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

Iacopo Iacopini [1,2,3] ✉, Márton Karsai [3,4] & Alain Barrat [5]

Representing social systems as networks, starting from the interactions between individuals, sheds light on the mechanisms governing their dynamics. However, networks encode only pairwise interactions, while most social interactions occur among groups of individuals, requiring higher-order network representations. Despite the recent interest in higher-order networks, little is known about the mechanisms that govern the formation and evolution of groups, and how people move between groups. Here, we leverage empirical data on social interactions among children and university students to study their temporal dynamics at both individual and group levels, characterising how individuals navigate groups and how groups form and disaggregate. We find robust patterns across contexts and propose a dynamical model that closely reproduces empirical observations. These results represent a further step in understanding social systems, and open up research directions to study the impact of group dynamics on dynamical processes that evolve on top of them.

Social interactions are the building blocks of our society[1]. Humans — and animals in general— form groups of different sizes[2] and have learnt the advantages of communicating and gathering in close social circles[3,4]. Our everyday social path is in fact a succession of these group events that involve different numbers of peers, from walking alone or having a coffee with a friend, to engaging in meetings or group conversations at work or during social gatherings. Networks provide a powerful tool to represent these complex social trajectories with the capacity to encode the structure and dynamics of interactions between individuals[5–8]. The use of network representations and of social network analysis tools[1], as well as the emerging field of temporal networks, have helped identifying the mechanisms that govern the formation and evolution of these structures[9–11]. Nevertheless, these conventional network descriptions are inherently limited to the description of pairwise interactions, which does not capture the full complexity of the social phenomena[12–15]. Considering interactions of higher-order is thus compulsory to represent and model how humans interact in groups[16,17] or how animals gather[18].

The structure and dynamics of group interactions, however, are complex[19]. Groups may have heterogeneous sizes[20–24], can change dynamically[25,26] or exhibit hierarchical and nested structures[27,28]. Possible driving mechanisms behind these characters include for example, simplicial closure[21] and homophily[24]. Most studies on group formation and structure, however, do not take into account the further temporal evolution of the underlying social systems—that is characterised by patterns of memory and burstiness, and exhibits a complex dynamics of merging and splitting of groups[22,29–31]. For instance, larger groups tend to have shorter durations[29] and exhibit shorter temporal correlations[31]; the dynamics of group formation and fragmentation exhibits a preferred temporal direction[22,31], and non-trivial recurrence of groups can emerge, driven by different contexts and geographical places of interactions and defining social circles[32]. These complex patterns are the results of microscopic individual level decisions, ultimately shaping the emergence of collective behaviours. Understanding these mechanisms is essential to better characterise the emerging group dynamics and their effects on processes such as

[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. ✉e-mail: iacopo.iacopini@nulondon.ac.uk

disease transmission or spread of information, social norms and behavioural patterns within and across group gatherings[7,33–35].

Here, we address this challenge by investigating empirical traces of group dynamics extracted from proximity data in different social and temporal contexts. Leveraging two data sets of temporally resolved human interactions among preschool children and freshmen students, we highlight complex mechanisms of group dynamics both at the individual and at the group level. Following the group membership of individuals across groups of different sizes, we find that the main dynamical patterns of group-change are independent from the context of interactions. The statistics of group durations exhibit as well robust properties, with a *long-gets-longer* effect[29,36] for all group sizes (i.e., the probability to change group decreases with the time spent in it). Furthermore, the dynamics of group aggregation and disaggregation show hierarchical and largely symmetrical properties of assembly and disassembly. Finally, we propose a dynamical model for temporal interactions −that takes groups explicitly into account− and show that it reproduces the empirical patterns.

Our results shed lights on how temporal patterns of group formation and evolution can result from microscopic choices at the level of individuals, by accounting for mechanisms of social and temporal memory. The proposed model can moreover serve as a synthetic structure for studying the impact of group interactions and their temporal properties on dynamical processes: indeed, recent works based on static hypergraphs have shown evidence that group interactions can induce critical mass effects in social contagion[37,38], amplify small initial opinion biases and accelerate the formation of consensus[23,39] and cooperation[40,41], but investigations on evolving structures are scarce[42]. Overall, our study provides a starting point for increasingly realistic modelling approaches to better characterise complex social systems and the phenomenology of attached processes.

Results
Extracting groups from real-world data
We consider records from two data-collection efforts that tracked social interactions at a university and in a preschool, yielding data sets in the form of temporal networks, in which each person is represented as a node and each interaction as a temporal edge (see *Methods*). Time is in each case discretized by the temporal resolution of the data collection setup. At each timestamp, we define as groups the maximal cliques (largest fully connected subgraphs) and build in this way a temporal hypergraph. At each timestamp, each node can thus be either isolated or part of one or several groups (hyperedges).

University. We use data collected by the Copenhagen Network Study (CNS)[43]. It is a temporally resolved data describing the proximity events of 706 freshmen students at the Technical University of Denmark, collected using the exchange of Bluetooth signals by their smartphones. We use the publicly available data describing these proximity events during four consecutive weeks and with a temporal resolution of 5 minutes[43]. We split the data into three different contexts, which might result in different interaction patterns. First, we treat all interactions taking place over the *weekends* as a separate set. Second, we divide interactions that happen during the workweek into *in-class* and *out-of-class* time. In this way, we do not mix the group dynamics emerging in unconstrained interactions during the free time of the lunch break and in-between classes with potentially constrained co-presence due to common attendance of classes and/ or seating configurations. For this data set moreover, we perform some data pre-processing before extracting the groups in each timestamp, to filter out very weak interactions (based on the Bluetooth Received Signal Strength Indication), smoothen intermittent patterns, remove spurious connections, and perform a standard triadic closure procedure with tailored parameters[32]. Additional

details on the data set and the pre-processing are described in the *Methods*.

