How Transitive Are Real-World Group Interactions? - Measurement and Reproduction (Online Appendix)

Sunwoo Kim
KAIST
kswoo97@kaist.ac.kr

Fanchen Bu
KAIST
boqvezen97@kaist.ac.kr

Minyoung Choe
KAIST
minyoung.choe@kaist.ac.kr

Jaemin Yoo
Carnegie Mellon University
jaeminyoo@cmu.edu

Kijung Shin
KAIST
kijungs@kaist.ac.kr

An online appendix provided in this document serves to complement and enhance the findings presented in our publication, titled **"How Transitive Are Real-World Group Interactions? - Measurement and Reproduction"**, which will be published in KDD 2023. In this document, we provide additional analyses that support and expand upon the main content of our paper.

1 DETAILED EXPLANATIONS OF EACH BASELINE MEASURE (B1-6)

In this section, we present formulae of baseline measures **B1-6**, which are naive-intuitive approaches (**B1** and **B2**) or extensions of measures that are proposed in other studies (**B3**, **B4**, **B5**, and **B6**) [3, 4, 9].

First, **B1** and **B2** are two simple and intuitive measures.

B1. Jaccard index: **B1** computes the Jaccard similarity between (1) the union of all candidate hyperedges in C and (2) the union of the two wings of w. Formally,

$$\mathcal{T}(w,C) = \frac{|(\bigcup_{e'\in C} e') \cap (L(w) \cup R(w))|}{|(\bigcup_{e'\in C} e') \cup (L(w) \cup R(w))|}.$$

B2. Ratio of covered interactions: **B2** computes the ratio of pair interactions in $P(w)$ that are covered by (included in) the candidate hyperedges. Formally,

$$\mathcal{T}(w,C) = \frac{1}{|P(w)|} \times \left| \bigcup_{e'\in C} \binom{e'}{2} \cap P(w) \right|.$$

Baseline methods **B3-6** are extensions of existing hypergraph transitivity measures [3, 4, 9]. Since no existing measures were defined at the hyperwedge level, we extend the concept of local measures (i.e., the local clustering coefficient of a node) to the hyperwedge level.

B3. Klamt et al. [4]: **B3** computes the proportion of candidate hyperedges that intersect with both wings out of those that intersect with at least one wing. Formally,

$$\mathcal{T}(w,C) = \frac{|M(L(w);C) \cap M(R(w);C)|}{|M(L(w);C) \cup M(R(w);C)|}, \quad (1)$$

where $M(V';C) = \{e' \in C : e' \cap V' \neq \emptyset\}$. Note that if $\nexists e \in C\ s.t. e \cap (L(w) \cup R(w)) \neq \emptyset$ then we let $\mathcal{T}(w,C) = 0$.

B4. Torres et al. [9]: **B4** computes the proportion of wing-nodes (i.e., nodes that belong to a wing of the target hyperwedge) that are in the same candidate hyperedge with a node in the other wing. Formally,

$$\mathcal{T}(w,C) = \frac{|L(w) \cap \mathcal{N}(R(w);C)| + |R(w) \cap \mathcal{N}(L(w);C)|}{|L(w)| + |R(w)|}, \quad (2)$$

where $\mathcal{N}(V',C) = \bigcup_{e'\in M(V';C)} e'$ with $M(V';C)$ is defined in **B3**.

B5. Gallagher et al. A [3]: **B5** computes the proportion of pairs of wing-nodes that co-exist in a candidate hyperedge that is disjoint with the body group $B(w)$ out of all pairs of wing-nodes. Formally,

$$\mathcal{T}(w,C) = \Pr[\exists e' \in \tilde{C} \text{ s.t. } v_1, v_2 \in e' : v_1, v_2 \in L(w) \cup R(w)],$$

where $\tilde{C} = \{e' \in C : e' \cap B(w) = \emptyset\}$.

B6. Gallagher et al. B [3]: **B6** is similar to **B5**, except that the candidate hyperedges that intersect with the body group $B(w)$ are considered, instead of those disjoint with $B(w)$.

$$\mathcal{T}(w,C) = \Pr[\exists e' \in \hat{C} \text{ s.t. } v_1, v_2 \in e' : v_1, v_2 \in L(w) \cup R(w)],$$

where $\hat{C} = \{e' \in C : e' \cap B(w) \neq \emptyset\}$.

2 PROOF OF THEOREMS AND PROPOSITIONS

2.1 Detailed descriptions of axioms

We briefly rewind the proposed measure and axioms. The proposed measure HyperTrans is defined as:

$$\mathcal{T}(w,C;f) = \sum_{\{v_1',v_2'\}\in P(w)} \frac{\max_{e\in C} f(w,e)\mathbb{1}[v_1',v_2' \in e]}{|P(w)|}. \quad (3)$$

We describe axioms, which are also presented in the main paper.

Axiom 1 (Minimum hyperwedge transitivity). *A hyperwedge transitivity of w is globally minimized if and only if there is no candidate hyperedge in C being an overlapping hyperedge. Formally,* $\mathcal{T}(w,C) = 0$ *(see Axiom 5)* $\Leftrightarrow C \cap \Omega(w) = \emptyset$.

Axiom 2. *In this axiom, we discuss how hyperwedge transitivity should change in different situations when we* ***include more hyperedges*** *in the candidate set C.*

Case 1: *(General) Whenever more hyperedges are included in the candidate set C, w's transitivity remains the same or increases. Formally,* $C \subseteq C' \subseteq E \Rightarrow \mathcal{T}(w,C) \leq \mathcal{T}(w,C')$.

Case 2: *(Only non-overlapping) When only non-overlapping hyperedges are further included in C, w's transitivity remains the same. Formally,* $(C \subseteq C' \subseteq E) \wedge \left((C' \setminus C) \cap \Omega(w) = \emptyset\right) \Rightarrow \mathcal{T}(w,C) = \mathcal{T}(w,C')$.

Case 3: *(More interactions covered in total) When some hyperedges are further included in C so that more interactions in* $P(w)$ *are* *covered, w's transitivity strictly increases. Formally,* $(C \subseteq C' \subseteq E) \wedge \left(\exists e' \in C' : (\binom{e'}{2} \setminus \bigcup_{e \in C} \binom{e}{2}) \cap P(w) \neq \emptyset\right) \Rightarrow \mathcal{T}(w, C) < \mathcal{T}(w, C')$.

Axiom 3. *In this axiom, we discuss how hyperwedge transitivity should change in different situations when some candidate hyperedges in C are **enlarged with wing-nodes**, i.e., replaced by their supersets where the new nodes are from the two wings.*[1]

Case 1: *(General) When each* $e \in C$ *is either kept the same or enlarged with wing-nodes, w's transitivity remains the same or increases. Formally,* $\left(\exists \text{ bijection } g : C \rightarrow C' \text{ s.t } \left(e \subseteq g(e) \subseteq (e \cup L(w) \cup R(w)), \forall e \in C\right)\right) \Rightarrow \mathcal{T}(w, C) \leq \mathcal{T}(w, C')$.

Case 2: *(Each candidate more interaction-covering) When each* $e \in C$ *is enlarged with wing-nodes so that it covers more interactions in* $P(w)$*, w's transitivity strictly increases. Formally,* $\left(\exists \text{ bijection } g : C \rightarrow C' \text{ s.t } \left(e \subsetneq g(e) \subseteq (e \cup L(w) \cup R(w)) \wedge (\binom{g(e)}{2} \setminus \binom{e}{2}) \cap P(w) \neq \emptyset, \forall e \in C\right)\right) \Rightarrow \mathcal{T}(w, C) < \mathcal{T}(w, C')$.

Remark 1. *Axiom 3 assumes a **bijection**, ensuring that the enlarged hyperedges do not become equivalent to any other hyperedges.*

Axiom 4 (Maximum hyperwedge transitivity). *When hyperwedge transitivity of w is globally maximized, there exists at least one* $e \in C$ *including all the nodes of two wings* $L(w)$ *and* $R(w)$*. Formally,* $\mathcal{T}(w, C) = 1 \Rightarrow \exists e \in C \text{ s.t } L(w) \cup R(w) \subseteq e$.

