



The frequency and position of stable associations offset their transitivity in a diversity of vertebrate social networks

Guillaume Péron

► To cite this version:

Guillaume Péron. The frequency and position of stable associations offset their transitivity in a diversity of vertebrate social networks. *Ethology*, 2023, 129 (1), pp.1-11. 10.1111/eth.13335. hal-04244309v1

HAL Id: hal-04244309

<https://hal.science/hal-04244309v1>

Submitted on 16 Oct 2023 (v1), last revised 5 Dec 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Research Article

The frequency and position of stable associations offset their transitivity in a diversity of vertebrate social networks

Péron Guillaume ^{1,*}

¹ CNRS, Université Lyon 1, 43 bd du 11 novembre 1918, 69622 VILLEURBANNE cedex, France; guillaume.peron@univ-lyon1.fr

* Correspondence: guillaume.peron@univ-lyon1.fr

Abstract: When the estimated strength of social associations corresponds to the proportion of time spent together, strong links, those that take up most of the recorded time of individuals, are compulsorily transitive and tend to occur in clusters. However, I describe three ways in which the frequency and position of strong associations apparently offset the expected transitivity of strong links in published association networks from 26 species of vertebrates. Instead of occurring in groups of three, strong links were mostly isolated. When they did occur in clusters, the clusters were small. The phenomena increased in intensity as the overall number of links of all strengths and the overall network transitivity increased. Since stable transitive motifs are beneficial to cooperation, these results can help explain why cooperative behaviors are not more frequent than they are in group-living vertebrates. Inversely, stable transitive motifs may be rare and small because the benefits of cooperation do not overcome the costs associated with these motifs. The summary statistics developed for this study captured information not conveyed by other network-level metrics; thus they may help quantify the socio-spatial structure of populations and potentially tease apart the environmental, species-specific, and individual drivers.

Keywords: clustering coefficient; graph theory; intransitivity; socio-spatial structure; peer of a peer; epidemiology; strength of weak ties

Significance statement:

- Stable transitive motifs are rare and small in 26 species of group-living vertebrates.
 - The article describes new network-level statistics of the frequency of strong links, the dissimilarity of link strengths within triangles, and the relative fragmentation of the subnetwork of strong links.
 - Results help explain why cooperation is not more frequent in group-living vertebrates
-

Introduction

Granovetter (1973) observed that strong friendships between humans tend to be transitive. In other words, when individual A has two best friends B and C, then B and C are also usually best friends. The proximate reasons for this transitivity include the “theory of cognitive balance” stating that good friends want their feelings towards third parties to be congruent (Heider, 1958), and “network homophily”, meaning that strong friendships tend to emerge from (and to further promote) similarity in cultural tastes etc. (McPherson et al., 2001). As a result, the networks feature clusters or cliques of individuals that are tightly-knit together. There are still weak ties between these cliques, which do not obey the same transitivity rules. These weak ties then perform a major role as the connections between, and the gateways into, the cliques: this is the “strength of weak ties paradigm” (Granovetter, 1973).

In animal studies, strong ties are also expected to be more transitive than weak ties, but firstly because of the way social ties are measured as the proportion of time that dyads spend together (Holekamp et al., 2012; Rubenstein et al., 2015). Indeed, if individual A spends most of its time with B and with C, then B and C also spend most of their time together. In other words, strong associations cannot be intransitive and should thus be more transitive on average than weak associations. The resulting stable clusters of closely-knit individuals promote the evolution of cooperative behaviors and represent a fitness advantage when cooperation is effective (Grinnell et al., 1995; Nowak, 2006; Silk et al., 2009; Teunissen et al., 2021). However, there are also costs to life in a tight group. Costs include the rapid spread of pathogens (Morrison et al., 2021), reproductive conflicts and other types of conflicts within the cliques (Datta, 1988; Holekamp et al., 2012), and impaired access to information during the periods when the weak links are not active (Artme et al., 2017). If these costs exceeded the benefits, the transitivity of strong links would operate as a constraint rather than an advantage.

The objective of the present paper is to apply several network statistics to decipher whether the transitivity of strong links is offset by the way strong links are distributed in animal association networks. For this purpose, I assembled a set of published animal association networks from 26 species (see methods). First, I verified that strong links were indeed more transitive than weak links. Second, I quantified three aspects of the distribution of link strength in the networks.

- (i) The network-level Gini coefficient of inequalities (Gini, 1936) (method section 1.5). This metric is an indicator of the overall rarity of strong links, i.e., the occurrence of a few very strong links amidst mostly weak links. For example, a group where mother-offspring bonds are much stronger than other types of associations would exhibit a high Gini coefficient if the offspring were few. The alternatives are that strong links are not very different from weak links, or that most of the links are strong.
- (ii) The triadic dissimilarity between the three links in each triangle (a new metric, see method section 1.6). This metric indicates whether strong links more often occur in groups of three, or in an iso-

lated fashion. For example, if breeding pairs travel together and they both avoid other pairs and evict single individuals from foraging spots (Black & Owen, 1989), strong links (breeding pairs) would be mostly isolated, leading to an excess of weak triplets closed by a single strong link (Péron, 2022).

- (iii) The fragmentation of the subnetworks of strong links (method section 1.7). This metric indicates whether, when strong links are not isolated, the clusters of strong links are relatively small or large. For example, if there is a core group of closely tied individuals amidst a cloud of individuals that spend most of their time alone, the subnetwork of strong links should appear less fragmented than the network as a whole. By contrast, if the network is made of family units in which recent offspring are more tightly linked to their mother than older offspring, then the subnetwork of strong links, made exclusively of the links between recent offspring and their mothers, would appear more fragmented than the network as a whole.