Preschool. We also consider another data set, collected in a preschool as part of the DyLNet project[44] to follow the social interactions of children of age 3 − 6 and their teachers and assistants. The data describes proximity social interactions between 174 children and 34 adults in seven classes, recorded by Radio Frequency Identification (RFID) Wireless Proximity Sensors carried by each participant. Interactions were recorded with a temporal resolution of 5 seconds, during periods of 5 consecutive days, for 10 consecutive months (overall 50 days of data collection) of a single academic year in a French preschool. For the purpose of our study we rely on a pre-processed data set shared in[44] and a temporal network reconstructed from the cleaned interaction signals as explained in[45], and we remove the data of interactions with and between adults, to focus on the childrens' group dynamics.

Similarly to the university setting, we also divide the data according to the contexts that may impose different constraints on the emergence of possible group interactions. We differentiate between *in-class* periods, during which the social grouping of children was strongly influenced by the teachers' instructions and scheduled activities, and *out-of-class* periods, when children could choose freely to interact with anyone from their own and potentially other classes. For more details on the data pre-processing, network reconstruction, and context selection, we refer to the *Methods* section.

The dynamics of group change
A first coarse summary of the complexity of group interactions is unveiled through the heterogeneity of groups sizes, already documented in a variety of studies[20–24,28]. We confirm this finding in the data sets considered here in Fig. 1A, B, with distributions of instantaneous group sizes having similar shape but spanning varying ranges of values: the interactions measured in the university (panel *A*) feature larger group sizes than preschool ones (panel *B*), possibly because of the longer range of the Bluetooth signals used in the data collection infrastructure. In addition, we observe a general tendency of gathering in smaller groups in contexts where students or children are free to interact (*out-of-class* and *weekend*).

The distribution of sizes is, however, by design, an aggregated observable that does not inform us about the dynamics of interactions: a given node might belong at different times to groups of very different sizes, just as a node in a temporal network might have very different numbers of neighbours or centrality values at different times[46–48]. We thus now investigate how the group membership of individuals evolves across various sizes (some example trajectories can be found in Supplementary Information, Figs. S1 and S2). In this regard, it is important to stress that whenever we refer here to a group change by a node, we interpret it in the most general sense, i.e., it does not necessarily mean that the node is actively changing from a group to another one. In fact, from the point of view of a given individual, a group change can also be due to another person joining or leaving their current group. Under this approach, adopted to avoid having to arbitrarily decide how group "labels" propagate whenever there is a change in one of the members, our analysis is purely observational, and agnostic with respect to the intention of individuals.

We build for each context a transition matrix $T \equiv \{T_{kk'}\}$ describing these changes as measured in the data: denoting by n_i^t the size of a group to which node i belongs at time t, each matrix element represents the conditional probability $P(n_i^{t+1} = k' | n_i^t = k)$ of finding a given node i in a group of size k' at time $t+1$ given that at time t it belonged to a different group of size k (see *Methods*). The results, displayed in Fig. 1C−G, show strikingly robust patterns across the different contexts, differing only in the cut-off associated with the largest group sizes observed: (*i*) at given group size k at t, the most probable group
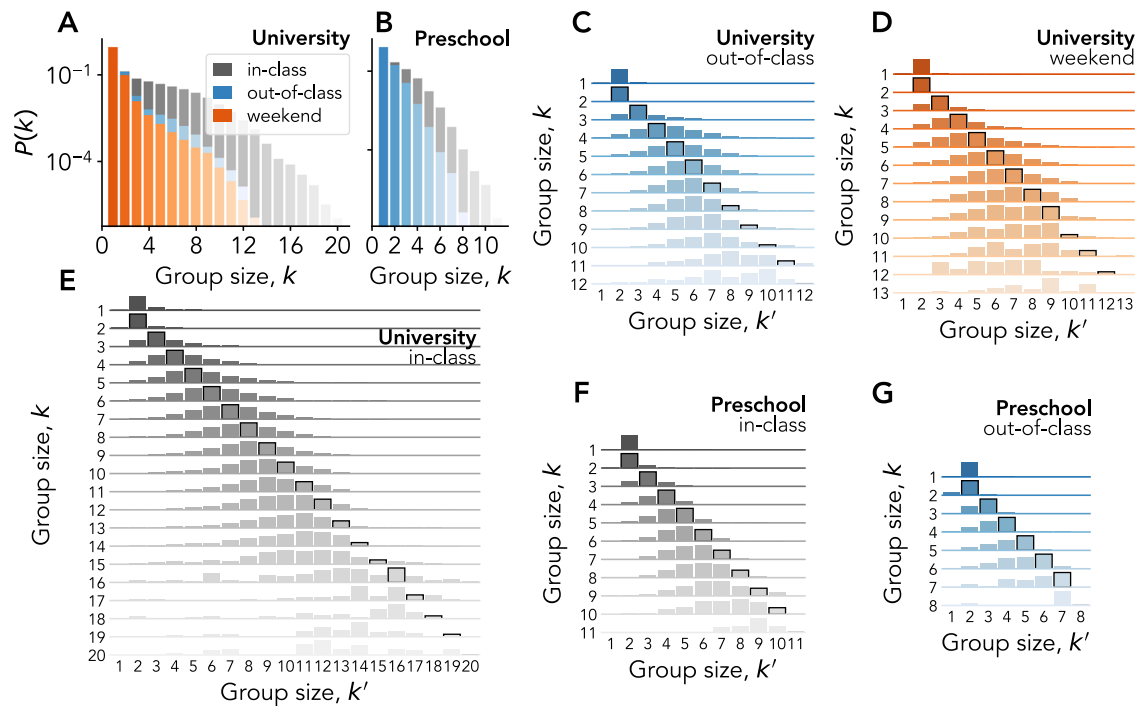
Fig. 1 | Group size distributions and node transition matrices. A, B Group size distributions for University (**A**) and Preschool (**B**) interactions that take place in-class, out-of-class, or during the weekend (see legend). **C–G** Node transition matrices for University (**C–E**) and Preschool (**F, G**), for interactions that take place during in-class (**E**), (**F**), out-of-class (**C, G**), or weekend (**D**) time. 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 between t and $t + 1$. 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.