Remark 2. *Since axioms focus on **group interaction**, Axiom 4 indicates that all elements in* $P(w)$ *should co-exist in a single hyperedge.*

Axiom 5 (Boundedness of hyperwedge transitivity). *A hyperwedge transitivity function* $\mathcal{T}$ *should be bounded. WLOG, we assume that the value is bounded within* $[0, 1]$*, i.e.,* $\mathcal{T}(w, C) \in [0, 1], \forall w \in W, \forall C \in 2^E \setminus \{\emptyset\}$.

We now propose two hypergraph-level axioms.

Axiom 6 (Reducibility to graph transitivity). *When the input hypergraph* $G = (V, E)$ *is a pairwise graph, i.e.,* $|e| = 2, \forall e \in E$*, the hypergraph transitivity* $T(G)$ *should be equal to (i.e., is reduced to) the graph transitivity [7] of* G.

Axiom 7 (Boundedness of hypergraph transitivity). *A hypergraph transitivity function* T *should be bounded. WLOG,* $T(G) \in [0, 1]$*, for every hypergraph* G.

2.2 Proof of Theorem 1

Note that Theorem 1 describes unconformity of all baseline measures **B1-9** where all baseline measures violate at least one of the proposed axioms. Formally, Theorem 1 is stated as

Theorem 1 (Unconformity of baseline measures). *Each baseline measure (**B1-9**) violates at least one among Axioms 1-7.*

We provide several counterexamples of baseline methods for each axiom. Here, we assume that **B7-9** (variants of HyperTrans) use a score function f that is defined in Eq (5) in the main paper, which is *good* (Proposition 1).

[1]Formally, for each hyperwedge w, a candidate hyperedge e is *enlarged with wing-nodes* (to e') if and only if $e \subseteq e'$ with $\emptyset \neq (e' \setminus e) \subseteq (L(w) \cup R(w))$.

In Lemma 1-7 we consider a hyperwedge $w = \{\{1, 2, 3\}, \{3, 4, 5\}\}$ and 15 hyperedges $e_1 = \{1, 2\}, e_2 = \{1, 3\}, e_3 = \{1, 4\}, e_4 = \{1, 5\}, e_5 = \{2, 4\}, e_6 = \{2, 5\}, e_7 = \{1, 3, 4\}, e_8 = \{1, 2, 4\}, e_9 = \{1, 2, 5\}, e_{10} = \{1, 4, 5\}, e_{11} = \{2, 4, 5\}, e_{12} = \{1, 2, 3, 4\}, e_{13} = \{1, 2, 3, 5\}, e_{14} = \{1, 3, 4, 5\}, e_{15} = \{2, 3, 4, 5\}$.

Lemma 1. *B1, B5, and B6 can not meet Axiom 1.*

Proof. Under the following setting, computed transitivity value of **B1**, **B5**, and **B6** are as follows, which are violations of Axiom 1.

$$\mathbf{B1} : \mathcal{T}(w, \{e_1\}) = 2/4 \neq 0,$$
$$\mathbf{B5} : \mathcal{T}(w, \{e_7\}) = 0/6 = 0,$$
$$\mathbf{B6} : \mathcal{T}(w, \{e_3\}) = 0/6 = 0.$$

□

Note that in Axiom 2 and 3, multiple cases are proposed respectively. Here, a violation of any of these two cases indicates the violation of the corresponding axiom. If certain result k is evidence of the violation of Case 2, we denote this as $k \xrightarrow{\mathbb{V}}$ Case 2. Note that $k \xrightarrow{\mathbb{V}}$ Case 2 used in the proof of Lemma 2 indicates that the Case 2 of Axiom 2 is violated.

Lemma 2. *B1, B3, B5, B6, B7, and B8 can not meet Axiom 2.*

Proof. Under the following setting, computed transitivity value of **B1**, **B3**, **B5**, **B6**, **B7**, and **B8** becomes as follows:

$\mathbf{B1} : \mathcal{T}(w, \{e_3\}) = 2/4 < \mathcal{T}(w, \{e_1, e_3\}) = 3/4 \xrightarrow{\mathbb{V}}$ **Case 2**,

$\mathbf{B3} : \mathcal{T}(w, \{e_3\}) = 1 > \mathcal{T}(w, \{e_1, e_3\}) = 1/2 \xrightarrow{\mathbb{V}}$ **Case 1, 2**,

$\mathbf{B5} : \mathcal{T}(w, \{e_7\}) = 1/6 = \mathcal{T}(w, \{e_4, e_7\}) = 1/6 \xrightarrow{\mathbb{V}}$ **Case 3**,

$\mathbf{B6} : \mathcal{T}(w, \{e_4\}) = 1/6 = \mathcal{T}(w, \{e_4, e_7\}) = 1/6 \xrightarrow{\mathbb{V}}$ **Case 3**,

$\mathbf{B7} : \mathcal{T}(w, \{e_3\}) = 1/16 > \mathcal{T}(w, \{e_1, e_3\}) = 1/24 \xrightarrow{\mathbb{V}}$ **Case 1, 2**,

$\mathbf{B8} : \mathcal{T}(w, \{e_8\}) = 1/2 = \mathcal{T}(w, \{e_4, e_8\}) = 1/2 \xrightarrow{\mathbb{V}}$ **Case 3**.

All these baseline methods violate at least one case of Axiom 2. □

Lemma 3. *B1, B2, B3, B4, B5, and B6 can not meet Axiom 3.*

Proof. Under the following setting, computed transitivity of **B1**, **B2**, **B3**, **B4**, **B5**, and **B6** becomes

$\mathbf{B1} : \mathcal{T}(w, \{e_3, e_6\}) = 1 = \mathcal{T}(w, \{e_9, e_{10}\}) = 1 \xrightarrow{\mathbb{V}}$ **Case 2**,

$\mathbf{B2} : \mathcal{T}(w, \{e_3, \cdots, e_6\}) = 1 = \mathcal{T}(w, \{e_8, \cdots, e_{11}\}) = 1 \xrightarrow{\mathbb{V}}$ **Case 2**,

$\mathbf{B3} : \mathcal{T}(w, \{e_3\}) = 1 = \mathcal{T}(w, \{e_8\}) = 1 \xrightarrow{\mathbb{V}}$ **Case 2**,

$\mathbf{B4} : \mathcal{T}(w, \{e_3, e_6\}) = 1 = \mathcal{T}(w, \{e_9, e_{10}\}) = 1 \xrightarrow{\mathbb{V}}$ **Case 2**,

$\mathbf{B5} : \mathcal{T}(w, \{e_2\}) = 0 = \mathcal{T}(w, \{e_7\}) = 0 \xrightarrow{\mathbb{V}}$ **Case 2**,

$\mathbf{B6} : \mathcal{T}(w, \{e_3\}) = 0 = \mathcal{T}(w, \{e_8\}) = 0 \xrightarrow{\mathbb{V}}$ **Case 2**.

All these baseline methods violate at least one case of Axiom 2. □

Now we show that transitivity values of some baseline measures are maximized even when hyperedges that contain both left and right wings (i.e., $L(w) \cup R(w) \subseteq e$) do not exist in a candidate set.

