A Proof of the theorem for strong relation patterns.

Theorem 1. *The KG embedding (without bias) defined by Definition 1 satisfies strong relation patterns.*

Definition 1. *Let $(\mathcal{P}, \mathcal{G})$ be the embedding space, where $\mathcal{P} = \mathbb{P}_{m,n,\kappa}$, and $\mathcal{G}$ is a group of mappings:*

$$g_{(\boldsymbol{a},\boldsymbol{A})}: \ \mathbb{P}_{m,n,\kappa} \to \ \mathbb{P}_{m,n,\kappa} \tag{1}$$

$$\boldsymbol{e} \mapsto \boldsymbol{a} \oplus_{\mathcal{P}}^{\kappa} (\boldsymbol{A} \cdot \boldsymbol{e}). \tag{2}$$

It is sufficient to prove the following claims.

Claim. The KG embedding defined by Definition 1 satisfies symmetric/asymmetric, inversion, and weak composition.

Proof. 1. **Symmetric** If $\mathbf{a}_r = \mathbf{0}$ be the original point and $\mathbf{A}_r \cdot \mathbf{A}_r = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, then we have

$$\begin{aligned} & g_r(\mathbf{e}_h) = \mathbf{e}_t \\ \Rightarrow\ & \mathbf{A}_r \cdot \mathbf{e}_h = \mathbf{e}_t \\ \Rightarrow\ & \mathbf{A}_r \cdot \mathbf{A}_r \cdot \mathbf{e}_h = \mathbf{A}_r \cdot \mathbf{e}_t \\ \Rightarrow\ & \mathbf{A}_r \cdot \mathbf{e}_h = \mathbf{e}_t \\ \Rightarrow\ & g_r(\mathbf{e}_h) = \mathbf{e}_t \end{aligned}$$

2. **Asymmetric** If $\mathbf{a}_r = \mathbf{0}$ and $\mathbf{A}_r = 2\mathbf{I}$, then $g_r(\mathbf{e}_h) = 2\mathbf{e}_h$. If $g_r(\mathbf{e}_h) = \mathbf{e}_t$ and $g_r(\mathbf{e}_t) = \mathbf{e}_h$, we must have $\mathbf{e}_h = \mathbf{e}_t = \mathbf{0}$.
3. **Inversion** If $\mathbf{a}_r = \mathbf{0}$ and $\mathbf{A}_r$ has inverse matrix $\mathbf{A}_r^{-1}$, then $g_r(\mathbf{e}) = \mathbf{A}_r \cdot \mathbf{e}$ has inverse map $g_{r'}(\mathbf{e}) = \mathbf{A}_r^{-1} \cdot \mathbf{e}$.
4. **Weak composition** Let $\mathbf{a}_i = \mathbf{0}$ for $i = 1, 2, 3$, and assume that $\mathbf{A}_1 \cdot \mathbf{A}_2 = \mathbf{A}_3$, then we have $g_2(g_1(\mathbf{e}_h)) = \mathbf{A}_1 \cdot \mathbf{A}_2 \cdot \mathbf{e}_h = \mathbf{A}_3 \cdot \mathbf{e}_h = g_3(\mathbf{e}_h), \forall \mathbf{e} \in \mathcal{M}$.

Claim. The KG embedding defined by Definition 1 satisfies strong composition.

Assume:

$$r \otimes \mathbf{x} = \frac{1}{\sqrt{\kappa}} \cdot tanh\left(r \cdot tanh^{-1}(\sqrt{\kappa}||\mathbf{x}||)\right) \frac{\mathbf{x}}{||\mathbf{x}||} \tag{3}$$

To prove the claim, we need the following results. Consider the *Einstein addition* $\oplus_E$ [6][Section 8] that satisfies

$$\mathbf{x} \oplus_E \mathbf{y} = 2 \otimes^{\kappa} \left((\frac{1}{2} \otimes^{\kappa} \mathbf{x}) \oplus^{\kappa} (\frac{1}{2} \otimes^{\kappa} \mathbf{y}) \right), \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}_{\kappa}^{d}.$$

$$\mathbf{x} \oplus^{\kappa} \mathbf{y} = \frac{1}{2} \otimes^{\kappa} \left((2 \otimes^{\kappa} \mathbf{x}) \oplus_E (2 \otimes^{\kappa} \mathbf{y}) \right), \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}_{\kappa}^{d}.$$

We call a function $f : \mathbb{H}_{\kappa}^{d} \to \mathbb{H}_{\kappa}^{d}$ *preserves Einstein distance* iff $\forall \mathbf{x}, \mathbf{y} \in \mathbb{H}_{\kappa}^{d}$, we have

$$||(-f(\mathbf{x})) \oplus_E f(\mathbf{y})|| = ||(-\mathbf{x}) \oplus_E \mathbf{y}||.$$