size at the next time step is $k' = k$, for small enough k (except for $k = 1$ in which case the next size is most often 2); (*ii*) the distribution $P(n_i^{t+1} = k'|n_i^t = k)$ extends to values around the diagonal $k' = k$, with both events of individuals undergoing a group change towards a larger or a smaller group but large differences between k and k' are rare; (*iii*) as k increases, the distribution shifts to the left of the diagonal, i.e., it becomes increasingly probable that a change of group leads an individual to a group of smaller size.

The approach just described follows the evolution between groups of different sizes from a purely individual standpoint and does not include any information about alter group members beside the considered ego. However, each transition from a group size k to k' could correspond to very different scenarios in terms of group members. To illustrate this point, we compute the overlap between consecutive groups of an ego, as measured by the Jaccard coefficient between their sets of members. The ego could for instance be at t and $t + 1$ in groups formed by totally distinct alters, leading to a low Jaccard coefficient. On the contrary, a group change could also result from another member of the group leaving, in which case the ego would see a strong overlap between the groups at successive times. Even at fixed k and k', the distributions of Jaccard similarity values between consecutive groups, shown in Supplementary Information, Fig. S3, confirm in fact that a broad range of intermediate situations are also encountered in terms of change of group members seen by the ego. Therefore, to get better insights on the underlying dynamics from a compositional point of view, we now shift the focus of our analysis from individuals to groups.

For each group, we define its times of birth and death respectively as the first and last appearance of the same set of members in consecutive time steps. The duration of a group is then naturally given by the temporal difference τ between its time of birth and death, and we investigate in Fig. 2 how the duration statistics depend on the group size. We find that the distributions of group duration $P_k(\tau)$ for groups

of size k present broad shapes for all k, with comparable patterns in shape, exponent values and size dependency across the very different contexts considered. In particular, whether students interact during classes, in the other spaces of the university, or elsewhere during the weekend, the distributions of their group interactions depend on the group size in a similar way: the heavy-tail distributions are broader for smaller group sizes, with longer averages and maximum observed durations. Group interactions at preschool show a similar pattern.

This phenomenon is closely linked to the one of burstiness, an important feature found in empirical temporal networks[10], where periods of node (or link) inactivity are heterogeneously distributed. Moving beyond pairwise interactions, node inactivity corresponds to groups of size 1, i.e., when a node is isolated. Figure 2 thus confirms the presence of bursty periods of inactivity at the node level. This is also illustrated by examples of node activity through time, as given by the temporal evolution of the size of the group to which a node belongs, reported for some selected nodes in Supplementary Information, Figs. S1 and S2. As expected, nodes display very heterogeneous levels of participation —and active periods featuring medium and large groups are inevitably correlated across different nodes. In addition to the inter-event time distribution for nodes given Fig. 2, we also find burstiness across the different contexts of interactions at the level of groups. This is deduced from the inter-event time distributions reported in Supplementary Information, Figs. S4 and S5 that are broadly distributed even after disaggregating by group size. This analysis extends the results described in[22,29] to very different contexts, showing that the strong robustness of statistical patterns of contacts goes beyond the one of pairwise interactions described in earlier works[49,50], and hinting at common robust mechanisms determining contact and group formation, duration, and evolution in different contexts.

To go further, we now investigate how groups change: indeed, the node transition matrices introduced above and shown in Fig. 1 give
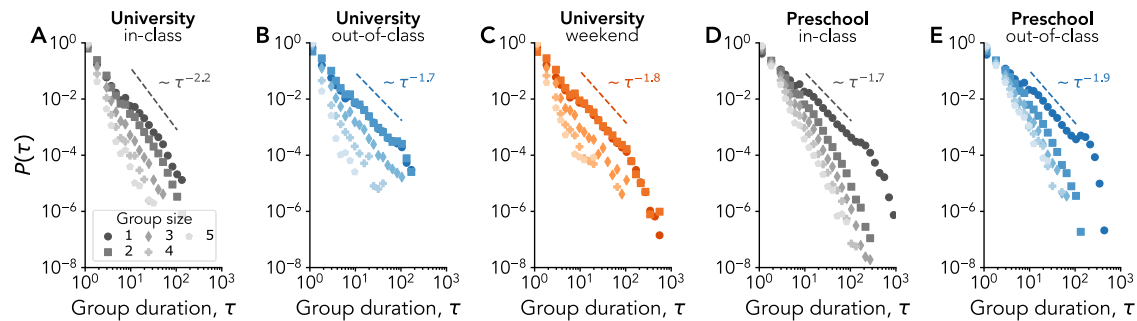
Fig. 2 | Distributions of group durations. Each panel reports the distribution of group duration τ for the CNS (**A–C**) and the DyLNet (**D–E**) data sets in different contexts: in-class (**A**, **D**), out-of-class (**B**, **E**) and weekend (**C**). Different symbols correspond to different group sizes. The distributions for group size 1 have been fitted using the method in Ref. [88], and can be characterized by the depicted exponent values.

