Reciprocity in Directed Hypergraphs: Measures, Findings, and Generators (Online Appendix)

NOTE: If a preview does not appear properly, please download this file.

1 Proof of Theorems

In this section, we formalize our proposed measure and provide proof of theorems in the main paper. We first introduce preliminaries of proofs, and prove why proposed reciprocity measure **HYPERREC** can satisfies all the **AXIOMS**. At last, we prove why several baseline measures fail in satisfying some **AXIOMS**.

1.1 Preliminaries of Proofs

In this subsection, we first give the general form of our proposed measure. Then, we introduce several important characteristics of *Jensen-Shannon Divergence* ($\mathrm{JSD}(\mathrm{P} \parallel \mathrm{Q})$) [1], which play key roles in proofs. After that, we examine how these concepts are applied to our measure.

The proposed reciprocity measure **HYPERREC** for hyperarc $r(e_i \mid C_i)$ and hypergraph $r(G)$ is defined as

$$r(e_i \mid C_i) := \max_{C_i \subseteq E} \left(\frac{1}{|C_i|}\right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right), \tag{1}$$

$$r(G) := \frac{1}{|E|} \sum_{i=1}^{|E|} r(e_i \mid C_i), \tag{2}$$

where $\mathcal{L}(p_h, p_h^*)$ denotes the Jensen-Shannon Divergence [1] between a transition probability distribution p_h and an optimal transition probability distribution p_h^*. Here, the transition probability of a target arc e_i with a candidate set that consists of a single candidate arc $\{e_i'\}$ is defined as

$$p_h(v) = \begin{cases} \frac{1}{|H_i'|} & \text{if } v \in H_i' \\ 0 & \text{otherwise} \end{cases}$$

This is a case where $v_h \in \{H_i \cap T_i'\}$. For each $v_h \in H_i \setminus T_i'$, $p_h(v) = 1$ if $v = v_{sunk}$ and 0 elsewhere. For the target arc e_j with an arbitrary number of candidate arcs $C_j = \{e_{j1}, \cdots e_{jk}\}$, the transition probability is defined formally as

$$p_h(v) = \begin{cases} p_{h,1}(v) & \text{if } v_h \in \bigcup_{e_s \in C_j} T_s, \\ p_{h,2}(v) & \text{otherwise} \end{cases}$$

where

$$p_{h,1}(v) = \frac{\sum_{e_s \in C_j} \left(\frac{\mathbf{1}[v_h \in T_s, v \in H_s]}{|H_s|}\right)}{\sum_{e_s \in C_j} (\mathbf{1}[v_h \in T_s])}, \quad p_{h,2}(v) = \begin{cases} 1 & \text{if } v = v_{sunk} \\ 0 & \text{otherwise,} \end{cases}, \quad \mathbf{1}[\text{TRUE}] = 1, \quad \text{and } \mathbf{1}[\text{FALSE}] = 0.$$

Throughout the proof, we use these formal expressions for both single-candidate-arc and multiple-candidate-arc cases.

We now discuss the theoretical aspect of $\mathrm{JSD}(P \parallel Q)$. The general form of $\mathrm{JSD}(P \parallel Q)$ for the discrete space is

$$\mathcal{L}(P, Q) = \sum_{i=1}^{|V|} \ell(p_i, q_i), \tag{3}$$

where

$$\ell(p_i, q_i) = \frac{p_i}{2} \log \frac{2p_i}{p_i + q_i} + \frac{q_i}{2} \log \frac{2q_i}{p_i + q_i}. \tag{4}$$

A. 1. For any two discrete probability distributions P and Q, $0 \leq \mathrm{JSD}(P \parallel Q) \leq \log 2$ holds, as shown in [1].

A. 2. For two discrete probability distributions P and Q where their non-zero probability domains do not overlap (i.e., $p_i q_i = 0, \forall i = \{1, \cdots, |V|\}$), $\mathrm{JSD}(P \parallel Q)$ is maximized as $\log 2$.

Proof: Let $\mathcal{X}_p$ be the domain where P has non-zero probabilities, and let $\mathcal{X}_q$ be the domain where Q has non-zero probabilities. As $\mathcal{X}_p$ and $\mathcal{X}_q$ do not overlap, Eq. (3) is rewritten as

$$\mathcal{L}(P,Q) = \sum_{i \in \mathcal{X}_p} \frac{p_i}{2} \log 2 + \sum_{i \in \mathcal{X}_q} \frac{q_i}{2} \log 2 = \frac{\log 2}{2} \left(\sum_{i \in \mathcal{X}_p} p_i + \sum_{i \in \mathcal{X}_q} q_i \right) = \frac{\log 2}{2} \times (1+1) = \log 2.$$

A. 3. Consider two discrete probability distributions P and Q. If there exists a value where both P and Q have non-zero probabilities, then $\mathrm{JSD}(P \parallel Q) < \log 2$ holds.

Proof: Let k be a value where $p_k q_k \neq 0$ holds. Then Eq. (3) is rewritten as

$$\mathcal{L}(P,Q) = \sum_{i \in \mathcal{X}_p \setminus k} \frac{p_i}{2} \log 2 + \sum_{i \in \mathcal{X}_q \setminus k} \frac{q_i}{2} \log 2 + \left(\frac{p_k}{2} \log \frac{2p_k}{p_k + q_k} + \frac{q_k}{2} \log \frac{2q_k}{p_k + q_k} \right). \tag{5}$$

In order to show that Eq. (3) is smaller than $\log 2$, it is required to show that

$$\begin{aligned} &\left(\frac{p_k}{2} \log 2 + \frac{q_k}{2} \log 2 \right) - \left(\frac{p_k}{2} \log \frac{2p_k}{p_k + q_k} + \frac{q_k}{2} \log \frac{2q_k}{p_k + q_k} \right) > 0 \\ \equiv\, &\frac{p_k}{2} \log \left(1 + \frac{q_k}{p_k} \right) + \frac{q_k}{2} \log \left(1 + \frac{p_k}{q_k} \right) > 0 \quad (\because p_k, q_k > 0). \end{aligned}$$

As the log function has a positive real value if its input is greater than 1, the last inequality holds. Thus, we can conclude that $\mathrm{JSD}(P \parallel Q) < \log 2$.

We use above three statements (**A. 1**, **A. 2**, and **A. 3**) to derive additional three statements as follows.

A. 4. If the target arc's head set and the tail sets of its candidate arcs do not overlap, the target arc's reciprocity becomes zero. Formally,

$$\text{If } \left| H_i \cap \bigcup_{e_k \in C_i} T_k \right| = 0 \quad \text{then} \quad r(e_i \mid C_i) = 0.$$

Proof: For this case, as mentioned in the main paper, the transition probability is heading toward sunken node v_{sunk}. On the other hand, optimal transition probability p^* is heading toward $v \in T_i$ where $v_{sunk} \notin T_i$. Thus, the non-zero probability domain of transition probability and optimal transition probability is not overlapped, where the probabilistic distance between them is maximized as $\mathcal{L}_{max}$ by **A. 2**. This happens for all H_i, in sum,

$$\begin{aligned} r(e_i \mid C_i) &= \left(\frac{1}{|C_i|} \right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}_{max}}{|H_i| \cdot \mathcal{L}_{max}} \right) \\ &= \left(\frac{1}{|C_i|} \right)^{\alpha} \left(1 - \frac{|H_i| \cdot \mathcal{L}_{max}}{|H_i| \cdot \mathcal{L}_{max}} \right) = \left(\frac{1}{|C_i|} \right)^{\alpha} \times 0 = 0 \end{aligned}$$

A. 5. If the target arc's tail set and candidate arc's head set do not overlap, the target arc's reciprocity becomes zero. Formally,

$$\text{If } |T_i \cap \bigcup_{e_k \in C_i} H_k| = 0 \quad \text{then} \quad r(e_i \mid C_i) = 0.$$

Proof: Similar to the **A. 4**, non-zero probability domain of transition probability and optimal probability is not overlapped since $|T_i \cap \bigcup_{e_k \in C_i} H_k| = 0$. Again, probabilistic distance is maximized as $\mathcal{L}_{max}$.

This happens for all H_i, in sum,

$$\begin{aligned} r(e_i \mid C_i) &= \left(\frac{1}{|C_i|} \right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}_{max}}{|H_i| \cdot \mathcal{L}_{max}} \right) \\ &= \left(\frac{1}{|C_i|} \right)^{\alpha} \left(1 - \frac{|H_i| \cdot \mathcal{L}_{max}}{|H_i| \cdot \mathcal{L}_{max}} \right) = \left(\frac{1}{|C_i|} \right)^{\alpha} \times 0 = 0 \end{aligned}$$

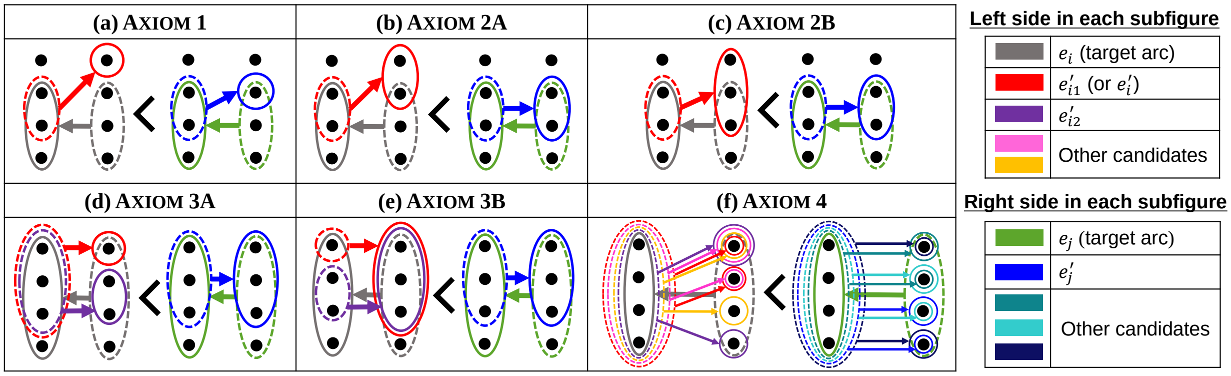


Figure 1: Examples for **AXIOMS** 1-4. In each subfigure, the reciprocity of the arc e_i on the left side should be smaller than that of the arc e_j on the right side.

A. 6. If the target arc's head set and candidate arc's tail set overlap and the target arc's tail set and candidate arc's head set overlap, the target arc's reciprocity is greater than zero. Formally,

$$\text{If} \sum_{e_k \in C_i} |H_i \cap T_k| \cdot |T_j \cap H_k| \geq 1 \quad \text{then} \quad r(e_i \mid C_i) > 0.$$

Proof: According to the statement, there exist at least one candidate arc e_k whose tail set overlaps to target arc's head set $|T_k \cap H_i| \geq 1$, and head set also overlaps to target arc's tail set $|H_k \cap T_i| \geq 1$. Thus, for $v_h \in H_i \cap T_k$, p_h and p^*_h have a common non-zero probability domain, which implies $\mathcal{L}(p_h, p^*_h) < \log 2$ by **A. 3**. In sum, we can derive the following inequality (**see the inequality sign between two middle terms**).

$$\begin{aligned} r(e_i \mid C_i) &= \left(\frac{1}{|C_i|}\right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p^*_h)}{|H_i| \cdot \mathcal{L}_{max}}\right) \\ &> \left(\frac{1}{|C_i|}\right)^{\alpha} \left(1 - \frac{|H_i| \cdot \mathcal{L}_{max}}{|H_i| \cdot \mathcal{L}_{max}}\right) = \left(\frac{1}{|C_i|}\right)^{\alpha} \times 0 = 0 \end{aligned}$$

1.2 Proof of Theorem 1.

For the **AXIOM** 1, we simply show that the former's reciprocity gets zero $r(e_i \mid C_i) = 0$, while the latter's reciprocity get's a positive value $r(e_j \mid C_j) > 0$. For the other **AXIOMS** , we first show the relationship between **AXIOMS** and the probabilistic distance (Step A). After, we show that a less reciprocal case has a higher probabilistic distance between observed transition probability and optimal transition probability for every head-set node of a target arc (Step B). An overview of **AXIOMS** is illustrated in Figure 1.

