Landry et al. *EPJ Data Science* (2025) 14:53
https://doi.org/10.1140/epjds/s13688-025-00557-7

EPJ Data Science
a SpringerOpen Journal

CORRECTION Open Access

Correction to: The simpliciality of higher-order networks

Nicholas W. Landry[1,2*], Jean-Gabriel Young[1,2] and Nicole Eikmeier[3]

The original article can be found online at https://doi.org/10.1140/epjds/s13688-024-00458-1

*Correspondence: nicholas.landry@uvm.edu
[1]Vermont Complex Systems Center, University of Vermont, 82 Innovation Pl, 05405 Burlington, USA
[2]Department of Mathematics and Statistics, University of Vermont, 82 Innovation Pl, 05405 Burlington, USA
Full list of author information is available at the end of the article

Correction to: EPJ Data Sci. 13, 17 (2024)
https://doi.org/10.1140/epjds/s13688-024-00458-1

Following publication of the original article [1], the authors identified errors in the equations in the 2.1 Measures section, Algorithm 2, 3, 4, Figure 1, 2, 3, 4, Table 1, 2 values and a value and wording in the 2.1 Measures, 3.3 Local measures of simpliciality section and the Acknowledgements section.

The incorrect and correct equations and wording in the 2.1 Measures section, Algorithm 2, 3, Table 1, 2, 3.3 Local measures of simpliciality section and Acknowledgements section are indicated hereafter.

The incorrect 2.1. sub-sections Edit Simpliciality and Face Edit Simpliciality:

Edit simpliciality The *edit simpliciality* (ES) is defined as the minimal number (or fraction, in the normalized case) of additional edges needed to make a hypergraph a simplicial complex.

Our formal definition uses the notion of an induced simplicial complex defined in Sect. 1.1. Given a hypergraph $\mathbf{H} = (V, E)$ for which we want to measure the ES, we find its maximal edges $\widetilde{E}$ and construct the simplicial complex $\mathbf{S} = (V, C)$ induced on $\mathbf{H}$, with $C = \bigcup_{e \in \widetilde{E}} \mathcal{P}(e)$. The edit simpliciality is then

$$\sigma_{\mathrm{ES}} = \frac{|E|}{|C|}, \tag{2}$$

again satisfying $\sigma_{\mathrm{ES}} \in [0, 1]$; see Fig. 1C. (We note that one can use the induced simplicial complex to define variants of the ES, e.g., a simplicial edit distance $d_{\mathrm{ES}} = |C| - |E|$ or a normalized distance $d_{\mathrm{NES}} = (|C| - |E|)/|C| = 1 - |E|/|C| = 1 - \sigma_{\mathrm{ES}}$.)

The ES answers a slightly different question than the SF does—it counts missing hyperedges that would make the dataset into a simplicial complex, rather than the edges that already satisfy downward closure. It thus offers a complementary, equally interpretable measure of simpliciality. However, the ES has the disadvantage of being sensitive to outliers, as a handful of large hyperedges with few inclusions will rapidly drive σ_{ES} towards

0. Indeed, a hyperedge of size m without any inclusion contributes one edge to $|E|$ but 2^m edges to $|C|$ in the denominator of Eq. (2).

Face edit simpliciality Finally, building upon the idea of edit simpliciality, we define a more localized notion of simpliciality, using the number of subfaces that must be added to the hypergraph to make a particular face a simplex.

Given a hyperedge e, the number of edges one must add to the hypergraph to make e a simplex is

$$d_{\mathrm{FES}}(e) = |\mathcal{P}(e)| - |c|,$$

where $c = \{f \in E \mid f \subseteq e\}$. We can think of this quantity as an edit distance, or *face edit distance*. We use this quantity to define an average

$$\bar{d}_{\mathrm{FES}} = \frac{1}{|F|}\sum_{e\in F} d_{\mathrm{FES}}(e),$$

where F is a set of edges—most commonly, $F = \widetilde{E}$ or E. We exclusively use $F = \widetilde{E}$ in this study. These quantities are on the scale of counts, and to define quantities analogous to previous simpliciality measures, we thus introduce a per-face normalization, either on a distance scale (meaning that the quantity grows as the dataset becomes less simplicial):

$$\bar{d}_{\mathrm{NFES}} = \frac{1}{|F|}\sum_{e\in F} \frac{d_{\mathrm{FES}}(e)}{|\mathcal{P}(e)|},$$

or, similarly to previous definitions, on a simpliciality scale:

$$\sigma_{\mathrm{FES}} = \frac{1}{|F|}\sum_{e\in F}\left(1 - \frac{d_{\mathrm{FES}}(e)}{\mathcal{P}(e)}\right). \tag{3}$$

We call this last measure the *face edit simpliciality* (FES).

The FES normalizes the face edit distance as a fraction of its maximal simpliciality. This normalization removes the dominance of large edges in the calculation of σ_{ES} and, in fact, exponentially down-weights the contribution of these edges. In addition, because this metric is computed on faces, this is an averaged local metric.

The correct 2.1. sub-sections Edit Simpliciality and Face Edit Simpliciality:

Edit simpliciality The *edit simpliciality* (ES) is defined as the minimal number (or fraction, in the normalized case) of additional edges needed to make a hypergraph a simplicial complex.

