Supplemental Information: The temporal dynamics of group interactions in higher-order social networks

Iacopo Iacopini[1,2,3,*], **Márton Karsai**[3,4], **and Alain Barrat**[5]

[1]Network Science Institute, Northeastern University London, London, E1W 1LP, United Kingdom
[2]Department of Physics, Northeastern University, Boston, MA 02115, USA
[3]Department of Network and Data Science, Central European University, A-1100 Vienna, Austria
[4]National Laboratory for Health Security, HUN-REN Alfréd Rényi Institute of Mathematics, 1053 Budapest, Hungary
[5]Aix Marseille Univ, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, Marseille, France
[*]corresponding author: iacopo.iacopini@nulondon.ac.uk

Contents

Supplementary Figures

List of Tables

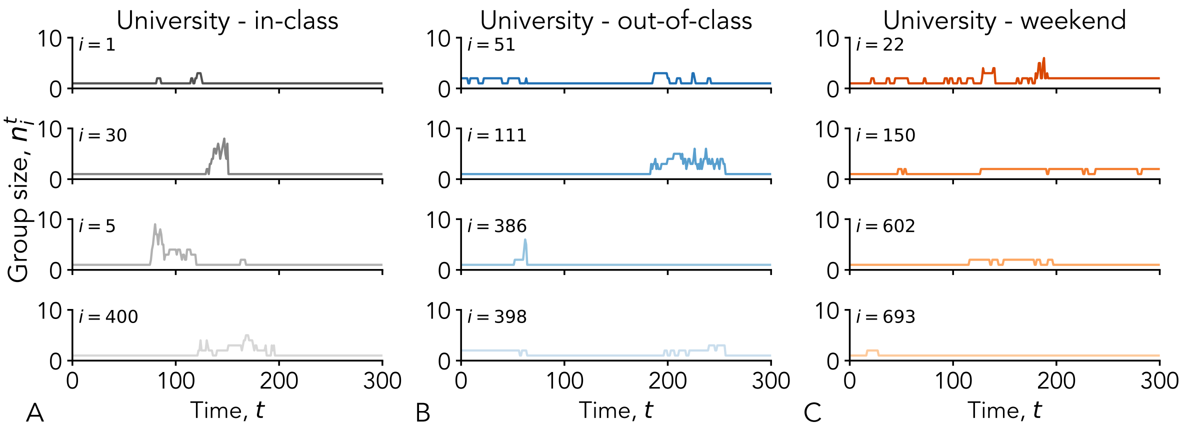

Figure S1. Example of timelines for group participation in CNS data divided by context of interaction: in-class (*A*), out-of-class (*B*), and weekend (*C*). Each panel shows a trajectory across different values of group size n_i^t as experienced by a node i at time t.

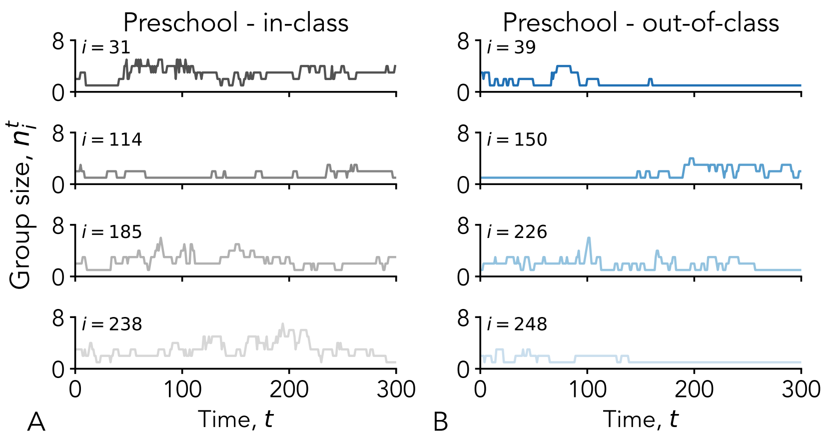

Figure S2. Example of timelines for group participation in DyLNet data divided by context of interaction: in-class (*A*), and weekend (*B*). Each panel shows a trajectory across different values of group size n_i^t as experienced by a node i at time t.

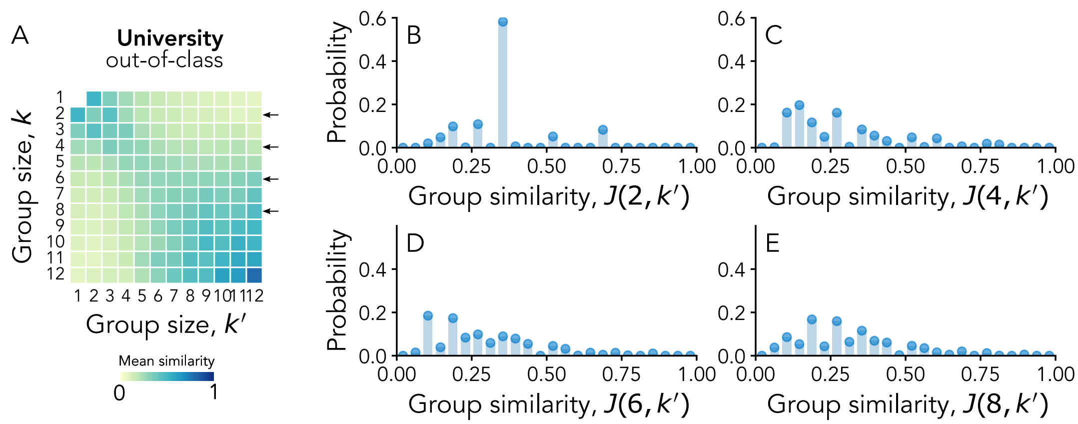

Figure S3. Similarity between successive groups of nodes that undergo a group change for CNS data, out-of-class interactions. (*A*) Heatmap of Jaccard similarity between successive groups of a given node, conditioned to the presence of a group change between a group of size k at time t and a group of size k' at time $t+1$. Each value of the matrix represents the similarity averaged across all nodes given two fixed values of k and k'. Full distributions (at fixed k but using all possible values of k') for four selected group sizes k are shown in the other panels: $k=2$ (*B*), $k=4$ (*C*), $k=6$ (*D*), and $k=8$ (*E*).