1.2.1 Proof of How HYPERREC Satisfies AXIOM 1

[AXIOM 1]: *Existence of Inverse Overlap*

Assume $C_i = \{e'_i\}$ and $C_j = \{e'_j\}$ such that

$$\min(|H_i \cap T'_i|, |T_i \cap H'_i|) = 0 \text{ and } \min(|H_j \cap T'_j|, |T_j \cap H'_j|) \geq 1$$

Then, the following inequality holds.

$$r(e_i \mid C_i) < r(e_j \mid C_j)$$

PROOF: We first show $r(e_i \mid C_i) = 0$. Term $\min(|H_i \cap T'_i|, |T_i \cap H'_i|) = 0$ implies at least one of $H_i \cap T'_i$ or $T_i \cap H'_i$ is an empty set. By the **A. 4** and **A. 5**, we can derive that reciprocity for this case is equal to zero $r(e_i \mid C_i) = 0$.

Now we show that $r(e_j \mid C_j) > 0$. As $\min(|H_j \cap T'_j|, |T_j \cap H'_j|) \geq 1$, we can guarantee that both $H_j \cap T'_j$ and $T_j \cap H'_j$ are non-empty sets. By using the **A. 6**, we prove the inequality $r(e_j \mid C_j) > 0$.

In conclusion as $r(e_i \mid C_i) = 0$ and $r(e_j \mid C_j) > 0$, we can derive $r(e_i \mid C_i) < r(e_j \mid C_j)$.

Note that from AXIOM 2-4, we assume two target arcs e_i and e_j have equal size $|H_i| = |H_j|$ and $|T_i| = |T_j|$.

1.2.2 Proof of How HYPERREC Satisfies AXIOM 2

AXIOM 2 is divided into two parts,

- **AXIOM 2A:** The candidate arc overlap more with the target arc in the e_j than e_i.
- **AXIOM 2B:** The candidate arc has less difference from the target arc in the e_j than e_i.

We prove two sub **AXIOMS** respectively as follows.

[AXIOM 2A]: *Degree of Inverse Overlap: More Overlap*

Let $C_i = \{e'_i\}$ and $C_j = \{e'_j\}$ where

$$|H'_i| = |H'_j| \text{ and } |T'_i| = |T'_j|$$
$$\text{(I)} \quad 0 < |H'_i \cap T_i| < |H'_j \cap T_j| \text{ and } 0 < |T'_i \cap H_i| \leq |T'_j \cap H_j| \quad \text{or}$$
$$\text{(II)} \quad 0 < |H'_i \cap T_i| \leq |H'_j \cap T_j| \text{ and } 0 < |T'_i \cap H_i| < |T'_j \cap H_j|$$

Then, the following inequality holds:

$$r(e_i \mid C_i) < r(e_j \mid C_j).$$

PROOF: As $|C_i| = |C_j| = 1$, cardinality penalty terms on both side can be discarded.

$$r(e_i \mid C_j) = \left(\frac{1}{|C_i|}\right)^{\alpha} \left(1 - \frac{\sum_{v_i \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) = 1 - \frac{\sum_{v_i \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} \quad \text{Same for } r(e_j \mid C_j) \text{ with subscript } j.$$

<u>Step A:</u> Here, we are trying to show $r(e_i \mid C_i) < r(e_j \mid C_j)$, which is equivalent to

$$\frac{\sum_{v_i \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} > \frac{\sum_{v_j \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}}$$

As $|H_i| = |H_j|$, denominator of both terms are identical. By removing them, inequality is rewritten as

$$\equiv \sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*) > \sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*) \tag{6}$$

Each head set can be divided into two parts ; $H_k \setminus T'_k$ and $H_k \cap T'_k$ $\forall k = i, j$.

For $H_k \setminus T'_k$, as described in **A. 4**, probabilistic distance gets maximized as $\mathcal{L}_{max} = \log 2$.

For $H_k \cap T'_k$, by using $T_k \cap H'_k \neq \emptyset$, we can derive, $\mathcal{L}(p_h, p_h^*) < \log 2$ satisfies for $\forall v_h \in H_k \cap T'_k$ where $\forall k = i, j$ by the **A. 6**. One more notable fact is that as there is a single candidate arc in both cases, $\mathcal{L}(p_h, p_h^*)$ is the same for every $v_h \in H_k \cap T'_k$. Here, let $p_{h,i}$ be the transition probability regarding a target arc e_i and its candidate set C_i. We summarize the main inequality (6) as follows.

$$\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*) > \sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*)$$
$$\equiv |H_i \setminus T'_i| * \log 2 + |H_i \cap T'_i| * \mathcal{L}(p_{h,i}, p_{h,i}^*) > |H_j \setminus T'_j| * \log 2 + |H_j \cap T'_j| * \mathcal{L}(p_{h,j}, p_{h,j}^*)$$

Further inequality can be developed differently according to the condition (whether it is (I) or (II)).

For the case of (I), the intersection of the target arc's head set and the candidate arc's tail set is larger in e_j than e_i. As a result, the inequality is re-written as follows.

$$|H_i \setminus T'_i| * \log 2 + |H_i \cap T'_i| * \mathcal{L}(p_{h,i}, p_{h,i}^*) > |H_j \setminus T'_j| * \log 2 + |H_j \cap T'_j| * \mathcal{L}(p_{h,i}, p_{h,i}^*)$$
$$\geq |H_j \setminus T'_j| * \log 2 + |H_j \cap T'_j| * \mathcal{L}(p_{h,j}, p_{h,j}^*)$$

For the case of (II), the intersection of the target arc's head set and the candidate arc's tail set can be identical, while the target arc's tail set and the candidate arc's head set intersection is greater in e_j than e_i. Then, the inequality is re-written as follows.

$$\begin{aligned}&|H_i \setminus T_i'| * \log 2 + |H_i \cap T_i'| * \mathcal{L}(p_{h,i}, p_{h,i}^*) \geq |H_j \setminus T_j'| * \log 2 + |H_j \cap T_j'| * \mathcal{L}(p_{h,i}, p_{h,i}^*)\\ &> |H_j \setminus T_j'| * \log 2 + |H_j \cap T_j'| * \mathcal{L}(p_{h,j}, p_{h,j}^*)\end{aligned}$$

Here, one can notice that if $\mathcal{L}(p_{h,i}, p_{h,i}^*) > \mathcal{L}(p_{h,j}, p_{h,j}^*)$ holds, then both cases (**I** and **II**) can be satisfied. Formally,

$$\begin{aligned}&|H_j \setminus T_j'| * \log 2 + |H_j \cap T_j'| * \mathcal{L}(p_{h,i}, p_{h,i}^*) > |H_j \setminus T_j'| * \log 2 + |H_j \cap T_j'| * \mathcal{L}(p_{h,j}, p_{h,j}^*)\\ &\equiv |H_j \cap T_j'| * \mathcal{L}(p_{h,i}, p_{h,i}^*) > |H_j \cap T_j'| * \mathcal{L}(p_{h,j}, p_{h,j}^*)\\ &\equiv \mathcal{L}(p_{h,i}, p_{h,i}^*) > \mathcal{L}(p_{h,j}, p_{h,j}^*)\end{aligned}$$

Step B: We only need to show that the probabilistic distance between transition probability and optimal transition probability is greater in e_i than e_j. Denote the intersection as $F_1 = |T_i \cap H_i'| \ < \ F_2 = |T_j \cap H_j'|$. We can decompose domain $v \in V$ into four parts as

$$v \in T_k \cap H_k' \quad v \in T_k \setminus H_k' \quad v \in H_k' \setminus T_k \quad v \in V \setminus \{H_k' \cup T_k\} \quad \forall k = i, j$$

For the last term, both transition probability and optimal of it have a value of zero $p_h(v) = p_h^*(v) = 0$, which result in no penalty. We only need to consider first three terms for comparison. Here, the probabilistic distance can be explicitly written as

$$\begin{aligned}\mathcal{L}(p_{h,i}, p_{h,i}^*) &= |F_1| * \ell\left(\frac{1}{T}, \frac{1}{A}\right) + (A - |F_1|) * \frac{1}{2A}\log 2 + (T - |F_1|) * \frac{1}{2T}\log 2\\ \mathcal{L}(p_{h,j}, p_{h,j}^*) &= |F_2| * \ell\left(\frac{1}{T}, \frac{1}{A}\right) + (A - |F_2|) * \frac{1}{2A}\log 2 + (T - |F_2|) * \frac{1}{2T}\log 2\end{aligned}$$

where ℓ denotes a single element comparison of $\mathrm{JSD}(P \parallel Q)$ explained in equation (4). Let $A = |S_i'| = |S_j'|$ and $T = |T_i| = |T_j|$. By showing $\mathcal{L}(p_{h,i}, p_{h,i}^*) - \mathcal{L}(p_{h,j}, p_{h,j}^*) > 0$, proof can be done. Formally,

$$\begin{aligned}&\mathcal{L}(p_{h,i}, p_{h,i}^*) - \mathcal{L}(p_{h,j}, p_{h,j}^*) > 0\\ &= (|F_1| - |F_2|) * \ell\left(\frac{1}{T}, \frac{1}{A}\right) + (|F_2| - |F_1|) * \left(\frac{1}{2A} + \frac{1}{2T}\right)\log 2 > 0\\ &\equiv (|F_2| - |F_1|) * \left(\frac{1}{2A}\log 2 + \frac{1}{2T}\log 2 - \ell\left(\frac{1}{T}, \frac{1}{A}\right)\right) > 0\\ &\equiv \left(\frac{1}{2A}\log 2 - \frac{1}{2A}\log\frac{2T}{A+T} + \frac{1}{2T}\log 2 - \frac{1}{2T}\log\frac{2A}{A+T}\right) > 0 \quad \because (|F_2| - |F_1|) > 0\\ &\equiv \frac{1}{2A}\log\frac{A+T}{T} + \frac{1}{2T}\log\frac{A+T}{A} > 0\end{aligned}$$

As $x > 1 \Rightarrow \log x > 0$, inequality holds. The proof is done.

[Axiom 2B]: *Degree of Inverse Overlap: Small difference*

Let $C_i = \{e_i'\}$ and $C_j = \{e_j'\}$ where

$$|H_i'| > |H_j'| \text{ and } |T_i'| = |T_j'| \quad 0 < |H_i' \cap T_i| = |H_j' \cap T_j| \text{ and } 0 < |T_i' \cap H_i| = |T_j' \cap H_j|$$

Then, the following inequality holds.

$$r(e_i \mid C_i) < r(e_j \mid C_j)$$

Proof: As $|C_i| = |C_j| = 1$, cardinality penalty term is being erased for both terms.