Lemma 4. *B1, B2, B3, B4, B5, and B6 can not meet Axiom 4.*

Proof. Under the following setting, computed transitivity of **B1**, **B2**, **B3**, **B4**, **B5**, and **B6** becomes as follows, which are violations of Axiom 4.

$$\begin{aligned}
&\mathbf{B1}: \mathcal{T}(w, \{e_3, e_6\}) = 1\\
&\mathbf{B2}: \mathcal{T}(w, \{e_3, \cdots e_6\}) = 1\\
&\mathbf{B3}: \mathcal{T}(w, \{e_2\}) = 1\\
&\mathbf{B4}: \mathcal{T}(w, \{e_3, e_6\}) = 1\\
&\mathbf{B5}: \mathcal{T}(w, \{e_8, \cdots, e_{11}\}) = 1\\
&\mathbf{B6}: \mathcal{T}(w, \{e_{12}, \cdots, e_{15}\}) = 1.
\end{aligned}$$

□

Lemma 5. *B9 can not meets Axiom 5.*

Proof. Let $w = \{\{1, 2, 3\}.\{3, 4, 5\}\}$ and $e = \{1, 2, 4, 5\}$. Since $f(w, e) = 1$ holds, resulting transitivity value becomes $\mathcal{T}(w, \{e\}) = 1 + 1 + 1 + 1 = 4$. The value lies outside $[0, 1]$, which is a violation of the condition of Axiom 5. Thus, **B9** violates Axiom 5. □

Lemma 6. *B6 can not meets Axiom 6.*

Proof. In this axiom, we assume a specific pair graph G where $G = (V = \{1, 2, 3\}, E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\})$. In this clique graph, $T(G) = 1$ should hold. By definition, candidate hyperedges with a size greater than 2 can make transitivity value of **B6** non-zero. However, since none of $e \in E$ include 3 nodes, $T(G)$ becomes zero with **B6**. □

Lemma 7. *B9 can not meets Axiom 7.*

Proof. In this axiom, we assume a specific hypergraph G where $G = (V = \{1, 2, 3, 4, 5\}, E = \{\{1, 2, 3\}, \{3, 4, 5\}, \{1, 2, 4, 5\}\})$. Here, $T(G)$ with **B9** becomes $T(G) = (4 + 2 + 2)/3 = 2.667$. The value lies outside $[0, 1]$, which is a violation of the condition of Axiom 5. Thus, Axiom 7 is violated. □

2.3 Proof of Theorem 2

Note that Theorem 2 states the soundness of HyperTrans by showing that HyperTrans satisfies all the axiom.

Theorem 2 (Soundness of HyperTrans). *HyperTrans (Eq (2)) with a good group interaction score function f satisfies Axiom 1-7.*

We decompose the proof of Theorem 2 by Lemma 8-17. Note that *good* function f should have the following properties.

Definition 1. *A group interaction function f is **good**, if f satisfies the following six properties for each w and e:*

(1) $f : (w, e) \in [0, 1], \forall w \in W(G), e \in E$.
(2) $e \in \Omega(w) \Rightarrow f(w, e) > 0$.
(3) $f(w, e) = 1 \Rightarrow L(w) \cup R(w) \subseteq e$.
(4) $L(w) \cup R(w) = e \Rightarrow f(w, e) = 1$.
(5) $e \subseteq e' \subseteq (e \cup L(w) \cup R(w)) \Rightarrow f(w, e) \leq f(w, e')$.
(6) $e \subsetneq e' \subseteq (e \cup L(w) \cup R(w)) \wedge e' \in \Omega(w) \Rightarrow f(w, e) < f(w, e')$.

We use the above Property 1-6 for proofs.

Lemma 8. *HyperTrans meets Axiom 1.*

Proof. We first show the following statement: $\mathcal{T}(w, C) = 0 \Rightarrow C \cap \Omega(w) = \emptyset$. By definition of HyperTrans and Property 1 (bounded condition), $\mathcal{T}(w, C) = 0$ is equivalent to

$$\max_{e \in C}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e]) = 0, \forall \{v'_1, v'_2\} \in P(w). \tag{4}$$

Let's assume $C \cap \Omega(w) \neq \emptyset$. Then, $\exists e \in C$ such that $e \in \Omega(w)$ holds, and the indicator function in Eq (4) yields a value of 1 for the corresponding e. In addition, according to Property 2, $f(w, e) > 0$ also holds. This results in $\max_{e \in C}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e]) \neq 0$, where contradiction occurs with Eq (4). Thus, Eq (4) induces $C \cap \Omega(w) = \emptyset$, which is our objective.

Now we show the other way around: $C \cap \Omega(w) = \emptyset \Rightarrow \mathcal{T}(w, C)$. Since $\nexists e \in C : e \in \Omega(w)$, the output of the indicator function for all hyperedges is zero. Thus, all scores are equal to zero, which is equivalent to our objective, $\mathcal{T}(w, C) = 0$. □

Lemma 9. *HyperTrans meets Case 1 of Axiom 2.*

Proof. Note that HyperTrans picks a maximum score $f(w, e)$ from $e \in C$. Since C is a subset of C', Eq (5) holds.

$$\begin{aligned}
&\max_{e \in C}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e])\\
&\leq \max_{e \in C'}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e]), \forall \{v'_1, v'_2\} \in P(w).
\end{aligned} \tag{5}$$

We rewrite the main inequality $\mathcal{T}(w, C) \leq \mathcal{T}(w, C')$ as follows:

$$\begin{aligned}
&\sum_{\{v'_1, v'_2\} \in P(w)} \frac{\max_{e \in C}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e])}{|L(w)| \times |R(w)|}\\
\leq &\sum_{\{v'_1, v'_2\} \in P(w)} \frac{\max_{e \in C'}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e])}{|L(w)| \times |R(w)|},\\
\equiv &\sum_{\{v'_1, v'_2\} \in P(w)} \max_{e \in C}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e])\\
\leq &\sum_{\{v'_1, v'_2\} \in P(w)} \max_{e \in C'}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e]).
\end{aligned} \tag{6}$$

By Eq (5), Eq (6) holds. Thus, $\mathcal{T}(w, C) \leq \mathcal{T}(w, C')$ also holds. □

Lemma 10. *HyperTrans meets Case 2 of Axiom 2.*

Proof. Since $\nexists e' \in (C' \setminus C) : e' \in \Omega(w)$, the indicator function yields 0 for all $e' \in C' \setminus C$, which is the minimum value of f (Property 1). Since HyperTrans picks a maximum score among the candidate set, e' can not change any transitivity value. Thus, $\mathcal{T}(w, C) = \mathcal{T}(w, C')$ always holds. □

Lemma 11. *HyperTrans meets Case 3 of Axiom 2.*

Proof. Denote $P(w, C)$ a set of total pair interactions by a candidate set C that intersects with $P(w)$ (i.e., $P(w, C) = \left(\bigcup_{e \in C} \binom{e}{2}\right) \cap P(w)$). Since $C \subseteq C'$ and there exists $e' \in C'$ such that $(\binom{e'}{2} \setminus (\bigcup_{e \in C} \binom{e}{2})) \cap P(w) \neq \emptyset$, the relation $P(w, C) \subsetneq P(w, C') \subseteq P(w)$ holds. At $\mathcal{T}(w, C)$, Eq (7) holds due to the indicator function:

$$\max_{e \in C}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e]) = 0, \forall \{v'_1, v'_2\} \in (P(w, C') \setminus P(w, C)). \tag{7}$$

Note that Eq 7 regards $e \in C$, not $e \in C'$. Here, due to Property 2 and the definition of $P(w, C')$, Eq (8) holds.

$$\exists \{v'_1, v'_2\} \in P(w) \setminus P(w, C) : \max_{e \in C'}(f(w, e) \times \mathbb{1}[v'_1, v'_2 \in e]) > 0. \tag{8}$$

Here, we rewrite $\mathcal{T}(w,C)$ as follows:

$$\sum_{\{v_1',v_2'\}\in(P(w,C))}\frac{\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])}{|L(w)|\times|R(w)|} \quad (9)$$

$$+\sum_{\{v_1',v_2'\}\in(P(w)\setminus P(w,C))}\frac{\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])}{|L(w)|\times|R(w)|}. \quad (10)$$

In the same way, we rewrite $\mathcal{T}(w,C')$ as follows:

$$\sum_{\{v_1',v_2'\}\in P(w,C)}\frac{\max_{e\in C'}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])}{|L(w)|\times|R(w)|} \quad (11)$$

$$+\sum_{\{v_1',v_2'\}\in(P(w)\setminus P(w,C))}\frac{\max_{e\in C'}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])}{|L(w)|\times|R(w)|}. \quad (12)$$