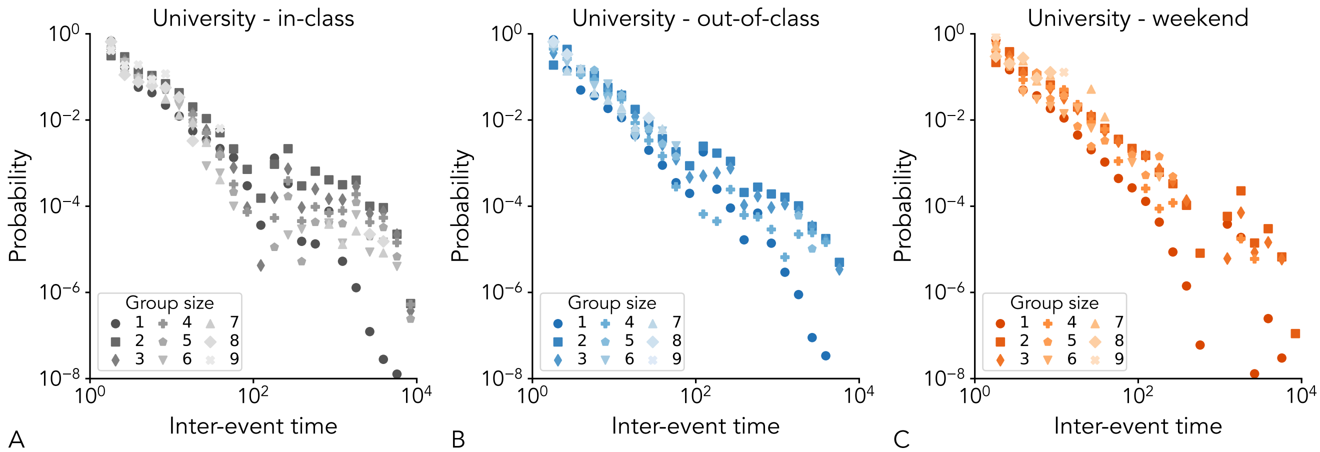

Figure S4. Distributions of group inter-event times (time between consecutive appearances of the same group) in the CNS data set for different contexts of interaction: in-class (*A*), out-of-class (*B*), and weekend (*C*). In each panel, different symbols correspond to different group sizes.

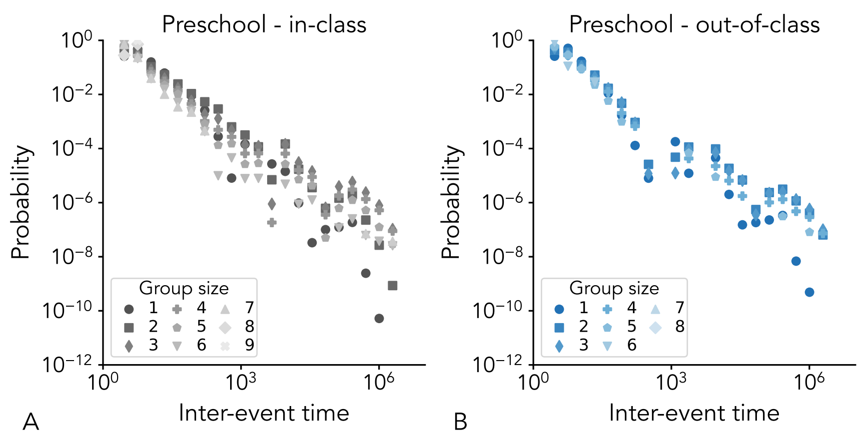

Figure S5. Distributions of group inter-event times (time between consecutive appearances of the same group) in the DyLNet data set for different contexts of interaction: in-class (*A*), and out-of-class (*B*). In each panel, different symbols correspond to different group sizes.

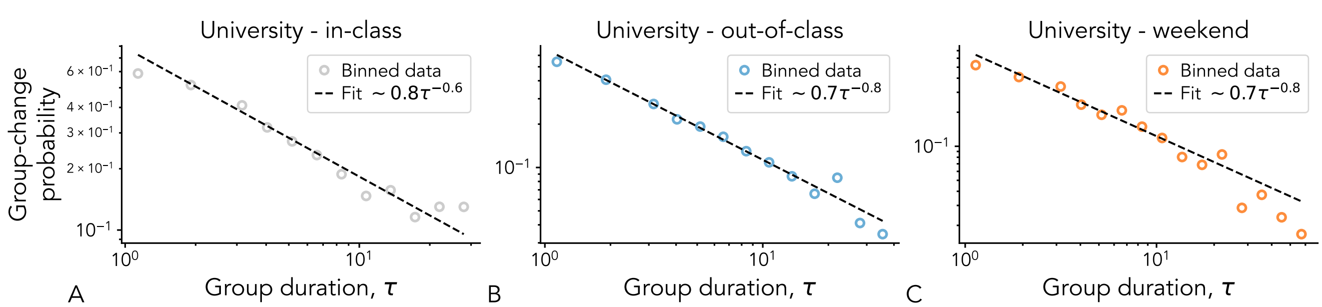

Figure S6. Fitting the group-change probability from the CNS data set by context of interaction: in-class (*A*), out-of-class (*B*), and weekend (*C*). We plot the probability for a node belonging to a group of any size to leave it for a different one after exactly τ timestamps. Points are binned empirical results, dashed lines represent a power-law fit of the form $b\tau^{\beta}$ —values reported in each panel. Compared to Fig. S8, group sizes have been aggregated up to size $k = 4$.

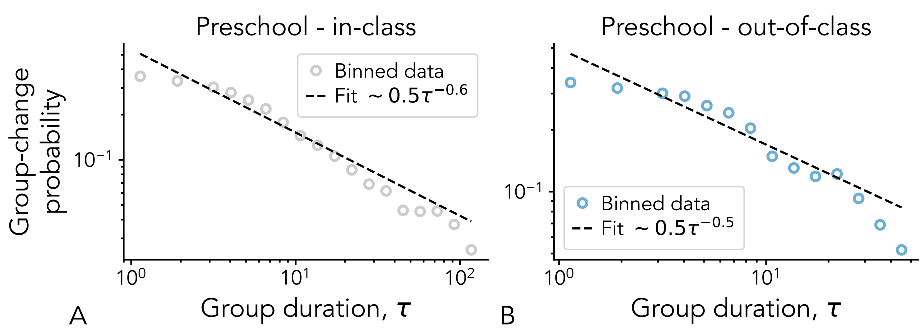

Figure S7. Fitting the group-change probability from the DyLNet data set by context of interaction: in-class (*A*) and out-of-class (*B*). We plot the probability for a node belonging to a group of any size to leave it for a different one after exactly τ timestamps. Points are binned empirical results, dashed lines represent a power-law fit of the form $b\tau^{\beta}$ —values reported in each panel. Compared to Fig. S9, group sizes have been aggregated up to size $k = 4$.