Step A: Overall inequality can be re-written as

$$r(e_i \mid C_i) < r(e_j \mid C_j) \equiv 1 - \frac{\sum_{v_i \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} < 1 - \frac{\sum_{v_j \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}}$$

Similar to the previous proof, $|H_i| = |H_j|$, and $\mathcal{L}(p_h, p_h^*)$ among different $v_h \in H_k \cap T_k'$ are all identical. In addition, $|H_i \cap T_i'| = |H_j \cap T_j'|$, $|T_i'| = |T_j'|$, number of target arc's head set nodes v_h that satisfy $\mathcal{L}(p_h, p_h^*) < \log 2$ are identical in both terms. Thus,

$$\begin{aligned}
&r(e_i \mid C_i) < r(e_j \mid C_j) \equiv \frac{\sum_{v_j \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}} < \frac{\sum_{v_i \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} \\
&\equiv |H_j \cap T_j'| * \mathcal{L}(p_{h,j}, p_{h,j}^*) < |H_i \cap T_i'| * \mathcal{L}(p_{h,i}, p_{h,i}^*) \\
&\equiv \mathcal{L}(p_{h,j}, p_{h,j}^*) < \mathcal{L}(p_{h,i}, p_{h,i}^*)
\end{aligned}$$

Step B: We only need to show that the probabilistic distance between transition probability and optimal transition probability is greater in e_i than e_j. Let $A = |H_i'| \; > \; B = |H_j'|$. We can decompose domain $v \in V$ into four parts as

$$v \in T_k \cap H_k' \quad v \in T_k \setminus H_k' \quad v \in H_k' \setminus T_k \quad v \in V \setminus \{H_k' \cup T_k\} \quad \forall k = i, j$$

Here, $\mathrm{JSD}(P \parallel Q)$ in the second term and fourth term are identical for both cases. That is, we only need to compare probabilistic distances which are related to the first and third parts of the above four domains.

$$\begin{aligned}
&\mathcal{L}(p_{h,j}, p_{h,j}^*) < \mathcal{L}(p_{h,i}, p_{h,i}^*) \\
&\equiv F * \ell\left(\frac{1}{B}, \frac{1}{T}\right) + \frac{B-F}{2B} \log 2 < F * \ell\left(\frac{1}{A}, \frac{1}{T}\right) + \frac{A-F}{2A} \log 2
\end{aligned}$$

where $F = |H_i' \cap T_i| = |H_j' \cap T_j|$ and $T = |T_i| = |T_j|$. Note that $A > B$. Inequality is rewritten as

$$\begin{aligned}
&\mathcal{L}(p_{h,i}, p_{h,i}^*) - \mathcal{L}(p_{h,j}, p_{h,j}^*) > 0 \\
&\equiv \frac{F}{2}\left(\frac{1}{B} - \frac{1}{A}\right) \log 2 > F * \left(\ell\left(\frac{1}{B}, \frac{1}{T}\right) - \ell\left(\frac{1}{A}, \frac{1}{T}\right)\right)
\end{aligned}$$

To simplify the equation, we unfold $\ell(p, q)$ as follows.

$$\begin{aligned}
&\frac{F}{2}\left(\frac{1}{B} - \frac{1}{A}\right) \log 2 > F * \left(\ell\left(\frac{1}{B}, \frac{1}{T}\right) - \ell\left(\frac{1}{A}, \frac{1}{T}\right)\right) \\
&\equiv \frac{1}{2}\left(\frac{1}{B} - \frac{1}{A}\right) \log 2 > \frac{1}{2T} \log \frac{2B}{B+T} + \frac{1}{2B} \log \frac{2T}{B+T} - \frac{1}{2T} \log \frac{2A}{A+T} - \frac{1}{2A} \log \frac{2T}{A+T} \\
&\equiv \left(\frac{1}{B} - \frac{1}{A}\right) \log 2 > \frac{1}{T} \log \frac{2B}{B+T} + \frac{1}{B} \log \frac{2T}{B+T} - \frac{1}{T} \log \frac{2A}{A+T} - \frac{1}{A} \log \frac{2T}{A+T} \\
&\equiv 0 > \frac{1}{T} \log \frac{B}{B+T} + \frac{1}{B} \log \frac{T}{B+T} - \frac{1}{T} \log \frac{A}{A+T} - \frac{1}{A} \log \frac{T}{A+T} \quad \because \text{pull 2 inside the log term out} \\
&\equiv 0 > \frac{1}{T} \log \frac{B(A+T)}{A(B+T)} - \frac{1}{B} \log \frac{B+T}{T} + \frac{1}{A} \log \frac{A+T}{T} \quad \text{multiply } T \text{ in both terms} \\
&\equiv 0 > \log \frac{AB+BT}{AB+AT} + \frac{T}{A} \log{(1 + \frac{A}{T})} - \frac{T}{B} \log{(1 + \frac{B}{T})}
\end{aligned}$$

We show the last inequality by splitting it to two terms ; $\log \frac{AB+BT}{AB+AT} < 0$ and $\frac{T}{A} \log{(1 + \frac{A}{T})} - \frac{T}{B} \log{(1 + \frac{B}{T})} < 0$. The first term is trivial since

$$\frac{AB+BT}{AB+AT} < 1 \quad \because B < A$$

Regarding the second term, note the form of the function.

$$\frac{T}{A} \log{(1 + \frac{A}{T})} - \frac{T}{B} \log{(1 + \frac{B}{T})} < 0 \tag{7}$$

Here, terms regarding A and B have forms of $f(x) = \frac{1}{x}\log(1+x)$. By using the fact that $A > B$, if a function $f(x)$ is a decreasing function at $x > 0$, inequality (7) is satisfied. For this, we show $f'(x) < 0 \; x > 0$,

$$
\begin{aligned}
f'(x) &= -\frac{1}{x^2}\log(x+1) + \frac{1}{x(x+1)} < 0 \quad \text{multiply } x^2(x+1) \text{ in both terms} \\
f'(x) &= -(x+1)\log(x+1) + x < 0 \quad \text{where } f'(0) = 0 \\
f''(x) &= -1 - \log(x+1) + 1 < 0 \quad \forall x > 0
\end{aligned}
$$

As $f''(x) < 0$, and $f'(x) = 0$ we can derive $f'(x) < 0$ for $x > 0$. This implies a function $f(x)$ is a decreasing function. Thus we have shown inequality (7) satisfies, The proof is done.

1.2.3 Proof of How HYPERREC Satisfies AXIOM 3

AXIOM 3 is divided into two parts,

- **AXIOM 3A:** The tail set of candidate arcs are identical while head sets do not overlap.
- **AXIOM 3B:** The head set of candidate arcs are identical while tail sets do not overlap.

We prove two sub **AXIOMS** respectively as follows.

[AXIOM 3A]: *Number of Candidate arcs differ where Tail Sets are Identical.*

Let $e'_k \subseteq_{(R)} e_k$ indicates $H'_k \subseteq T_k$ and $T'_k \subseteq H_k$.
Assume $C_i = \{e'_{i1}, e'_{i2}\}$ and $C_j = \{e'_j\}$ where

$$
\begin{gathered}
e'_{i1} \subseteq_{(R)} e_i \quad e'_{i2} \subseteq_{(R)} e_i \quad e'_j \subseteq_{(R)} e_j \quad T'_{i1} = T'_{i2} \quad |T'_{i1}| = |T_j| \\
H'_{i1} \cap H'_{i2} = \emptyset \quad \text{and} \quad |(H'_{i1} \cup H'_{i2}) \cap T_i| = |H'_j \cap T_j|
\end{gathered}
$$

Then, the following inequality holds.

$$r(e_i \mid C_i) < r(e_j \mid C_j)$$

PROOF: Unlike previous cases where there exists only a single hyperarc in a candidate set, there are two arcs in the former case. Thus, cardinality penalty terms are not discarded. Two reciprocity terms $r(e_i \mid C_i)$ and $r(e_j \mid C_j)$ becomes

$$
\begin{aligned}
r(e_i \mid C_i) &= \left(\frac{1}{2}\right)^{\alpha}\left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) \\
r(e_j \mid C_j) &= \left(1 - \frac{\sum_{v_h \in H_j} \mathcal{L}((p_h, p_h^*))}{|H_j| \cdot \mathcal{L}_{max}}\right)
\end{aligned}
$$

<u>**Step A:**</u> From the inequality $r(e_i \mid C_i) < r(e_j \mid C_j)$, we can derive that

$$
\begin{aligned}
& r(e_i \mid C_i) < r(e_j \mid C_j) \\
&= \left(\frac{1}{2}\right)^{\alpha}\left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) < \left(1 - \frac{\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}}\right) \\
&\equiv \left(\frac{1}{2}\right)^{\alpha}\left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) < \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) \leq \left(1 - \frac{\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}}\right)
\end{aligned}
$$

The last less or equal relation can be induced because $\alpha > 0$. In sum, proof is rewritten as

$$\frac{\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}} \leq \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} \quad \text{as } |H_i| = |H_j|\text{, this can be re-written as}$$

$$\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*) \leq \sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*) \tag{8}$$

For the target arc e_i, as $T'_{i1} = T'_{i2}$, each $p_h \forall v_h \in H_i$ has a same distribution. For the target arc e_j, there is only one single candidate arc, still each $p_h\ \forall v_h \in H_i$ has a same distribution. That is, inequality (8) is rewritten as

$$\begin{aligned}
&\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*) \leq \sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*) \\
&\equiv |H_j \cap T'_j| * \mathcal{L}(p_{h,j}, p^*_{h,j}) \leq |H_i \cap T'_{i1}| * \mathcal{L}(p_{h,i}, p^*_{h,i})
\end{aligned}$$

Since $e'_{i1} \subseteq_{(R)} e_i,\ e'_{i2} \subseteq_{(R)} e_i,\ e'_j \subseteq_{(R)} e_j$ and $|T'_j| = |T'_{i1}|$, last inequality can be written by

$$\begin{aligned}
&|H_j \cap T'_j| * \mathcal{L}(p_{h,j}, p^*_{h,j}) \leq |H_i \cap T'_{i1}| * \mathcal{L}(p_{h,i}, p^*_{h,i}) \\
&\equiv |T'_j| * \mathcal{L}(p_{h,j}, p^*_{h,j}) \leq |T'_{i1}| * \mathcal{L}(p_{h,i}, p^*_{h,i}) \\
&\equiv \mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i})
\end{aligned}$$

The proof can be done by showing the last inequality, $\mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i})$.