Since $C\subseteq C'$, Eq (9) $\leq$ Eq (10) holds because of Eq (5). In addition, according to Eq (7) and Eq (8), Eq (9) $<$ Eq (10) holds. Putting these two together, $\mathcal{T}(w,C)<\mathcal{T}(w,C')$ holds. □

Lemma 12. *HyperTrans meets Case 1 of Axiom 3.*

Proof. Let $e^M_{\{v_1',v_2'\}}$ denotes $\arg\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])$. Then, according to the setting of an axiom, Eq (13) holds.

$$\exists e'\in C':(e^M_{\{v_1',v_2'\}}\subseteq e')\wedge((e'\setminus e^M_{\{v_1',v_2'\}})\subseteq(L(w)\cup R(w))). \quad (13)$$

Denote $e'_{\{v_1',v_2'\}}\in C'$ that satisfies the condition expressed in Eq (13). Then, Eq (14) holds due to Property 5.

$$f(w,e^M_{\{v_1',v_2'\}})\leq f(w,e'_{\{v_1',v_2'\}}). \quad (14)$$

By definition, $f(w,e^M_{\{v_1',v_2'\}})=\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])$ holds. By using this equality and Eq (14), we derive Eq (15).

$$\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])\leq\max_{e'\in C'}(f(w,e')\times\mathbb{1}[v_1',v_2'\in e']). \quad (15)$$

Since the result of Eq (15) holds for an arbitrary $\{v_1',v_2'\}\in P(w)$, this can be extended to $\forall\{v_1',v_2'\}\in P(w)$. Thus, $\mathcal{T}(w,C)\leq\mathcal{T}(w,C')$ is induced from Eq (15). □

Lemma 13. *HyperTrans meets Case 2 of Axiom 3.*

Proof. Let $e^M_{\{v_1',v_2'\}}$ denotes $\arg\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])$. Then, according to the setting of axiom, Eq (16) holds.

$$\exists e'\in C':(e^M_{\{v_1',v_2'\}}\subset e')\wedge((e'\setminus e^M_{\{v_1',v_2'\}})\subseteq(L(w)\cup R(w))). \quad (16)$$

Denote $e'_{\{v_1',v_2'\}}\in C'$ that satisfies the condition expressed in Eq (16). Since $e'\in\Omega(w)$, we induce Eq (17) due to Property 6.

$$f(w,e^M_{\{v_1',v_2'\}})<f(w,e'_{\{v_1',v_2'\}}). \quad (17)$$

By definition, $f(w,e^M_{\{v_1',v_2'\}})=\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])$ holds. By using this equality and Eq (17), Eq (18) holds.

$$\max_{e\in C}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])<\max_{e'\in C'}(f(w,e')\times\mathbb{1}[v_1',v_2'\in e']). \quad (18)$$

Since the result of Eq (18) holds for an arbitrary $\{v_1',v_2'\}\in P(w)$, this can be extended to $\forall\{v_1',v_2'\}\in P(w)$. Thus, $\mathcal{T}(w,C)<\mathcal{T}(w,C')$ is induced from Eq (18). □

Lemma 14. *HyperTrans meets Axiom 4.*

Proof. By definition of HyperTrans and Property 1, $\mathcal{T}(w,C)=1$ is equivalent to Eq (19).

$$\max_{e\in C'}(f(w,e)\times\mathbb{1}[v_1',v_2'\in e])=1,\forall\{v_1',v_2'\}\in P(w). \quad (19)$$

This indicates there exists $e\in C$ that satisfies $f(w,e)\times\mathbb{1}[v_1',v_2'\in e]=1$. To this end, $f(w,e)=1$ and $\mathbb{1}[v_1',v_2'\in e]=1$ should hold. By Property 3, $e\in C$ such that $f(w,e)=1$ should satisfy $L(w)\cup R(w)\subseteq e$. Thus, Eq (19) induces $\exists e\in C:L(w)\cup R(w)\subseteq e$, which is a condition of the axiom. □

Lemma 15. *HyperTrans meets Axiom 5.*

Proof. Here, the proof is straightforward as follows.

$$\begin{aligned}
&0\leq f(w,e)\leq 1,\forall e\in E,\quad \because \text{Property 1}.\\
&\equiv 0\leq\max_{e\in C}f(w,e)\leq 1, C\subseteq E.\\
&\equiv 0\leq\max_{e\in C}f(w,e)\times\mathbb{1}[v_1',v_2']\leq 1, C\subseteq E, v_1',v_2'\in V.\\
&\equiv 0\leq\sum_{\{v_1',v_2'\}\in P(w)}\max_{e\in C}f(w,e)\times\mathbb{1}[v_1',v_2'\in e]\leq|P(w)|\\
&\equiv 0\leq\sum_{\{v_1',v_2'\}\in P(w)}\frac{\max_{e\in C}f(w,e)\times\mathbb{1}[v_1',v_2'\in e]}{|L(w)|\times|R(w)|}\leq 1. \quad (20)\\
&\because |L(w)|\times|R(w)|=|P(w)|
\end{aligned}$$

Note that the Eq (20) is identical to the form of HyperTrans. □

Lemma 16. *HyperTrans meets Axiom 7.*

Proof. Since $|e_i|=2,\forall i=\{1,\cdots,|E|\}$, the size of all wings of G are 1 (i.e., $|L(w)|=|R(w)|=1,\forall w\in W(G)$). Let $L(w)=\{v_1'\}$ and $R(w)=\{v_2'\}$. We rewrite $\mathcal{T}(w,C)$ as

$$\mathcal{T}(w,C)=\max_{e\in C}f(w,e)\times\mathbb{1}[v_1',v_2'\in e]. \quad (21)$$

Here, we rewrite $\mathbb{1}[v_1',v_2'\in e]$ as follows:

$$\mathbb{1}[v_1',v_2'\in e]\equiv\mathbb{1}[\{v_1',v_2'\}=e]\quad\because|e|=2. \quad (22)$$

By Eq (22) and Property 4, we rewrite Eq (21) as follows:

$$\mathcal{T}(w)=\begin{cases}1 & \text{if }\exists e\in E:e=\{v_1',v_2'\}\\ 0 & \text{else}\end{cases} \quad (23)$$

Eq (23) indicates that for a hyperwedge $w=\{e_i,e_j\}, e_i=\{v_1',v_3'\}\in E$ and $e_j=\{v_2',v_3'\}\in E$, $\mathcal{T}(w)=1$ holds if there exists an edge $e_k=\{v_1',v_2'\}\in E$. In such a case, $e_i,e_j,e_k\in E$, and this is identical to the triangle between $v_1',v_2',$ and v_3'. Here, for $w'=\{e_i,e_k\}$ and $w''=\{e_j,e_k\}$, $\mathcal{T}(w')=\mathcal{T}(w'')=1$ also holds. In sum, for a triangle formed by $e_i,e_j,$ and e_k, $\mathcal{T}(w)+\mathcal{T}(w')+\mathcal{T}(w'')=3$ holds. Thus, a hypergraph-level measure of HyperTrans is identical to the number of triangles in a hypergraph multiplied by the 3 and divided by the number of hyperwedges, which is identical to that of graph transitivity measure [10]. □

Lemma 17. *HyperTrans meets Axiom 7.*

Proof. According to Lemma 15, $0\leq\mathcal{T}(w)\leq 1$ holds. Thus,

$$0\leq\sum_{w\in W}\mathcal{T}(w)\leq|W|\equiv 0\leq\frac{1}{|W|}\sum_{w\in W}\mathcal{T}(w)\leq 1. \quad (24)$$

Note that Eq (24) is the condition of Axiom 7. □

2.4 Proof of Proposition 1

Note that Proposition 1 states that Eq (25) (Eq (2) in the main paper) is *good* (i.e., has PROPERTY 1-6).

