.424	0.497	0.532	0.495	0.594
QuatE	0.564	0.503	0.685	0.460	0.425	0.519	0.249	0.156	0.459
T-QuatE	0.694	0.675	0.722	0.500	0.481	0.531	0.358	0.285	0.525
ST-QuatE	0.690	0.675	0.714	0.502	0.485	0.528	0.515	0.478	0.576
DyHE	0.563	0.503	0.684	0.460	0.426	0.518	0.243	0.152	0.448
T-DyHE	**0.715**	**0.684**	0.775	**0.516**	**0.487**	0.564	0.377	0.318	0.517
ST-DyHE	0.704	0.665	**0.779**	0.485	0.427	**0.583**	**0.568**	**0.550**	**0.599**
ST-NewDE	0.708	0.682	0.758	**0.536**	**0.500**	**0.598**	**0.572**	**0.556**	**0.603**

Table 4. Location/Time completion results. The best results are written bold.

Location Completion $(h, r, t, ?, \tau)$									
	YAGO3K			DBpedia34K			WikiData53K		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
ST-TransE	0.349	0.083	0.903	0.245	0.045	0.633	0.321	0.008	0.838
ST-NewDE	**0.352**	**0.114**	**0.920**	**0.354**	**0.094**	**0.797**	**0.582**	**0.231**	**0.970**
Time Completion $(h, r, t, l, ?)$									
	YAGO3K			DBpedia34K			WikiData53K		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
ST-TransE	0.249	0.024	0.620	0.271	0.082	0.571	0.075	0.006	0.198
ST-NewDE	**0.580**	**0.241**	**0.958**	**0.432**	**0.151**	**0.792**	**0.162**	**0.020**	**0.514**

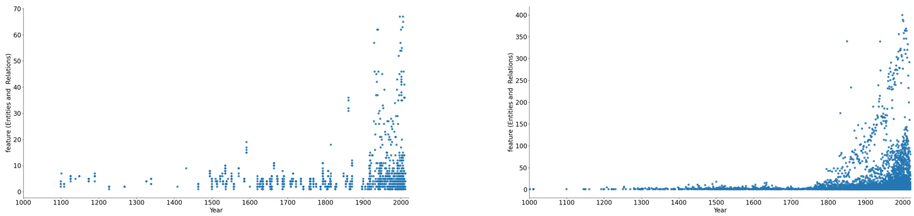

Fig. 3. Distribution of temporal information in YAGO3K and WikiData53K.

Table 5. Hyperparameter settings. Models without prefix consume triples. In this Table d = emb dimension, B = Batch size, LR = Learning rate, R = regularization parameter.

	YAGO3K				DBpedia34K				WikiData53K			
	d	B	LR	R	d	B	LR	R	d	B	LR	R
TransE	100	100	0.001	10e-11	100	100	0.001	0.001	100	100	0.001	10e-11
T-TransE	100	100	0.001	0.01	100	100	0.001	0.01	100	100	0.001	0
ST-TransE	100	100	0.001	0	100	100	0.001	0.001	100	100	0.001	0.001
RotatE	100	100	0.001	0.1	100	100	0.001	0.1	100	100	0.001	0.001
T-RotatE	100	100	0.001	0.01	100	100	0.001	0.1	100	100	0.001	0.001
ST-RotatE	100	100	0.001	0.1	100	100	0.001	0.1	100	100	0.001	0
ComplEx	100	100	0.001	0.01	100	100	0.001	0.01	100	100	0.001	10e-11
T-ComplEx	100	100	0.001	0.01	100	100	0.001	0.1	100	100	0.001	0.001
ST-ComplEx	100	100	0.001	0.001	100	100	0.001	0	100	100	0.001	0.001
QuatE	100	100	0.001	0.1	100	100	0.001	0.01	100	100	0.001	0.01
T-QuatE	100	100	0.001	0.001	100	100	0.001	0.1	100	100	0.001	0.001
ST-QuatE	100	100	0.001	0.01	100	100	0.001	0	100	100	0.001	0.001
DyHE	100	100	0.001	0.01	100	100	0.001	0.1	100	100	0.001	0.1
T-DyHE	100	100	0.001	0.01	100	100	0.001	0.1	100	100	0.001	0.1
ST-DyHE	100	100	0.001	0	100	100	0.001	0.1	100	100	0.001	0.01
ST-NewDE	100	100	0.001	0.001	100	100	0.001	0.001	100	100	0.001	0

Table 6. Example of ablation study results on predicted answers for incomplete quintuples over three datasets of YAGO3K, DBpedia34K, and WikiData53K. The correct objects are written **bold**.