Step B: In order to show $\mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i})$, we should clarify the transition probability of e_i, since it is different from the previous arcs with single candidate arc. That is

$$p_{h,i}(v) = \begin{cases} \frac{1}{2|H'_{i1}|} & \text{if } v \in H'_{i1} \\ \frac{1}{2|H'_{i2}|} & \text{if } v \in H'_{i2} \\ 0 & \text{otherwise} \end{cases}$$

Since $e'_{i1} \subseteq_{(R)} e_i,\ e'_{i2} \subseteq_{(R)} e_i,\ e'_j \subseteq_{(R)} e_j$, probabilistic domain can be separated as

$$\begin{aligned}
&v \in H'_{i1} \quad v \in H'_{i2} \quad v \in T_i \setminus \{H'_{i1} \cup H'_{i2}\} \quad v \in V \setminus T_i \quad \text{for } e_i \\
&v \in H'_j \quad v \in T_j \setminus H'_j \quad v \in V \setminus T_j \quad \text{for } e_j
\end{aligned}$$

As $|T_i| = |T_j|$, the last term is identical for both cases. By using this information, we can formally express $\mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i})$ by the below inequality. Since $|H_{i1} \cup H_{i2}| = |H_j|$ where $|H_{i1} \cap H_{i2}| = 0$, we derive $|H_{i1}| + |H_{i2}| = |H_j|$. By using it, we simplify the size of each set as follows ; $|H_j| = A, |H_{i1}| = B, |H_{i2}| = A - B$, and $|T_i| = |T_j| = T$.

$$\begin{aligned}
&\mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i}) \\
&\equiv A * \ell\left(\frac{1}{T}, \frac{1}{A}\right) + \frac{T-A}{2T}\log 2 \leq B * \ell\left(\frac{1}{T}, \frac{1}{2B}\right) + (A-B) * \ell\left(\frac{1}{T}, \frac{1}{2(A-B)}\right) + \frac{T-B-(A-B)}{2T}\log 2 \\
&\equiv A * \ell\left(\frac{1}{T}, \frac{1}{A}\right) \leq B * \ell\left(\frac{1}{T}, \frac{1}{2B}\right) + (A-B) * \ell\left(\frac{1}{T}, \frac{1}{2(A-B)}\right) \\
&\equiv \frac{A}{2T}\log\frac{2A}{A+T} + \frac{A}{2A}\log\frac{2T}{A+T} \quad \text{unfolding } \ell(p,q) \\
&\leq \frac{B}{2T}\log\frac{4B}{2B+T} + \frac{B}{4B}\log\frac{2T}{2B+T} + \frac{(A-B)}{2T}\log\frac{4(A-B)}{2(A-B)+T} + \frac{(A-B)}{4(A-B)}\log\frac{2T}{2(A-B)+T}
\end{aligned}$$

If we pull out $\log 2$ terms from both sides

$$\left(\frac{A}{2T} + \frac{A}{2A}\right)\log 2 \qquad \left(\frac{B}{2T} + \frac{B}{4B} + \frac{A-B}{2T} + \frac{B-A}{4(B-A)}\right)\log 2$$

LHS and RHS have identical terms. By erasing them, original inequality can be re-written as

$$\begin{aligned}
&\equiv \frac{A}{2T}\log\frac{A}{A+T} + \frac{A}{2A}\log\frac{T}{A+T} \\
&\leq \frac{B}{2T}\log\frac{2B}{2B+T} + \frac{B}{4B}\log\frac{T}{2B+T} + \frac{(A-B)}{2T}\log\frac{2(A-B)}{2(A-B)+T} + \frac{(A-B)}{4(A-B)}\log\frac{T}{2(A-B)+T}
\end{aligned}$$

Instead of directly comparing every term, we partially compare inequality as

$$\frac{A}{2T}\log\frac{A}{A+T} \leq \frac{B}{2T}\log\frac{2B}{2B+T} + \frac{(A-B)}{2T}\log\frac{2(A-B)}{2(A-B)+T} \tag{9}$$

$$\frac{1}{2}\log\frac{T}{A+T} \leq \frac{1}{4}\log\frac{T}{2B+T} + \frac{1}{4}\log\frac{T}{2(A-B)+T} \tag{10}$$

We first show the inequality (9). By multiplying $2T$ on both sides, we get

$$\begin{aligned}
&\frac{A}{2T}\log\frac{A}{A+T} \leq \frac{B}{2T}\log\frac{2B}{2B+T} + \frac{(A-B)}{2T}\log\frac{2(A-B)}{2(A-B)+T} \\
\equiv\ & A\log\frac{A}{A+T} \leq B\log\frac{2B}{2B+T} + (A-B)\log\frac{2(A-B)}{2(A-B)+T}
\end{aligned}$$

Here, we prove this inequality by using the functional form of $f(B) = B\log\frac{2B}{2B+T} + (A-B)\log\frac{2(A-B)}{2(A-B)+T}$ where B lies in $0 < B < A$.

$$\begin{aligned}
f(B) &= B[\log(2B) - \log(2B+T)] + (A-B)[\log(2(A-B)) - \log(2(A-B)+T)] \\
\frac{\partial f(B)}{\partial B} &= \log(2B) - \log(2B+T) + B\left(\frac{2}{2B} - \frac{2}{2B+T}\right) \\
&\quad - \log(2(A-B)) + \log(2(A-B)+T) - (A-B)\left(\frac{2}{2(A-B)} - \frac{2}{2(A-B)+T}\right)
\end{aligned}$$

By cancelling out several identical terms, we can derive

$$\begin{aligned}
f'(B) &= \log(2B) - \log(2B+T) - \frac{2B}{2B+T} - \log(2(A-B)) + \log(2(A-B)+T) + \frac{2(A-B)}{2(A-B)+T} \\
&\equiv \log\frac{2B}{2B+T} - \frac{2B}{2B+T} - \log\frac{2(A-B)}{2(A-B)+T} + \frac{2(A-B)}{2(A-B)+T}
\end{aligned}$$

Note that the functional form of the last equation is $f'(B) = \log x - x - (\log y - y)$. Denote $x = \frac{2B}{2B+T}$ and $y = \frac{2(A-B)}{2(A-B)+T}$ who lie in $0 < x, y < 1$. Here, we derive the functional from of $f(B)$ by using following facts.

- If we plug in $B = \frac{A}{2}$, we have $f'(B) = 0$.
- $\log x - x$ is an increasing function for $0 < x < 1$.
- If $0 < B < \frac{A}{2}$, then $(\log y) - y > (\log x) - x$, which implies $f'(B) < 0$.
- If $\frac{A}{2} < B < A$, then $(\log x) - x > (\log y) - y$, which implies $f'(B) > 0$.

In sum, we can derive that

$$f'(B) = \begin{cases} < 0 & \text{if } 0 < B < \frac{A}{2} \\ = 0 & \text{if } B = \frac{A}{2} \\ > 0 & \text{if } \frac{A}{2} < B < A \end{cases}$$

Thus, we infer that $f(B)$ is a convex function at $0 < B < A$. Value of $f(B = \frac{A}{2}) = \frac{A}{2}\log\frac{A}{A+T} + \frac{A}{2}\log\frac{A}{A+T} = A\log\frac{A}{A+T}$. Rewind that our original goal is to show the below inequality.

$$A\log\frac{A}{A+T} \leq B\log\frac{2B}{2B+T} + (A-B)\log\frac{2(A-B)}{2(A-B)+T}$$

That is, the RHS term of inequality is a convex function that has its minimum value at $B = A/2$, and its minimum value is equal to the LHS term. Thus we can guarantee the inequality (9) is satisfied. The overview of this result is illustrated in Figure 2.

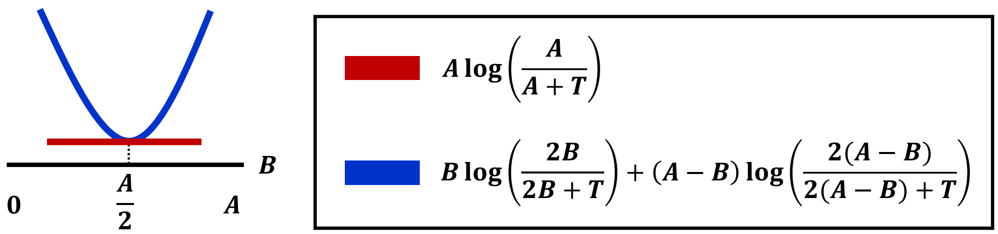


Figure 2: Result of the inequality (9)

Now we show the inequality (10).

$$
\begin{aligned}
&\frac{1}{2}\log\frac{T}{A+T} \le \frac{1}{4}\log\frac{T}{2B+T} + \frac{1}{4}\log\frac{T}{2(A-B)+T} \\
&\equiv \frac{1}{4}\log\frac{T}{A+T} + \frac{1}{4}\log\frac{T}{A+T} \le \frac{1}{4}\log\frac{T}{2B+T} + \frac{1}{4}\log\frac{T}{2(A-B)+T} \\
&\equiv \frac{1}{4}\log\frac{T^2}{(A+T)(A+T)} \le \frac{1}{4}\log\frac{T^2}{(2B+T)(2(A-B)+T)}
\end{aligned}
$$

This can be shown by comparing denominator of both terms, we rewrite last inequality by

$$
\begin{aligned}
&\equiv (2B+T)(2(A-B)+T) \le (A+T)(A+T) \\
&= 4B(A-B) + 2AT + T^2 \le A^2 + 2AT + T^2 \\
&\equiv A^2 - 4AB + 4B^2 = (A-2B)^2 \ge 0
\end{aligned}
$$

where equality holds at $B = A/2$. As both inequality (9) and (10) hold, where equality occur at $B = A/2$ in both terms, $\mathcal{L}(p_{h,j}, p^*_{h,j}) \le \mathcal{L}(p_{h,i}, p^*_{h,i})$ is satisfied, and the proof is done.

[AXIOM 3B]: *Number of Candidate Arcs differ where Head Sets are Identical.*

Let $e'_k \subseteq_{(R)} e_k$ indicates $H'_k \subseteq T_k$ and $T'_k \subseteq H_k$.
Assume $C_i = \{e'_{i1}, e'_{i2}\}$ and $C_j = \{e'_j\}$ where

$$
\begin{gathered}
e'_{i1} \subseteq_{(R)} e_i \quad e'_{i2} \subseteq_{(R)} e_i \quad e'_j \subseteq_{(R)} e_j \quad H'_{i1} = H'_{i2} \quad |H'_{i1}| = |H_j| \\
T'_{i1} \cap T'_{i2} = \emptyset \quad |(T'_{i1} \cup T'_{i2}) \cap H_i| = |T'_j \cap H_j|
\end{gathered}
$$

Then, the following inequality holds.

$$r(e_i \mid C_i) < r(e_j \mid C_j)$$

PROOF: Setting is all the same as the proof of **AXIOM 3A** until the inequality (8).

Step A: Following the proof of **AXIOM 3A**, below two inequalities are equivalent.

$$r(e_i \mid C_i) < r(e_j \mid C_j) \equiv \sum_{v_h \in H_i} \mathcal{L}(p_h, p^*_h) \le \sum_{v_h \in H_j} \mathcal{L}(p_h, p^*_h)$$

Here, note that as $H'_{i1} = H'_{i2}$, two candidate arcs of e_i have the same transition probability. Furthermore, as $T'_{i1} \cap T'_{i2} = \emptyset$, transition probability for every $v_h \in H_i \cap \{T'_{i1} \cup T'_{i2}\}$ are identical. Thus, let transition probability of e_i in this case as $p_i(v)$, which is

$$p_i(v) = p_{h,i}(v) \quad \forall v_h \in H_i \cap \{T'_{i1} \cup T'_{i2}\}$$

As $|H_i \cap \{T'_{i1} \cup T'_{i2}\}| = |H_j \cap T'_j|$, above final inequality is rewritten as

$$\mathcal{L}(p_{h,i}, p^*_{h,i}) \le \mathcal{L}(p_{h,j}, p^*_{h,j})$$

Step B: Note that $e'_{i1} \subseteq_{(R)} e_i$ $e'_{i2} \subseteq_{(R)} e_i$ $e'_j \subseteq_{(R)} e_j$ and $|H'_{i1}| = |H'_j|$, probabilistic distance at e_i and e_j are identical.

$$\mathcal{L}(p_{h,i}, p^*_{h,i}) = \mathcal{L}(p_{h,j}, p^*_{h,j})$$

Although their probabilistic distance are identical, penalty term $\left(\frac{1}{|C_i|}\right)^\alpha = \left(\frac{1}{2}\right)^\alpha < 1 \; \forall \alpha > 0$ enables the inequality $r(e_i \mid C_i) < r(e_j \mid C_j)$. The proof is done.