If strong links were rare (high Gini coefficient), isolated (high triadic dissimilarity), and if any cluster of strong links was small (high fragmentation of the subnetwork of strong links), then I concluded that the expected transitivity of strong links was in effect counter-balanced by the distribution of strong links in the network. In addition, I tested whether these patterns occurred more often or more intensely in networks that have many links and many transitive motifs compared to networks that have few links and few transitive motifs. If that was the case, this would suggest a functional response, i.e., a change in the probability to create specific patterns in the social network with a change in the availability of social partners. For example, the number of social partners could influence the perceived benefits obtained from stable partners, reinforce or weaken the effect of existing relationships on the probability to create new relationships, or the attraction exerted on bonded pairs by other bonded pairs. In practice, I first verified the expected transitivity of strong links. Next, I developed and applied the aforementioned three network-level statistics and correlated them to the edge density.

1. Material and methods

1.1. Definitions

Association networks (sometimes termed contact networks) refer to undirected social networks where connections occur through proximity between individuals. This definition excludes directed networks, such as dominance relationships, grooming, etc. The links (or ties, or edges) can be weighted by the relative dyadic association frequency (Holekamp et al., 2012; Rubenstein et al., 2015), hereafter termed the *link strength*.

Following previous authors (Sah et al., 2019), I considered three types of associations in this study: (i) *physical contact* or staying any amount of time within touching distance of associates, (ii) *close proximity*, a category in which I pooled nearest neighbor data and data documenting the time spent within a given radius of

associates (with cutoff distance and duration defined by the original authors of the source studies), and (iii) *shared group membership*, which mostly applies in a fission-fusion context.

Triplets correspond to situations where one individual A is associated with two different individuals B and C. In a *transitive* association network, most triplets are closed by a link between B and C, thereby forming a *triangle* (Granovetter, 1973; Newman et al., 2002). The unweighted *transitivity score* of an association network, hereafter denoted C_0 , corresponds to the proportion of closed triangles among all the triplets (Newman et al., 2002). Several adjacent triangles make a *transitive motif* or a transitive cluster or a clique. Examples of completely intransitive networks include grid-like and tree-like networks (Newman, 2008). Some researchers use the clustering coefficient instead of the transitivity score which differ from the transitivity score because the clustering coefficient averages an individual transitivity score whereas the transitivity score averages a triplet score (Barrat et al., 2004). Importantly, the interpretation of the transitivity score depends on whether the links are undirected, as is the case in this study, or directed. In dominance networks, the links are directed, and the transitivity score measures the linearity of the social hierarchy (McDonald & Shizuka, 2013). In association networks which are the topic of the present study, the links are not directed, and the transitivity score measures the frequency and size of transitive motifs.

The triplets can be also weighed according to the strength of the links inside of them, in order to generate a *weighted transitivity score* (Opsahl & Panzarasa, 2009), hereafter denoted C_1 . If C_1 is larger than C_0 , the probability that a triplet is closed increases with the strength of the links inside of it, and inversely if C_1 is smaller than C_0 then weak triplets are on average more likely to be closed than strong triplets (Opsahl & Panzarasa, 2009).

Another approach to the quantification of transitive clusters is to consider them as modules of individuals that interact more among themselves than with the rest of the network. Strong links would then correspond to within-module links and weak links would correspond to cross-module links. The network *modularity* quantifies how distinct these modules are, i.e., how rare and weak are the links between modules (Newman et al., 2002; Pons & Latapy, 2005). In practice, I delineated the modules using the short random walk community-finding algorithm (routine `cluster_walktrap` from `igraph`; Pons & Latapy, 2005). Except when explicitly stated otherwise, I took into account the weight of the links when delineating the modules (argument `weights = E(graph)$weight`). I then computed the modularity score following the usual formula (Newman, 2008).

Lastly the *edge density*, denoted D_0 , corresponds to the overall number of links in the network divided by the maximum possible number of links if all individuals were connected to each other. The edge density measures the overall probability that a link exists between any two individuals.

1.2. Literature search for network association data

This study uses published data only. Original data were collected in accordance with relevant institutional and national guidelines, as explained in the source articles listed in Table 1.

I focused on vertebrates that always or mostly forage in groups and/or always or mostly roost in groups, but without any criteria regarding the occurrence of cooperative behaviors within those groups. The data needed to be collected from free-ranging groups with naturally occurring kinship structure, and to document one of the aforementioned association types (see section 1.1). I first searched open data repositories: dryad.com and <https://bansallab.github.io/asnr/about.html> (“A social network repository” or ASNR; Sah et al., 2019) on Jan 5, 2022. Next, I used the search engine googlescholar.com with the keywords “animal” and “social network”. I searched the citation network of the first 60 hits upward and downward. The cutoff number 60 was chosen as a tradeoff between the risk of missing a poorly-cited study and the risk of oversight due to the sheer task at hand. If the title or abstract indicated that data corresponding to the above criterion existed, I elicited data sharing over email. This procedure yielded data from 26 species and three taxonomic classes.

These original studies varied in the way they quantified dyadic association rates. The most frequent method was to use the co-occurrence frequency relative to each associate’s own frequency in the dataset. Several studies however reported indexes derived from activity time budget analyses, and a few reported a discretized index of association strength. To standardize the link strengths across datasets, I rescaled the link strengths between 0 and 1 using a logit-transformation so that the median point between the weakest and strongest links of each network was attributed strength 0.5.

The seasonal timing of data collection was decided by the original authors. I did not select data according to that criterion.

1.3. Verifying the natural transitivity of strong links

I used two tests of the natural transitivity of strong links. First, I compared the unweighted transitivity score \mathcal{C}_0 and the weighted transitivity score \mathcal{C}_1 (Opsahl & Panzarasa, 2009). If strong links are more transitive than weak links, then triplets made of strong links are more likely to be closed than triplets made of weak links, and thus I expect $\mathcal{C}_1 > \mathcal{C}_0$.

Second, I manipulated the networks by removing an increasing proportion of the weakest links. I computed the quantity $b(Q) = \frac{\mathcal{C}_0(Q) - \mathcal{D}_0(Q)}{1 - \mathcal{D}_0(Q)}$ where Q is the proportion of remaining links and \mathcal{D}_0 is the edge density. If $\mathcal{D}_0(Q) = 1$, then also $\mathcal{C}_0(Q) = 1$ and $b(Q) = 1$. Note that this is the only section where \mathcal{C}_0 and \mathcal{D}_0 depend on Q . In the rest of the paper, I report the values corresponding to $Q = 100\%$, i.e., unmanipulated networks, meaning that $\mathcal{D}_0 = \mathcal{D}_0(100\%)$ and $\mathcal{C}_0 = \mathcal{C}_0(100\%)$.