Then by [5][Theorem 3.26], we have:

Theorem 2. *All function $f : \mathbb{H}^d_\kappa \to \mathbb{H}^d_\kappa$ that preserves Einstein distance iff f are of the form*

$$f(\boldsymbol{x}) = \boldsymbol{a} \oplus_E \boldsymbol{A} \cdot \boldsymbol{x},$$

for some $\boldsymbol{a} \in \mathbb{H}^d_\kappa$, $\boldsymbol{A} \in \mathbf{O}(n)$.

Moreover, we have

Lemma 1. *If $A \in \mathbf{O}(n)$, then we have $r \otimes^\kappa (\boldsymbol{A} \cdot \boldsymbol{y}) = \boldsymbol{A} \cdot (r \otimes^\kappa \boldsymbol{y}), \forall \boldsymbol{x} \in \mathbb{H}^d_\kappa$.*

Proof. If $\mathbf{A} \in \mathbf{O}(n)$, then $||\mathbf{y}|| = ||\mathbf{A} \cdot \mathbf{y}||$ for all $\mathbf{y} \in \mathbb{H}^d_\kappa$. Therefore, we have

$$\begin{aligned} r \otimes^\kappa (\mathbf{A} \cdot \mathbf{y}) &= \frac{1}{\sqrt{\kappa}} \cdot tanh\left(r \cdot tanh^{-1}(\sqrt{\kappa}||\mathbf{Ay}||\right) \frac{\mathbf{Ay}}{||\mathbf{Ay}||} \\ &= \frac{1}{\sqrt{\kappa}} \cdot tanh\left(r \cdot tanh^{-1}(\sqrt{\kappa}||\mathbf{y}||\right) \frac{\mathbf{Ay}}{||\mathbf{y}||} \\ &= \mathbf{A} \cdot (r \otimes^\kappa \mathbf{y}). \end{aligned}$$

Now, we start to prove Theorem 1.

Proof (proof of Theorem 1). Without loss of generality, we may assume that $\oplus^\kappa_{\mathcal{P}} = \oplus$.

Let $p : \mathbb{H}^d_\kappa \to \mathbb{H}^d_\kappa$ be the mapping defined by $p(u) = 2 \otimes^\kappa \mathbf{x}$. Then, the inverse map p^{-1} of p is defined by $p^{-1}(\mathbf{x}) = \frac{1}{2} \otimes^\kappa \mathbf{x}$. Next, we denote the composition of mappings $f_1, \ldots, f_n$ as $f_1 \circ \ldots \circ f_n$. That is, $f_1 \circ \ldots f_{n-1} \circ f_n(\mathbf{x}) = f_1(\ldots f_{n-1}(f_n(\mathbf{x})) \ldots)$.

Then, for each $g_{(\mathbf{a},\mathbf{A})} \in \mathcal{G}$, where

$$\begin{aligned} g_{(\mathbf{a},\mathbf{A})} : \ \mathbb{H}^d_\kappa &\to \ \mathbb{H}^d_\kappa \\ \mathbf{e} &\mapsto \ \mathbf{a} \oplus^\kappa (\mathbf{A} \cdot \mathbf{e}). \end{aligned}$$

We have

$$\begin{aligned} g_{(\mathbf{a},\mathbf{A})}(\mathbf{e}) &= \mathbf{a} \oplus^\kappa (\mathbf{A} \cdot \mathbf{e}) \\ &= \frac{1}{2} \otimes^\kappa \left((2 \otimes^\kappa \mathbf{a}) \oplus_E (2 \otimes^\kappa (\mathbf{A} \cdot \mathbf{e})) \right) \\ &= \frac{1}{2} \otimes^\kappa \left((2 \otimes^\kappa \mathbf{a}) \oplus_E (\mathbf{A} \otimes^\kappa (2 \cdot \mathbf{e})) \right) \quad \text{(by Lemma 1)} \\ &= p^{-1}\left((2 \otimes^\kappa \mathbf{a}) \oplus_E (\mathbf{A} \otimes^\kappa p(\mathbf{e})) \right) \\ &= (p^{-1} \circ f_{(\mathbf{a},\mathbf{A})} \circ p)(\mathbf{e}), \end{aligned}$$

where

$$\begin{aligned} f_{(\mathbf{a},\mathbf{A})} : \ \mathbb{H}^d_\kappa &\to \ \mathbb{H}^d_\kappa \\ \mathbf{e} &\mapsto \ (2 \otimes^\kappa \mathbf{a}) \oplus_E (\mathbf{A} \cdot \mathbf{e}). \end{aligned}$$