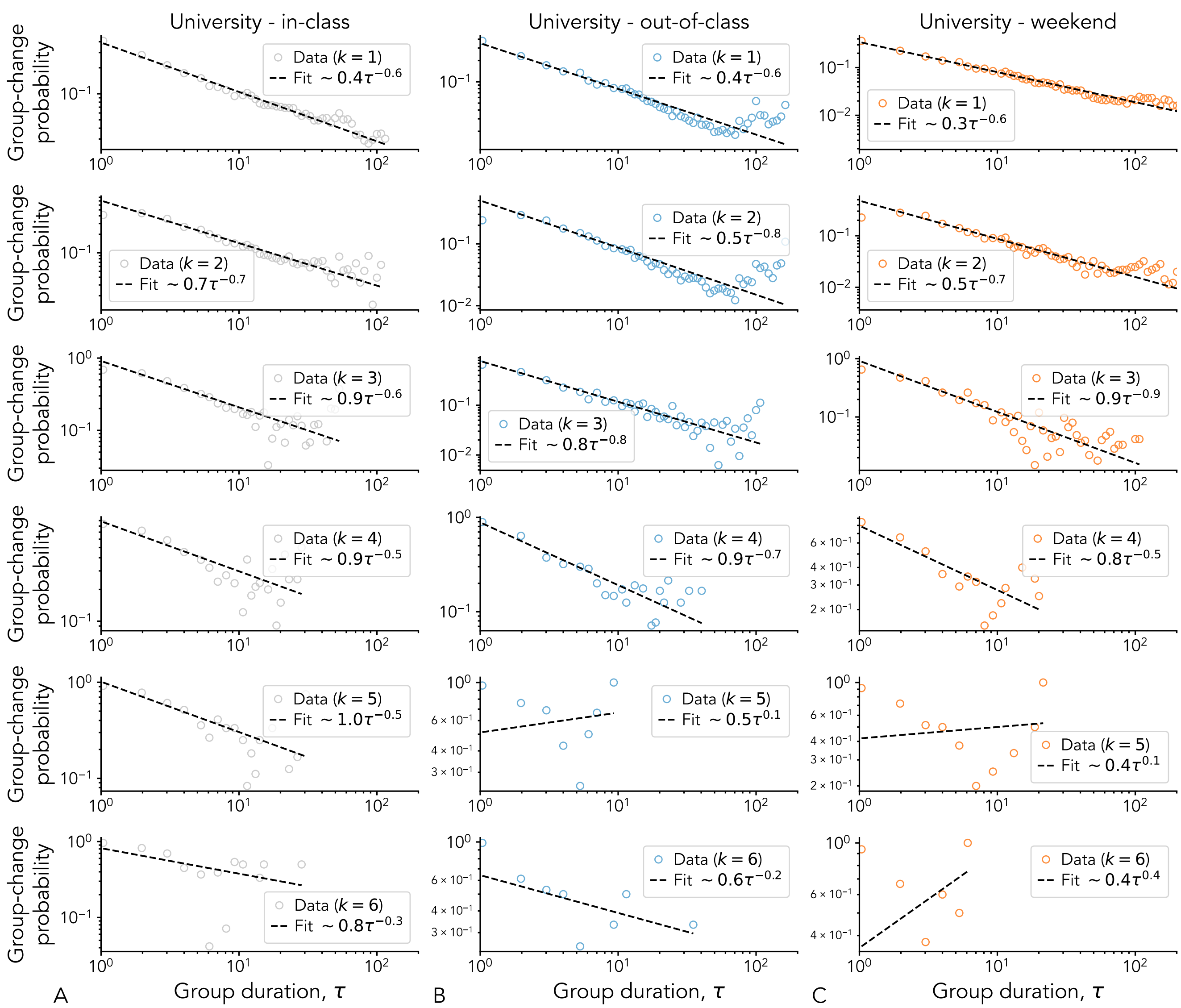

Figure S8. Fitting the group-change probability from the CNS data set by group size k and by context of interaction: in-class (A), out-of-class (B), and weekend (C). Each row corresponds to a different size k of the group interaction. We plot the probability for a node belonging to a group of size k to leave it for a different one after exactly τ timestamps. Points are binned empirical results, dashed lines represent a power-law fit of the form $b\tau^{\beta}$ —values reported in each panel.