The quantity $b(Q)$ takes value 0 under the null hypothesis that the manipulated network is not more transitive than expected from the overall probability of a link occurring anywhere in the manipulated network. Otherwise, $b(Q)$ varies between $-\frac{\mathcal{D}_0(Q)}{1 - \mathcal{D}_0(Q)}$ for a completely intransitive network and +1 as the upper

boundary value for completely transitive networks. I expected b to increase as Q decreased, i.e., as I removed links. I regressed the logit-transformed b against logit-transformed Q in interaction with the initial edge density, using a nonparametric spline model (function `gam` in R-package `mgcv`).

1.4. Cross-species regressions

In sections 1.5 to 1.7 below, I perform regressions between population-specific metrics. A potential source of spurious inference in such regressions is phylogenetic inertia, i.e., when the dependent variable is conserved along the phylogeny, and thereby pairs of closely related taxa essentially duplicate one another. If not corrected for, this mechanism can lead to over-estimate the effect sizes or the statistical significance of the results. Phylogenetic generalized models, which force the residuals of related species to be more similar than those of non related species, have become the default method to address the issue (Paradis et al., 2004). However, this statistical method requests adequate sample sizes to perform, and is only necessary if there is indeed a problematic phylogenetic structure in the data. The present study has 26 species for 23 genera, and only 2.4 species per taxonomic order on average (median: 1). In that situation, deep phylogenetic divergences expectedly take precedence. The potential bias can be either major if e.g., all mammals or all primates exhibited the same values, or undetectable if e.g., the variance between the 8 primates was larger than the variance across taxonomic orders. Given these constraints and predictions, and given that preliminary analyses indicated that the second scenario was most likely (Table 1), I replaced the correlated error structure by a simple random effect of the taxonomic order. Results were qualitatively similar if removing that effect altogether.

1.5. The rarity of strong links

I computed the network-level Gini coefficient of inequalities between link strengths as $G = \frac{1}{2n^2\bar{x}} \sum_{i=1}^n \sum_{j=1}^n |x_j - x_i|$ where the x_i are the link strengths and n is the total number of links. The null hypothesis that strong links are not rare was represented by the uniform distribution of link strength, which corresponds to $G = 0.33$. If a few strong links were much stronger than the bulk of the links, G would tend towards 1. If all links were of equal strength, $G = 0$. I tested the prediction that strong links should be rarer in dense and transitive networks using a linear model with the logit-transformed Gini index as dependent variable and the logit-transformed edge density of the networks as predictor.

1.6. Dissimilarity between the three links in each triadic relationship

The objective here is to compare the three link strengths in each triad. Hereafter, a triad is any three individuals with at least one non-zero link between two of the three individuals. Triads therefore include closed triangles, open triplets, but also “false triplets” corresponding to a dyad plus a disconnected individual.

The coefficient of variation is inadequate for this purpose because it captures the average pairwise difference, and not the variability between the pairwise differences. Many different combinations of three link strengths can yield the same coefficient of variation. Therefore I derived new triad-level dissimilarity scores. For each triad, I denoted a , b , and c the three link strengths in increasing order. They varied between 0 (no

link) and 1 (maximum recorded strength). I measured the dissimilarity between them with ω the rescaled sum of proportional pairwise differences, and g the triad-level Gini coefficient (Péron, 2022):

$$\text{Eq. 1a} \quad \omega = \frac{a-c}{a} + \frac{b-c}{a} - 1 = \frac{b-2c}{a}$$

$$\text{Eq. 1b} \quad g = \frac{1}{2} \frac{(a-b) + (b-c) + (a-c)}{a+b+c} = \frac{a-c}{a+b+c}$$

These formulae reduces the 3D space $\{a, b, c\}$ into a 2D space $\{\omega, G\}$. For ease of interpretation, I suggest to divide the $\{\omega, g\}$ space into four quadrants corresponding to different triad configurations (explained in Figure 3).

Next, I computed graph-level triadic dissimilarity scores based on ω and g . They are simply the average of ω and g over all the triads in the network, yielding the new network-level statistics \mathcal{C}_ω and \mathcal{C}_g .

$$\text{Eq. 2a} \quad \mathcal{C}_\omega = \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} \omega_t$$

$$\text{Eq. 2b} \quad \mathcal{C}_g = \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} g_t$$

The index t refers to the triad number (\tilde{T} triads in total). Eqs. 2a-2b are superficially similar to the weighed transitivity score (Opsahl & Panzarasa, 2009). However I replaced the triplet strength by the triplet dissimilarity scores. The code to compute \mathcal{C}_ω and \mathcal{C}_g from the adjacency matrix is provided (Appendix S1).

I then devised some permutation procedures to assess whether the observed dissimilarities exceeded the expectation at random (code in Appendix S1). *Test A*: I generated full random networks by drawing 10 groups with individual probability of occurrence proportional to their frequency in the original data and computing the link strength as the co-occurrence frequency in the 10 groups. Thus, the expected association strength between two individuals A and B was the frequency of A times the frequency of B. *Test B*: To represent the edge density of the original network, I removed some links at random. To do that, I drew an Erdős–Rényi graph with the same edge density as the original network. I then assigned link strengths line by line in the adjacency matrix, by drawing them from the observed link strengths departing from the focal individual. I then rescaled the link strengths so that the sum of the link strengths in each line was proportional to the observation frequency of the focal individual in the original dataset. This process generated a random matrix with an excess of zeros but the same magnitude of across- and within-row variation in link strength as the original data. Then I drew 100 groups by first drawing an individual and then its associates based on that matrix. From these 100 groups, I computed dyadic co-occurrence frequencies. These frequencies represented the final link strength of the simulation. Compared to the initial Erdős–Rényi graph, some links were created and some disappeared during the group drawing process. In both tests, I computed the Mahalanobis distance between the simulated and observed \mathcal{C}_ω and \mathcal{C}_g scores. The Mahalanobis distance M measures the distance between a point P (here, the two observed dissimilarity scores \mathcal{C}_ω and \mathcal{C}_g) and a distribution characterized

by a mean P_0 and a variance-covariance matrix S , according to the formula $M = \sqrt{(P - P_0)^T \cdot S^{-1} \cdot (P - P_0)}$. I computed P_0 and S from the cloud of simulated C_ω and C_g values. I then assessed the Mahalanobis distance against the chi-squared distribution with two degrees of freedom.