only partial information regarding the actual group dynamics. The individual point of view adopted is useful to understand how individual group membership evolves between different sizes, but the impact of these individual changes on the sizes of the groups needs an independent analysis. For instance, while Fig. 2 shows that larger groups tend to have shorter lives, how they break up is still to uncover. Similarly, a group might appear due to the fusion of two pre-existing groups of comparable sizes −like water droplets that merge after overlapping due to surface tension, or from a gradual process of the integration of one individual at a time. To investigate this issue, we follow the members of each group, before the group's birth and after its break-up. Moreover, we pool together the results of groups of the same size, to check whether groups of different sizes undergo different aggregation and disaggregation dynamics. For each group size k, we show in Fig. 3 the heatmaps (one for each context) of the size distributions of the largest sub-set of group members observed just before the birth or just after the death of a group (see *Methods*). For small group sizes, both group aggregation and disaggregation tend to happen gradually from or to groups of similar sizes. This is in agreement with previous results for the formation of 3-body interactions[22]. For increasingly larger groups, the picture evolves in a slightly context-dependent way. The general picture is that no merging from (or splitting into) equally sized groups is observed, as could be expected from a purely combinatorial point of view. Medium- and large-sized groups tend to be created from a group of slightly smaller size joined by one or few small ones, and symmetrically lose members in a small chunk, remaining of mid to large size. This points towards a partially hierarchical dynamical mechanism according to which individuals first engage in small groups and the small to mid and large groups aggregate to form even larger ones. A symmetric process takes place when large groups dismantle −first into smaller medium-sized subgroups and then loosing members one at the time.

A data-driven dynamical model for groups' evolution

Most models describing the temporal evolution of social interactions consider network representations, i.e., are based on mechanisms governing how pairwise interactions are established and successively broken[36,51–55]. Here, we describe instead a model that explicitly integrates how individuals form groups of arbitrary sizes[29,56,57]: at each time step an individual can decide to stay in their current group, leave and join a different one, or become isolated.

The model is inspired by the one put forward in[29,56]: we consider N agents that interact in groups over time. For simplicity, we assume that each agent $i = 1, 2, ..., N$ participates in only one group at a time (as it happens for most of the empirically observed interactions, see Supplementary Information, Table S1). We call $\mathcal{K}^t$ the set of groups present at time t. We denote by $\sigma_i^t \in \mathcal{K}^t$ the −single− group to which agent i participates at time t, of size $n_i^t \equiv |\sigma_i^t|$. If $n_i^t = 1$, the agent i is isolated, or inactive (cf. Supplementary Information, Figs. S1 and S2). We note that

many models devoted to describing the evolution of interactions between individuals are based on successive pairwise interactions (groups of size 2)[11,51,53,55]. As time is aggregated, the result of these binary interactions is a social network between a set of nodes representing individuals[1], where a link denotes the fact that two individuals have been in contact (each link can be weighted, e.g., by the cumulated time of interactions). In the present case, interactions between agents are described by groups of various sizes. The adequate tool to describe the temporally aggregated picture is then not anymore a network, but a hypergraph between the set of vertices $\mathcal{V}$ representing the agents[12,58]. This hypergraph is formed by hyperedges between these vertices, where a k-hyperedge $\sigma \in \mathcal{K}$ is a set of k vertices representing a group interaction of size k, which can be weighted by the total time this group interaction has taken place.

The model evolves through iterations where each time step t corresponds to an epoch, during which each one of the N agents is selected in a random order. Whenever an agent i is selected, currently a member of group σ_i^t, the model then evolves according to two sequential mechanisms. With the first one the agent decides to either stay in the same group, depending on the time spent there and the group size, or alternatively leave it for a different one. If the agent stays, nothing happens. If the agent instead leaves its current group, a second mechanism is triggered, corresponding to the choice of its next group: this choice is based on the acquaintances made until that time. We note that in the original pairwise model[29,56], individuals leaving a group became automatically isolated, and isolated individuals could join groups of any size, which implies that the shape of the empirical transition matrix of Fig. 1 could not be reproduced. In the next paragraphs we leverage qualitative insights and direct measures from data to define these two mechanisms in more details.

We first take into account that the probability for an agent to have a group change decreases with the time τ the agent has already spent in that group, i.e., a "long-gets-longer" effect. Evidence for such effect has been found empirically in pairwise interactions[36,53]. The case of larger group sizes and the potential dependency on the group size have however not been investigated. We thus measure the group-change probabilities in our data, and how they depend on the group sizes. Specifically, we compute the probability $P^\gamma(k, \tau)$ for a node belonging to a group of size k to have a change of group after τ timestamps (see *Methods*). Figure 4A, D shows the resulting probability distribution as a function of the group duration τ aggregated over all group sizes k (the distributions shown in the figure correspond to interactions that take place out of class, and results for the other contexts are shown in the Supplementary Information, Figs. S6 and S7). These empirical results suggest that the "old-gets-older" mechanism observed in pairwise interactions remains valid for groups: the probability for an agent to change group decreases with the time they have spent in that group. In other words, the longer a group has been established, the smaller the probability that it will break apart. Similar trends can be found for
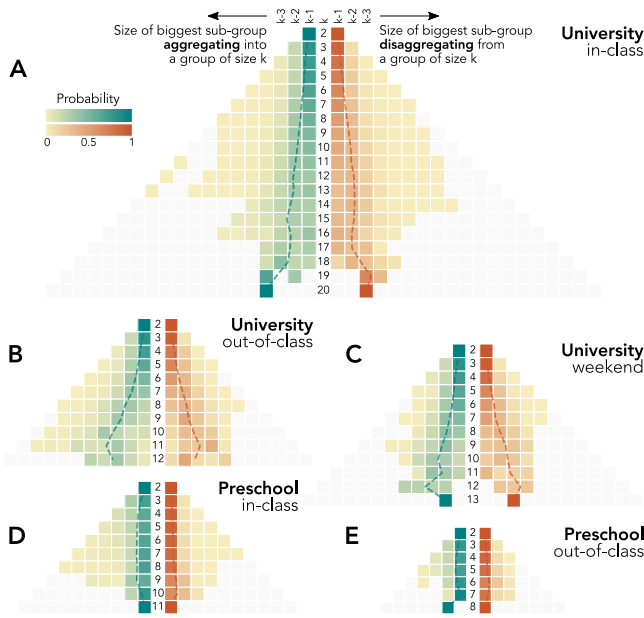
Fig. 3 | Group dynamics of aggregation and disaggregation. The panels report results for University (**A**–**C**) and Preschool (**D**–**E**) interactions that take place during in-class (**A**, **D**), out-of-class (**B**, **E**), or weekend (**C**) time. Each side of the pyramidal heatmaps shows the 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.