Then, for any $\mathbf{a}, \mathbf{a}' \in \mathbb{H}^d_\kappa, \mathbf{A}, \mathbf{A}' \in \mathbf{O}(n)$, we have

$$g_{(\mathbf{a},\mathbf{A})} \circ g_{(\mathbf{a}',\mathbf{A}')} = (p^{-1} \circ f_{(\mathbf{a},\mathbf{A})} \circ p) \circ (p^{-1} \circ f_{(\mathbf{a}',\mathbf{A}')} \circ p) = p^{-1} \circ f_{(\mathbf{a},\mathbf{A})} \circ f_{(\mathbf{a}',\mathbf{A}')} \circ p.$$

By Theorem 2, we know $f_{(\mathbf{a},\mathbf{A})}, f_{(\mathbf{a}',\mathbf{A}')}$ both preserve Einstein distance. Then, their composition $f_{(\mathbf{a},\mathbf{A})} \circ f_{(\mathbf{a}',\mathbf{A}')}$ still preserves Einstein distance. By Theorem 2, we can find $\mathbf{b} \in \mathbb{H}^d_\kappa, B \in \mathbf{O}(n)$ such that

$$f_{(\mathbf{a},\mathbf{A})} \circ f_{(\mathbf{a}',\mathbf{A}')}(\mathbf{x}) = \mathbf{b} \oplus_E (\mathbf{B} \cdot \mathbf{x}) = f_{(\frac{1}{2} \otimes^\kappa \mathbf{b}, \mathbf{B})}(\mathbf{x}).$$

Notice that, if $\mathbf{A}, \mathbf{A}' \in \mathbf{SO}(n)$, then $f_{(\mathbf{a},\mathbf{A})}$ and $f_{(\mathbf{a}',\mathbf{A}')}$ preserve the *orientation* of the manifold $\mathbb{H}^d_\kappa$, so as their composition $f_{(\mathbf{a},\mathbf{A})} \circ f_{(\mathbf{a}',\mathbf{A}')}$. Therefore, we have $\mathbf{B} \in \mathbf{SO}(n)$. Finally, we have

$$g_{(\mathbf{a},\mathbf{A})} \circ g_{(\mathbf{a}',\mathbf{A}')} = p^{-1} \circ f_{(\frac{1}{2} \otimes^\kappa \mathbf{b}, \mathbf{B})} \circ p = g_{(\frac{1}{2} \otimes^\kappa \mathbf{b}, \mathbf{B})}.$$

This proves the Theorem 1.

B Proof of claim about the comparison with RotatE and TransE

Claim. The *KG* embedding *TranE* [2], *RotatE* [4] are special cases of HolmE.

Proof. If we require $\mathbf{A} = \mathbf{I}$ and the curvature $\kappa = 0$, then we have $g_{(\mathbf{a},\mathbf{A})} = \mathbf{a} \oplus^\kappa_{\mathcal{P}} (\mathbf{A} \cdot \mathbf{e}) = \mathbf{a} + \mathbf{e}$. Then HolmE is identical to the *TranE* model.

If we require $\mathbf{a} = \mathbf{0}$, the curvature $\kappa = 0$, and

$$\mathbf{A} = diag[\mathbf{R}^n(\theta_{r,1}), \ldots, \mathbf{R}^n(\theta_{r,d/n})]$$

be the rotation matrix. Then HolmE is identical to *RotatE* model.

C Proof of claim about the comparison with Murp and AttH

Claim. The *KG* embedding MurP [1], AttH [3] (without bias) cannot satisfy strong composition.

Proof. For a given tuple (h, r, t), the score function in MurP (without bias) of (h, r, t) is defined by:

$$s(h, r, t) = -d^\kappa(\mathbf{R}\mathbf{e}_h, \mathbf{e}_t \oplus^\kappa \mathbf{a}_r),$$

where $\mathbf{R}$ is a diagonal matrix, $\mathbf{e}_h, \mathbf{e}_t, \mathbf{a}_r \in \mathbb{H}^{d,\kappa}$.