If test A was not significant, I concluded that the dissimilarities were explained by variation in individual frequencies in the dataset, i.e., the dissimilarities were not large. If test B was not significant, I concluded that the dissimilarities were large but mostly explained by missing links, rather than by the variation in the strength of existing links. If both tests A and B were significant, I concluded that there was either an excess of weak triplets closed by one strong link or an excess of balanced triangles made of three equally strong links, which an examination of the ω and g scores could decipher.

I also devised a test that focused on the lack of “forbidden triads” *sensu* Granovetter, i.e., strong triplets closed by a weak link. For this test C, I simply shuffled the link strengths while conserving other network aspects (Opsahl et al., 2008). I recomputed the ω and g scores after the shuffle. I used a chi-squared test with four degrees of freedom to determine whether the observed quadrant distribution differed from the expectation from shuffled networks. I expected a significant lack of triads in quadrant 3 (“forbidden triads”). This could also complete tests A and B by pointing out whether quadrant 1 (balanced triangles made of three equally strong links) or quadrant 2 (weak triplets closed by one strong link) were in excess.

I compared the new transitivity scores across studies using linear mixed models with the logit-transformed C_ω and C_g as dependent variables and the logit-transformed edge density as predictor. Because the occurrence of modules of individuals that interact more amongst themselves than with the rest of the network could be the cause of triad-level dissimilarities, I also considered the additive effect of the network modularity.

1.7. The fragmentation of the subnetwork of strong ties

After having delineated the aforementioned modules (see section 1.1), I computed the *fragmentation* score as the overall network size divided by the average size of the modules that contained more than one individual. If the score is high, the network is fragmented into a large number of small modules, and inversely. I used notation \mathcal{F}_w and \mathcal{F}_0 for the weighted and unweighted versions of the fragmentation score, respectively. \mathcal{F}_w is computed when the community-finding algorithm accounts for link strength and \mathcal{F}_0 is computed when it does not. I then computed the fragmentation ratio $\frac{\mathcal{F}_w - \mathcal{F}_0}{\mathcal{F}_0}$. The ratio quantifies how much more fragmented the subnetwork of strong links is, compared to the network as a whole (Appendix S2 for a simulation). A large ratio indicates that strong links occur in many small modules, e.g., there is more than one module of strong links in each module of weak links (a module can contain a single link and two individuals). A small ratio indicates that there is on average about one module of strong links embedded in each module of weak links. A negative ratio indicates that some of the modules do not feature any strong link. In other words, the frag-

mentation ratio $\frac{\mathcal{F}_w - \mathcal{F}_0}{\mathcal{F}_0}$ increases when the subnetwork made of only the stronger ties is more fragmented than the network as a whole. I predicted that this ratio should be positive and increase with the edge density. I tested these predictions using a linear model with the fragmentation ratio as the dependent variable and the logit-transformed edge density as predictor.

2. Results

2.1. Verifying the natural transitivity of strong links

The weighted transitivity score \mathcal{C}_1 was on average larger than the unweighted score \mathcal{C}_0 but only by 2% (\pm SD: 6%) (Table 1). This suggests that strong triplets were too rare to influence the computation of \mathcal{C}_1 (cf. next section).

In the second test, as I removed an increasing proportion of the weakest links, there was moderate evidence that the transitivity increased more than the edge density (Fig. 1a; McFadden's $r^2 = 0.06$). Statistically speaking the effect was however significant (likelihood ratio test against the intercept-only model: $df = 8.3$, deviance = 6.3, F-test $P < 0.001$). The evidence mostly came from networks of intermediate initial edge density (Fig. 1a: grey polygon; likelihood ratio test against the model without the interaction: $df = 7.7$, deviance = 5.8, $P < 0.001$, McFadden r^2 for the interaction = 0.04). The predicted increase in b from about 0.45 to 0.6 (grey curve) would translate in approximately a +0.08 gain in transitivity. Thus, the increase in transitivity with Q was moderate on average across studies. However, the increase was clear-cut in at least some of the studies (e.g., Fig. 1b).

2.2. The rarity of strong links

56% of the studies reported distributions of link strength that were more inegalitarian than a uniform distribution (Fig. 2). However, most of the dense networks exhibited inegalitarian properties, as the Gini coefficient increased with the edge density (Wald's $Z = 4.8$, McFadden's $r^2 = 0.3$, ANOVA: $P = 0.002$; Fig. 2). In other words, strong links were rarer in dense than sparse networks.

2.3. Dissimilarity between the three links in each triadic relationship

As the edge density increased, the networks dissimilarity scores went from quadrant 4 (mostly false triplets) to quadrant 1 (mostly balanced triangles) through quadrant 2 (mostly weak triplets closed by one strong link) while avoiding quadrant 3 (mostly strong triplets closed by one weak link; "forbidden triad") (Fig. 3). Triplets in quadrants 1 and 2 were more frequent and quadrant 3 was rarer than expected from shuffled networks (test C) (Fig. 3). There was strong evidence that the excess of quadrants 1 and 2 was due to the occurrence of distinct modules, because both the \mathcal{C}_w and \mathcal{C}_g score correlated with the network modularity, even after accounting for the edge density (respectively, $Z = -0.5$, $r^2 = 0.01$, $P = 0.001$ and $Z = 6.2$, $r^2 = 0.43$, $P < 0.001$; Table 1). The \mathcal{C}_w and \mathcal{C}_g scores nevertheless captured information neither conveyed by the modularity score nor by the usual transitivity score (Appendix S3).