Dataset	Query on head or tail parts	ST-NewDE	ComplEx
WikiData53K	(Philip_Guston, creatorOf, ?, United_States, 1972)	**Late_Afternoon**	Jules_Olitski
DBpedia34K	(Santiago_Calatrava, architectOf, ?, Maroussi, 1982)	**Olympic_Stadium_(Athens)**	SoFi_Stadium
YAGO3K	(?, created, Ulysses_(movie), Italy, 1955)	**Ennio_de_Concini**	Lasse_Hallstroem
Dataset	Query on location or time parts	ST-NewDE	ST-ComplEx
WikiData53K	(Edward_Witten, awardReceived, Alan_T._Waterman_Award, ?, 1982)	**United States**	Spain
DBpedia34K	(Nikolai_Nekrasov, authorOf, Korobeiniki_(poem), ?, 1861)	**Russia**	Serbia
YAGO3K	(Richard_Harvey, wroteMusicFor, Animal_Farm_(movie), United_States, ?)	**1999**	2010

does not consume spatial or temporal parts. For the queries about the spatial or temporal parts, we compared our model ST-NewDE against ST-ComplEx. In all of the studied cases, the ST-NewDE model predicts the correct matches.

Result Analysis: Regarding Table 3, for YAGO3K in terms of MRR and Hits@1 our model T-DyHE performed better than others by achieving MRR and Hits@1 scores of 0.715 and 0.684 respectively. Our other model ST-NewDE achieved a very similar score (0.682). By observing the results on DBpedia34K, it can be stated that, our model T-DyHE performed better by achieving higher MRR (0.516) and Hits@1 (0.487). ST-NewDE performed even better and outperformed others in terms of these two metrics by achieving MRR score of 0.500. In terms of Hits@10 ST-NewDE outperformed the other state of the art models by scoring 0.598 Hits@10. ST-DyHE and ST-NewDE also

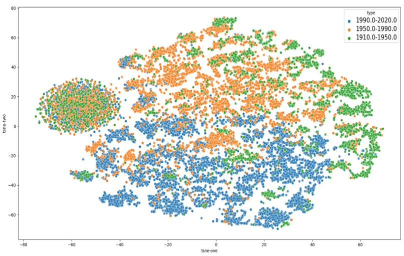

Fig. 4. Entity clustering w.r.t time.

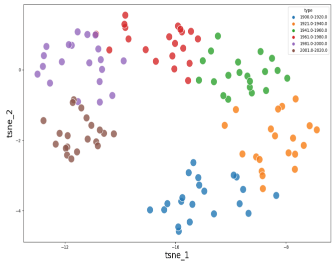

Fig. 5. Time clustering w.r.t year.

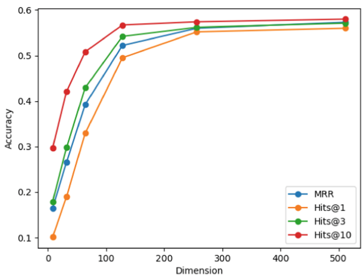

Fig. 6. Effect of d on ST-NewDE.

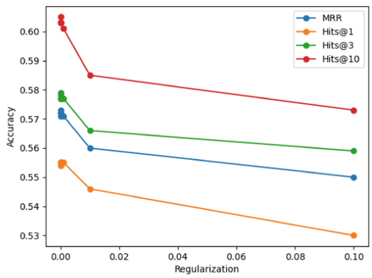

Fig. 7. Effect of R on ST-NewDE.