1.2.4 Proof of How HYPERREC Satisfies AXIOM 4

[AXIOM 4]: *Bias in the Candidate Arcs*

For two target arcs e_i and e_j,

$$|C_i| = |C_j| = |T_i| = |T_j| \quad |T_i'| = |T_j'|$$
$$T_i' = H_i \ , \ H_i' \subset T_i \ , \ |H_i'| = 2, \forall e_i' \in C_i$$
$$T_j' = H_j \ , \ H_j' \subset T_j \ , \ |H_j'| = 2, \forall e_j' \in C_j$$

$$\exists u, v \in T_i \quad \text{where} \quad |\{u \in e_i' \mid e_i' \in C_i\}| \neq |\{v \in e_i' \mid e_i' \in C_i\}| \tag{11}$$
$$\forall u, v \in T_j \quad \text{where} \quad |\{u \in e_j' \mid e_j' \in C_j\}| = |\{v \in e_j' \mid e_j' \in C_j\}| \tag{12}$$

Then, the following inequality holds.

$$r(e_i \mid C_i) < r(e_j \mid C_j)$$

PROOF: This **AXIOM** implies if there exists a bias in the candidate arcs, (roughly speaking, candidate arcs are concentrated on specific nodes of the target arc only) it is less reciprocal than the case where candidate arcs are uniformly pointing the target arc's tail set nodes.

Step A: General form can be written as

$$\begin{aligned} & r(e_i \mid C_i) < r(e_j \mid C_j) \\ & = \left(\frac{1}{|C_i|}\right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) < \left(\frac{1}{|C_j|}\right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}}\right) \end{aligned}$$

As $|C_i| = |C_j|$, $|H_i| = |H_i|$, $T_i' = H_i$, and $T_j' = H_j$, above inequality is rewritten as

$$\begin{aligned} & r(e_i \mid C_i) < r(e_j \mid C_j) \\ & \equiv \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) < \left(1 - \frac{\sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*)}{|H_j| \cdot \mathcal{L}_{max}}\right) \\ & \equiv \sum_{v_h \in H_j} \mathcal{L}(p_h, p_h^*) < \sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*) \\ & \equiv \mathcal{L}(p_{h,j}, p_{h,j}^*) < \mathcal{L}(p_{h,i}, p_{h,i}^*) \end{aligned}$$

Step B: Here, we prove the **AXIOM** by showing that given e_j is an optimally reciprocal case from the **AXIOM**'s setting, and other cases are inevitably less reciprocal than it. For the target arc $e_k \ \forall k = i, j$, corresponding candidate arcs are having head set of size 2, and given candidate set has cardinality of $|C_k| = |T_k|$. In this setting, Eq (12) (about e_j) implies that every node in target arc's tail set $v \in T_i$ is included in candidate arcs' head set twice. Thus,

- There are $|T_j|$ number of candidate arcs, whose heat set sizes are all 2.
- Every tail set node of target arc is being involved in the candidate arc's head set twice. $|T_j| * 2$
- Every candidate arc's head set is a subset of the target arc's tail set.

From above three statements, the transition probability can be induced as

$$p_{h,j} = \begin{cases} \frac{1}{2|T_j|} + \frac{1}{2|T_j|} = \frac{1}{|T_j|} & \text{if } v \in T_j \\ 0 & \text{otherwise} \end{cases}$$

Note that this is identical to the optimal transition probability.

Now, consider the case of e_i. Here, as there exists at least one inequality between the number of inclusion of each target arc's tail set node to the candidate arcs' head set, the transition probability cannot be uniform as in the case of e_j.

Let's assume $v_{j1}' \in T_i$ belongs to the candidate arcs' head set $K \neq 2$ times. Then, the transition probability at v_{j1}' is defined as $p(v_{j1}') = \frac{1}{2|T_i|} * K \neq \frac{1}{|T_i|}$. This result indicates the transition probability of e_i is not optimal at all. Thus, we can guarantee the following inequality. The proof is done.

$$\mathcal{L}(p_{h,j}, p_{h,j}^*) < \mathcal{L}(p_{h,i}, p_{h,i}^*) \quad \forall v_{h,i} \in T_i' \ \forall v_{h,j} \in T_j'$$

1.2.5 Proof of How HYPERREC Satisfies AXIOM 5

[AXIOM 5]: *Upper and Lower Bounds of Hyperarc Reciprocity*

For every hyperarc $e_i \in E$ with $C_i \subseteq E$, following range should holds.

$$0 \leq r(e_i \mid C_i) \leq 1$$

PROOF: We begin from the $\mathcal{L}(p, q)$.

$$\begin{aligned}
&0 \leq \frac{\mathcal{L}(p_h, p_h^*)}{\mathcal{L}_{max}} \leq 1 \quad \because 0 \leq \mathcal{L}(p, q) \leq \log 2 \\
&\equiv 0 \leq \sum_{v_h \in H_i} \frac{\mathcal{L}(p_h, p_h^*)}{\mathcal{L}_{max}} \leq |H_i| \\
&\equiv 0 \leq \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} \leq 1 \\
&\equiv 0 \leq 1E - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}} \leq 1 \\
&\equiv 0 \leq \left(\frac{1}{|C_i|}\right)^{\alpha} \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right) \leq 1 \quad \because \alpha > 0
\end{aligned}$$

The proof is done.

1.2.6 Proof of How HYPERREC Satisfies AXIOM 6

[AXIOM 6]: *Inclusion of Graph Reciprocity*

Digraph reciprocity is a special case of directed hypergraph reciprocity. That is, for $G = (V, E)$ if every hyperarc's head set and tail set size are equal to 1, $|H_i| = |T_i| = 1 \; \forall e_i \in E$ (i.e., digraph), following equality should hold.

$$r(G) = \frac{|E^{\leftrightarrow}|}{|E|} \quad \text{where } E^{\leftrightarrow} = \{e_i \mid \exists e_k = \langle H_k = T_i, T_k = H_i \rangle, (e_i, e_k \in E)\} \tag{13}$$

Note that this reciprocity measure is identical to the normal digraph reciprocity [2, 3].

PROOF: Hypergraph level reciprocity of **HYPERREC** is defined as

$$r(G) = \frac{1}{|E|} \sum_{i=1}^{|E|} r(e_i \mid C_i)$$

As every hyperarc's head set size is equal to 1, each $r(e_i \mid C_i)$ can be written as

$$r(e_i \mid C_i) = \max_{C_i \subseteq E} \left(\frac{1}{|C_i|}\right)^{\alpha} (\mathcal{L}(p_h, p_h^*)) \quad \text{where } \{v_h\} = H_i$$

Here, the optimal transition probability is defined as

$$p_h^*(v) = \begin{cases} 1 & \text{if } \{v\} = T_i \\ 0 & \text{otherwise} \end{cases}$$

One can notice that only one arc is required for the candidate set since there is no partially covered arc (i.e., $T_i' \subset H_i,\ H_i' \subset T_i$ does not exist), and we only need to decide whether an arc inversely-overlap or not, because max term will filter arcs which are not necessary.

For $C_i = \{e_k\}$ where $T_k \neq H_i$, by the **A. 4**, reciprocity for such hyperarc gets zero.

For $C_i = \{e_k\}$ where $T_k = H_i$, transition probability for the target arc e_i is defined as

$$p_h^*(v) = \begin{cases} 1 & \text{if } \{v\} = H_k \\ 0 & \text{otherwise} \end{cases}$$

Here, the transition probability and the optimal transition probability become identical if and only if $H_k = T_i$. Otherwise, as their non-zero probability domain does not overlap, reciprocity gets zero.

In sum, $r(e_i \mid C_i)$ is formally re-written as

$$r(e_i \mid C_i) = \begin{cases} 1 & \text{if } H_k = T_i \text{ and } T_k = H_i \\ 0 & \text{otherwise} \end{cases} \quad \text{where } C_i = \{e_k\}$$

Thus, $r(e_i \mid C_i)$ works as an indicator function which gives a value of 1 if there exists its inverse-pair, $\exists e_k \in E$ where $e_k = \langle H_k = T_i, T_k = H_i \rangle$ and zero elsewhere. Formally,

$$\begin{aligned} r(G) &= \frac{1}{|E|} \sum_{i=1}^{|E|} r(e_i \mid C_i) \\ &= \frac{1}{|E|} \sum_{i=1}^{|E|} \mathbb{1}((\exists e_k = \langle H_k = T_i, T_k = H_i \rangle), (e_k \in E)) \quad \text{where } \mathbb{1}(\text{TRUE}) = 1 \quad \mathbb{1}(\text{FALSE}) = 0 \\ &= \frac{|E^{\leftrightarrow}|}{|E|} \end{aligned}$$

which is identical to the digraph reciprocity measure. The proof is done.

1.2.7 Proof of How HYPERREC Satisfies AXIOM 7

[AXIOM 7]: *Upper and Lower Bounds of Hypergraph Reciprocity*

The reciprocity of any hypergraph should be within a fixed range. Specifically, for any hypergraph G, $0 \leq r(G) \leq 1$ should hold.

PROOF: Hypergraph level reciprocity of **HYPERREC** is defined as

$$r(G) = \frac{1}{|E|} \sum_{i=1}^{|E|} r(e_i \mid C_i)$$

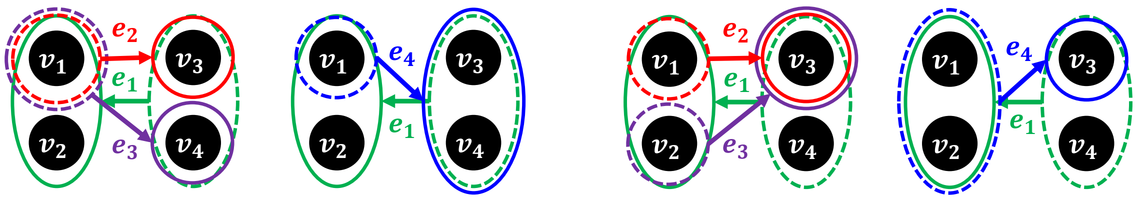
Here, by the **AXIOM 5**, we verify that an arbitrary hyperarc reciprocity $r(e_i \mid C_i)$ lies in $[0, 1]$. From this fact, we can derive that

$$\begin{aligned} & 0 \leq r(e_i \mid C_i) \leq 1 \quad \forall i = \{1, \cdots, |E|\} \\ & \equiv 0 \leq \sum_{i=1}^{|E|} r(e_i \mid C_i) \leq |E| \\ & \equiv 0 \leq \frac{1}{|E|} \sum_{i=1}^{|E|} r(e_i \mid C_i) \leq 1 \end{aligned}$$

The proof is done.