In most studies, including all the studies that involved more than 35 individuals, both permutation tests A and B were positive (Table 1). This confirmed that the excess of quadrants 1 and 2 was not due to links missing at random or to variation in individual frequency of occurrence. Yet, in a few studies, the dissimilarities were as expected under one of the null models. In networks that were both small and dense, I found no evidence that the triads were any different from those of a full network (Table 1: *Poecila reticulata*, *Macaca assamensis*, *Macaca fuscata*: $P_A > 0.05$). In the proximity logs of barn swallows *Hirundo rustica*, of one of the subpopulations of elk *Cervus canadensis*, and in the Nilgiri langur *Trachypitecus johnii*, the triadic dissimilarities could apparently be created by links missing at random ($P_B > 0.05$).

2.4. The fragmentation of the subnetwork of strong ties

Accounting for link strength increased the fragmentation score on average by 35% (\pm SD 50%) (Fig. 4 and Fig. 1b: average fragmentation ratio 0.35). The fragmentation ratio increased with edge density ($Z = 2.3$, $r^2 = 0.09$, $P = 0.003$; Fig. 4). These results mean that most modules featured at least one strong link; and that within each module, strong links tended to be isolated or to form several small clusters rather than to occur as a single cluster of strong links per module. The phenomenon increased in intensity as the network edge density increased.

3. Discussion

By reanalyzing 26 published datasets of animal association networks, I first retrieved a major result of the field, namely that strong links are more transitive than weak links (Fig. 1). However, the effect was maybe not as strong as expected. My explanation is that, in many of the studied networks, even the strongest associations were not strong enough to mechanistically force transitivity. In addition, a few of the networks that I included feature minimal variation in link strength. Nevertheless, in the studies that were not affected by any of these two issues, the increase in transitivity with link strength was clear-cut (e.g., Fig. 1b). Next, I observed three ways in which the frequency and position of strong links offset their natural transitivity. First, strong links became rarer as the edge density and overall transitivity increased (Fig. 2). Second, the three links in each triangle were mostly dissimilar, more dissimilar than expected, and increasingly dissimilar as the edge density and overall transitivity increased (Fig. 3). Third, the subnetworks of strong links were increasingly fragmented as the edge density and the overall transitivity increased (Fig. 4). These results do not challenge the strength of weak ties paradigm (Granovetter, 1973). Indeed the paradigm described some of the networks very well (Fig. 1b). However, in many of the included studies, most ties were weak, making the strength of weak ties a somewhat tautological property.

These results may help explain why cooperative behaviors are not more widespread than they are among group-living vertebrates. First, cooperation requires stable relationships (Nowak, 2006; Teunissen et al., 2021). I found these to be increasingly rare as the overall gregariousness, the number of association partners, increased

(Fig. 2). This suggests that vertebrate social groups can either be dense or stable. Upper limits on the number of simultaneous partners, or a negative effect of existing relationships on the probability to create new ones, might be involved (Dávid-Barrett & Dunbar, 2013; Stadtfeld et al., 2020). Second, I found that stable transitive motifs were fewer and smaller than what they could have been. Compared to intransitive networks, stable transitive motifs offer more pathways for the benefits of collaboration to be collected, e.g., via indirect reciprocity (Block, 2015) and contributions towards common goods (Mielke et al., 2019). Stable transitive motifs might also facilitate the policing and coercion of selfish associates for the same reasons. The scarcity and small size of stable transitive motifs would not help cooperative behaviors to emerge. Overall, my observations help explain why many group-living vertebrates do not exhibit advanced cooperative behaviors, or why cooperative behaviors are rarely expressed. Note however the alternative interpretation that stable transitive motifs may be rare and small *because* cooperation does not bring enough benefits. In addition, as mentioned earlier, a functional response is also possible, in which individuals navigate their social environment in a way that is analogous to the way resource selection occurs in the physical environment (Holling, 1959). More precisely, the marginal benefit of an additional stable relationship might decrease with the number of pre-existing stable relationships, leading to a concave relationship between the number of potential partners (as proxied by the edge density) and the selection of specific partners. Formal tests would require dynamic data rather than aggregated data, so that a dynamic model of link creation, stability, and activation can be fitted, while taking individual attributes such as kin relationships into account (e.g. Snijders et al., 2010).

Importantly, the patterns that I report do not necessarily emerge from individual decisions alone. Environmental and demographic variation clearly contribute to network structure. For example, the fragmentation ratio changed from -0.3 to +2.1 between a dense subpopulation and a sparse subpopulation of elk *Cervus canadensis* (data: Webber & Vander Wal, 2020). On the other hand, the observed patterns can be created by relatively simple social behaviors. Kin-biased associations in particular, especially between mother and offspring, and associations between breeding male and female, can cause an excess of weak triplets closed by one strong link (Péron, 2022). This point is important because different species have different hard-wired socio-spatial structures, e.g., in most mammals, offspring form strong bonds with their mothers that may last past weaning age, whereas in birds the dominant bond may be between the male and female of a pair. In other words, the patterns that I report stem from a wide variety of mechanisms. Some are environmental constraints, some are species-specific evolved social strategies, and others represent individual reactions norms.

In terms of caveats, a major one is that the datasets are not representative of all vertebrate social systems. First, the literature is biased towards small group sizes, which are easier to monitor. Besides that, social network methods are not compulsorily used on the most social species. For example, the association network of lionesses *Panthera leo* can appear trivial because association is almost obligate within a pride and almost forbidden across prides. Yet, it is in this context that the strength of weak ties paradigm applies the most (Craft et

al., 2011), that the transitivity expectedly increases the most with link strength, and that cooperation is expectedly the most beneficial.