outperformed other models by achieving better accuracy in terms of MRR, Hits@1 and Hits@10. In case of our models, the best performing accuracy are marked in **bold**. In case of other models we have highlighted the better scores by underlining. Overall, in most of the cases, considering temporal part improves the performance of the triple-based models. In other words, for the queries of the form $(h, r, ?)$, $(?, r, t)$, adding temporal information leads to a more accurate prediction. Similarly, adding the spatial and temporal information simultaneously, improves the accuracy of head or tail prediction even higher than other results (in most of the cases). In addition, this shows the way that time and location information is formulated in a model, has a direct impact on the results. Embedding temporal and spatial information in Dihedron space obtains a higher performance than Quaternion space, although both spaces are 4D. For example, on WikiData53K, ST-DyHE outperformed ST-QuatE (e.g. 0.515 vs 0.568 Hits@1). Moreover, in Dihedron space, ST-NewDE outperforms ST-DyHE in most of the cases. This is due to a more elegant formulation in the same space (Dihedron). While combining various transformations such as translation/rotation in Dihedron space (DyHE) obtains a higher performance than other models and spaces, our main formulation (ST-NewDE) obtains the highest accuracy in most cases. This confirms that the inclusion of all four parts namely entity (head and tail), relation, location and time provides orthogonal bases of the Dihedron space and computes the scores based on Dihedron product. This further reinforces the hypothesis that not only Spatio-temporal transformations are important, but also the *Spatio-temporal geometric representation* has a high impact on efficiently exploiting the location/time information. This design of the KGE models

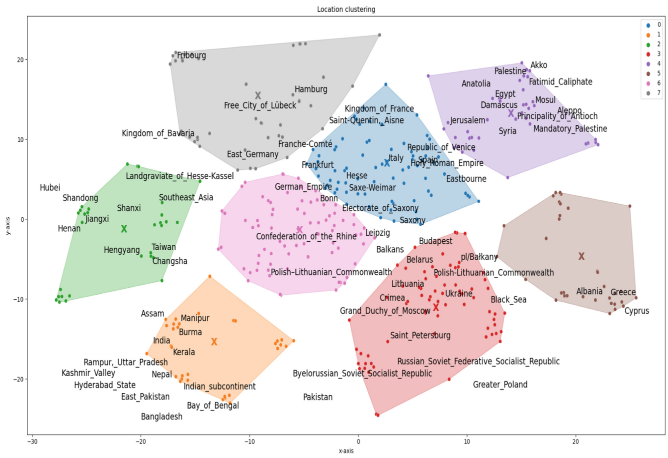

Fig. 8. Location clustering on YAGO3K.

leads to a more efficient formulation of the score function that works for such spatio-temporal data and outperforms other models. Such formulation has strong mathematical and geometric interpretation (discussed in method section). In addition to entity completion, we report the results of our model ST-NewDE and the best performing competitor ST-TransE. As shown in the Table 4, for completing the queries of the form $(h, r, t, l, ?)$ and $(h, r, t, ?, \tau)$, our model outperformed ST-TransE in all the used metrics. Such observation shows the efficiency of our model in head, tail, location, and time predictions. Note that like most of existing KGE models, our models can perform predictions on entities, relations, locations, and times that are seen during the training process. We consider inductive Spatio-temporal KGEs for the prediction of elements that are not given during training as future work. Regarding the number of parameters, it is noteworthy that ST-NEwDE has slightly fewer parameters than ST-TransE, and ST-DyHE has slightly more parameters than ST-TransE. Overall, the models have a very close number of parameters and their differences mainly come from the varying number of relation parameters which is significantly lower than the number of entity parameters.

Clustering Figure 4 depicts an ablation study on the clustering ability of our model. The quintuples are divided into three time categories of old (1910 - 1950) in orange, medium-old (1959 - 1990) in green, and recent (1990 - 2020) in blue. The results are generally "reasonable" and distinct while the overlapping parts belong to the cases with same head or tail entities appearing in different time periods. The dense cluster in the left side is caused by a high number of overlaps in the head or tail of the triples that belong to all the time categories.

The Fig. 5 illustrates the clustering ability of ST-NewDE over temporal part of the quintuples. For visualization purposes, each block of subsequent 20 years are grouped in one interval cluster. By using t-SNE, we visualize time embeddings in 2D space. First, our model puts all time points of the same 20-years block (1900–1920) in the same cluster and embeds the neighboring temporal intervals closely while the temporal order is preserved. Therefore, together with high performance in accuracy and the distinguished clustering results, we conclude that ST-NewDE efficiently captures the similarity of temporal part. As of ST-NewDE, it captured the location information or spatio-temporal part. As shown in Fig. 8, our model mainly embeds neighboring locations closely e.g., India, Nepal, Pakistan as well as cities in Germany (Freiburg, Hamburg).