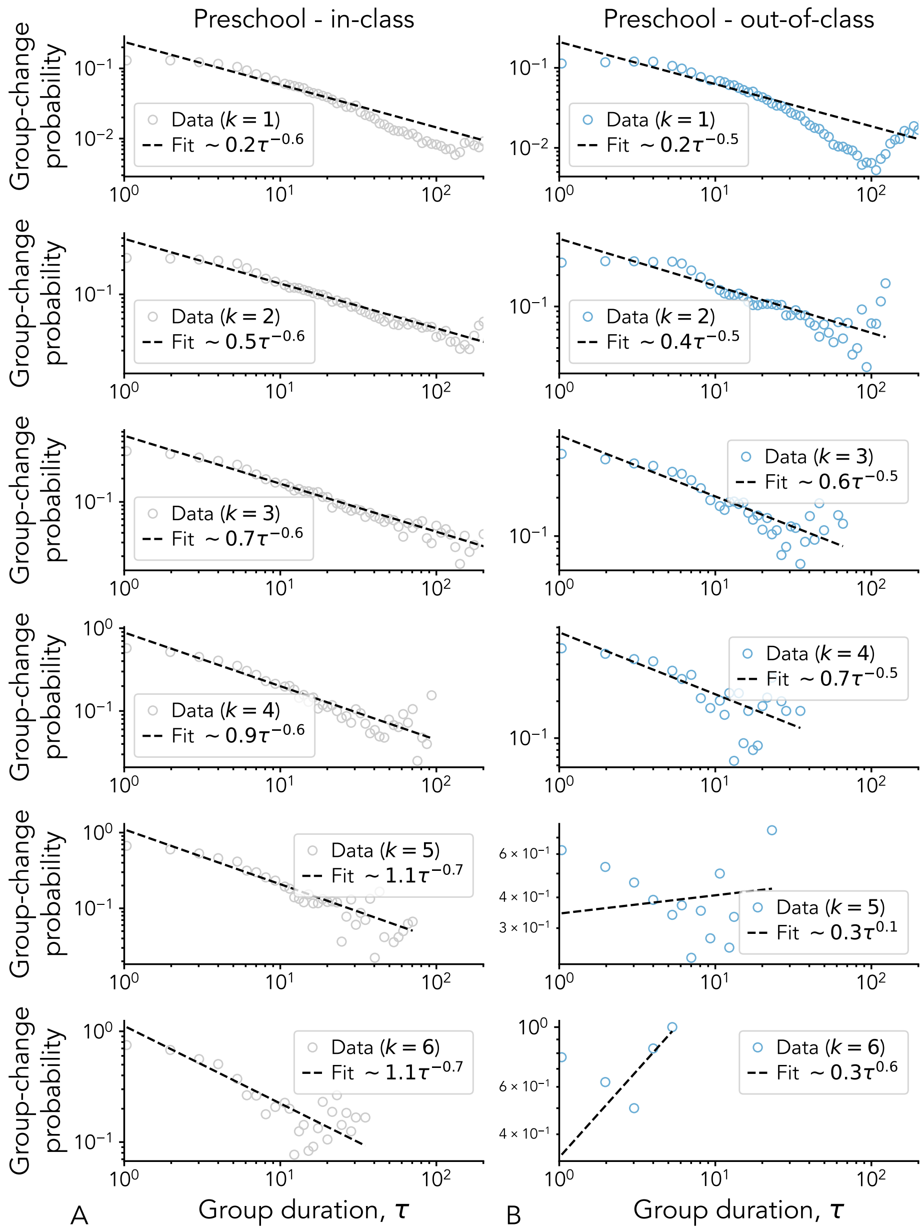

Figure S9. Fitting the group-change probability from the DyLNet data set by group size k and by context of interaction: in-class (*A*) and out-of-class (*B*). Each row corresponds to a different size k of the group interaction. We plot the probability for a node belonging to a group of size k to leave it for a different one after exactly τ timestamps. Points are binned empirical results, dashed lines represent a power-law fit of the form $b\tau^{\beta}$ —values reported in each panel.

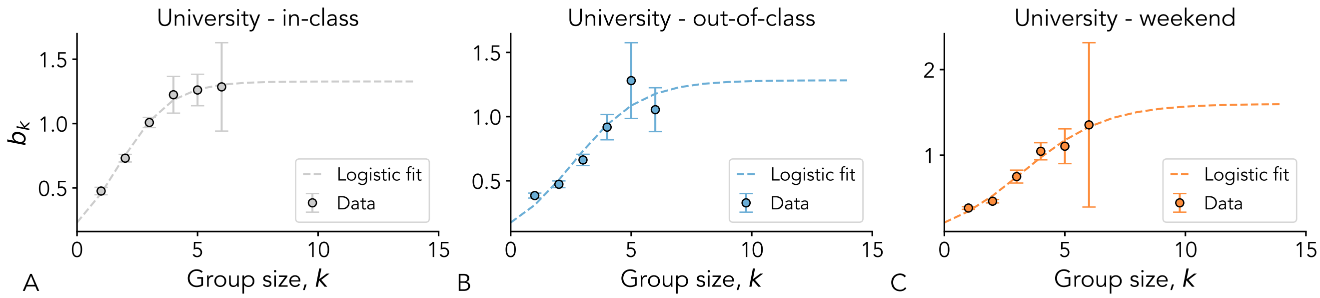

Figure S10. Estimation of the group-size-dependent constants b_k, as given in Eq. (1) of the main text, from the CNS data set by context of interaction: in-class (*A*), out-of-class (*B*), and weekend (*C*). For each context, the fitted exponent β for the group-change probability reported in Fig. S6 is used to estimate the values of b_k for different k with power-law fits. Estimated b_k (points and 95% CI error bars) are plotted together with a logistic fit (dashed lines).

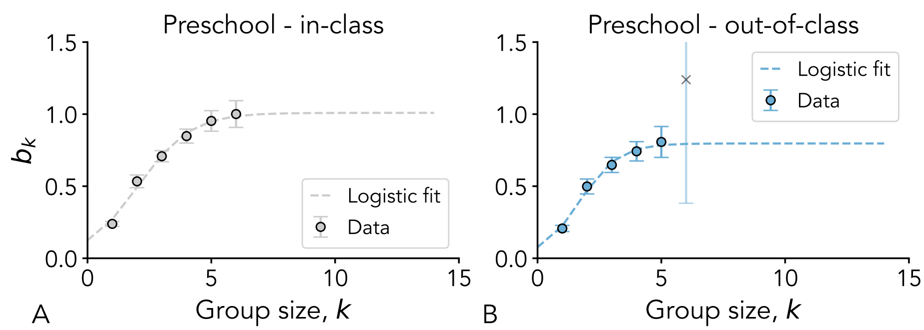

Figure S11. Estimation of the group-size-dependent constants b_k, as given in Eq. (1) of the main text, from the DyLNet data set by context of interaction: in-class (*A*) and out-of-class (*B*). For each context, the fitted exponent β for the group-change probability reported in Fig. S7 is used to estimate the values of b_k for different k with power-law fits. Estimated b_k (points and 95% CI error bars) are plotted together with a logistic fit (dashed lines).

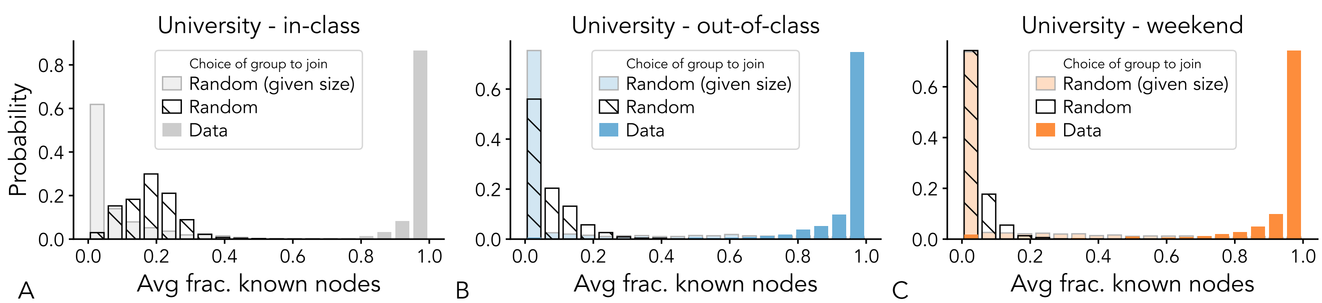

Figure S12. Measuring signals of social memory in group changes for the CNS data set in different contexts of interaction: in-class (*A*), out-of-class (*B*), and weekend (*C*). For each context, whenever we register a group change, we compute the fraction of nodes in the new group that were previously known to the focal node; we also compute the same quantity using a random group of the chosen size available at the considered time, or a random group without any constraints. The resulting distributions of values, averaged over the different time steps, are plotted by comparing the data with the random scenario.

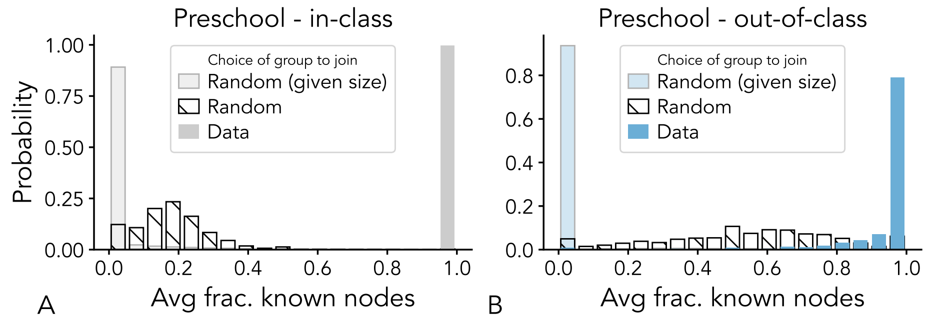

Figure S13. Measuring signals of social memory in group changes from the DyLNet data set in different contexts of interaction: in-class (*A*), and out-of-class (*B*). For each context, whenever we register a group change, we compute the fraction of nodes in the new group that were previously known to the focal node; we also compute the same quantity using a random group of the chosen size available at the considered time, or a random group without any constraints. The resulting distributions of values, averaged over the different time steps, are plotted by comparing the data with the random scenario.