1.2.8 Proof of How HYPERREC Satisfies AXIOM 8

[AXIOM 8]: *Surjection of Reciprocity*



(a) Axiom 3A (b) Axiom 3B

Figure 3: Counterexamples for baselines. Left hypergraph of each subfigure is G_1 and right hypergraph of each subfigure is G_2.

The maximum reciprocity, which is 1 by the **AXIOM 7**, should be reachable from any hypergraph by adding specific hyperarcs. That is, for every $G = (V, E)$, there exists $G^* = (V, E^*)$ with $E \subseteq E^*$ such that $r(G^*) = 1$.

PROOF: From an arbitrary hypergraph G, let a subset of hyperarcs whose perfect reciprocal opponents do not exist in the same hypergraph i.e., $E^{'} = \{e_k \mid (\langle T_k, H_k\rangle \notin E), (e_k \in E)\}$. Let a set of additional hyperarcs that consist of perfectly reciprocal opponents of E' as E^{opt}.

$$E^{opt} = \bigcup_{e_k \in E'} \langle T_k, H_k\rangle$$

Here, let $E^* = E \cup E^{opt}$. For E^*, there always exists perfectly reciprocal opponent for each element (hyperarc).

$$\text{Let } C_i = \{\langle T_i, H_i\rangle\} \quad \text{Then} \quad r(e_i \mid C_i) = \left(1 - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p_h^*)}{|H_i| \cdot \mathcal{L}_{max}}\right)$$
$$= \left(1 - \frac{0 + \cdots + 0}{|H_i| \cdot \mathcal{L}_{max}}\right) = 1 \quad \because p_h = p_h^*$$

As $C_i = \{\langle T_i, H_i\rangle\}$, the target arc e_i can always achieves maximum hyperarc reciprocity. In addition, the corresponding candidate set can always be found from E^* due to the max searching. As every hyperarc reciprocity $r(e|C)$ is equal to 1, its hypergraph reciprocity $r(G)$ is also 1. The proof is done.

1.3 Violations of Axioms of Baseline Measures

In this subsection, we prove why several baseline measures fail in satisfying some **AXIOMS** .

COROLLARY 1.1. All the baseline measures violate at least one of **AXIOM** 1-8.

PROOF: We list up violated **AXIOMS** and show which baselines fail in satisfying corresponding **AXIOM** .

1.3.1 Violations of Axiom 3

B1. Pearcy et al. [4]: They use clique expansion which bi-cliques a hypergraph to the weighted digraph (see related work section of the main paper). We provide a simple example of the **AXIOMS** where hypergraphs become indistinguishable after clique expansion (see Figure 3). Note that as [4] does not propose the arc level reciprocity, we compare cases of $r(G_1)$ and $r(G_2)$ where $G_1 = (V_1 = \{v_1, v_2, v_3, v_4\}, E_1 = \{e_1, e_2, e_3\})$ and $G_2 = (V_2 = \{v_1, v_2, v_3, v_4\}, E_2 = \{e_1, e_4\})$. We show the counterexample by proving $r(G_1) = r(G_2)$.

Here for Figure 3 (a), resulting clique expansion $\bar{A}_1$ and $\bar{A}_2$ becomes

$$\bar{A}_1 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \bar{A}_2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

One can notice that resulting clique expanded adjacency matrices are identical. Here, perfect reciprocal hypergraphs $\bar{A}'_1$ and $\bar{A}'_2$ are also the same.

$$\bar{A}'_1 = \begin{pmatrix} 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{pmatrix} \quad \bar{A}'_2 = \begin{pmatrix} 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{pmatrix}$$

Thus, resulting reciprocity $r(G_1) = \frac{tr(\bar{A}_1^2)}{tr(\bar{A}_1'^2)}$ and $r(G_2) = \frac{tr(\bar{A}_2^2)}{tr(\bar{A}_2'^2)}$ are surely identical. That is, **AXIOM 3A** which should be $r(G_1) < r(G_2)$ has been violated.

For Figure 3 (b), resulting clique expansion $\bar{A}_1$ and $\bar{A}_2$ become

$$\bar{A}_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \bar{A}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

Similar to the previous case, resulting clique expanded adjacency matrices are identical. Furthermore, perfect reciprocal hypergraphs $\bar{A}'_1$ and $\bar{A}'_2$ are also the same.

$$\bar{A}'_1 = \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 2 & 2 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \bar{A}'_2 = \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 2 & 2 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

Here again, resulting reciprocity $r(G_1) = \frac{tr(\bar{A}_1^2)}{tr(\bar{A}_1'^2)}$ and $r(G_2) = \frac{tr(\bar{A}_2^2)}{tr(\bar{A}_2'^2)}$ are identical. Thus, **AXIOM 3B** which should be $r(G_1) < r(G_2)$ has been violated.

In sum, **B1. Pearcy et al. [4]** satisfies neither **AXIOM 3A** nor **AXIOM 3B**.

<u>B2. Ratio of Covered Pairs:</u> Unlike the previous case, we compare hyperarc level reciprocity. Here, we show that this metric results in $r(e_1 \mid C_1 = \{e_2, e_3\}) = r(e_1 \mid C_1 = \{e_4\})$, which is again a violation of **AXIOM 3**. For Figure 3 (a), each pair interaction can be defined as

$$R(e_1) = \{\langle v_1, v_3\rangle, \langle v_1, v_4\rangle, \langle v_2, v_3\rangle, \langle v_2, v_4\rangle\}.$$
$$R(e_2) = \{\langle v_3, v_1\rangle\} \quad R(e_3) = \{\langle v_4, v_1\rangle\} \quad R(e_4) = \{\langle v_3, v_1\rangle, \langle v_4, v_1\rangle\}$$
$$R^{-1}(e_2) = \{\langle v_1, v_3\rangle\} \quad R^{-1}(e_3) = \{\langle v_1, v_4\rangle\} \quad R^{-1}(e_4) = \{\langle v_1, v_3\rangle, \langle v_1, v_4\rangle\}$$

Thus, each reciprocity is defined as

$$r(e_1 \mid C_1 = \{e_2, e_3\}) = \frac{2}{4} = 0.5$$
$$r(e_1 \mid C_1 = \{e_4\}) = \frac{2}{4} = 0.5$$

In conclusion, we have shown $r(G_1) = r(G_2)$, which is a violation of the **AXIOM 3A**. Similarly, for Figure 3 (b), interaction pairs are defined as

$$R(e_1) = \{\langle v_1, v_3\rangle, \langle v_1, v_4\rangle, \langle v_2, v_3\rangle, \langle v_2, v_4\rangle\}.$$
$$R(e_2) = \{\langle v_3, v_1\rangle\} \quad R(e_3) = \{\langle v_3, v_2\rangle\} \quad R(e_4) = \{\langle v_3, v_1\rangle, \langle v_3, v_2\rangle\}$$
$$R^{-1}(e_2) = \{\langle v_1, v_3\rangle\} \quad R^{-1}(e_3) = \{\langle v_2, v_3\rangle\} \quad R^{-1}(e_4) = \{\langle v_1, v_3\rangle, \langle v_2, v_3\rangle\}$$

Similarly, each reciprocity is defined as

$$r(e_1 \mid C_1 = \{e_2, e_3\}) = \frac{2}{4} = 0.5$$
$$r(e_1 \mid C_1 = \{e_4\}) = \frac{2}{4} = 0.5$$

Here again, identical reciprocity values on both examples, which is a violation of the **Axiom 3B**. We conclude that **B2. Ratio of Covered Pairs** also violates **Axiom 3B**.

B5. HyperRec w/o Size Penalty: While proving general **HyperRec**, original inequality $r(e_i \mid C_i) < r(e_j \mid C_j)$ has been relaxed to $\mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i})$ with the help of the size penalty term $(1/|C_i|)^\alpha$. On the other hand, if we remove this size penalty term, the equality in $\mathcal{L}(p_{h,j}, p^*_{h,j}) \leq \mathcal{L}(p_{h,i}, p^*_{h,i})$ cannot guarantees the original reciprocity inequality. However, we have shown several cases where $\mathcal{L}(p_{h,j}, p^*_{h,j}) = \mathcal{L}(p_{h,i}, p^*_{h,i})$ occur, which result in the violation of **Axiom 3**. In sum, **HyperRec** without the size penalty term $(1/|C_i|)^\alpha$ cannot satisfies the **Axiom 3**.

1.3.2 Violations of Axiom 4

B2. Ratio of Covered Pairs: If every pair of target arc has been covered, then reciprocity for such a case is always equal to 1. That is, how many times each node has been pointed is indistinguishable from this metric. Here, reciprocity for both cases become $r(e_i \mid C_i) = r(e_j \mid C_j)$, which violates the **Axiom 4**.

B3. Penalized Ratio of Covered Pairs: Although size penalty has been added to **B2**, cardinality of candidate sets are identical for both cases. Thus, this also has the same issue, $r(e_i \mid C_i) = r(e_j \mid C_j)$, which violates the **Axiom 4**.

1.3.3 Violations of Axiom 5

B4. HyperRec w/o Normalization: If we have similar steps without $|H_i|$ in the denominator, we get

$$
\begin{aligned}
&0 \leq \frac{\mathcal{L}(p_h, p^*_h)}{\mathcal{L}_{max}} \leq 1 \quad \because 0 \leq \mathcal{L}(p,q) \leq \log 2 \\
&\equiv 0 \leq \sum_{v_h \in H_i} \frac{\mathcal{L}(p_h, p^*_h)}{\mathcal{L}_{max}} \leq |H_i| \\
&\equiv 0 \leq |H_i| - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p^*_h)}{\mathcal{L}_{max}} \leq |H_i| \\
&\equiv 0 \leq \left(\frac{1}{|C_i|}\right)^\alpha \left(|H_i| - \frac{\sum_{v_h \in H_i} \mathcal{L}(p_h, p^*_h)}{\mathcal{L}_{max}}\right) \leq |H_i| \quad \because \alpha > 0
\end{aligned}
$$

We can see the reciprocity lies between $0 \leq r(e_i \mid C_i) \leq |H_i|$, which is a violation of the **Axiom 5**. This contains the problem that it is hard to compare different hyperarcs' reciprocity who have different head set sizes, since the hyperarc reciprocity's range depends on the size of its head set.

1.3.4 Violations of Axiom 6

B1. Pearcy et al. [4]: We provide a simple counterexample. See the case illustrated in Figure 4

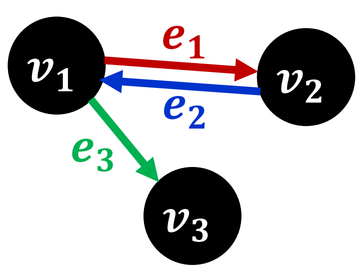


Figure 4: Counterexample why certain baselines fail in **Axiom 6**.

Digraph reciprocity for Figure 4 can be computed as

$$E = \{e_1, e_2, e_3\},\ E^{\leftrightarrow} = \{e_1, e_2\} \ \rightarrow\ r(G) = \frac{2}{3}$$

In this case, clique expanded adjacency matrix of actual data and its optimal case are

$$\bar{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \bar{A}' = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

The reciprocity becomes $\frac{2}{4} = 0.5$ as $tr(\bar{A}^2) = 2$ and $tr(\bar{A}'^2) = 4$. Since $\frac{2}{4} \neq \frac{2}{3}$, **B1. Pearcy et al. [4]** violates the **AXIOM 6**.