Ablation Study on Dimension Here, we analyze the effect of dimension on the performance of ST-NewDE. In Table 3, we set $d = 100$ for all models. To have a fair comparison, we divided the used dimension relative to the space in which the models were designed. For Dihedron- and Quaternion-based models (4D algebra), we divided the dimension by four i.e. combination of all four parts will be 100 dimension. Similarly, for the models designed in complex vector space (e.g., ComplEx), we divided the dimension by two. For the ones in real space (e.g., TransE), we divided the dimension by one. Therefore, the performance improvement is assured not to be affected by other factors but only the core formulation. We experimented ST-NewDE with $d = \{8, 32, 64, 128, 256, 512\}$, and the other hyperparameters have been fixed. Figure 6 shows that by increasing the dimension on Wikidata53K, the performance improves quickly and then converges.

Ablation Study Regularization As can be seen in Fig. 7, we provided a study on the effect of regularization on the performance of ST-NewDE in Wikidata53K. According to Table 5, we noticed that the choice of regularization is important for improving the model performance. For some of the models such as DyHE, using a high value of regularization, improved the performance. In term of other models such as ST-NewDE, choice of a smaller regularization led to a better performance. We experimented the model with the following values for regularization $[0.1, 0.01, 0.001, 0.0001, 0.00001, 0.000001,\ 0.0000001]$. As shown, smaller regularization values obtain better performance than higher values. This shows that our model works well even without the N3 regularization.

6 Conclusion

In this paper, we address the problem of current KGE models performing on spatio-temporal KGs. We specifically proposed a family of embedding models which take advantage of the Dihedron algebra. The models were analysed in terms of the mathematical and *geometric interpretations*. We showed that our model facilitates *Spatio-temporal coordinate* representation and captures *Spatio-temporal and relational dependencies*. With these characteristics, the Dihedron-based KGE approach is capable of efficiently embedding spatio-temporal information into a rich geometric space. We additionally adapted the already exiting models to be able to encode spatio-temporal data. Our experiments on the subset of three public knowledge bases YAGO3K, DBpedia34K and WikiData53K showed that our approach usually achieves significantly higher performance than the extended state-of-the-art KGE models with time- or location-specific information. While our models predict entities, location, and time points which have been seen during training, we consider the prediction of unseen elements as future work.

Acknowledgement. We acknowledge the support of the following projects: SPEAKER (BMWi FKZ 01MK20011A), JOSEPH (Fraunhofer Zukunftsstiftung), the EU projects Cleopatra (GA 812997), PLATOON(GA 872592), TAILOR(EU GA 952215), CALLISTO(101004152), the BMBF projects MLwin(01IS18050) and the BMBF excellence clusters ML2R (BmBF FKZ 01 15 18038 A/B/C) and ScaDS.AI (IS18026A-F).

References

1. Balažević, I., Allen, C., Hospedales, T.M.: Tucker: Tensor factorization for knowledge graph completion. arXiv preprint (2019). arXiv:1901.09590
2. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data. In: Neural Information Processing Systems (NIPS), pp. 1–9 (2013)
3. Chami, I., Wolf, A., Juan, D.-C., Sala, F., Ravi, S., Ré, C.: Low-dimensional hyperbolic knowledge graph embeddings. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 6901–6914 (2020)
4. Dasgupta, S.S., Ray, S.N., Talukdar, P.: Hyte: hyperplane-based temporally aware knowledge graph embedding. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pp. 2001–2011 (2018)
5. Dettmers, T., Minervini, P., Stenetorp, P., Riedel, S.: Convolutional 2d knowledge graph embeddings. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32 (2018)
6. García-Durán, A., Dumančić, S., Niepert, M.: Learning sequence encoders for temporal knowledge graph completion. arXiv preprint (2018). arXiv:1809.03202
7. Hobbs, J., Blythe, J., Chalupsky, H., Russ, T.A.: A survey of geospatial resources, representation and reasoning. Public Distribution of the University of Southern California (2006)
8. 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 Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (vol. 1: Long Papers), pp. 687–696 (2015)
9. Ji, S., Pan, S., Cambria, E., Marttinen, P., Philip, S.Y.: A survey on knowledge graphs: representation, acquisition, and applications. IEEE Trans. Neural Netw. Learn. Syst. (2021)
10. Jin, W., Zhang, C., Szekely, P., Ren, X.: Recurrent event network for reasoning over temporal knowledge graphs. arXiv preprint (2019). arXiv:1904.05530
11. Kingma, D.P., Adam, J.Ba.: A method for stochastic optimization. arXiv preprint (2014). arXiv:1412.6980
12. Lacroix, T., Obozinski, G., Usunier, N.: Tensor decompositions for temporal knowledge base completion. arXiv preprint (2020). arXiv:2004.04926
13. Lacroix, T., Usunier, N., Obozinski, G.: Canonical tensor decomposition for knowledge base completion. In: International Conference on Machine Learning, pp. 2863–2872. PMLR (2018)
14. Leblay, J., Chekol, M.W.: Deriving validity time in knowledge graph. In: Companion Proceedings of the Web Conference 2018, pp. 1771–1776 (2018)
15. Lehmann, J., et al.: Dbpedia-a large-scale, multilingual knowledge base extracted from wikipedia, vol. 6, pp. 167–195 (2015, IOS Press)
16. Leidner, J.L.: A survey of textual data & geospatial technology. In: Werner, M. (ed.) Handbook of Big Geospatial Data, pp. 429–457. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-55462-0_16
17. Li, Z., Liu, H., Zhang, Z., Liu, T., Xiong, N.N.: Learning knowledge graph embedding with heterogeneous relation attention networks. IEEE Trans. Neural Netw. Learn. Syst. (2021)