PROPOSITION 1. *The function f defined in Eq* (25) *(Eq (2) in the main paper) is good.*

Note that the Eq (25) is

$$f(w,e) = \frac{|L(w) \cap e)| \times |R(w) \cap e|}{|L(w) \cup (e \setminus R(w))| \times |R(w) \cup (e \setminus L(w))|}. \tag{25}$$

We decompose $f(w,e)$ of Eq (25) into two parts:

$$\frac{|L(w) \cap e|}{|L(w) \cup (e \setminus R(w))|}, \tag{26}$$

$$\frac{|R(w) \cap e|}{|R(w) \cup (e \setminus L(w))|}. \tag{27}$$

To prove that Eq (25) is *good*, we show Eq (25) meets all PROPERTY 1-6 by proving LEMMA 18-24.

LEMMA 18. *Eq (25) meets PROPERTY 1*

PROOF. Note that $L(w) \cap e \subseteq L(w) \cup (e \setminus R(w))$ and $R(w) \cap e \subseteq L(w) \cup (e \setminus R(w))$ hold by definition of $L(w)$ and $R(w)$. In addition, the denominator of Eq (26) and (27) are non-empty. Thus, $0 \leq$ Eq (26) ≤ 1 and $0 \leq$ Eq (27) ≤ 1 hold. Since Eq (25) is a product of Eq (26) and (27), $0 \leq$ Eq (25) ≤ 1 also holds, which is our objective. □

LEMMA 19. *Eq (25) meets PROPERTY 2*

PROOF. When $e \in \Omega(w)$, numerator of both Eq (26) and Eq (27) become greater than 0. Thus, Eq (25) > 0 holds. □

For the ease of proof of LEMMA 21-22, we provide an additional lemma and its proof.

LEMMA 20. *For f = Eq (25), $f(w,e) = 1$ iff $e = L(w) \cup R(w)$.*

PROOF. We first proof $f(w,e) = 1 \Rightarrow e = L(w) \cup R(w)$. Since the numerator is a subset of the denominator at both Eq (26) and Eq (27), $f(w,e) = 1$ hold *iff* the numerator and denominator are identical. Formally,

$$\begin{aligned} f(w,e) \Leftrightarrow &(L(w) \cap e = L(w) \cup (e \setminus R(w))) \wedge \\ &(R(w) \cap e = R(w) \cup (e \setminus L(w))). \end{aligned} \tag{28}$$

Let's first assume that $e \cap (V \setminus (L(w) \cup R(w))) \neq \emptyset$. Then, Eq (28) is violated. Thus, Eq (29) should hold.

$$e \cap (V \setminus (L(w) \cup R(w))) = \emptyset \Rightarrow e \subseteq L(w) \cup R(w). \tag{29}$$

Now let's assume that $e \subsetneq (L(w) \cup R(w))$. Then, $(L(w) \cap e = L(w) \cup (e \setminus R(w)))$ or $(L(w) \cap e = L(w) \cup (e \setminus R(w)))$ might not hold, which is a violation of Eq (28). Thus, $e = L(w) \cup R(w)$ is induced from the aforementioned contradiction.

Now we show $L(w) \cup R(w) = e \Rightarrow f(w,e) = 1$. In such a case, both Eq (26) and Eq (27) become 1, which result in $f(w,e) = 1$. Thus, $L(w) \cup R(w) = e \Rightarrow f(w,e) = 1$ also holds. Since bidirectional relation holds, *iff* condition is satisfied. □

LEMMA 21. *Eq (25) meets PROPERTY 3*

PROOF. By LEMMA 20, $f(w,e) = 1 \Rightarrow e = L(w) \cup R(w)$ holds. Note that $e = L(w) \cup R(w)$ is a special case of $L(w) \cup R(w) \subseteq e$, and the proof is done. □

LEMMA 22. *Eq (25) meets PROPERTY 4*

PROOF. By LEMMA 20, $e = L(w) \cup R(w) \Rightarrow f(w,e) = 1$ holds. Note that this is a statement of PROPERTY 4 and the proof is done. □

LEMMA 23. *Eq (25) meets PROPERTY 5*

PROOF. Let a set difference between e and e' as V'. Formally,

$$V' = (e' \setminus e) \subseteq L(w) \cup R(w). \tag{30}$$

Here, $f(w,e') = f(w, e \cup V')$ is expressed as

$$\begin{aligned} f(w,(e \cup V')) = &\frac{|L(w) \cap (e \cup V'))|}{|L(w) \cup ((e \cup V') \setminus R(w))|} \\ &\times \frac{|R(w) \cap (e \cup V')|}{|R(w) \cup ((e \cup V') \setminus L(w))|}. \end{aligned} \tag{31}$$

By definition of V', denominators of $f(w,e)$ and $f(w,e')$ are identical. Thus, $f(w,e) \leq f(w,e')$, which is our main objective, is equivalent to Eq (32).

$$\begin{aligned} \equiv &|L(w) \cap e| \times |R(w) \cap e| \\ \leq &|L(w) \cap (e \cup V')| \times |R(w) \cap (e \cup V')|. \end{aligned} \tag{32}$$

By definition of V', Eq (32) holds. □

LEMMA 24. *Eq (25) meets PROPERTY 6*

PROOF. Let a set difference between e and e' as V'. Formally,

$$V' = (e' \setminus e) \subseteq L(w) \cup R(w),\ V' \neq \emptyset. \tag{33}$$

Eq (32) is similar to the main inequality of PROPERTY 6, except for $\leq$, which should be $<$. Formally,

$$\begin{aligned} &|L(w) \cap e| \times |R(w) \cap e| \\ < &|L(w) \cap (e \cup V')| \times |R(w) \cap (e \cup V')|. \end{aligned} \tag{34}$$

By definition of V' and the fact that $e' \in \Omega(w)$ holds, the following inequality (Eq (35)) also holds.

$$|L(w) \cap (e \cup V')| \times |R(w) \cap (e \cup V')| > 0 \because \text{LEMMA 19}. \tag{35}$$

By definition of V' and Eq (35), Eq (34) is guaranteed. □

2.5 Proof of Theorem 3

Note that Theorem 3 states the exactness of FAST-HYPERTRANS.

THEOREM 3 (EXACTNESS). *Given any w, C, f, FAST-HYPERTRANS outputs $\mathcal{T}(w, C; f)$ as defined in Eq (3).*

PROOF. Note that $\max_{e \in C} f(w,e) \mathbb{1}[v_i', v_j' \in e]$, a term that is inside the summation of Eq (3), is equivalent to $\max_{e \in C_{i,j}} f(w,e)$ where $C_{i,j}$ denotes a set of hyperedges that include both v_i and v_j since $0 \leq f(w,e) \leq 1$. For each $e \in C$, FAST-HYPERTRANS sequentially compares a score mapped by $\{v_i', v_j'\} \in (P(w) \cap \binom{e}{2})$ with the score of the corresponding hyperedge $f(w,e)$. The existing value is replaced with $f(w,e)$ if it is smaller than $f(w,e)$. When the algorithm terminates, a value mapped by a key $\{v_i', v_j'\}$ is identical to the $\max_{e \in C_{i,j}} f(w,e)$ since this process is equivalent to the common maximum-searching algorithm. Thus, computing an average of every mapped value is identical to computing Eq (3) in turn. □

Algorithm 1: NAIVE-HYPERTRANS: a direct computation algorithm for HYPERTRANS

Input: Hyperwedge w, candidate set C, and score function f.
Output: Hyperwedge transitivity $\mathcal{T}(w, C)$

1 $\mathbb{T} \leftarrow 0$
2 **foreach** $\{v_1', v_2'\} \in P(w)$ **do**
3     $t = 0$
4   **foreach** $e \in C$ **do**
5   **if** $\{v_1', v_2'\} \subseteq e$ **then**
6             $t = \max(t, \mathcal{T}(w, e))$
7     $\mathbb{T} \leftarrow \mathbb{T} + t$
8 **return** $\mathbb{T}/|P(w)|$

2.6 Description of NAIVE-HYPERTRANS and Proof of Theorem 4

We first describe NAIVE-HYPERTRANS, which computes HYPERTRANS as expressed in formula Eq (3). The Pseudocode of NAIVE-HYPERTRANS is provided in Algorithm 1. For each pairwise elements $\{v_1', v_2'\} \in P(w)$ (line 2), it selects which candidate hyperedge $e \in C$ includes the corresponding $\{v_1', v_2'\}$. Then, it chooses the maximum $f(w, e)$ among scores of such candidate hyperedges (line 4-6).