B6. HYPERREC with All Arcs as Candidates: As we put all the hyperarcs in C_i, this measure cannot achieves hyperarc reciprocity of 1 even in the case where there exists its perfectly reciprocal opponent.

In Figure 4, although e_2 has its perfectly reciprocal arc i.e., e_1, it takes e_3 also in consideration, which results in the transition probability of $p_1 = [v_1, v_2, v_3, v_{sunk}] = [0, 0.5, 0.5, 0]$. Here, final reciprocity becomes $r(e_2 \mid C_2 = \{e_1, e_3\}) = 0.7842 \neq 1$, which implies this metric cannot works as a proper indicator function. Thus, **AXIOM 6** cannot be achieved.

1.3.5 Violations of Axiom 8

B6. HYPERREC with All the Arcs as Candidates: Even though there exists a perfect reciprocal opponent of a specific hyperarc, transition probability cannot be identical to the optimal transition probability if there exist another inversely overlapping hyperarcs (see target arc e_2's case in Figure 4).

B7. HYPERREC with Inversely Overlapping Arcs as Candidates: Similar to the previous case, if there exist multiple inversely overlapping hyperarcs, it inevitably uses all of them as a candidate set. As a result, for such hyperarcs, cardinality penalty gets smaller than 1 i.e., $(1/|C_i|)^{\alpha} < 1$, resulting in $r(e_i \mid C_i) < 1$. Consequently, overall hypergraph reciprocity becomes smaller than 1, which cannot be fully reciprocal.

2 Data Description

In this section, we provide the sources of the considered datasets and describe how we preprocess them.

Metabolic datasets: We use two metabolic hypergraphs, **iAF1260b** and **iJO1366**, which are provided by Yadati et al. [5]. They are provided in the form of directed hypergraphs, and they do not require any pre-processing. We remove one hyperarc from each dataset since their head set or tail set is abnormally large. Specifically, the size of their head sets is greater than 20, while the second largest one is 8. Each node corresponds to a gene, and each hyperarc indicates a metabolic reaction among them. Specifically, a hyperarc e_i indicates that a reaction among the genes in the tail set T_i results in the genes in the head set H_i.

Email datasets: We use two email hypergraphs, **email-enron** and **email-eu**. The **Email-enron** dataset is provided by Chodrwo et al. [6]. We consider each email as a single hyperarc. Specifically, the head set is composed of the receiver(s) and cc-ed user(s), and the tail set is composed of the sender. The **Email-eu** dataset is from SNAP [7]. The original dataset is a dynamic graph where each temporal edge from a node u to a node v at time t indicates that u sent an email to v at time t. The edges with the same source node and timestamp are replaced by a hyperarc, where the tail set consists only of the source node and the head set is the set of destination nodes of the edges. Note that every hyperarc in these datasets has a unit tail set, i.e., $|T_i| = 1, \forall i = \{1, \cdots, |E|\}$.

Citation datasets: We use two citation hypergraphs, **citation-data mining** and **citation-software**, which we create from pairwise citation networks, as suggested by Yadati et al. [8]. Nodes are the authors of publications. Assume that a paper A, which is co-authored by $\{v_1, v_2, v_3\}$, cited another paper B, which is co-authored by $\{v_4, v_5\}$. Then, this citation leads to a hyperarc where the head set is $\{v_4, v_5\}$ and the tail set is $\{v_1, v_2, v_3\}$. As pairwise citation networks, we use subsets of a DBLP citation dataset [9]. The subsets consist of papers published in the venues of data mining and software engineering, respectively.[1] In addition, we filter out all papers co-authored by more than 10 authors to minimize the impact of such outliers.

Question answering datasets: We use two question answering hypergraphs, **qna-math** and **qna-server**. We create directed hypergraphs from the log data of a question answering site, *stack exchange*, provided at [11]. Among various

[1] We use the venues listed at [10].

Table 1: Hypergraph reciprocity $r(G)$ of 11 datasets when $\alpha \approx 0$, $\alpha = 0.5$, and $\alpha = 1$.

		metabolic		email		citation		qna		bitcoin		
		iAF1260b	iJO1366	enron	eu	data mining	software	math	server	2014	2015	2016
$r(G)$	$\alpha \approx 0$	21.455	22.533	59.001	79.416	12.078	15.316	9.608	13.219	10.829	6.923	3.045
	$\alpha = 0.5$	17.756	18.497	49.321	65.405	10.840	13.984	9.283	13.196	10.654	6.845	2.988
	$\alpha = 1.0$	16.654	17.385	44.299	58.069	10.585	13.704	9.236	13.193	10.606	6.828	2.977

Table 2: Hypergraph reciprocity $r(G)$ is robust to the choice of α. Although their absolute value may differ (see Table 1), their relative values are not sensitive to the to the choice of α, as supported by the fact that all the measured Pearson correlation coefficients and Spearman rank correlation coefficients are greater than 0.99.

		Pearson Correlation	**Spearman Rank Correlation**
$r(G)$	$\alpha \approx 0 \leftrightarrow \alpha = 0.5$	0.999	1.0
	$\alpha \approx 0 \leftrightarrow \alpha = 1.0$	0.999	0.991
	$\alpha = 0.5 \leftrightarrow \alpha = 1.0$	0.999	0.991

domains, we choose *math-overflow*, which covers mathematical questions, and *server-fault*, which treats server related issues. The original log data contains the posts of the site, and one questioner and one or more answerers are involved with each post. We ignore all posts without any answerer. We treat each user as a node, and we treat each post as a hyperarc. For each hyperarc, the questioner of the corresponding post composes the head set, and the answerer(s) compose the tail set. Note that every hyperarc in these datasets has a unit head set, i.e., 1 $|H_i| = 1, \forall i = \{1, \cdots, |E|\}$.

Bitcoin transaction dataset: We use three bitcoin transaction hypergraphs, **bitcoin-2014**, **bitcoin-2015**, and **bitcoin-2016**. The original datasets are provided by Wu et al. [12], and they contain first 1,500,000 transactions in 11/2014, 06/2015, and 01/2016, respectively. We model each account as a node, and we model each transaction as a hyperarc. As multiple accounts can be involved in a single transaction, the accounts from which the coins are sent compose the tail set, and the accounts to which the coins sent compose the head set. We remove all transactions where the head set and the tail set are exactly the same.

3 Experiments

In this section, we provide full experimental results, which are omitted in the main paper due to the space limit.

3.1 Reciprocity and the Choice of α

In Section IV-B of the main paper, we demonstrate that hypergraph reciprocity $r(G)$ is robust to the choice of α (i.e., size penalty scalar of candidate sets). Specifically, we show that, while their absolute values vary depending on α, their ranks in real-world hypergraphs remain almost the same. Moreover, we demonstrate that arc-level reciprocity values obtained using different α values are highly correlated. Regarding the observations, we provide the full results obtained from all the 11 considered datasets in Tables 1, 2, and 3.

3.2 Observation 1

In Section IV-B of the main paper, we report that hypergraph reciprocity is several orders of magnitude greater in real-world hypergraphs than in corresponding null hypergraphs. To statistically verify this statement, we perform a Z-test for each dataset as follows:

1. We create 30 randomized hypergraphs, $r(G_{null,i}), \forall i = \{1, \cdots 30\}$
2. Using the 30 randomized hypergraphs, we measure the average $\overline{r(G_{null})}$ and standard deviation $sd(r(G_{null}))$ of their hypergraph reciprocity.

Table 3: Arc-level reciprocity $r(e|C)$ is robust to the choice of α. Although their absolute values may differ, their relative values are not sensitive to the choice of α, as supported by the fact that the measured Pearson correlation coefficients and Spearman rank correlation coefficients are at least 0.678 and in many cases even close to 1.

		metabolic		email		citation		qna		bitcoin		
		iAF1260b	iJO1366	enron	eu	data mining	software	math	server	2014	2015	2016
Pearson	$\alpha \approx 0 \leftrightarrow \alpha = 0.5$	0.961	0.957	0.928	0.836	0.984	0.986	0.994	0.999	0.997	0.998	0.997
	$\alpha \approx 0 \leftrightarrow \alpha = 1.0$	0.916	0.913	0.828	0.678	0.973	0.977	0.992	0.999	0.995	0.997	0.996
	$\alpha = 0.5 \leftrightarrow \alpha = 1.0$	0.985	0.986	0.975	0.967	0.998	0.998	0.999	0.999	0.999	0.999	0.999
Spearman Rank	$\alpha \approx 0 \leftrightarrow \alpha = 0.5$	0.973	0.969	0.947	0.817	0.998	0.998	0.999	0.999	0.999	0.999	0.999
	$\alpha \approx 0 \leftrightarrow \alpha = 1.0$	0.925	0.918	0.869	0.721	0.996	0.996	0.999	0.999	0.999	0.999	0.999
	$\alpha = 0.5 \leftrightarrow \alpha = 1.0$	0.975	0.973	0.978	0.983	0.999	0.999	0.999	0.999	0.999	0.999	0.999

3. Based on the central limit theorem, we approximate Eq. (14) to the standard normal distribution, where $r(G)$ is a real-world hypergraph's reciprocity, and $\xrightarrow{d}$ indicates convergence in distribution.

$$Z_{(real-null)} = \frac{\overline{r(G_{null,obs})} - r(G)}{\frac{sd(r(G_{null}))}{\sqrt{30}}} \xrightarrow{d} N(0,1) \tag{14}$$

4. We perform a hypothesis test, where the null hypothesis μ_0 and the alternative hypothesis μ_a are as follows:

$$\mu_0 : \overline{r(G_{null})} = r(G)$$
$$\mu_a : \overline{r(G_{null})} < r(G)$$

5. From the computed $Z_{(real-null)}$, we derive the p-value as $P(z < Z_{(real-null)})$, where $P(z)$ is the cumulative probability distribution of the standard normal distribution.

We set the significance level of the testing to 0.01. For all the datasets, we adopt the alternative hypothesis, which implies that $\overline{r(G_{null})}$ is statistically-significantly smaller than $\overline{r(G)}$. The detailed numerical results of the tests are provided in Table 4. In summary, we demonstrate that the hypergraph reciprocity is statistically-significantly greater in real-world hypergraphs than in corresponding null hypergraphs.

Table 4: P-value testing results on the 11 considered datasets. The null hypotheses are all rejected, which implies that real-world hypergraphs are significantly more reciprocal than null hypergraphs. A p-value smaller than 0.00001 is denoted by 0.0000^{*}.

	metabolic		email		citation		qna		bitcoin		
	iAF1260b	iJO1366	enron	eu	data mining	software	math	server	2014	2015	2016
Z-stat	-1502.52	-1789.79	-241.13	-3835.98	-17548.20	-9605.19	-8884.98	-88965.12	-691316.77	-555709.95	-325308.06
P-value	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
Null hypothesis	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject

3.3 Observation 2

In Section IV-B of the main paper, we present that in real-world hypergraphs, arcs with non-zero reciprocity tend to have a higher head set out-degree and tail set in-degree than arcs with zero reciprocity. Regarding the observation, we provide the full results obtained from all the 11 considered datasets in Figure 5.