Assume $h_{\mathbf{b}}(\mathbf{e}) = \mathbf{e} \oplus^\kappa \mathbf{b}$. In our setting, MurP is equivalent to the model on the embedding space $(\mathcal{M}, \mathcal{G}, s)$ with $\mathcal{M} = \mathbb{H}^{d,\kappa}$, and $\mathcal{G}$ is a group of mappings:

$$g_{(\mathbf{b},\mathbf{R})} : \mathbb{H}^{d,\kappa} \to \mathbb{H}^{d,\kappa} \quad (4)$$
$$\mathbf{e} \mapsto h^{-1}_{\mathbf{b}} (\mathbf{R} \cdot \mathbf{e}). \quad (5)$$

Proof by Counterexample. We show MurP cannot satisfy strong composition by a contract example. Next, assume that $d = 2, \kappa = 1$. Let $\mathbf{b}_1 = (1, 0), \mathbf{R}_1 = \mathbf{I}$ and $\mathbf{b}_2 = (0, 0), \mathbf{R}_2 = 2 \cdot \mathbf{I}$. We show that there do not exists $\mathbf{b}_3, \mathbf{R}_3$ such that $g_{(\mathbf{b}_3,\mathbf{R}_3)} = g_{(\mathbf{b}_2,\mathbf{R}_2)} \circ g_{(\mathbf{b}_1,\mathbf{R}_1)}$.

If not, assume $\mathbf{b}_3, \mathbf{R}_3$ satisfies $g_{(\mathbf{b}_3,\mathbf{R}_3)} = g_{(\mathbf{b}_2,\mathbf{R}_2)} \circ g_{(\mathbf{b}_1,\mathbf{R}_1)}$. Then, for any $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in \mathbb{H}^{d,\kappa}$, we have :

$$\mathbf{e}_2 = g_{(\mathbf{b}_1,\mathbf{R}_1)}(\mathbf{e}_1),\ \mathbf{e}_3 = g_{(\mathbf{b}_2,\mathbf{R}_2)}(\mathbf{e}_2) \quad \Rightarrow \quad \mathbf{e}_3 = g_{(\mathbf{b}_3,\mathbf{R}_3)}(\mathbf{e}_1).$$

That is (recall that $\mathbf{b}_1 = (1, 0), \mathbf{R}_1 = \mathbf{I}$ and $\mathbf{b}_2 = (0, 0), \mathbf{R}_2 = 2 \cdot \mathbf{I}$):

$$\mathbf{e}_2 \oplus^\kappa \mathbf{b}_1 = \mathbf{e}_1, \quad \mathbf{e}_3 = 2\mathbf{e}_2 \quad \Rightarrow \quad \mathbf{e}_3 \oplus^\kappa \mathbf{b}_3 = \mathbf{R}_3\mathbf{e}_1.$$

Then we have

$$(0.5\mathbf{e}_3) \oplus^\kappa \mathbf{b}_1 = \mathbf{e}_1 \quad \Rightarrow \quad \mathbf{e}_3 \oplus^\kappa \mathbf{b}_3 = \mathbf{R}_3\mathbf{e}_1.$$

Therefore

$$\mathbf{e}_3 \oplus^\kappa \mathbf{b}_3 = \mathbf{R}_3\Big((0.5\mathbf{e}_3) \oplus^\kappa \mathbf{b}_1\Big), \forall \mathbf{e}_3 \in \mathbb{H}^{d,\kappa}.$$

Let $\mathbf{e}_3 = (0, 0)$, we have $\mathbf{b}_3 = \mathbf{R}_3\mathbf{b}_1$. Assume $e_3 = (x, y), \mathbf{R}_3 = diag(a, b)$, then $\mathbf{b}_3 = (a, 0)$. We have:

$$(x, y) \oplus^\kappa (a, 0) = diag(a, b) \cdot \Big((0.5x, 0.5y) \oplus^\kappa (1, 0)\Big).$$

Then, by comparing the first coordinate of each side, we have:

$$\frac{(1 - 2ax - a^2)x + a(1 + x^2 + y^2)}{1 - 2ax + a^2(x^2 + y^2)} = a\frac{1 + \frac{1}{4}(-x^2 + y^2)}{1 - x + \frac{1}{4}(x^2 + y^2)}, \quad \forall x, y \in \mathbb{R}. \quad (6)$$

Let $x = 0$, we have

$$\frac{a(1 + y^2)}{1 + a^2y^2} = a, \forall y \in \mathbb{R}.$$

Then $a = 0$, then by equation (6), we have $x = 0, \forall x \in \mathbb{R}$. Contradiction!