the two data sets, both when aggregating over group sizes −as just shown− and for the distributions $P^\curvearrowright(k, \tau)$ separated by group size, as shown in Supplementary Information, Figs. S8 and S9 for the CNS and the DyLNet data sets, respectively.

We thus assume in the model that the probability that a node i in a group of size k has a change of group decreases with the time τ that node i has spent in that group (residence time) as

$$p_k(\tau) = \frac{b_k}{\tau^\beta/N}, \tag{1}$$

where β is a real-valued exponent that modulates the impact of the residence time, which we obtain by fitting the empirical distributions (dashed lines in Fig. 4A, D). With probability $1 - p_k(\tau)$ the agent i stays in the same group.

In Equation (1), b_k is a constant that depends on the size of the current group of node i. It is indeed reasonable to assume that the probability of leaving a group also depends on the size of the group itself, as size is a crucial factor that determines a group's sociological form[59], and its ability to sustain a single conversation −leading to the phenomenon known as *schisming*[60]. To gain insights on the dependency of b_k on the group size k, we fit each $P^\curvearrowright(k, \tau)$ using a power-law function of the form $b_k\tau^{-\beta}$ (with β taken from the fit shown in Fig. 4A, D). The resulting values of b_k are reported in Fig. 4B, E. They show a monotonic increase with k, which we fit to a logistic form

$$b_k = \frac{1}{1 + e^{-\alpha(k-k_0)}}. \tag{2}$$

Similar results across the different contexts and data sets are reported in Supplementary Information, Figs. S10 and S11 for the CNS and DyLNet data sets, respectively. In all cases, the logistic fit falls within the confidence intervals of the empirical measures, justifying the choice of a logistic function.

Let us now focus on the second mechanism of the model, which controls for the selection of the new group after the group change. Previous empirical and modelling investigations for temporal networks support the idea that individuals have a preference for repeating interactions with people already met, i.e., a mechanism of *social memory*[36,53,55]. Nevertheless, this explicit signal at the level of groups has never been measured. Hence, we check the presence of social memory in the empirical data by looking, whenever a node i undergoes a group change at any time t, at the fraction $\chi_{i,\omega}(t)$ of known nodes in the new group ω of node i. We then take the average over all the group changes and plot the associated distributions for the two considered data sets in Fig. 4C, F. We also compare the results with two baseline scenarios in which the group choice is performed uniformly at random, among the available groups at the moment of the change, or at random but restricted to those having the same size of the new group in the empirical data (see *Methods*). The results of Fig. 4C, F clearly show that both data sets display a strong signal of social memory as compared to their random counterparts. The same holds for the other contexts and data sets, as reported in Supplementary Information, Figs. S12 and S13.

We thus take into account this mechanism of social memory in the following way in the model. In case of a group change, we denote by σ_i^t the group of i at time t, and by $\omega \in \Omega_t = \{\mathcal{K}^t \setminus \sigma_i^t \cup \emptyset\}$ the new group after the group change. Note that, by including the empty set $\emptyset$ among the possible target groups, we account for the possibility that i becomes isolated. Specifically, we include in the ensemble Ω_t of possible groups to join multiple copies of $\emptyset$: this multiplicity, controlled by a parameter ϵ, makes it possible to tune in the model the willingness of agents to become isolated upon leaving a group. Among all possible groups $\omega \in \Omega_t$, i selects the one to join via the second behavioural mechanism, which involves the memory of previous interactions. Namely, the probability to join ω is proportional to the fraction $\chi_{i,\omega}(t)$ of agents in ω that at time t have already interacted with i in the past. Let us know define a slightly different quantity $\chi_{i,\omega}^\dagger(t)$ as

$$\chi_{i,\omega}^\dagger(t) = \frac{1 + \left[\omega \cap \bigcup_{t'=1}^{t} \sigma_i^{t'}\right]}{1 + |\omega|}. \tag{3}$$

Note that, differently from the simple fraction of known agents, node i itself is included in the computation of $\chi_{i,\omega}^\dagger(t)$ in order to have a non-zero probability for i to join either an empty group or a group of previously unmet individuals. In other words, $\chi_{i,\omega}^\dagger(t)$ is the density of agents in the group which are known to i right after joining. Altogether, the probability for an agent i belonging to group σ_i^t at time t to be found in a different group ω at time $t + 1$ is given by

$$\text{Prob}(\sigma_i^{t+1} = \omega|\{\sigma_i^t\}_t) = \frac{\chi_{i,\omega}^\dagger(t)}{\sum_{\omega'} \chi_{i,\omega'}^\dagger(t)}. \tag{4}$$