3.4 Observation 3

In Section IV-B of the paper, we show that, in real-world hypergraphs, nodes with balanced in-degree and out-degree tend to be involved in higher reciprocity hyperarcs. Regarding this observation, the full results from all the 11 considered datasets are provided in Figure 6. The detailed procedures for obtaining the line plots in Figure 6 are as follows:

1. We quantify the degree balance of each node $v \in V$ as follows:

$$Bal(v) = log(d_{in}(v)) - log(d_{out}(v)).$$

Note that, if $d_{in}(v) \approx d_{out}(v)$, then $Bal(v) \approx 0$. If $d_{in}(v) \gg d_{out}(v)$ then $Bal(v \gg 0$, and if $d_{in}(v) \ll d_{out}(v)$, then $Bal(v) \ll 0$.

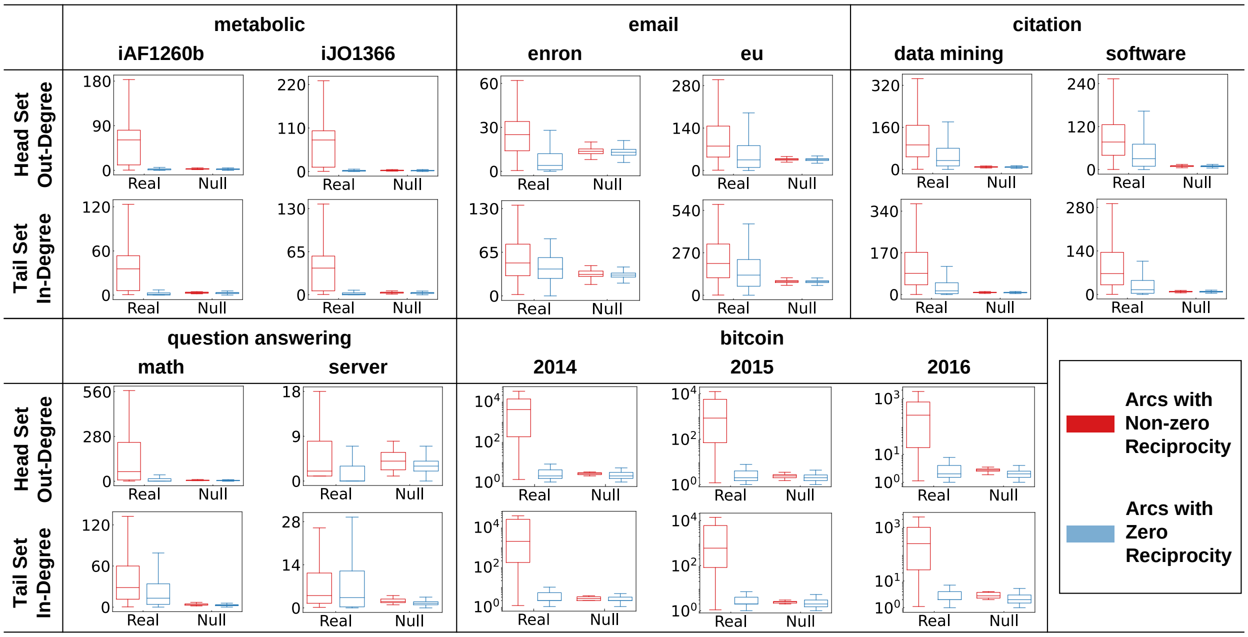


Figure 5: Full results regarding Observation 2. Arcs with non-zero reciprocity tend to have higher head set out-degrees and tail set in-degrees than arcs with zero reciprocity in all the datasets except for the question answering server dataset. However, such a phenomenon is not observed in null hypergraphs.

2. We quantify the node-level reciprocity of each node $v \in V$ as follows:

$$r(v) = \frac{1}{|E_v|} \sum_{e_k \in E_v} r(e_k)$$

where $E_v = \{e_k : v \in (H_k \cup T_k)\}$ is the set of hyperarcs where v is included in its head set or tail set.

3. We plot how the average node-level reciprocity depends on the degree balance of nodes after applying the Savitzky–Golay filter [13] for smoothing the curves.

3.5 REDI's Reproducibility of Observation 2

In Section V-B of the main paper, we demonstrate that hypergraphs generated by the proposed generator **REDI** exhibit our second empirical pattern (i.e., Observation 2) in the real-world hypergraphs. That is, arcs with non-zero reciprocity tend to have higher (a) head set out-degree and (b) tail set in-degree than arcs with zero reciprocity. Regarding this finding, the full experimental results obtained from all the 11 considered datasets are given in Figure 7. Noticeably, the tail set in-degree distribution of the question answering server dataset is similar between non-zero-reciprocity arcs and zero-reciprocity arcs (see Figure 5), and the tendency is reproduced by the hypergraphs generated by REDI generated question answering server dataset (see Figure 7).

3.6 REDI's Reproducibility of Observation 3

In Section V-B of the main paper, we show that our third empirical pattern (i.e., Observation 3) in the real-world hypergraphs is reproduced by hypergraphs generated by the proposed generator **REDI**. Regarding this finding, the full experimental results obtained from all the 11 considered datasets are given in Figure 6. To quantitatively compare how well the three methods (**REDI**, null, and baseline) reproduce the pattern, we also measure the difference between the plot from each generated (or null) hypergraph and the plot from the corresponding real-world hypergraph. Specifically, we measure the *mean-gap* between two plots, which is defined for any two plots f and f' as follows:

$$\textit{mean-gap}(f, f') = \frac{1}{|D|} \sum_{x \in D} (f(x) - f'(x))^2, \tag{15}$$

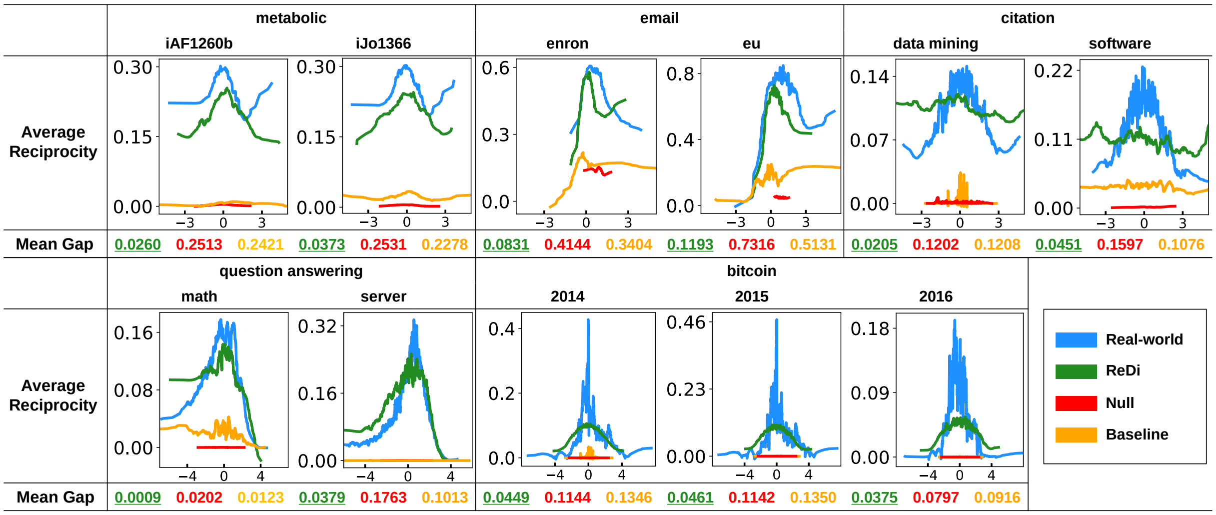


Figure 6: Full results regarding Observation 3 that show the superiority of **ReDi** over the baseline approach. In the real-world hypergraphs, node-level reciprocity tends to increase as nodes' in- and out-degrees become balanced. Similar patterns are observed in the hypergraphs generated by **ReDi** in all the datasets except for the citation datasets. However, null hypergraphs and those generated by the baseline approach do not exhibit such a tendency.

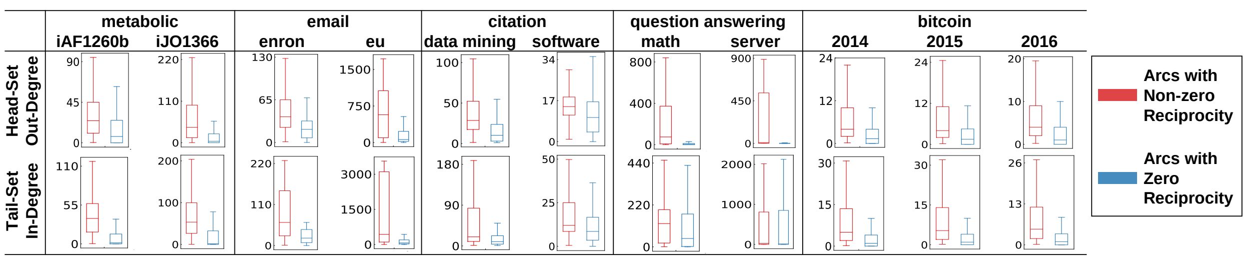


Figure 7: Observation 2 is reproduced by **ReDi**. As in the real-world hypergraphs, arcs with non-zero reciprocity tend to have tend to have higher head set out-degrees and tail set in-degrees than arcs with zero reciprocity in most of the hypergraphs generated by **ReDi**.

where D is the intersection of the domains of f and f'. As shown from the mean-gaps provided in Figure 6, among the three compared methods, the distance from the real-world hypergraph is smallest for the hypergraphs generated by **ReDi**. This result implies that **ReDi** reproduces Observation 3 best among them.

References

[1] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information theory*, vol. 37, no. 1, pp. 145–151, 1991.

[2] M. E. Newman, S. Forrest, and J. Balthrop, "Email networks and the spread of computer viruses," *Physical Review E*, vol. 66, no. 3, p. 035101, 2002.

[3] D. Garlaschelli and M. I. Loffredo, "Fitness-dependent topological properties of the world trade web," *Physical review letters*, vol. 93, no. 18, p. 188701, 2004.

[4] N. Pearcy, J. J. Crofts, and N. Chuzhanova, "Hypergraph models of metabolism," *International Journal of Biological, Veterinary, Agricultural and Food Engineering*, vol. 8, no. 8, pp. 752–756, 2014.

[5] N. Yadati, V. Nitin, M. Nimishakavi, P. Yadav, A. Louis, and P. Talukdar, "Nhp: Neural hypergraph link prediction," in *CIKM*, 2020.

[6] P. Chodrow and A. Mellor, "Annotated hypergraphs: Models and applications," *Applied network science*, vol. 5, no. 1, pp. 1–25, 2020.

[7] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, Jun. 2014.

[8] N. Yadati, T. Gao, S. Asoodeh, P. Talukdar, and A. Louis, "Graph neural networks for soft semi-supervised learning on hypergraphs," in *PAKDD*, 2021.

[9] A. Sinha, Z. Shen, Y. Song, H. Ma, D. Eide, B.-J. Hsu, and K. Wang, "An overview of microsoft academic service (mas) and applications," in *WWW*, 2015.

[10] "Computer Science Conference wikipedia." https://en.wikipedia.org/wiki/List_of_computer_science_conferences.

[11] "Stack Exchange Data Dump question answering dataset of stack exchange inc." https://archive.org/details/stackexchange.

[12] J. Wu, J. Liu, W. Chen, H. Huang, Z. Zheng, and Y. Zhang, "Detecting mixing services via mining bitcoin transaction network with hybrid motifs," 2020.

[13] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures." *Analytical chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.