Now we give a proof of Theorem 4. Note that Theorem 4 states that the time complexity of FAST-HYPERTRANS is upper bounded by that of NAIVE-HYPERTRANS.

THEOREM 4 (TIME COMPLEXITY). *Given any w, C, f, $TC_{fast}(w, C) = O(TC_{naive}(w, C))$, where $TC_{naive}(w, C, f)$ is the time complexity of NAIVE-HYPERTRANS, and $TC_{fast}(w, C, f)$ is that of FAST-HYPERTRANS.*

PROOF. We first compute the complexity of NAIVE-HYPERTRANS. For a specific $\{v_1', v_2'\} \in P(w)$, NAIVE-HYPERTRANS should check which candidate hyperedges $e \in C$ is containing $\{v_1', v_2'\}$ and picks a maximum score among them. This process has a complexity of $\Theta(|C|)$. The above procedure is performed for every $\{v_1', v_2'\} \in P(w)$. Consequently, its final complexity with big-Θ notation becomes

$$\Theta(|L(w)| \times |R(w)| \times |C|) = \Theta\left(\sum_{e \in C} |L(w)| \times |R(w)|\right). \tag{36}$$

Now we describe the complexity of FAST-HYPERTRANS. For a specific $e \in C$, FAST-HYPERTRANS should first get $\binom{e}{2} \cap P(w)$, which requires $\min(|L(w)|, |e|) + \min(|R(w)|, |e|)$ complexity. After finding the intersection, FAST-HYPERTRANS goes through every $\binom{e}{2} \cap P(w)$ and compares the existing value and $f(w, e)$. This requires the complexity of $|L(w) \cap e| \times |R(w) \cap e|$. In sum, the overall complexity regarding the specific candidate hyperedge $e \in C$ is as follows:

$$\min(|L(w)|, |e|) + \min(|R(w)|, |e|) + |L(w) \cap e| \times |R(w) \cap e|. \tag{37}$$

Note that this complexity occurs for each $e \in C$. Now we find the upper bound of Eq (37). Since $|L(w) \cap e| \leq \min(|L(w)|, |e|)$ and $|R(w) \cap e| \leq \min(|R(w)|, |e|)$ holds, Eq (37) $\leq$ Eq (38) holds.

$$\min(|L(w)|, |e|) + \min(|R(w)|, |e|) + \min(|L(w)|, |e|) \times \min(|R(w)|, |e|). \tag{38}$$

Algorithm 2: THERA: Transitive hypergraph generator

Input: (1) Number of nodes n, hyperedge size distribution S
(2) Community size C, intra-community hyperedge ratio p
(3) Level-sampling coefficient α, level-size coefficient β
Output: Generated hypergraph $G' = (V', E')$

$idx \leftarrow 1$; $T \leftarrow 0$; $\Psi_L(T') \leftarrow \emptyset, \forall T' \in \mathbb{N}^+$; $E' \leftarrow \emptyset$; $m \leftarrow \sum_k S(k)$
$AE(1) \leftarrow 0$; $AE(i) \leftarrow 1, \forall i = \{2, \cdots, n\}$; $\Psi_L(0) \leftarrow \{v_1\}$
while $\text{sum}(AE) < m$ **do**
    $a \sim$ discrete-uniform$(\{2, \cdots n\})$; $AE(a) \leftarrow AE(a) + 1$
while $idx < n$ **do**
    $T \leftarrow T + 1$ ▹ level in the hierarchy
    $\Psi_L(T) \leftarrow \{v_{(idx+i)}\}_{1 \leq i \leq \min(CT^\beta, n-idx)}$ ▹ node set at level T
    for $i = 1$ ***to*** $\min(CT^\beta, n - idx)$ **do**
        $idx \leftarrow idx + 1$
        for $j = 1$ ***to*** $AE(idx)$ **do**
            $e' \leftarrow \{v_{idx}\}$; $s \sim S$; $q \sim$ uniform$(0, 1)$
            if $q < p$ **then**
                $e' \leftarrow$ **IntraCommunityGenerate**$(e', idx, C, T, s, \Psi_L)$
            if $|e'| < s$ **then**
                $e' \leftarrow$ **HierarchicalGenerate**(e', T, s, Ψ_L)
            $E' \leftarrow E' \cup \{e'\}$
return $G' = (V' = \{v_1, \cdots, v_n\}, E')$

IntraCommunityGenerate$(e', idx, C, T, s, \Psi_L)$
    $idx' = C \times \lceil (idx - 2)/C \rceil + 1$ ▹ beginning index of community
    $V_C \leftarrow \Psi_L(T) \cap \{v_{(idx'+1)}, \cdots, v_{(idx'+C)}\} \setminus e'$ ▹ community
    $V' \leftarrow$ uniformly sample $\min(s - 1, |V_C|)$ nodes from V_C
    return $e' \cup V'$

HierarchicalGenerate(e', T, s, Ψ_L)
    while $|e'| < s$ **do**
        $\ell \leftarrow$ sample a level from $[0, 1, \cdots, T]$ proportional to $[|\Psi_L(0)|, \alpha^{-1}|\Psi_L(1)|, \cdots, \alpha^{-T}|\Psi_L(T)|]$ respectively
        $v' \leftarrow$ sample a node from $\Psi_L(\ell)$ uniformly at random
        $e' \leftarrow e' \cup \{v'\}$
    return e'

By using big-O notation, we rewrite Eq (38) as follows:

$$\begin{aligned} &O(\min(|L(w)|, |e|) + \min(|R(w)|, |e|) \\ &+ \min(|L(w)|, |e|) \times \min(|R(w)|, |e|)) \\ =&O(\min(|L(w)|, |e|) \times \min(|R(w)|, |e|)). \end{aligned} \tag{39}$$

Thus, the total complexity of every $e \in C$ becomes as follows:

$$O\left(\sum_{e \in C} \min(|L(w)|, |e|) \times \min(|R(w)|, |e|)\right). \tag{40}$$

Here, the relationship between terms inside the summation of Eq (36) and (40) is as follows:

$$\min(|L(w)|, |e|) \times \min(|R(w)|, |e| \leq |L(w)| \times |R(w)|, \forall e \in C. \tag{41}$$

From Eq (41) and the fact that Eq (36) is about big-Θ and Eq (40) is about big-O, it is clear that TC_{full} is upper bounded by TC_{naive}. □

Figure 1: Relations between node degree and node transitivity on the entire real-world hypergraphs (observation 3).