The model reproduces the higher-order dynamical features

We now explore the ability of the model defined above to capture the key empirical features we have uncovered in the dynamics of group interactions. As empirical results are robust across data sets and contexts, we consider as an example the University interactions taking place during out-of-class time. We thus run the model initialised with $N = 700$ agents for $T = 2000$ time steps, using different parameter values for ϵ, α, and k_0 −while β is set to 0.8 as measured in Fig. 4A. Each realisation of the model generates a sequence of temporally-ordered hypergraphs that we can analyse as per the empirical data, obtaining in particular group size distributions and group size transition matrices. As described in *Methods*, we can thus jointly fit the model on these two observables.
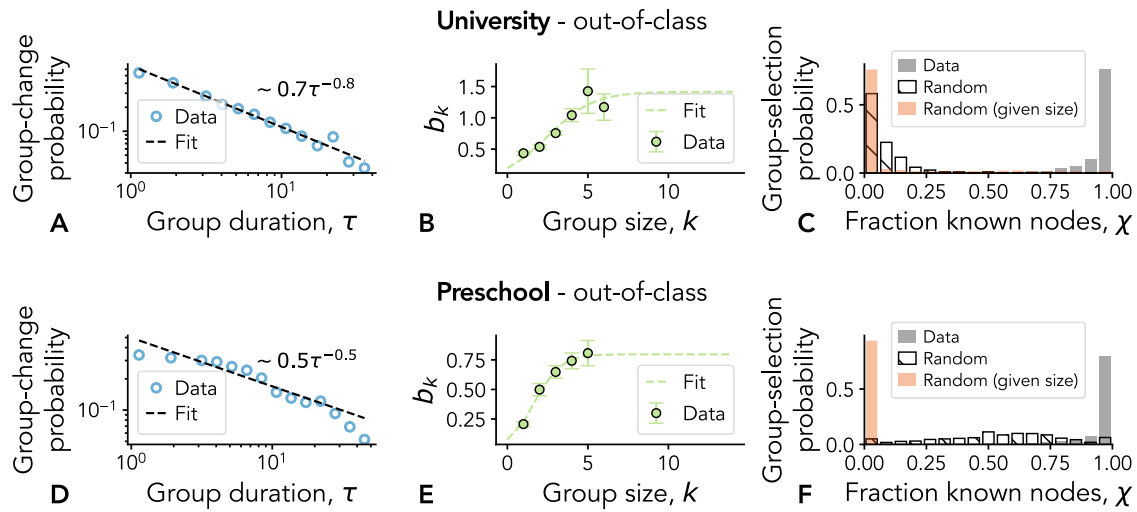
Fig. 4 | Fitting the empirical group-change probability and measuring the signal of social memory. The different panels refer to out-of-class interactions from the CNS (**A**–**C**) and the DyLNet (**D**–**F**) data sets. **A**, **D** Probability for a node belonging to a group of any size to leave it 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. Fitted exponents β are then used to estimate the group-size-dependent constants b_k, as given in Eq. (1), with another power-law fit. In **B**, **E** the resulting values for b_k are plotted (points and 95% CI error bars) together with a logistic fit (dashed lines). **C**, **F** Distribution of fraction of nodes–composing a newly chosen group–that were previously known to the focal node. The resulting distributions of values, averaged over the different time steps, are plotted in comparison to two null scenarios where the group to join is chosen at random, or at random given the target size.
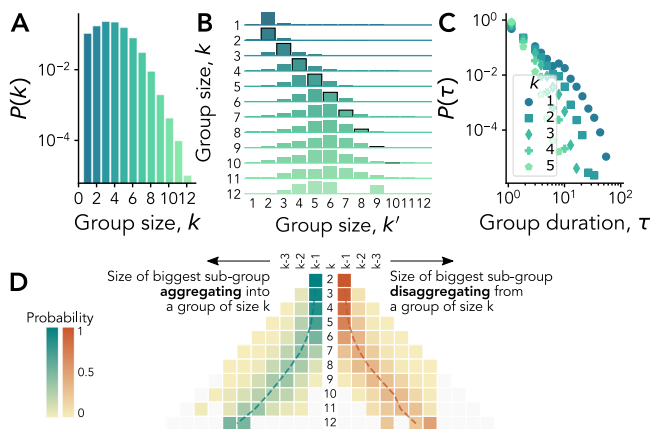

Fig. 5 | The synthetic model reproduces the higher-order dynamical features found in empirical data. Simulated distribution of group size (**A**), node transition matrix (**B**), group duration for different group sizes k (**C**), and pyramidal heatmap associated to the aggregation and disaggregation dynamics (**D**) generated by the proposed temporal hypergraph model.

We show the results of the best-performing model in Fig. 5. All obtained results are in line with the empirical data analysed. The group-size distribution (Fig. 5A) spans a range of values comparable with the empirical observation in Fig. 1A. The group size transition matrix for the dynamics of group changes from the node point of view, shown in Fig. 5B, has similar symmetric patterns for small sizes and a biased transition towards smaller sizes due to the cut-off effect for larger groups, as in Fig. 1C. It is important to note that other group properties, albeit not taken into account for the exploration of the model parameters, are also reproduced. Indeed, the group duration distributions (Fig. 5C) display broad tails, with a similar group size dependency as in the empirical observations of Fig. 2B. More importantly, even complex dynamical characters such as the group disaggregation and aggregation probability distributions, displayed in Fig. 5D, resembles the empirical findings (Fig. 3B), showing the excellent capacity of the proposed model to account for

and reproduce the complex phenomenology of the dynamics of group interactions.