Our formal definition uses the notion of an induced simplicial complex defined in Sect. 1.1. Given a hypergraph $\mathbf{H} = (V, E)$ for which we want to measure the ES, we find its maximal edges $\widetilde{E}$ and construct the simplicial complex $\mathbf{S} = (V, C)$ induced on $\mathbf{H}$, with $C = \bigcup_{e\in\widetilde{E}} \mathcal{P}(e)$. The edit simpliciality is then

$$\sigma_{\mathrm{ES}} = \frac{|E| - |\widetilde{E}|}{|C| - |\widetilde{E}|}, \tag{2}$$

again satisfying $\sigma_{\mathrm{ES}} \in [0,1]$; see Fig. 1C. We subtract the number of maximal edges from the total number of edges so that the edit simpliciality is zero in the case of a hypergraph with no subfaces. (We note that one can use the induced simplicial complex to define variants of the ES, e.g., a simplicial edit distance $d_{\mathrm{ES}} = |C| - |E|$ or a normalized distance $d_{\mathrm{NES}} = (|C| - |E|)/(|C| - |\widetilde{E}|) = 1 - (|E| - |\widetilde{E}|)/(|C| - |\widetilde{E}|) = 1 - \sigma_{\mathrm{ES}}$.)

The ES answers a slightly different question than the SF does—it counts missing hyperedges that would make the dataset into a simplicial complex, rather than the edges that already satisfy downward closure. It thus offers a complementary, equally interpretable measure of simpliciality. However, the ES has the disadvantage of being sensitive to outliers, as a handful of large hyperedges with few inclusions will rapidly drive σ_{ES} towards 0. Indeed, a hyperedge of size m without any inclusion contributes one edge to $|E|$ but 2^m edges to $|C|$ in the denominator of Eq. (2).

Face edit simpliciality Finally, building upon the idea of edit simpliciality, we define a more localized notion of simpliciality, using the number of subfaces that must be added to the hypergraph to make a particular face a simplex.

Given a hyperedge e, the number of edges one must add to the hypergraph to make e a simplex is

$$d_{\mathrm{FES}}(e) = |\mathcal{P}(e)| - |c|,$$

where $c = \{f \in E \mid f \subseteq e\}$. We can think of this quantity as an edit distance, or *face edit distance*. We use this quantity to define an average

$$\bar{d}_{\mathrm{FES}} = \frac{1}{|F|} \sum_{e \in F} d_{\mathrm{FES}}(e),$$

where F is a set of edges—most commonly, $F = \widetilde{E}$ or E. We exclusively use $F = \widetilde{E}$ in this study. These quantities are on the scale of counts, and to define quantities analogous to previous simpliciality measures, we thus introduce a per-face normalization, either on a distance scale (meaning that the quantity grows as the dataset becomes less simplicial):

$$\bar{d}_{\mathrm{NFES}} = \frac{1}{|F|} \sum_{e \in F} \frac{d_{\mathrm{FES}}(e)}{|\mathcal{P}(e)| - 1},$$

or, similarly to previous definitions, on a simpliciality scale:

$$\sigma_{\mathrm{FES}} = \frac{1}{|F|} \sum_{e \in F} \left(1 - \frac{d_{\mathrm{FES}}(e)}{|\mathcal{P}(e)| - 1}\right). \tag{3}$$

We call this last measure the *face edit simpliciality* (FES). We subtract one in the denominators of both expressions so that when an edge has no subfaces, its normalized face edit distance is one.

The FES normalizes the face edit distance as a fraction of its maximal simpliciality. This normalization removes the dominance of large edges in the calculation of σ_{ES} and, in fact, exponentially down-weights the contribution of these edges. In addition, because this metric is computed on faces, this is an averaged local metric.

The incorrect Table 1:

Table 1 Properties of empirical datasets and their simpliciality. $|V|$, $|E|$, $\langle k\rangle$, $\langle s\rangle$, σ_{SF}, σ_{ES}, and σ_{FES} denote the number of nodes, the number of hyperedges, the mean degree, the mean edge size, the simplicial fraction (SF), edit simpliciality (ES), and the face edit simpliciality (FES), respectively

Dataset	$\lvert V\rvert$	$\lvert E\rvert$	$\langle k\rangle$	$\langle s\rangle$	σ_{SF}	σ_{ES}	σ_{FES}
Proximity datasets							
`contact-primary-school`	242	12,704	52.50	2.42	0.85	0.92	0.94
`contact-high-school`	327	7,818	23.91	2.33	0.81	0.93	0.92
`hospital-lyon`	75	1,824	24.32	2.43	0.91	0.95	0.97
Email datasets							
`email-enron`	143	1,442	10.08	2.97	0.31	0.05	0.50
`email-eu`	967	23,729	24.54	3.12	0.32	0.05	0.52
Biological datasets							
`diseasome`	516	314	0.61	3.00	0.00	0.05	0.04
`disgenenet`	1,982	760	0.38	5.14	0.00	0.00	0.01
`ndc-substances`	2,740	4,754	1.74	5.16	0.02	0.01	0.07
Other							
`congress-bills`	1,715	58,788	34.28	4.95	0.03	0.01	0.10
`tags-ask-ubuntu`	3,021	145,053	48.01	3.43	0.15	0.25	0.46

The correct Table 1:

Table 1 Properties of empirical datasets and their simpliciality. $|V|$, $|E|$, $\langle k\rangle$, $\langle s\rangle$, σ_{SF}, σ_{ES}, and σ_{FES} denote the number of nodes, the number of hyperedges, the mean degree, the mean edge size, the simplicial fraction (SF), edit simpliciality (ES), and the face edit simpliciality (FES), respectively