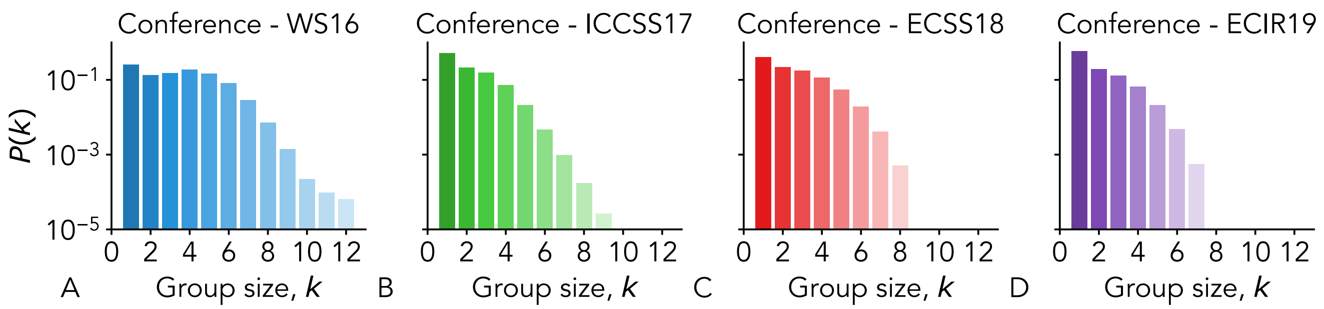

Figure S14. Group size distributions from the GESIS data set for interactions taking place during different events: WS16 (*A*), ICCSS17 (*B*), ECSS18 (*C*), and ECIR19 (*D*).

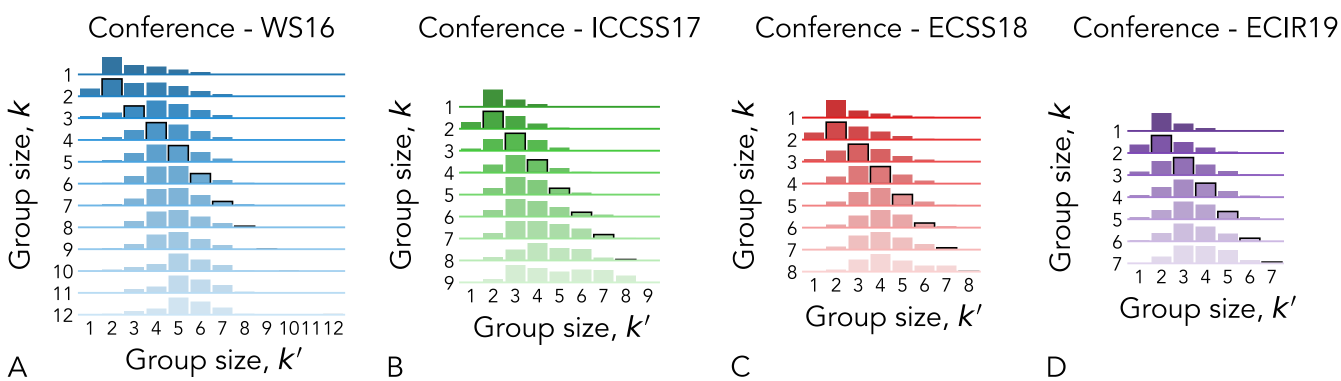

Figure S15. Node transition matrices from the GESIS data set for interactions taking place during different events: WS16 (*A*), ICCSS17 (*B*), ECSS18 (*C*), and ECIR19 (*D*). The elements of each matrix represent the conditional probability that a node that is member of a group of size k at time t is next member of a different group of size k' at time $t+1$ —given that it undergoes a group change. Probability values are given by the height of each element (normalized by row). Note that the scales on the y-axes —one for each matrix row— vary for visualization purposes.

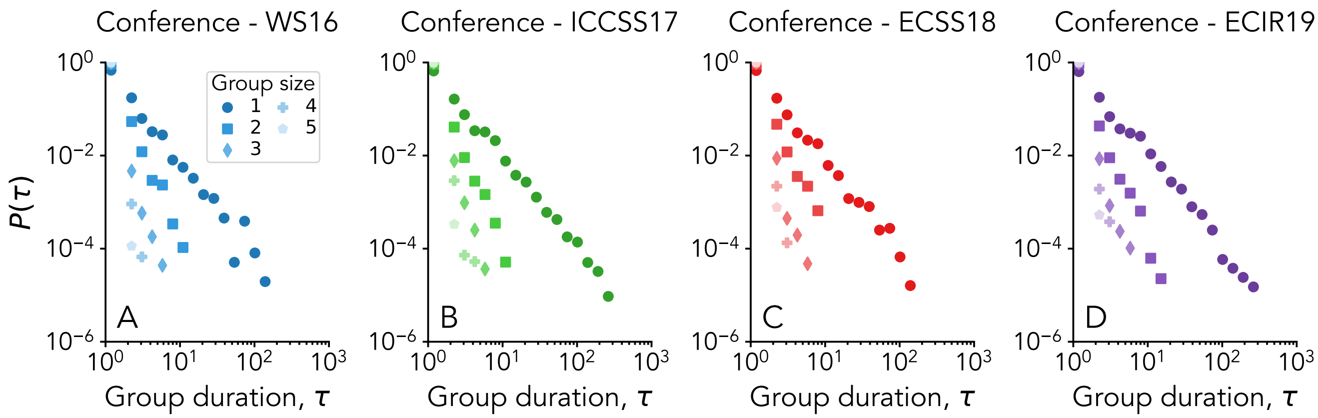

Figure S16. Distributions of group durations τ for the GESIS data set in different contexts: WS16 (*A*), ICCSS17 (*B*), ECSS18 (*C*), and ECIR19 (*D*). In each panel, different symbols correspond to different group sizes.