Now, we prove AttH [3] (without bias) cannot satisfy strong composition in the same manner. For simplicity, we assume that the embedding manifold is of dimension 2. Then the scoring function of AttH is defined by:

$$s(h, r, t) = -d^\kappa(Q(h, r) \oplus^\kappa \mathbf{a}_r, \mathbf{e}_t),$$

when $\kappa = 0$, $Q(h, r) = (\lambda Rot(\theta) + (1-\lambda)Ref(\phi)) \cdot \mathbf{e}_h$, where λ, θ, ϕ are parameters determined by r s.t. $\lambda \in (0, 1), \theta, \phi \in \mathbb{R}$ and

$$Rot(\theta) := \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} \quad (7)$$

$$Ref(\phi) := \begin{bmatrix} cos(\phi) & sin(\phi) \\ sin(\phi) & -cos(\phi) \end{bmatrix} \tag{8}$$

Next, assume $\kappa = 0$. In this case, the Riemannian manifold is identical to the linear space, Let $\mathbf{a}_r = \mathbf{a}_s = \mathbf{0}$, $\lambda \neq \lambda', \lambda, \lambda' \in (0, 1)$, and

$$Q(h, r) = (\lambda Rot(0) + (1 - \lambda)Ref(0)) \cdot \mathbf{e}_h,$$

$$Q(h, s) = (\lambda' Rot(0) + (1 - \lambda')Ref(\pi)) \cdot \mathbf{e}_h.$$

We show that there does not exists $\lambda'', Q(h.t)$ such that

$$Q(Q(h, r), s) = Q(h, t) \oplus^{\kappa} \mathbf{a}_r. \tag{9}$$

Proof by Counterexample. Otherwise, assume Equation (9) is true for

$$Q(h, s) = (\lambda'' Rot(\theta) + (1 - \lambda'')Ref(\phi)) \cdot \mathbf{e}_h.$$

Then by taking $\mathbf{e}_h = (0, 0)$, we have $\mathbf{a}_r = (0, 0)$. Therefore, we have

$$(\lambda' Rot(0)+(1-\lambda')Ref(\pi))\cdot(\lambda Rot(0)+(1-\lambda)Ref(0))\cdot\mathbf{e}_h = (\lambda'' Rot(\theta)+(1-\lambda'')Ref(\phi))\cdot\mathbf{e}_h.$$

That is:

$$\mathbf{R}_0 \cdot \mathbf{e}_h = \mathbf{R}(\theta, \phi) \cdot \mathbf{e}_h, \quad \forall \mathbf{e}_h \in \mathbb{R}^2, \tag{10}$$

where

$$\mathbf{R}_0 := \begin{bmatrix} 1 - 2\lambda' & 0 \\ 0 & 1 - 2\lambda \end{bmatrix}, \quad \lambda \neq \lambda', \lambda, \lambda' \in (0, 1), \text{ and} \tag{11}$$

$$\mathbf{R}(\theta, \phi) := \lambda'' \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} + (1 - \lambda'') \begin{bmatrix} cos(\phi) & sin(\phi) \\ sin(\phi) & -cos(\phi) \end{bmatrix} \tag{12}$$

Then $\mathbf{R}_0 = \mathbf{R}(\theta, \phi)$, and thus $\lambda'' \cdot (-sin(\theta)) = -(1 - \lambda'')sin(\phi) = \lambda'' \cdot (-sin(\theta))$. Therefore, they are all 0. There are two cases:

1. $\lambda'' \in \{0, 1\}$, then $\mathbf{R}(\theta, \phi) = Ref(\phi)$ or $Rot(\theta)$. For any θ, ϕ, none of them equals $\mathbf{R}_0$
2. $sin(\theta) = sin(\phi) = 0$, then $\{cos(\theta), cos(\phi)\} \subseteq \{1, -1\}$. However, in any cases, we have $|\mathbf{R}(\theta, \phi)_{ii}| = 1$ for some $i \in \{1, 2\}$. Then, $\mathbf{R}(\theta, \phi) \neq \mathbf{R}_0$ as $0 < (\mathbf{R}_0)_{11}, (\mathbf{R}_0)_{22} < 1$.

Finally, in both cases, $\mathbf{R}(\theta, \phi) \neq \mathbf{R}_0$. This violated Equation (10).

References

1. Balažević, I., Allen, C., Hospedales, T.: Multi-relational poincaré graph embeddings. Advances in Neural Information Processing Systems **32** (2019)
2. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data **26** (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. Sun, Z., Deng, Z.H., Nie, J.Y., Tang, J.: RotatE: Knowledge graph embedding by relational rotation in complex space. In: International Conference on Learning Representations (2018)
5. Ungar, A.: Beyond Pseudo-Rotations in Pseudo-Euclidean Spaces. Academic Press (2018)
6. Ungar, A.A.: Möbius transformation and einstein velocity addition in the hyperbolic geometry of bolyai and lobachevsky. In: Nonlinear analysis, pp. 721–770. Springer (2012)