Dataset	$\lvert V\rvert$	$\lvert E\rvert$	$\langle k\rangle$	$\langle s\rangle$	σ_{SF}	σ_{ES}	σ_{FES}
Proximity datasets							
`contact-primary-school`	242	12,704	52.50	2.42	0.85	0.88	0.94
`contact-high-school`	327	7,818	23.91	2.33	0.81	0.91	0.92
`hospital-lyon`	75	1,824	24.32	2.43	0.91	0.94	0.97
Email datasets							
`email-enron`	143	1,442	10.08	2.97	0.31	0.04	0.50
`email-eu`	967	23,729	24.54	3.12	0.32	0.04	0.52
Biological datasets							
`diseasome`	516	314	0.61	3.00	0.00	0.02	0.04
`disgenenet`	1,982	760	0.38	5.14	0.00	0.00	0.01
`ndc-substances`	2,740	4,754	1.74	5.16	0.02	0.00	0.07
Other							
`congress-bills`	1,715	58,788	34.28	4.95	0.03	0.00	0.10
`tags-ask-ubuntu`	3,021	145,053	48.01	3.43	0.15	0.11	0.46

The incorrect 3.3 Local measures of simpliciality section:

The sentence currently reads:

(The correlation drops to $\rho = 0.69$ when comparing the SF and ES).

The sentence should read:

(The correlation drops to $\rho = 0.6$ when comparing the SF and ES).

The sentence currently reads:

For tags-ask-ubuntu, FES is weakly assortative, whereas the other two measures are weakly disassortative.

The sentence should read:

For tags-ask-ubuntu, FES is weakly assortative, SF is weakly disassortative, and ES is strongly disassortative.

The incorrect Table 2:

Table 2 The simplicial assortativity of each dataset filtered to only include interactions of sizes two and three for computational tractability

Dataset	ρ_{SF}	ρ_{ES}	ρ_{FES}
Proximity datasets			
`contact-primary-school`	0.15	0.15	0.14
`contact-high-school`	0.22	0.34	0.24
`hospital-lyon`	−0.02	−0.02	−0.01
Email datasets			
`email-enron`	0.29	0.29	0.24
`email-eu`	0.19	0.16	0.16
Biological datasets			
`ndc-substances`	0.57	0.65	0.72
`diseasome`	N/A	0.46	0.75
`disgenenet`	N/A	0.55	0.89
Other			
`congress-bills`	0.78	0.48	0.75
`tags-ask-ubuntu`	−0.03	−0.08	0.04

The correct Table 2:

Table 2 The simplicial assortativity of each dataset filtered to only include interactions of sizes two and three for computational tractability

Dataset	ρ_{SF}	ρ_{ES}	ρ_{FES}
Proximity datasets			
`contact-primary-school`	0.15	0.17	0.14
`contact-high-school`	0.22	0.37	0.24
`hospital-lyon`	−0.02	−0.01	−0.01
Email datasets			
`email-enron`	0.29	0.29	0.24
`email-eu`	0.19	0.16	0.16
Biological datasets			
`ndc-substances`	0.56	0.54	0.69
`diseasome`	N/A	0.28	0.68
`disgenenet`	N/A	0.28	0.78
Other			
`congress-bills`	0.78	0.33	0.75
`tags-ask-ubuntu`	−0.03	−0.24	0.04

The incorrect Algorithm 2:

Algorithm 2: Exhaustive edit simpliciality

Input: K, a set of edge sizes
  m, the minimum acceptable simplex size
  $\mathbf{H} = (V, E)$, a hypergraph
  T, a trie constructed from the edges in $\mathbf{H}$
Output: σ_{ES}

```
σ_ES = 0
// Construct the set of maximal faces.
F = {e ∈ E | e ∉ f, ∀f ∈ E, |e| ≥ m}
// D stores the unique missing subfaces.
D = ∅, is a set of sets
// Iterate over all maximal faces.
for f ∈ F do
    // For each maximal face of the hypergraph, add all of
       its missing subfaces not already present in the
       global set of missing faces.
    for e ∈ P_K(f) do
        if e ∉ T then
            D ← D ∪ e
        end
    end
end
// The number of edges in the minimal simplicial complex
   is the sum of the number of edges in the original
   hypergraph and the number of missing subfaces.
σ_ES = |E|/(|E| + |D|)
return σ_ES
```

The correct Algorithm 2:

Algorithm 2: Exhaustive edit simpliciality

Input: K, a set of edge sizes
m, the minimum acceptable simplex size
$\mathbf{H} = (V, E)$, a hypergraph
T, a trie constructed from the edges in $\mathbf{H}$

Output: σ_{ES}

```
σ_ES = 0
E = {e ∈ E | |e| ∈ K, |e| > m}
// Construct the set of maximal faces.
F = {e ∈ E | e ∉ f, ∀f ∈ E}
// D stores the unique missing subfaces.
D = ∅, is a set of sets
// Iterate over all maximal faces.
for f ∈ F do
    // For each maximal face of the hypergraph, add all of
       its missing subfaces not already present in the
       global set of missing faces.
    for e ∈ P_K(f) do
        if e ∉ T then
            D ← D ∪ e
        end
    end
end
// The number of edges in the minimal simplicial complex
   is the sum of the number of edges in the original
   hypergraph and the number of missing subfaces.
σ_ES = (|E| − |Ẽ|)/(|E| − |Ẽ| + |D|)
return σ_ES
```