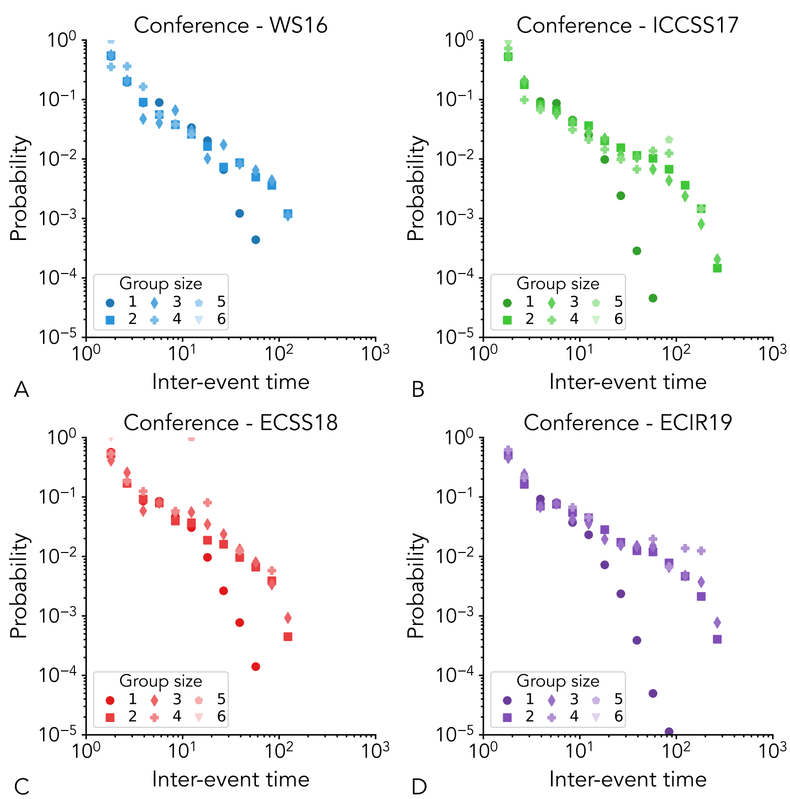

Figure S17. Distributions of group inter-event times (time between consecutive appearances of the same group) in the GESIS data set for different contexts of interaction: WS16 (*A*), ICCSS17 (*B*), ECSS18 (*C*), and ECIR19 (*D*). In each panel, different symbols correspond to different group sizes.

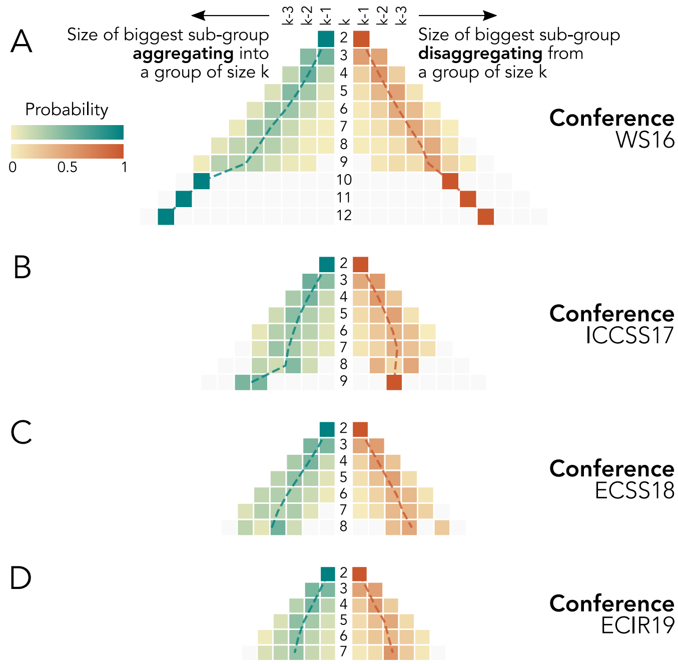

Figure S18. Group dynamics of aggregation and disaggregation from the GESIS data set for interactions taking place during different events: WS16 (*A*), ICCSS17 (*B*), ECSS18 (*C*), and ECIR19 (*D*). Each side of the pyramidal heatmaps shows the probability distribution associated to the size for the largest sub-group joining and the largest subgroup leaving a group of size *k*. The central column reports the considered group size *k*, while the probability distributions on its left-hand side and right-hand side respectively corresponds to group aggregation and disaggregation. Dashed lines refer to the distribution average.

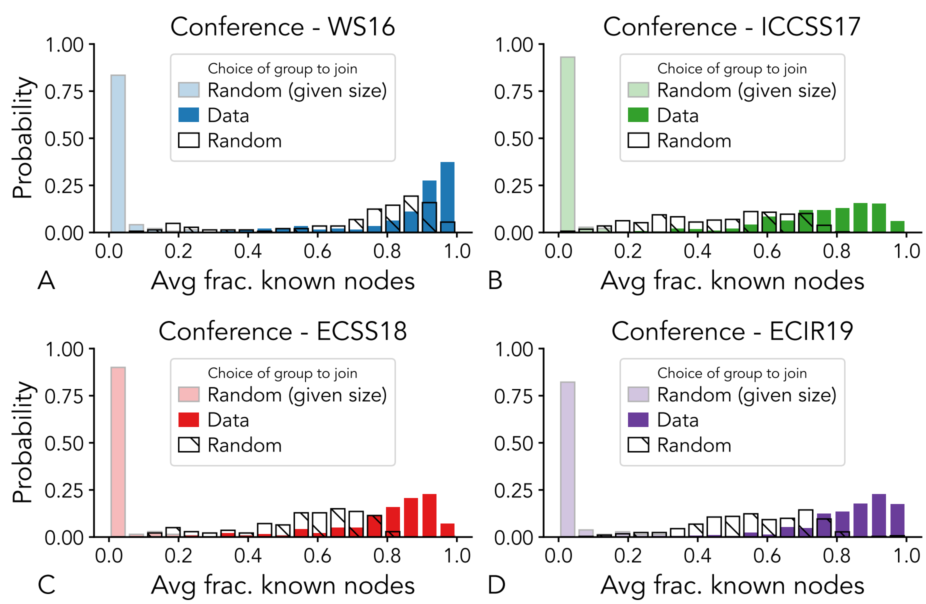

Figure S19. Measuring signals of social memory in group changes for GESIS data collected in different context of interaction: WS16 (*A*), ICCSS17 (*B*), ECSS18 (*C*), and ECIR19 (*D*). For each context, whenever we register a group change, we compute the fraction of nodes in the new group that were previously known to the focal node; we also compute the same quantity using a random group of the chosen size available at the considered time, or a random group without any constraints. The resulting distributions of values, averaged over the different time steps, are plotted by comparing the data with the random scenario.