18. Lin, Y., Liu, Z., Sun, M., Liu, Y., Zhu, X.: Learning entity and relation embeddings for knowledge graph completion. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 29 (2015)
19. Ma, Y., Tresp, V., Daxberger, E.A.: Embedding models for episodic knowledge graphs. J. Web Semant. **59**, 100490 (2019)
20. Mahdisoltani, F., Biega, J., Suchanek, F.: Yago3: a knowledge base from multilingual wikipedias. In: 7th Biennial Conference on Innovative Data Systems Research. CIDR Conference (2014)
21. Nayyeri, M., Vahdati, S., Aykul, C., Lehmann, J.: 5* knowledge graph embeddings with projective transformations. arXiv preprint (2020). arXiv:2006.04986
22. Qian, T., Liu, B., Nguyen, Q.V.H., Yin, H.: Spatiotemporal representation learning for translation-based poi recommendation. ACM Trans. Inf. Syst. (TOIS) **37**(2), 1–24 (2019)
23. Rotman, J.J.: Advanced Modern Algebra, vol. 114. American Mathematical Soc, New York (2010)
24. Sun, Z., Deng, Z.-H., Nie, J.-Y., Tang, J.: Rotate: knowledge graph embedding by relational rotation in complex space. arXiv preprint (2019). arXiv:1902.10197
25. Pellissier Tanon, T., Weikum, G., Suchanek, F.: YAGO 4: a reason-able knowledge base. In: Harth, A. (ed.) ESWC 2020. LNCS, vol. 12123, pp. 583–596. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49461-2_34
26. Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., Bouchard, G.: Complex embeddings for simple link prediction. In: International Conference on Machine Learning, pp. 2071–2080. PMLR (2016)
27. Vrandečić, D., Krötzsch, M.: Wikidata: a free collaborative knowledgebase. Commun. ACM **57**(10), 78–85 (2014)
28. Wang, Z., Zhang, J., Feng, J., Chen, Z.: Knowledge graph embedding by translating on hyperplanes. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 28 (2014)
29. Wildberger, N.J.: The geometry of the dihedrons (and quaternions). https://www.youtube.com/watch?v=qcmH0iKRF2w&t=1569s. Accessed 11 June 2021
30. Xu , C., Li, R.: Relation embedding with dihedral group in knowledge graph. arXiv preprint (2019). arXiv:1906.00687
31. Xu, C., Nayyeri, M., Alkhoury, F., Yazdi, H.S., Lehmann, J.: A time-aware knowledge graph embedding via temporal rotation. arXiv preprint (2020). arXiv:2010.01029
32. Yang, B., Yih, W.-T., He, X., Gao, J., Deng, L.: Embedding entities and relations for learning and inference in knowledge bases. arXiv preprint (2014). arXiv:1412.6575
33. Zhang, S., Tay, Y., Yao, L., Liu, Q.: Quaternion knowledge graph embeddings. arXiv preprint (2019). arXiv:1904.10281
34. Zhao, L., Deng, H., Qiu, L., Li, S., Hou, Z., Sun, H., Chen, Y.: Urban multi-source spatio-temporal data analysis aware knowledge graph embedding. Symmetry **12**(2), 199 (2020)