Another caveat is that the criteria to determine what constituted an association undoubtedly influenced the recorded network structures (Gazda et al., 2015). For example, the network drawn from the ritualized embraces of brown spider-monkeys *Ateles hybridus* (Rimbach et al., 2015) was one of the few that scored almost into quadrant 3 (the “forbidden triad”; Fig. 3). This specific type of association lasts less than a minute (Rimbach et al., 2015). This would expectedly relax the physical constraint that strong links are compulsorily transitive. This suggests that the triadic dissimilarity scores C_ω and C_g can perform as an indicator of the intransitivity of strong links. Ritualized embraces may obey a principle of preferential attachment to keystone individuals (Range & Noë, 2005; Schino, 2001), which would create strong triplets featuring a keystone in the middle, closed by weak or no links between the subordinates. The rudimentary permutation tests A and B are however not suited to formally test such hypotheses. Here also, formal tests would require dynamic data rather than aggregated data. Nevertheless, the statistics developed for this study captured information not summarized in other network-level metrics (Appendix S3). They can therefore help quantify the socio-spatial structure of different species and populations, the extent to which this structure favors the evolution of cooperation, and the social network response to changes that influence the costs and benefits of sociality.

Supplementary Materials: Appendix S1: R script to compute the dissimilarity-weighted transitivity scores and to perform the quadrant and distance tests. Appendix S2: Simulation study supporting the use of the fragmentation ratio to quantify whether the subnetwork of strong links is more fragmented than the network as a whole. Appendix S3: Principal component analysis.

Author Contributions: NA

Funding: This research received no external funding.

Ethical statement: This study uses published data only. Original data were collected in accordance with relevant institutional and national guidelines, as explained in the source articles listed in Table 1.

Data Availability Statement: This manuscript does not use new data. All data sources are listed in Table 1.

Acknowledgments: I thank all the people who contributed to the collection and curation of the association network data.

Conflicts of Interest: The authors declare no conflict of interest

References

- Aplin, L. M., Major, R. E., Davis, A., & Martin, J. M. (2021). A citizen science approach reveals long-term social network structure in an urban parrot, *Cacatua galerita*. *Journal of Animal Ecology*, 90(1), 222–232. <https://doi.org/10.1111/1365-2656.13295>
- Artime, O., Ramasco, J. J., & San Miguel, M. (2017). Dynamics on networks: Competition of temporal and topological correlations. *Scientific Reports*, 7, 41627. <https://doi.org/10.1038/srep41627>
- Barocas, A., Ilany, A., Koren, L., Kam, M., & Geffen, E. (2011). Variance in Centrality within Rock Hyrax Social Networks Predicts Adult Longevity. *PLoS ONE*, 6(7), e22375. <https://doi.org/10.1371/journal.pone.0022375>
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11), 3747–3752. <https://doi.org/10.1073/pnas.0400087101>
- Black, J. M., & Owen, M. (1989). Agonistic behaviour in barnacle goose flocks: assessment, investment and reproductive

- 397 success. *Animal Behaviour*, 37(2), 199–209. [https://doi.org/10.1016/0003-3472\(89\)90110-3](https://doi.org/10.1016/0003-3472(89)90110-3)
- 398 Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks*, 40, 163–173.
399 <https://doi.org/10.1016/j.socnet.2014.10.005>
- 400 Craft, M. E., Volz, E., Packer, C., & Meyers, L. A. (2011). Disease transmission in territorial populations: the small-world
401 network of Serengeti lions. *Journal of The Royal Society Interface*, 8(59), 776–786. <https://doi.org/10.1098/rsif.2010.0511>
- 402 Cross, P. C., Lloyd-Smith, J. O., Bowers, J. A., Hay, C. T., Hofmeyr, M., & Getz, W. M. (2004). Integrating association data
403 and disease dynamics in a social ungulate: Bovine tuberculosis in African buffalo in the Kruger National Park.
404 *Annales Zoologici Fennici*, 41(6), 879–892.
- 405 Datta, S. (1988). The acquisition of dominance among free-ranging rhesus monkey siblings. *Animal Behaviour*, 36(3),
406 754–772. [https://doi.org/10.1016/S0003-3472\(88\)80159-3](https://doi.org/10.1016/S0003-3472(88)80159-3)
- 407 Dávid-Barrett, T., & Dunbar, R. I. M. (2013). Processing power limits social group size: computational evidence for the
408 cognitive costs of sociality. *Proceedings. Biological Sciences*, 280(1765). <https://doi.org/10.1098/RSPB.2013.1151>
- 409 Farine, D. R., & Milburn, P. J. (2013). Social organisation of thornbill-dominated mixed-species flocks using social network
410 analysis. *Behavioral Ecology and Sociobiology*, 67(2), 321–330. <https://doi.org/10.1007/s00265-012-1452-y>
- 411 Franz, M., Altmann, J., & Alberts, S. C. (2015). Knockouts of high-ranking males have limited impact on baboon social
412 networks. *Current Zoology*, 61(1), 107–113. <https://doi.org/10.1093/czoolo/61.1.107>
- 413 Gazda, S., Iyer, S., Killingback, T., Connor, R., & Brault, S. (2015). The importance of delineating networks by activity type
414 in bottlenose dolphins (*Tursiops truncatus*) in Cedar Key, Florida. *Royal Society Open Science*, 2(3), 140263.
415 <https://doi.org/10.1098/rsos.140263>
- 416 Gini, C. (1936). On the Measure of Concentration with Special Reference to Income and Statistics. *Colorado College*
417 *Publication, Colorado Springs, General Series*, 208, 73–79.
- 418 Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360–1380.
- 419 Grant, T. R. (1973). Dominance and association among members of a captive and a free-ranging group of grey kangaroos
420 (*Macropus giganteus*). *Animal Behaviour*, 21(3), 449–456. [https://doi.org/10.1016/S0003-3472\(73\)80004-1](https://doi.org/10.1016/S0003-3472(73)80004-1)
- 421 Griffin, R. H., & Nunn, C. L. (2012). Community structure and the spread of infectious disease in primate social networks.
422 *Evolutionary Ecology*, 26(4), 779–800. <https://doi.org/10.1007/s10682-011-9526-2>
- 423 Grinnell, J., Packer, C., & Pusey, A. E. (1995). Cooperation in male lions: kinship, reciprocity or mutualism? *Animal*
424 *Behaviour*, 49(1), 95–105. [https://doi.org/10.1016/0003-3472\(95\)80157-X](https://doi.org/10.1016/0003-3472(95)80157-X)
- 425 Heider, F. (1958). The Psychology of Interpersonal Relations. In *The Psychology of Interpersonal Relations*. Wiley.
426 <https://doi.org/10.4324/9780203781159>
- 427 Holekamp, K. E., Smith, J. E., Strelhoff, C. C., Van Horn, R. C., & Watts, H. E. (2012). Society, demography and genetic
428 structure in the spotted hyena. *Molecular Ecology*, 21(3), 613–632. <https://doi.org/10.1111/j.1365-294X.2011.05240.x>
- 429 Holling, C. S. (1959). The Components of Predation as Revealed by a Study of Small-Mammal Predation of the European
430 Pine Sawfly. *The Canadian Entomologist*, 91(5), 293–320. <https://doi.org/10.4039/Ent91293-5>
- 431 Hunt, T. N., Allen, S. J., Bejder, L., & Parra, G. J. (2019). Assortative interactions revealed in a fission–fusion society of
432 Australian humpback dolphins. *Behavioral Ecology*, 30(4), 914–927. <https://doi.org/10.1093/beheco/arz029>
- 433 Krause, S., Wilson, A. D. M., Ramnarine, I. W., Herbert-Read, J. E., Clément, R. J. G., & Krause, J. (2017). Guppies occupy
434 consistent positions in social networks: mechanisms and consequences. *Behavioral Ecology*, 28(2), 429–438.
435 <https://doi.org/10.1093/BEHECO/ARW177>
- 436 Levin, I. I., Zonana, D. M., Fosdick, B. K., Song, S. J., Knight, R., & Safran, R. J. (2016). Stress response, gut microbial
437 diversity and sexual signals correlate with social interactions. *Biology Letters*, 12(6), 20160352.
438 <https://doi.org/10.1098/rsbl.2016.0352>
- 439 McDonald, D. B., & Shizuka, D. (2013). Comparative transitive and temporal orderliness in dominance networks.
440 *Behavioral Ecology*, 24(2), 511–520. <https://doi.org/10.1093/BEHECO/ARS192>
- 441 McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of*
442 *Sociology*, 27, 415–444.
- 443 Mielke, A., Crockford, C., & Wittig, R. M. (2019). Snake alarm calls as a public good in sooty mangabeys. *Animal Behaviour*,
444 158, 201–209. <https://doi.org/10.1016/j.ANBEHAV.2019.10.001>
- 445 Morrison, R. E., Mushimiyimana, Y., Stoinski, T. S., & Eckardt, W. (2021). Rapid transmission of respiratory infections
446 within but not between mountain gorilla groups. *Scientific Reports* 2021 11:1, 11(1), 1–12.
447 <https://doi.org/10.1038/s41598-021-98969-8>
- 448 Murphy, D., Mumby, H. S., & Henley, M. D. (2020). Age differences in the temporal stability of a male African elephant
449 (*Loxodonta africana*) social network. *Behavioral Ecology*, 31(1), 21–31. <https://doi.org/10.1093/BEHECO/ARZ152>
- 450 Newman, M. E. J. (2008). The mathematics of networks. In L. E. Blume & S. N. Durlauf (Eds.), *The New Palgrave*
451 *Encyclopedia of Economics* (pp. 312–334). Palgrave Macmillan.
- 452 Newman, M. E. J., Watts, D. J., & Strogatz, S. H. (2002). Random graph models of social networks. *Proceedings of the*
453 *National Academy of Sciences of the United States of America*, 99, 2566–2572.

- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563. https://doi.org/10.1126/SCIENCE.1133755/SUPPL_FILE/NOWAK.SOM.PDF
- Opsahl, T., Colizza, V., Panzarasa, P., & Ramasco, J. J. (2008). Prominence and Control: The Weighted Rich-Club Effect. *Physical Review Letters*, 101(16), 168702. <https://doi.org/10.1103/PhysRevLett.101.168702>
- Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks*, 31(2), 155–163. <https://doi.org/10.1016/j.socnet.2009.02.002>
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20(2), 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Péron, G. (2022). Transitivity scores that account for triadic edge weight similarity in undirected graphs. <https://doi.org/10.1101/2022.01.11.475816>
- Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3733 LNCS, 284–293. https://doi.org/10.1007/11569596_31
- Puga-Gonzalez, I., Ostner, J., Schülke, O., Sosa, S., Thierry, B., & Sueur, C. (2018). Mechanisms of reciprocity and diversity in social networks: a modeling and comparative approach. *Behavioral Ecology*, 29(3), 745–760. <https://doi.org/10.1093/beheco/ary034>
- Range, F., & Noë, R. (2005). Can simple rules account for the pattern of triadic interactions in juvenile and adult female sooty mangabeys? *Animal Behaviour*, 69(2), 445–452. <https://doi.org/10.1016/j.anbehav.2004.02.025>
- Rimbach, R., Bisanzio, D., Galvis, N., Link, A., Di Fiore, A., & Gillespie, T. R. (2015). Brown spider monkeys (*Ateles hybridus*): a model for differentiating the role of social networks and physical contact on parasite transmission dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1669), 20140110. <https://doi.org/10.1098/rstb.2014.0110>
- Rubenstein, D. I., Sundaresan, S. R., Fischhoff, I. R., Tantipathananandh, C., & Berger-Wolf, T. Y. (2015). Similar but Different: Dynamic Social Network Analysis Highlights Fundamental Differences between the Fission-Fusion Societies of Two Equid Species, the Onager and Grevy's Zebra. *PLOS ONE*, 10(10), e0138645. <https://doi.org/10.1371/journal.pone.0138645>
- Sah, P., Méndez, J. D., & Bansal, S. (2019). A multi-species repository of social networks. *Scientific Data*, 6(1), 1–6. <https://doi.org/10.1038/s41597-019-0056-z>
- Schino, G. (2001). Grooming, competition and social rank among female primates: A meta-analysis. *Animal Behaviour*, 62(2), 265–271. <https://doi.org/10.1006/anbe.2001.1750>
- Shizuka, D., Barve, S., Johnson, A. E., & Walters, E. L. (2022). Constructing social networks from automated telemetry data: A worked example using within- and across-group associations in cooperatively breeding birds. *Methods in Ecology and Evolution*, 13(1), 133–143. <https://doi.org/10.1111/2041-210X.13737>
- Shizuka, D., Chaine, A. S., Anderson, J., Johnson, O., Laursen, I. M., & Lyon, B. E. (2014). Across-year social stability shapes network structure in wintering migrant sparrows. *Ecology Letters*, 17(8), 998–1007. <https://doi.org/10.1111/ele.12304>
- Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., Wittig, R. M., Seyfarth, R. M., & Cheney, D. L. (2009). The benefits of social capital: close social bonds among female baboons enhance offspring survival. *Proceedings of the Royal Society B: Biological Sciences*, 276(1670), 3099–3104. <https://doi.org/10.1098/rspb.2009.0681>
- Smith, J. E., Gamboa, D. A., Spencer, J. M., Travenick, S. J., Ortiz, C. A., Hunter, R. D., & Sih, A. (2018). Split between two worlds: automated sensing reveals links between above- and belowground social networks in a free-living mammal. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1753), 20170249. <https://doi.org/10.1098/rstb.2017.0249>
- Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44–60. <https://doi.org/10.1016/j.socnet.2009.02.004>
- Stadtfeld, C., Takács, K., & Vörös, A. (2020). The Emergence and Stability of Groups in Social Networks. *Social Networks*, 60, 129–145. <https://doi.org/10.1016/J.SOCNET.2019.10.008>
- Teunissen, N., Kingma, S. A., Fan, M., Roast, M. J., & Peters, A. (2021). Context-dependent social benefits drive cooperative predator defense in a bird. *Current Biology*, 31(18), 4120–4126.e4. <https://doi.org/10.1016/J.CUB.2021.06.070>
- Webber, Q. M. R., & Vander Wal, E. (2020). Heterogeneity in social network connections is density-dependent: implications for disease dynamics in a gregarious ungulate. *Behavioral Ecology and Sociobiology*, 74(6), 77. <https://doi.org/10.1007/s00265-020-02860-x>
- Weiss, M. N., Franks, D. W., Balcomb, K. C., Ellifrit, D. K., Silk, M. J., Cant, M. A., & Croft, D. P. (2020). Modelling cetacean morbillivirus outbreaks in an endangered killer whale population. *Biological Conservation*, 242, 108398. <https://doi.org/10.1016/j.biocon.2019.108398>