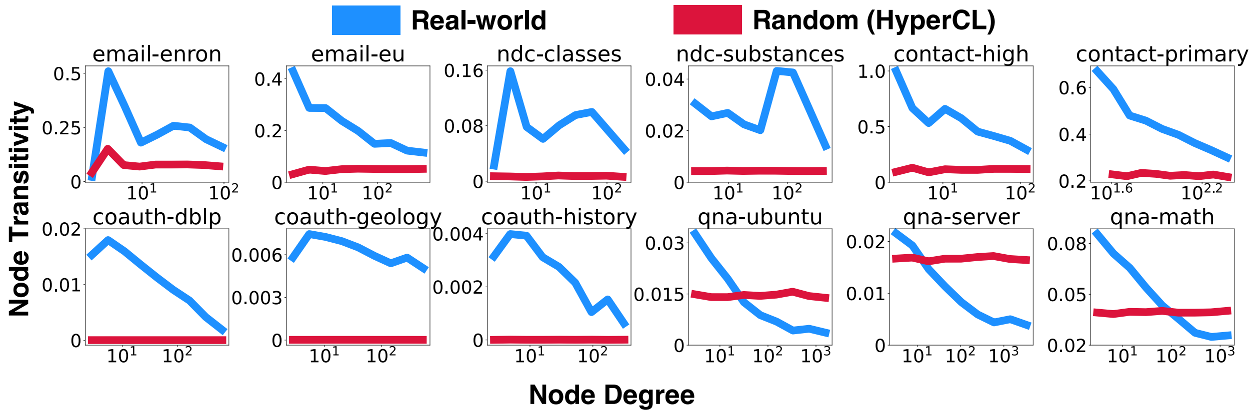
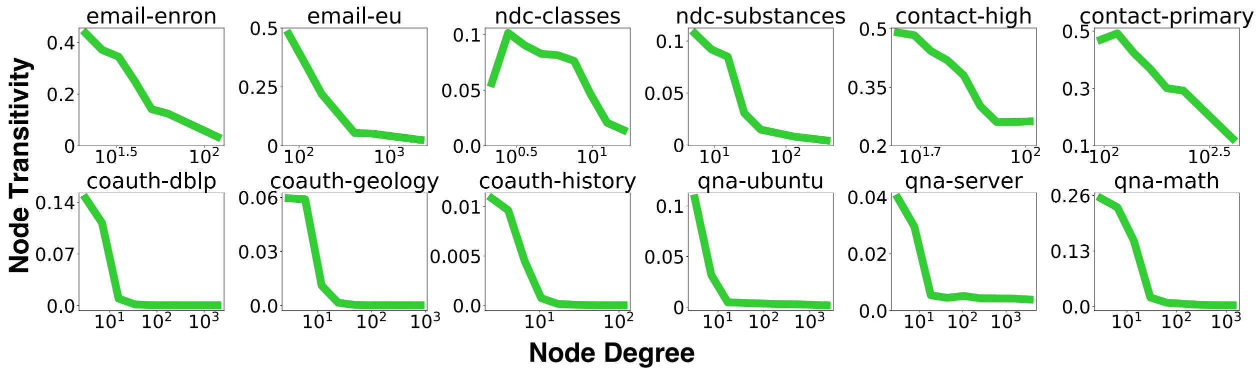


Figure 2: Relations between node degree and node transitivity on the entire generated hypergraphs by THERA (observation 3).



2.7 Proof of Proposition 2

Note that Proposition 2 indicates that the nodes of lower-level are likely to have higher degrees. To enhance the reader's understanding, We present a pseudo-code of THERA (Algorithm 2), which is also presented in the main paper.

PROPOSITION 2 (NEGATIVE CORRELATION BETWEEN LAYER INDEX AND NODE DEGREE). *For any v_1, v_2 with $L(v_1) > L(v_2)$ and $\alpha \geq 1$, in a hypergraph generated by THERA (Algorithm 2), the expected degree of v_1 is smaller than that of v_2, i.e., $\mathbb{E}[d(v_1)] < \mathbb{E}[d(v_2)]$.*

PROOF. We assume that every layer is fully filled, i.e., each layer ℓ consists of exactly ℓ^{β} communities, and each contains C nodes. Let T_{max} be the maximum layer index. This assumption does not affect the statement since only the last layer is possibly incomplete and this can only make the nodes at the last layer, i.e., with the highest level index, have lower expected node degrees. For ease of presentation, let $v_{\ell,x,y}$ denote the y-th node in the x-th community at the ℓ-th layer. For each node $v_{\ell,x,y}$, it is included in a hyperedge in two cases: (1) it can be included in an intra-community hyperedge by **IntraCommunityGenerate**, or (2) it can be included in a hierarchical hyperedge by **HierarchicalGenerate**. Note that for all the nodes, the distribution of $AE(\cdot)$ and hyperedge sizes (following S) are identical. Formally, let $S_{\ell,x,y}$ be the sequence of the s values for $v_{\ell,x,y}$ as in Line 11 in Algorithm 2, and the distribution of $S_{\ell,x,y}$ for all the nodes $v_{\ell,x,y}$'s is identical up to permutations. For case (1), the expected number of edges for each node $v_{\ell,x,y}$ is the same, which is $\mathbb{E}_{S_{\ell,x,y}}[1 + (C-1)q \sum_{s \in S_{\ell,x,y}} \binom{c-1}{s-2} / \binom{c-1}{s-1}]$. Hence, it suffices to check case (2) only. For case (2), for each node $v_{\ell,x,y}$, the expected number of hyperedges is

$$\mathbb{E}_{S_{\ell,x,y}}[\sum_{\ell'=\ell}^{T_{max}} \sum_{x'=1}^{\ell'^{\beta}} \sum_{i'=1}^{C} \sum_{s \in S_{\ell,x,y}} (1-q)(1-(1-\frac{\alpha^{-\ell}}{\sum_{\ell''=1}^{\ell'} \alpha^{\ell} \ell''^{\beta}})^{s})],$$

where the first summation $\sum_{\ell'=\ell}^{T_{max}}$ is strictly decreasing (ℓ' has a smaller range) with ℓ increasing, and all the inner summation is positive and non-increasing with ℓ increasing when $\alpha \geq 1$. □

3 EXPERIMENTAL RESULTS

3.1 Additional experimental results regarding Observation 3

In the main paper, we discuss the relationship between the node degree and node transitivity. Specifically, the average transitivity of nodes decreases as the node degree increase, while such trends cannot be observed on their random counterpart (refer to Observation 3 in the main paper). In addition, THERA has successfully reproduced this pattern (refer to Section 5.2 of the main paper). In this subsection, we present additional experimental results that demonstrate such patterns in the entire dataset.

Table 1: Detailed statistics regarding the reproducibility of various hypergraph properties of our proposed generator THera and four baseline generative models (HyperPA [2], HyperFF [5], HyperLap [6], and HyperLap+ [6]). We use formulae of each hypergraph property that are explained in Choe et al. [1]. For statistics that can be summarized as a single scalar, we report the absolute statistics (density and diameter). For statistics that are represented by a distribution, we report Kolmogorov–Smirnov statistic (D-stat) [8], a distributional distance between the distribution of the corresponding statistic of a real-world hypergraph and that of the generated hypergraph (hyperedge size, degree, and intersection size). Note that the proposed generator THera's reproducibility of various hypergraph properties is comparable to that of other baseline hypergraph generative models.