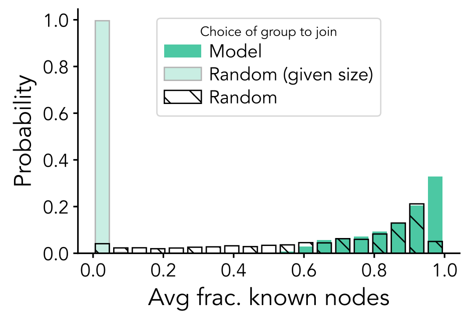

Figure S20. Measuring signals of social memory in group changes in a simulation of the proposed temporal hypergraph model. For each time step of the simulation, whenever we register a group change, we compute the fraction of nodes in the new group that were previously known to the focal node; we also compute the same quantity using a random group of the chosen size available at the considered time, or a random group without any constraints. The resulting distributions of values, averaged over the different time steps, are plotted by comparing the data with the random scenario.

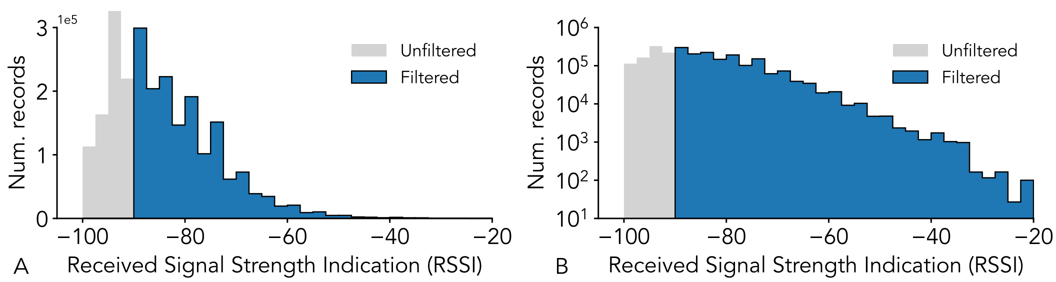

Figure S21. Filtering bluetooth signals from the CNS data set. (*A*) Number of recorded interactions as a function of the associated Received Signal Strength Indication (RSSI) [dBm]. The y-axis in panel (*B*) is in logarithmic scale. The volume of interactions retained after the filtering is displayed in blue.

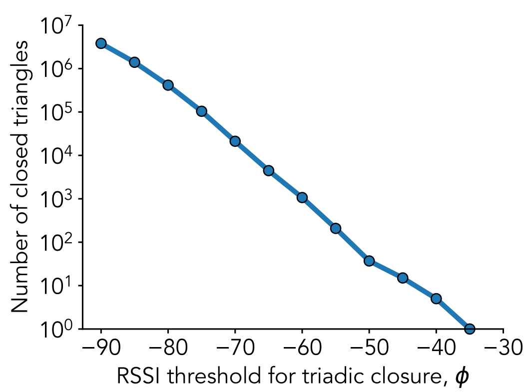

Figure S22. Pre-processing of the CNS dataset: triadic closure. Number of links added via triadic closure as a function of their approximated Received Signal Strength Indication (RSSI).

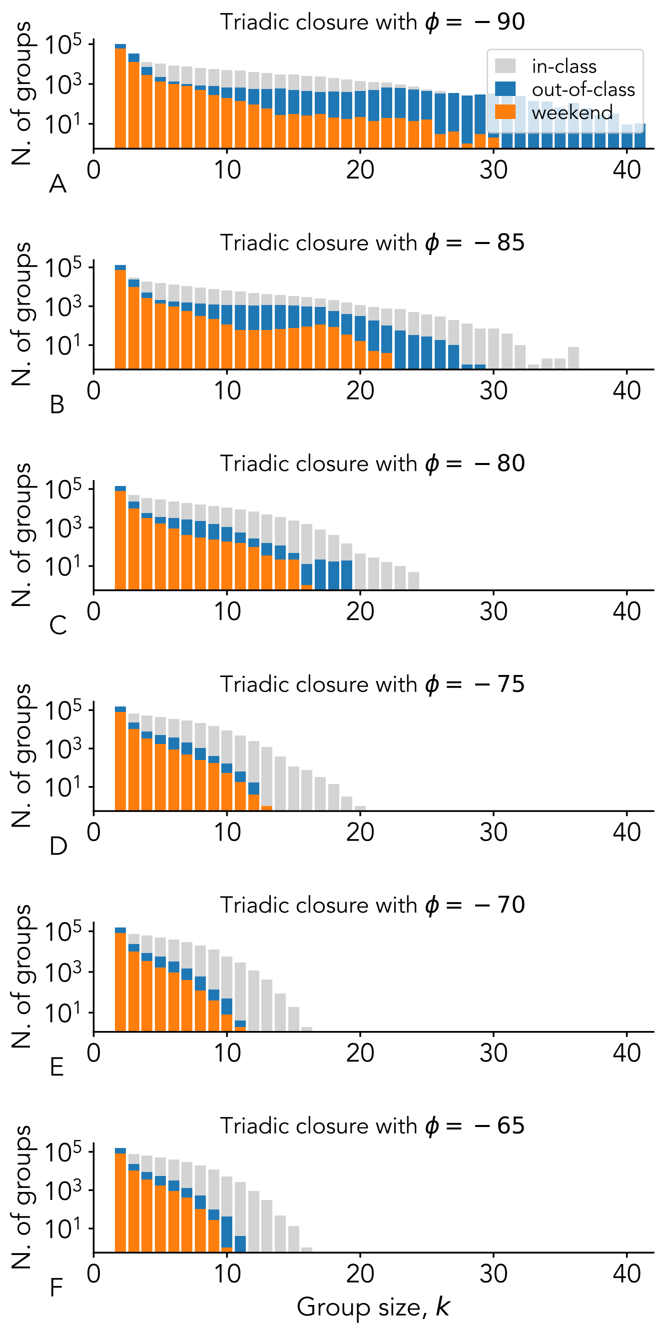

Figure S23. Pre-processing of the CNS dataset: triadic closure with threshold. Impact of thresholding the newly added link —due to the triadic closure— on the group size distributions for different contexts of interaction. Each panel shows the number of groups, divided by context [in-class (gray), out-of-class (blue), weekend (orange)], as a function of the group size k. Different panels (rows) correspond to different value of the threshold ϕ used to filter out the links added with the triadic closure. Links featuring a Received Signal Strength Indication (RSSI) [dBm] lower than ϕ are thus removed.