Discussion

We have here analysed human interactions under the lens of group dynamics in two data sets, collected respectively among preschool children and university freshmen students. Despite the inherently different nature of their interactions due to age, contexts, and setting constraints such as class schedules, and despite the differences in data collection techniques, we have uncovered strikingly similar group dynamics both at the individual and group level. In particular, we have observed similar group size and duration distributions, and more importantly, consistent dynamical patterns of individual group transitions and group formation and dissolution phenomena in the two settings and at times corresponding to different activity types. Strikingly, we show in the Supplementary Information that these signatures can also be found in data collected in other contexts, such as social interactions that took place during four different scientific conferences organised by GESIS[61] (see Supplementary Information, Note 1 and Figs. S14, S15, S16, S17, S18, S19). The individuals whose interactions are reported in those data are more heterogeneous than in the university and preschool scenarios analysed here: conference participants cover indeed a broad range of ages, with interactions involving different levels of seniority. Even the context can be considered as a mixture of the in-class and out-of-class settings, as conference schedules usually provide more freedom than classes –with participants often not attending all sessions. Despite all these differences, very similar patterns of group statistics and group changes are observed. It would of course be interesting to extend these results even further by considering still other contexts of human interactions to confirm the generality of such group dynamics and patterns of group size change among humans.

Our analysis and results contribute to the obtention of a more complete representation of social dynamics than the ones limited to pairwise interactions. We have accordingly proposed a synthetic model describing how nodes representing individuals form groups and navigate between groups of different sizes. The model includes mechanisms of short-term memory ("long gets longer") and long-term

social memory (higher probability to join a group including individuals already encountered, see Supplementary Information, Fig. S20), and is able to reproduce the non-trivial dynamics of group changes, both from a node-centric point of view and from the point of view of group formation and break-up. Note that, both when discussing the robustness of the patterns obtained in different empirical contexts, and when validating the model, we have remained at a phenomenological level, for several reasons. First, there is no clearly recognized quantitative measure of distance or similarity between two temporal networks or temporal hypergraphs. Second, the data sets we study have different sizes and maximal group sizes, so that the matrices we look at have different sizes, and the distributions have different cut-offs. Third, we prefer to avoid claims about specific shapes of functions (especially for distributions with broad shapes) or of universality, as such claims are notoriously difficult to establish or disprove. Finally, even if we were to define an ad hoc quantitative measure of difference combining the various observables in an arbitrary manner, we would not have a reference value to compare it to.

Thanks to its realistic group dynamics, our model could be used to generate synthetic substrates for studying the impact of higher-order temporal interactions on dynamical processes. Indeed, while the impact of higher-order interactions on various dynamical processes has been well assessed[13,37], studies on structures undergoing a realistic temporal evolution are scarce[42,62]. The interplay with the dynamics of groups might prove relevant for a wide variety of processes of interest in many contexts, such as the dynamics of adoption[63] and opinion formation[64], but also in synchronisation[65,66], cooperation[40,41,67] and other evolutionary dynamical processes[68]. Overall, our results call for the development of more modelling approaches that explicitly take both the temporal and the many-body nature of social interactions into account, both to understand the mechanisms from which the complex group dynamics emerges, and to investigate the consequences of such group gatherings in collective dynamics. For instance, the model could also be extended to other forms of memory, as explored in pairwise interactions[55]. In addition, the model we implemented relies on a set of minimalistic mechanisms that shape the behaviour of the nodes. Future work may further enrich these rules −for the choice of group change and selection− by integrating homophily-driven decisions and mechanisms of opinion dynamics that would co-evolve together with the social structure[69]. Notably, all these approaches should not be limited to humans, as non-human animals have also shown to be sensible to higher-order social effects[18,70,71] and, at the pairwise interaction level, complex features similar to the ones of human interactions have been observed[72]. Further studies would however be required in order to integrate behavioural response to non-pairwise interactions with additional environmental[73], cultural[74], and ecological factors −like splitting for resource competition, or grouping as a defensive strategy against predators[75]. Indeed, there are cases in which the drivers of animal grouping can have genetic roots[76]. Alternatively, one might try to devise a microfounded principle that could explain the observed temporal evolution of group sizes in term of balancing costs and benefits −akin to evolutionary models used for collective action problems[77]. Along this line, a game-theoretic interpretation has been given for the different group sizes of static hypergraphs constructed from scientific collaborations[78].

Our results inevitably rely on the given definition of group interaction, as constructed from pairwise data, which represents a proxy for real-world group encounters that also depend on the temporal resolution of the data. Given the current lack of a longitudinal data collection effort specifically designed to track group interactions, several research directions could be explored. For example, it would be interesting to check the robustness of the empirical results with respect to other definitions of groups or hyperedges from data obtained by measuring pairwise interactions, such as the Bayesian inference approach to distinguish hyperedges from combinations of lower-order interactions[79], or extraction of statistically significant hyperedges[80]. Even the hardcore definition of group that we used could be challenged, using instead less stringent conditions[81] together with the possibility of having nodes that display multi-membership.

To conclude, our study contributes towards a better understanding of human behaviour in terms of the formation and disaggregation of groups, and of the navigation between social groups. We expect that the analysis presented can support researchers working at the intersection of social and behavioural sciences, while the proposed model can directly be used to inform more realistic simulations of social contagion, norm emergence, and spreading phenomena[34] in interacting populations.

Methods
Data description and pre-processing
Copenhagen network study. We use data collected via the Copenhagen Network Study (CNS)[43] that represents a temporally resolved proximity data collected through the Bluetooth signal of cellular phones carried by 706 freshmen students at the *Technical University of Denmark*. The publicly available data corresponds to the data recorded during four weeks of a semester, and describes proximity of students with a temporal resolution of 5 minutes. The raw data (already pre-processed in Ref. 43) contains 5,474,289 records. Each entry contains a timestamp, the ID of one user (ego), the ID of another user (alter), and the associated Received Signal Strength Indication (RSSI) measured in dBm. The data is already processed to neglect the directionality of each interaction (which device is scanning). Empty scans (no other device found) are reported with a 0 RSSI, which corresponds to isolated nodes.