Statistic	Generator	coauthorship			q&a			email		ndc		contact		Average
		dblp	geology	history	ubuntu	math	server	enron	eu	classes	substances	high	primary	ranking
Density (Absolute Value)	**Real-world**	1.18	0.83	0.50	1.33	2.59	1.46	10.20	24.87	0.91	1.82	23.91	52.50	NA
	THera	1.18	1.00	1.00	1.00	2.59	1.46	10.20	24.87	1.00	1.82	23.91	52.50	1.25
	HyperPA [2]	-	-	-	2.47	-	-	9.79	25.32	2.66	-	181.81	145.81	4.50
	HyperFF [5]	3.24	3.23	3.24	3.23	3.23	3.23	3.49	11.09	3.62	3.20	3.02	2.95	4.08
	HyperLap [6]	1.43	1.04	0.71	1.72	2.83	1.73	10.20	24.99	0.99	1.87	23.91	52.50	1.58
	HyperLap+ [6]	1.47	1.09	0.72	1.89	3.42	1.98	10.20	25.54	1.08	2.15	23.91	52.50	2.50
Diameter (Absolute Value)	**Real-world**	6.81	7.09	11.45	4.74	3.75	4.29	2.38	2.78	4.65	3.59	2.63	1.88	NA
	THera	5.90	5.99	6.25	7.46	4.62	4.63	1.86	1.89	3.35	2.11	2.41	1.87	2.67
	HyperPA [2]	-	-	-	3.77	-	-	1.95	1.97	2.60	-	1.83	1.79	4.08
	HyperFF [5]	7.14	7.01	6.83	6.37	5.95	6.56	4.00	3.05	4.73	5.01	5.11	4.91	2.83
	HyperLap [6]	5.32	5.08	7.17	6.48	4.23	4.83	4.00	3.05	4.73	5.01	5.11	4.91	2.58
	HyperLap+ [6]	15.66	12.10	35.68	6.75	4.79	4.94	2.78	3.41	3.51	3.70	3.11	2.73	2.83
Hyperedge Size (D-stat)	**THera**	0.003	0.002	0.001	0.306	0.037	0.003	0.123	0.257	0.024	0.011	0.249	0.281	3.67
	HyperPA [2]	-	-	-	0.301	-	-	0.123	0.257	0.024	0.011	0.249	0.281	3.92
	HyperFF [5]	0.381	0.397	0.063	0.306	0.056	0.104	0.143	0.186	0.470	0.561	0.106	0.186	4.17
	HyperLap [6]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.00
	HyperLap+ [6]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.00
Degree (D-stat)	**THera**	0.500	0.555	0.564	0.489	0.709	0.499	0.378	0.569	0.691	0.667	0.373	0.347	3.91
	HyperPA [2]	-	-	-	0.225	-	-	0.343	0.322	0.064	-	0.878	0.550	3.93
	HyperFF [5]	0.360	0.430	0.605	0.566	0.411	0.466	0.532	0.359	0.348	0.158	0.856	0.963	3.75
	HyperLap [6]	0.244	0.266	0.326	0.308	0.251	0.264	0.119	0.043	0.244	0.141	0.024	0.029	1.25
	HyperLap+ [6]	0.258	0.292	0.331	0.369	0.408	0.366	0.084	0.111	0.323	0.320	0.046	0.050	1.92
Intersection Size (D-stat)	**THera**	0.675	0.643	0.412	0.200	0.508	0.250	0.231	0.195	0.167	0.414	0.417	0.250	3.00
	HyperPA [2]	-	-	-	0.300	-	-	0.255	0.234	0.280	-	0.250	0.250	3.83
	HyperFF [5]	0.267	0.275	0.308	0.200	0.444	0.167	0.385	0.249	0.482	0.250	0.333	0.500	3.08
	HyperLap [6]	0.233	0.315	0.227	0.300	0.333	0.233	0.192	0.222	0.357	0.214	0.250	0.250	1.92
	HyperLap+ [6]	0.259	0.315	0.227	0.333	0.278	0.233	0.231	0.195	0.167	0.414	0.417	0.250	2.42

As shown in Figure 2, Observation 3 in the main paper has been discovered in 10 out of 12 datasets, where such a pattern is not clear in NDC-classes and NDC-substances datasets. In addition, as shown in Figure 2, THera has successfully reproduced Observation 3 in datasets where such a pattern is clearly observed.

3.2 Reproduction of other properties

In this subsection, we present detailed experimental results on how well the proposed generator THera and other baseline generators including HyperPA [2], HyperFF [5], HyperLAP [6], and HyperLAP+ [6] reproduce various properties of real-world hypergraphs. We examine density, diameter, hyperedge size, degree, and intersection size. For density and diameter, we have reported the value of each statistic in real-world and generated hypergraphs, and the ranks of each generator have been measured according to the absolute difference between the statistic of real-world hypergraphs and that of generated hypergraphs. For hyperedge size, degree, and intersection size, since they are distribution, we have reported D-statistic between the distribution of the real-world and that of the generated hypergraphs. Here, ranks are measured according to how small the corresponding D-statistic is (i.e., how close the distribution of generated hypergraphs' statistics are to that of real-world hypergraphs). Detailed results are reported in Table 1. As explained in Appendix C of the main paper, THera shows outperforming or competitive results on various hypergraph properties, compared to other (including both incremental and static) hypergraph generators. In addition, it outperforms other incremental generators including HyperPA and HyperFF in most properties.

4 COMPLEXITY ANALYSIS

In this subsection, we provide a detailed complexity analysis of used baseline generators. Refer to Appendix C.3 and Table 10 of the main paper regarding the complexity.

Complexity of HyperPA [2]: HyperPA is an incremental generator that has its basis on preferential attachment. It samples a group of nodes proportional to the group degree and composes a hyperedge with the sampled group. Here, since there may exist $\binom{|V|}{|e|-1}$ number of possible candidates that can be sampled[2], the

[2] A newly introduced node should be included, and this results in sampling $|e| - 1$ number of nodes

time complexity of HyperPA for generating a size $|e|$ hyperedge is equivalent to $O(\log_2 \binom{|V|}{|e|})$.[3] In sum, the time complexity of HyperPA in generating a hypergraph with given statistics is equivalent to $O\left(\sum_{e\in E}\log_2\binom{|V|}{|e|}\right)$.

Regarding memory, it should save all subsets of each generated hyperedge. Thus, memory complexity of HyperPA is equivalent to $O\left(\sum_{e\in E} 2^{|e|}\right)$. Due to its exponential memory complexity, HyperPA often fails in replicating large-scale hypergraphs (see Table 7 of the main paper).

Complexity of HyperFF [5]: **HyperFF** is an incremental generator that selects an ambassador from the existing nodes and includes neighboring nodes of the ambassador in a stochastic way (spec., according to two hyperparameters p and q). In the worst case, to make a hyperedge with a size of $|e|$, it may visit every node as many times as $|e|$. Thus, the time complexity of HyperFF for generating a size $|e|$ hyperedge is equivalent to $O(|V| \times |e|)$. In sum, the time complexity of HyperFF in generating a hypergraph with given statistics is equivalent to $O\left(|V| \times \sum_{e\in E}|e|\right)$.

Regarding memory complexity, HyperFF only requires space for generated nodes and hyperedges. Thus, the memory complexity of HyperFF is equivalent to $O(|V| + \sum_{e\in E}|e|)$.

Complexity of HyperLap and HyperLap+ [6]: Regarding the time and memory complexity of HyperLap and HyperLap+, We follow the complexity analysis provided by Lee et al. [6].

REFERENCES

[1] Minyoung Choe, Jaemin Yoo, Geon Lee, Woonsung Baek, U Kang, and Kijung Shin. 2022. Midas: Representative sampling from real-world hypergraphs. In *WWW*.

[2] Manh Tuan Do, Se-eun Yoon, Bryan Hooi, and Kijung Shin. 2020. Structural patterns and generative models of real-world hypergraphs. In *KDD*.

[3] Suzanne Renick Gallagher and Debra S Goldberg. 2013. Clustering coefficients in protein interaction hypernetworks. In *BCB*.

[4] Steffen Klamt, Utz-Uwe Haus, and Fabian Theis. 2009. Hypergraphs and cellular networks. *Plos Computational Biology* 5, 5 (2009), e1000385.

[5] Jihoon Ko, Yunbum Kook, and Kijung Shin. 2022. Growth patterns and models of real-world hypergraphs. *Knowledge and Information Systems* 64, 11 (2022), 2883–2920.

[6] Geon Lee, Minyoung Choe, and Kijung Shin. 2021. How do hyperedges overlap in real-world hypergraphs?-patterns, measures, and generators. In *WWW*.

[7] Mark EJ Newman, Steven H Strogatz, and Duncan J Watts. 2001. Random graphs with arbitrary degree distributions and their applications. *Physical review E* 64, 2 (2001), 026118.

[8] Nickolay Smirnov. 1948. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics* 19, 2 (1948), 279–281.

[9] Leo Torres, Ann S Blevins, Danielle Bassett, and Tina Eliassi-Rad. 2021. The why, how, and when of representations for complex systems. *SIAM Rev.* 63, 3 (2021), 435–485.

[10] Stanley Wasserman, Katherine Faust, et al. 1994. Social network analysis: Methods and applications. (1994).

[3]Note that weighted sampling requires $O(\log_2 n)$.