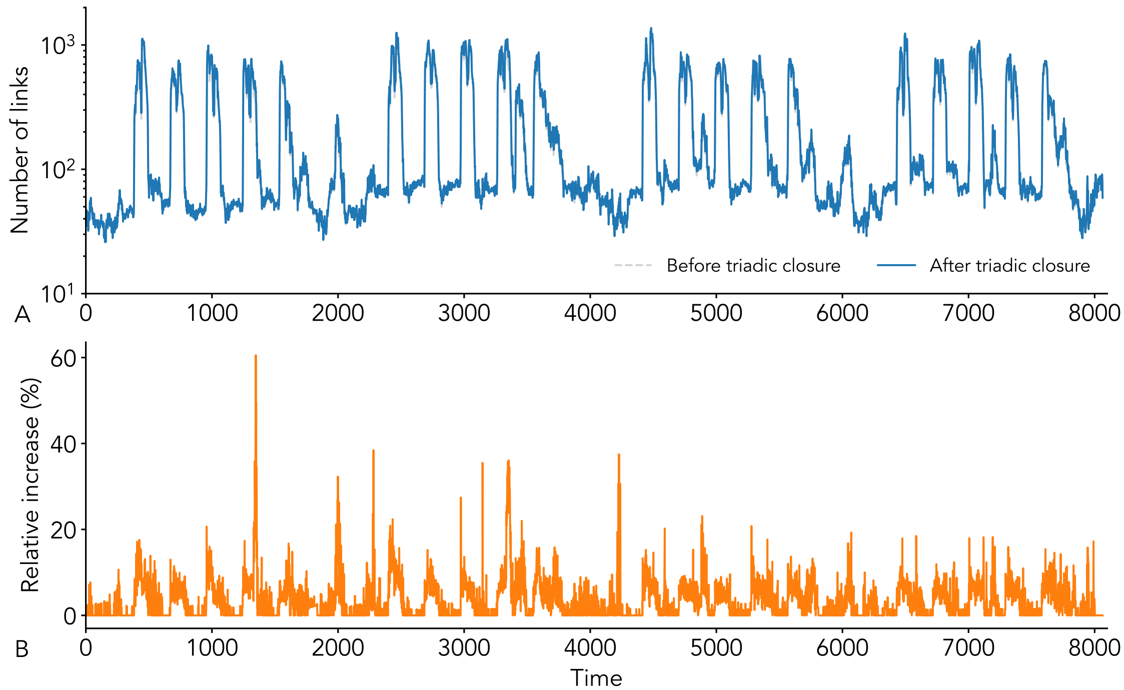

Figure S24. Preprocessing of the CNS dataset: impact of the triadic closure on the number of raw pairwise interactions in time (*A*), and relative change (*B*).

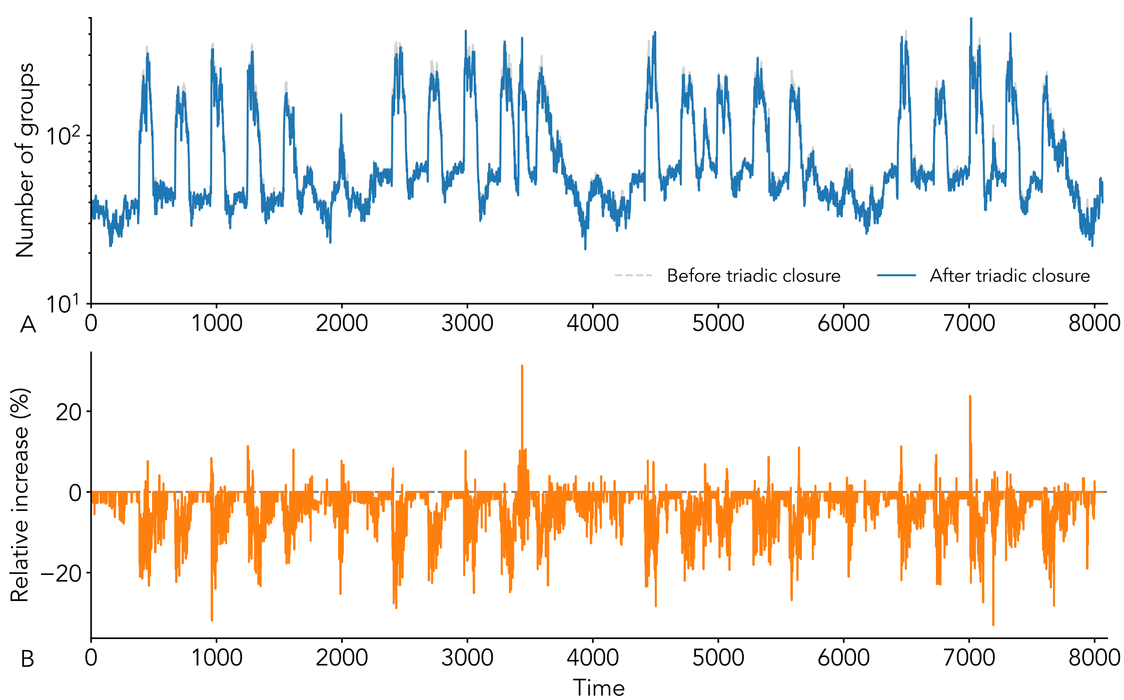

Figure S25. Preprocessing of the CNS dataset: impact of the triadic closure on the number of groups in time (*A*), and relative change (*B*).

Data set	Context	Unique membership (frac.)
University	in-class	0.75
University	out-of-class	0.96
University	weekend	0.98
Preschool	in-class	0.80
Preschool	out-of-class	0.94

Table S1. Fractions of single-membership interactions.

1 Supplementary Note 1 – GESIS data sets on interactions at conferences

We use public data collected by the SocioPatterns collaboration[1] during four different scientific conferences organised by GESIS, the Leibniz Institute for Social Sciences, Germany: the 3rd GESIS Computational Social Science Winter Symposium in 2016 (WS16), the International Conference on Computational Social Science in 2017 (ICCSS17), the Eurosymposium on Computational Social Science in 2018 (ECSS18), and the 41st European Conference on Information Retrieval in 2019 (ECIR19). These conferences were strongly interdisciplinary, covering a wide range of topics from computer to social and natural sciences. The data consists in face-to-face interactions among participants of the different conferences as measured by wearable RFID sensors with a temporal resolution of 20 seconds. The details of the data collection, including the collection technique and the analysis of the interactions, are presented in the original Ref.[2]. The raw data for each conference contains 153,371 (WS16), 229,536 (ICCSS17), 96,362 (ECSS18), and 132,949 (ECIR19) entries, respectively. We aggregate these records, as per the University data set, using a temporal window of 5 minutes, ending up with 38,832 (WS16), 58,745 (ICCSS17), 24,637 (ECSS18), and 34,217 (ECIR19) entries. We then add isolated nodes to the data. In particular, for each conference, we consider all the nodes that ever appear and add a record of the node being isolated for each time step after their first appearance. At this point the data are composed of 46,915 (WS16), 98,124 (ICCSS17), 35,124 (ECSS18), and 61,371 (ECIR19) entries.

2 Supplementary Note 2 – *

References

1. SocioPatterns Collaboration. http://www.sociopatterns.org/.

2. Génois, M. *et al.* Combining sensors and surveys to study social interactions: A case of four science conferences. *Pers. Sci.* **4**, 1–24, DOI: https://doi.org/10.5964/ps.9957 (2023).