The incorrect Algorithm 3:

Algorithm 3: Memory-efficient edit simpliciality

Input: K, a set of edge sizes
  m, the minimum acceptable simplex size
  $\mathbf{H} = (V, E)$, a hypergraph
  T, a trie constructed from the edges in $\mathbf{H}$
Output: σ_{ES}
$\sigma_{\text{ES}} = 0$
`// Construct the set of maximal faces.`
$F = (e \in E \mid e \notin f, \forall f \in E, |e| \geq m)$
$d = 0$
`// Iterate over all enumerated maximal faces.`
for $i = 1 \ldots |F|$ **do**
  $f = F_i$
  `// First, calculate the number of missing faces for a given maximal face.`
  $\widetilde{d} = |\mathcal{P}_K(f)|$
  for $e \in \mathcal{P}_K(f)$ **do**
    if $e \in T$ **then**
      $\widetilde{d} \leftarrow \widetilde{d} - 1$
    end
  end
  `// Update the total number of missing subfaces`
  $d \leftarrow d + \widetilde{d}$
  `// Calculate the number of redundant missing subfaces counted for the maximal faces already seen. To prevent looping over all previous maximal edges, we iterate only over the previous maximal faces, which are also neighbors of the current maximal face.`
  $D = \emptyset$
  for $j = \{1 \ldots i-1\} \cap \{k \mid e_k \cap f \neq \emptyset\}$ **do**
    `// For each prior maximal face, we add the missing edges formed by the powerset of the intersection of that edge and the current maximal edge to the complete set of redundant missing edges.`
    $e = F_j$
    $g = e \cap f$
    for $h \in \mathcal{P}_{K \cup |g|}(g)$ **do**
      if $g \notin T$ **then**
        $D \leftarrow D \cup g$
      end
    end
  end
  `// Subtract the redundant missing subfaces`
  $d \leftarrow d - |D|$
end
$\sigma_{\text{ES}} = |E|/(|E| + d)$
return σ_{ES}

The correct Algorithm 3:

Algorithm 3: Memory-efficient edit simpliciality

Input: K, a set of edge sizes
m, the minimum acceptable simplex size
$\mathbf{H} = (V, E)$, a hypergraph
T, a trie constructed from the edges in $\mathbf{H}$
Output: σ_{ES}
$\sigma_{\text{ES}} = 0$
$E = \{e \in E \mid |e| \in K,\ |e| > m\}$
`// Construct the set of maximal faces.`
$F = \{e \in E \mid e \notin f,\ \forall f \in E\}$
$d = 0$
`// Iterate over all enumerated maximal faces.`
for $i = 1 \ldots |F|$ **do**
  $f = F_i$
  `// First, calculate the number of missing faces for a given maximal face.`
  $\widetilde{d} = |\mathcal{P}_K(f)|$
  for $e \in \mathcal{P}_K(f)$ **do**
    if $e \in T$ **then**
      $\widetilde{d} \leftarrow \widetilde{d} - 1$
    end
  end
  `// Update the total number of missing subfaces`
  $d \leftarrow d + \widetilde{d}$
  `// Calculate the number of redundant missing subfaces counted for the maximal faces already seen. To prevent looping over all previous maximal edges, we iterate only over the previous maximal faces, which are also neighbors of the current maximal face.`
  $D = \emptyset$
  for $j = \{1 \ldots i-1\} \cap \{k \mid e_k \cap f \neq \emptyset\}$ **do**
    `// For each prior maximal face, we add the missing edges formed by the powerset of the intersection of that edge and the current maximal edge to the complete set of redundant missing edges.`
    $e = F_j$
    $g = e \cap f$
    for $h \in \mathcal{P}_{K \cup |g|}(g)$ **do**
      if $g \notin T$ **then**
        $D \leftarrow D \cup g$
      end
    end
  end
  `// Subtract the redundant missing subfaces`
  $d \leftarrow d - |D|$
end
`// The number of edges in the minimal simplicial complex is the sum of the number of edges in the original hypergraph and the number of missing subfaces.`
$\sigma_{\text{ES}} = (|E| - |\widetilde{E}|)/(|E| - |\widetilde{E}| + d)$
return σ_{ES}

The incorrect Algorithm 4:

Algorithm 4: Face edit simpliciality

Input: K, a set of edge sizes
 m, the minimum acceptable simplex size
 $\mathbf{H} = (V, E)$, a hypergraph
 T, a trie constructed from the edges in $\mathbf{H}$

Output: σ_{FES}

```
σ_FES = 0
// Construct the set of maximal faces.
F = {e ∈ E | e ∉ f, ∀f ∈ E, |e| ≥ m}
σ_FES = 0
// Iterate over all maximal faces.
for f ∈ F do
    // For each maximal face, calculate the fraction of
       missing faces.
    s = 0
    for e ∈ P_K(f) do
        if e ∈ T then
            s ← s + 1/|P_K(f)|
        end
    end
    // Update the running average.
    σ_FES ← σ_FES + s/|F|
end
return σ_FES
```

The correct Algorithm 4:

Algorithm 4: Face edit simpliciality

Input: K, a set of edge sizes
    m, the minimum acceptable simplex size
    $\mathbf{H} = (V, E)$, a hypergraph
    T, a trie constructed from the edges in $\mathbf{H}$
Output: σ_{FES}
$\sigma_{\text{FES}} = 0$
$E = \{e \in E \mid |e| \in K,\ |e| > m\}$
`// Construct the set of maximal faces.`
$F = \{e \in E \mid e \notin f,\ \forall f \in E\}$
$\sigma_{\text{FES}} = 0$
`// Iterate over all maximal faces.`
for $f \in F$ **do**
    `// For each maximal face, calculate the fraction of missing faces.`
    $s = 0$
    for $e \in \mathcal{P}_K(f)$ **do**
        if $e \in T$ **then**
            $s \leftarrow s + 1/(|\mathcal{P}_K(f)| - 1)$
        end
    end
    `// Update the running average.`
    $\sigma_{\text{FES}} \leftarrow \sigma_{\text{FES}} + s/|F|$
end
return σ_{FES}

The incorrect Figure 1:

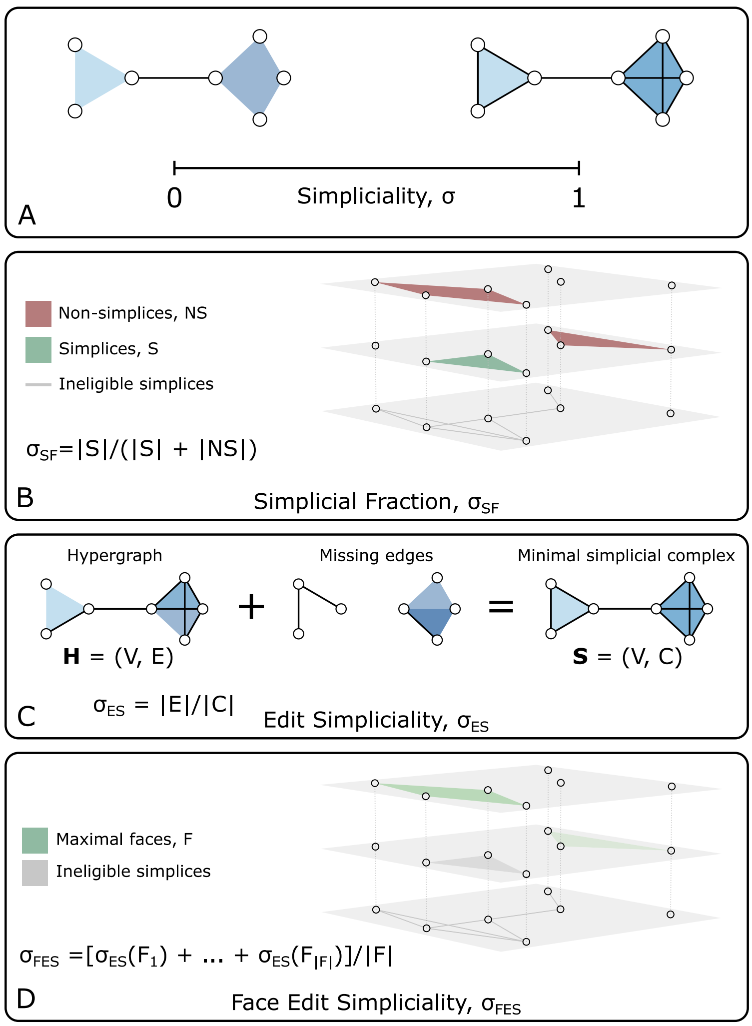

Figure 1 The Ego Network Model, with the names and expected sizes of each subgroup for social networks of humans

The correct Figure 1:

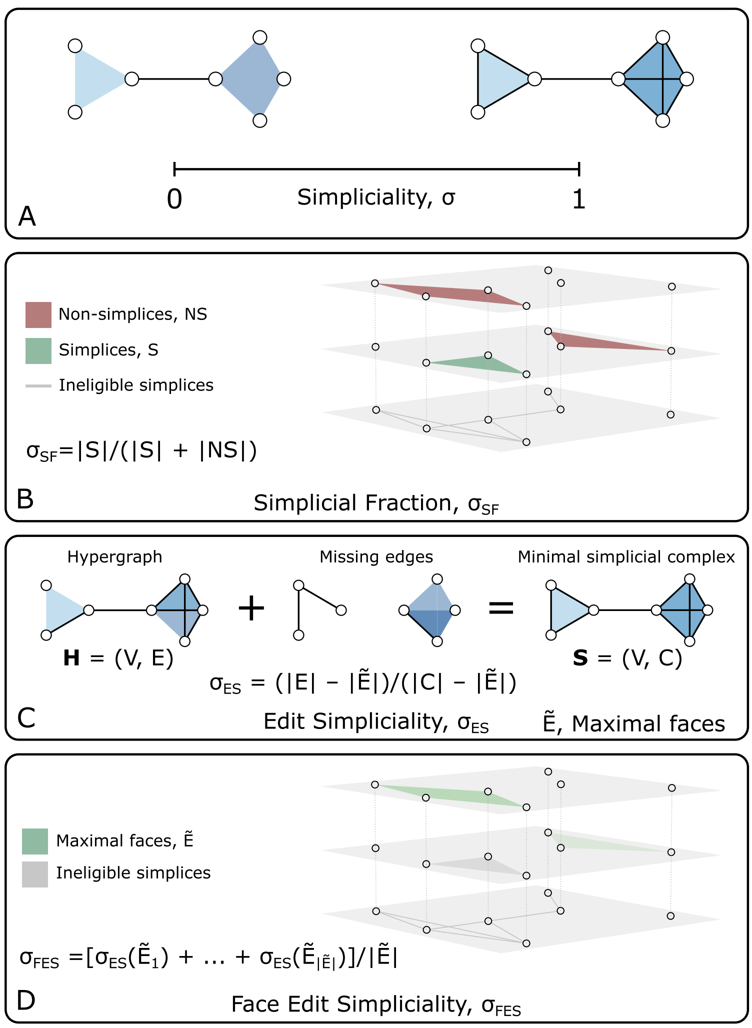

Figure 1 The Ego Network Model, with the names and expected sizes of each subgroup for social networks of humans

The incorrect Figure 2:

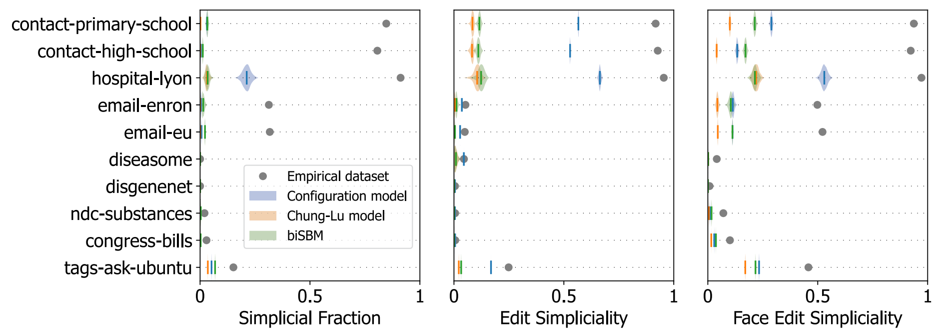

Figure 2 All four possible signed triads, as per Structural Balance Theory. The subscript number following the "T" corresponds to the number of positive connections for that triad

The correct Figure 2:

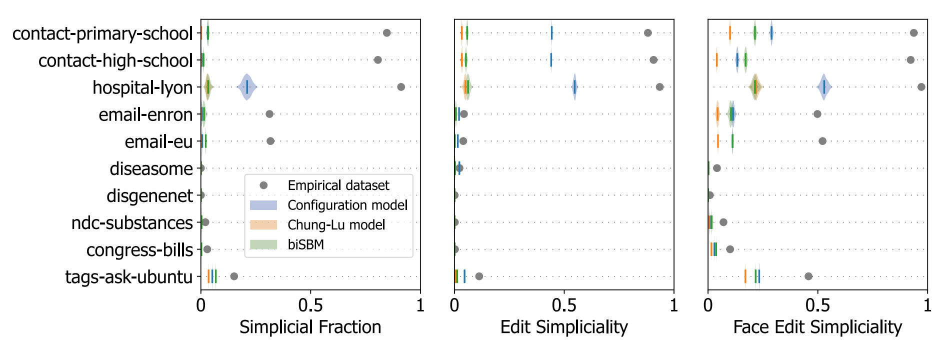

Figure 2 All four possible signed triads, as per Structural Balance Theory. The subscript number following the "T" corresponds to the number of positive connections for that triad

The incorrect Figure 3:

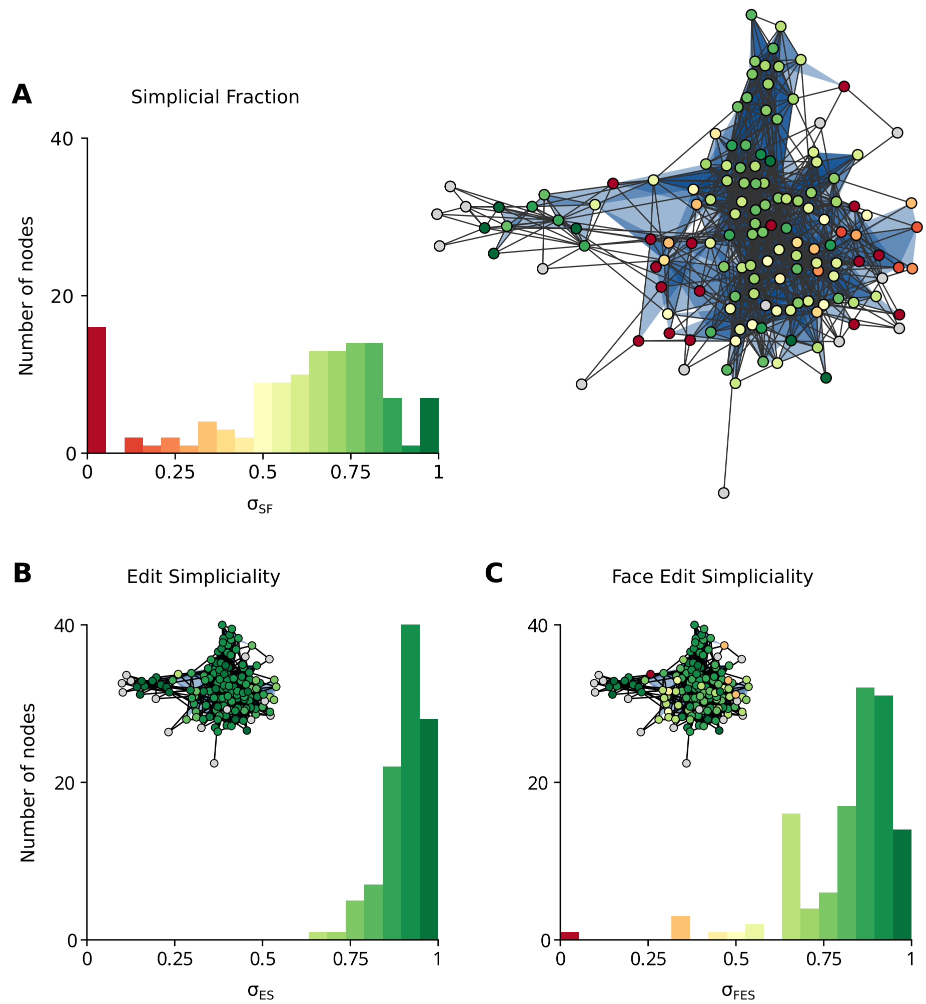

Figure 3 Percentages of positive, neutral and negative interaction labels estimated by each model (95% confidence intervals)

The correct Figure 3:

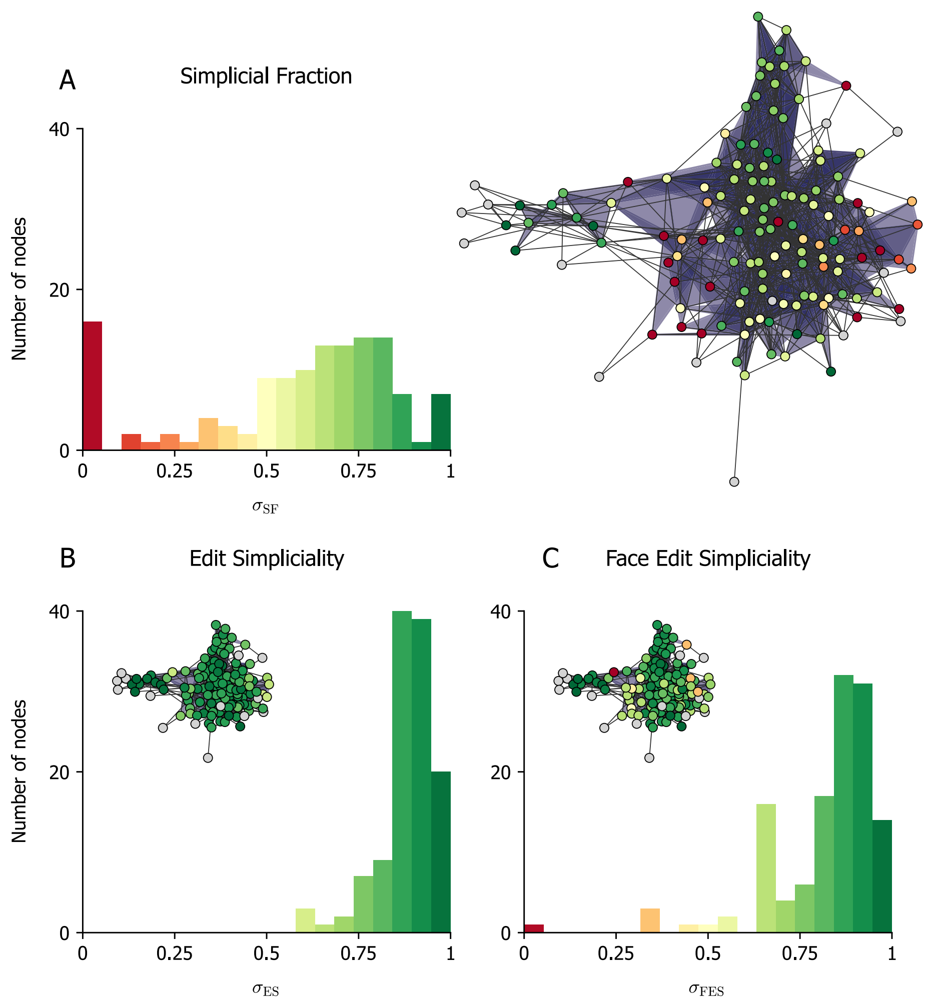

Figure 3 Percentages of positive, neutral and negative interaction labels estimated by each model (95% confidence intervals)

The incorrect Figure 4:

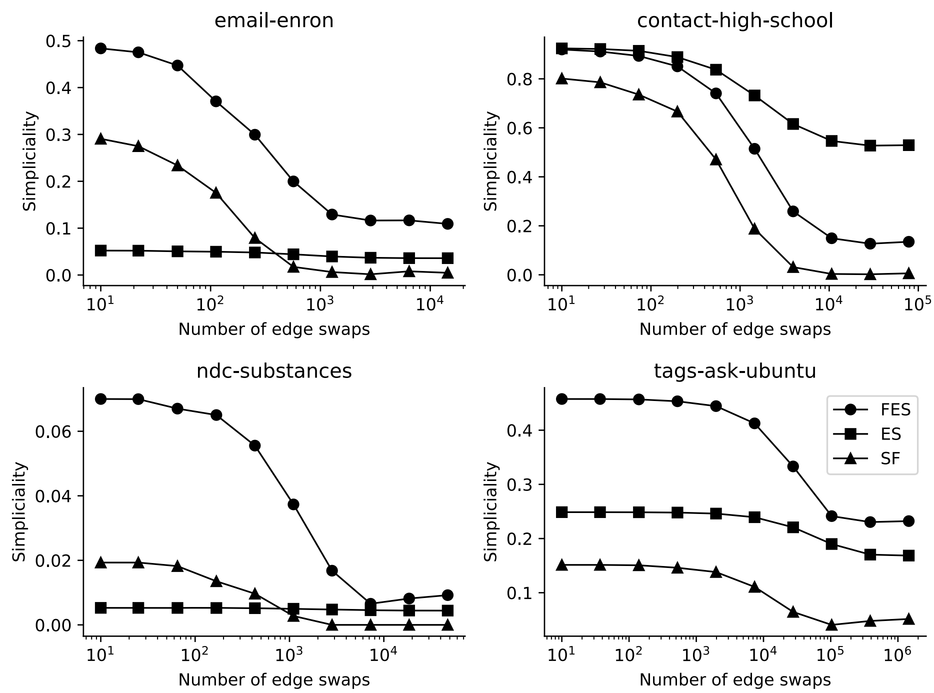

Figure 4 Percentages of positive and negative relationship labels estimated by each model (95% confidence intervals)

The correct Figure 4:

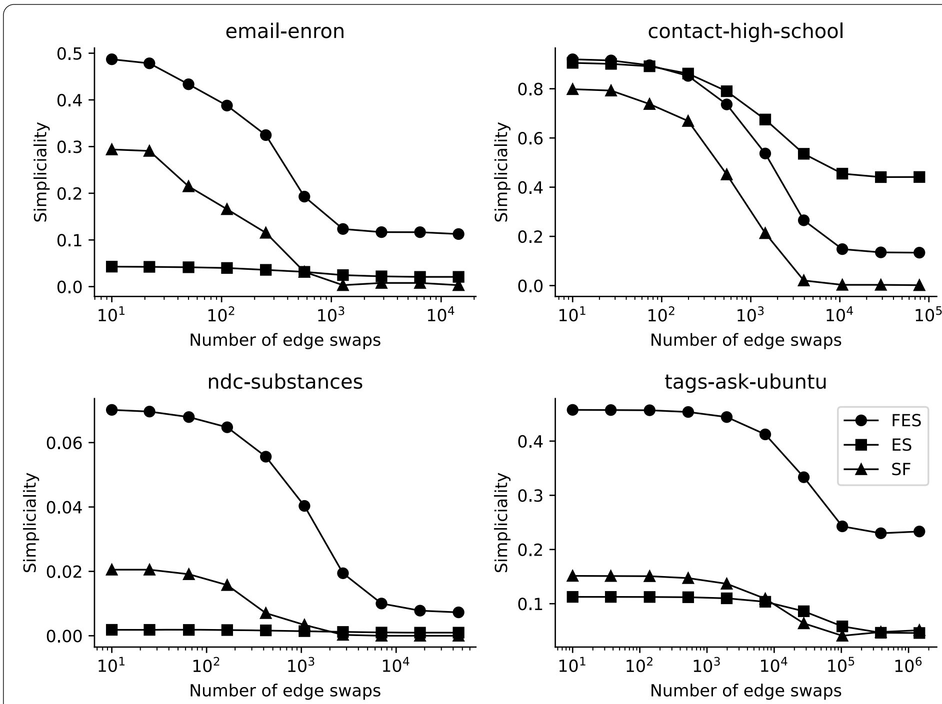

Figure 4 Percentages of positive and negative relationship labels estimated by each model (95% confidence intervals)

Ref. 63 is updated and now reads: Landry, N.: nwlandry/the-simpliciality-of-higher-order-networks: v0.3. https://doi.org/10.5281/zenodo.15707834

The incorrect Acknowledgements section reads:

N.W.L. would like to acknowledge the participants of the"Workshop on Modelling and Mining Complex Networks as Hypergraphs"at Toronto Metropolitan University and Tim LaRock for helpful conversations. N.W.L. would also like to thank Tzu-Chi Yen for lending his expertise on the biSBM inference.

The correct Acknowledgements section should read:

N.W.L. would like to acknowledge the participants of the"Workshop on Modelling and Mining Complex Networks as Hypergraphs"at Toronto Metropolitan University and Tim LaRock for helpful conversations. N.W.L. would also like to thank Tzu-Chi Yen for lending his expertise on the biSBM inference. We thank François Théberge for pointing out inconsistencies between the code and notation in the earlier versions of this manuscript.

Equations in the 2.1 Measures section, Algorithm 2, 3, 4, Figure 1, 2, 3, 4, Table 1, 2, 3.3 Local measures of simpliciality section and Acknowledgements section have been updated above and the original article [1] has been corrected.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Vermont Complex Systems Center, University of Vermont, 82 Innovation Pl, 05405 Burlington, USA. [2]Department of Mathematics and Statistics, University of Vermont, 82 Innovation Pl, 05405 Burlington, USA. [3]Department of Computer Science, Grinnell College, 1116 8th Ave, 50112 Grinnell, USA.

Published online: 22 July 2025

References
1. Landry NW, Young JG, Eikmeier N (2024) The simpliciality of higher-order networks. EPJ Data Sci 13:17. https://doi.org/10.1140/epjds/s13688-024-00458-1