511 Wolf, J. B. W., Mawdsley, D., Trillmich, F., & James, R. (2007). Social structure in a colonial mammal: unravelling hidden
512 structural layers and their foundations by network analysis. *Animal Behaviour*, 74(5), 1293–1302.
513 <https://doi.org/10.1016/j.anbehav.2007.02.024>
514

Tables

Table 1: Overall results of the weighted network analysis for the 26 species. Given are the taxonomic order, the type of association (ty = GM: shared subgroup membership, PR: spatial proximity, CO: physical contact), the number of study individuals N , the edge density \mathcal{D}_0 , the modularity \mathcal{M} , the variance in rescaled edge weights Var , the Gini coefficient of inequality in link strength G , the fragmentation ratio FR quantifying the fragmentation of the subnetwork of strong links relative to the network as a whole, the unweighted transitivity coefficient \mathcal{C}_0 , the transitivity coefficient weighted for triplet strength \mathcal{C}_1 , and the triadic dissimilarity scores \mathcal{C}_ω and \mathcal{C}_g . “q.” gives the quadrants that were more frequent than expected in test C. “Ref.” is the original study and “ASNR” is the reference number in “A social network repository” (Sah et al., 2019; <https://bansallab.github.io/asnr/about.html>).

Figure legends

Fig. 1: (a) Model predictions illustrating that, when an increasing proportion of the weakest links is removed (x-axis), the network transitivity \mathcal{C}_0 is increasingly larger than expected from the edge density \mathcal{D}_0 (y-axis). (b) Raw data for a single population. The data document group membership in a killer whale (*Orcinus orca*) population (Weiss et al. 2020). Thick black lines indicate the 10% strongest links.

Fig. 2: The Gini coefficient (y-axis) increases with the edge density \mathcal{D}_0 of the network (x-axis). The color scale corresponds to the transitivity score \mathcal{C}_0 . The Gini coefficient is a measure of dispersion based on the pairwise differences in link strength among all the connected dyads in the network. The highest Gini coefficient in the present study (acorn woodpecker *Melanerpes formicivorus*; data: Shizuka et al., 2022) means that 95% of the recorded associations occurred between 5% of the connected dyads. The dashed line corresponds to the null hypothesis that the distribution of link strength is uniform.

Fig. 3: (a) Graph-level triadic dissimilarity scores \mathcal{C}_g and \mathcal{C}_ω in 26 species (see methods for the definition). The color scale corresponds to the edge density \mathcal{D}_0 . The grey outline delineates the domain of possible (ω, g) values. The pictograms represent the six extreme cases of triads. *Triplets* correspond to triads where one individual A is associated with two different individuals B and C. If the triplet is closed by a link between B and C, it is called a *triangle*. If the three links in the triangle are equally strong, the triad falls in the lowerleft corner of the plot. Different triangular configurations correspond to different other sections of the plot, a delineated by the four quadrants 1 to 4. A *false triplet* is a triad made of a dyad plus a disconnected singleton, corresponding to the rightmost part of the plot. (b) Triad-level dissimilarity scores summarized over two categories of edge density and over the four quadrants. The black outline indicates the expected distribution from test C, demonstrating the lack of “forbidden triads” (quadrant 3) and the excess of triangles where the closing link is strong (quadrant 1 and quadrant 2).

Fig. 4: The fragmentation ratio (y-axis) increases with the edge density \mathcal{D}_0 of the network (x-axis). The color scale corresponds to the transitivity score \mathcal{C}_0 . The fragmentation ratio is the proportional difference between the weighted and unweighted fragmentation scores. The highest fragmentation ratio in the present study (a low density population of elk *Cervus canadensis*; data: Webber & Vander Wal, 2020) means that the weighted network was 3 times more fragmented than the unweighted network.