We split the data records into three main periods according to the hour of the day and the day of the week. Even though the released data[43] do not contain precise information on time and date, these can be easily inferred by cross-checking activity patterns in the temporal sequences with the official timetables for Bachelor studies at DTU[82]. The resulting contexts are:

- *Workweek (in-class)*: Monday to Friday, 8 a.m. to 5 p.m.
- *Workweek (out-of-class)*: Monday to Friday, 12 a.m. (midnight) to 8 a.m. and 5 p.m. to 11.59 p.m.
- *Weekend*: Saturday and Sunday.

We further clean the data in the following way. First, we remove external users by deleting all records in which the device of a participant scanned a device that did not take part in the experiment (resulting in 4,646,415 records). We then retain only records with an RSSI higher than -90dBm [see Supplementary Information, Fig. S21]. This is slightly less restrictive than the threshold of −80dBm used in[83], which was used to select interactions occurring within a radius of 2 meters (a typical distance for social interactions among close acquaintances[84]). After doing this, we have 3,824,052 records divided into 1,603,916 pairwise interactions and 2,220,136 empty scans. We treat the latter as isolated nodes.

We then perform three pre-processing steps as in Ref. 32. First, in order to smooth the pairwise interactions, we look for all the gaps composed by pairwise interactions that are present at times $t-1$ and $t+1$ but not at time t. We fill the resulting 163,349 gaps by using the mean RSSI of the adjacent timestamps (eventually replacing, if present, a record of an empty scan from one of the two interacting nodes).

Second, we filter out spurious interactions by removing all the 130,935 pairwise signals that are present solely at time t but not at times $t-1$ and $t+1$, leaving us with 3,855,139 records. This is also in line with the procedure performed in Ref. 32, which is based on the convention developed by the *Rochester Interaction Record*[85], according to which an encounter needs to last 10 minutes or longer to be classified as meaningful.

Third, we perform triadic closure. Namely, if at time t a user i scans a user j and user j scans a user k, then we also add a record (if not already present) of an active scan between user i and user k. We assign to this interaction the minimum RSSI between the other two. One of the potential pitfalls of performing triadic closure is the addition of many links to events that have already a low RSSI. In particular, if we filter the number of newly added links —due to the triadic closure— by RSSI, we notice that this number scales as a power law with the RSSI [see Supplementary Information, Fig. S22]. In order to avoid closing triangles associated to "weak" events, we select an additional threshold of $-75 dBm$ for the RSSI of the newly added links. This is chosen as the lowest threshold that preserves the group-size distribution across the different contexts [see Supplementary Information, Fig. S23]. Figures Supplementary Information, Figs. S24 and S25 show the impact of the triadic closure —with the chosen threshold— on the number of links and groups in time, respectively. When observed through time, the added links by themselves do not significantly affect the number of tracked links, but help reducing the number of groups —merging together components that would be disconnected otherwise.

As a final step, we check whether the links removed during the procedure involved were the only interactions of an involved node at that particular time (also considering the triadic closure). In this case, we add back the node to the records and declare it as isolated. We finally end up with a pre-processed data set of 3,991,329 records.

DyLNet study. The DyLNet data set was collected with the purpose of observing longitudinally the co-evolution of social network and language development of children in pre-school age. The data collection was carried out in a French preschool by recording the proxy social interactions and voice of 174 children between age 3 to 6 and their teachers and assistants. In this study we rely on data openly shared in[44] and focus on the proximity social interaction data that was recorded over 9 sessions (5 morning and 4 afternoon periods —there are no classes in France on Wednesdays) per week, in 10 consecutive months during a single academic year. The data collection was carried out by using autonomous Radio Frequency Identification (RFID) Wireless Proximity Sensors (employing IEEE 802.15.4 low-rate wireless standard to communicate) installed on participants. Ground-truth data was collected in situ or via controlled experimental settings. Badges broadcasted a 'hello' packet with 0 dBm transmission power for 384 μs every 5 seconds, otherwise they were in listening mode to record the badge ID and RSSI of other proxy badges if the received signal reached the minimum sensitivity value of -94 dBm. Mutually observed badges were paired to indicate proximity interactions, and were further pre-processed to finally obtain an undirected temporal network[45]. Interactions in this network indicate face-to-face proximity of participants within 2 meters with temporal resolution of 5 seconds.

Taking the reconstructed network[44] as a starting point, we remove teachers from the data, restricting our attention to children. The data are also enriched with information. For example, the record of each pairwise interaction comes with a 5 digits label that tracks the context category of each of the two individuals at the beginning and end of the interaction. Leveraging this information, as well as the identity and class membership of each individual, as per the CNS, allows us to split interactions into two categories:

- *in-class*: interactions among children belonging to the same class that starts and finish during class time. Spurious interactions of children belonging to different classes during class time or interactions that start in class and end during the free time are thus removed;
- *out-of-class*: interactions among children of any class that start and finish during the free time. Spurious interactions that start in class and end during free time or viceversa are thus removed.

Differently from the CNS data set, no data collection was performed over the weekends. Although the resolution of the original data is 5 seconds, since there was no central unit synchronising the clocks of the badges attached to each participant (they were in-sync only once per a day), we remove interactions that last for less than 10 seconds. Finally, differently from the CNS, the DyLNet data records do not explicitly include isolated participants, i.e., a child that a given timestamp does not participate to any interaction. We thus "add back" isolated records for each child in all those timestamps in which that child does not interact with other nodes, but such that the child had at least one interaction during the same school session.

Computing node transition matrices

The node transition matrices measure the conditional probabilities of moving across groups of different sizes. Each matrix is constructed —from a node-centric point of view— by counting, for each node, the number of observed transitions at consecutive times across two different groups of sizes k and k'. Let us denote by σ_i^t the group where node i belongs at time t, and by $n_i^t = |\sigma_i^t|$ its size. Let us consider now a random node x at a time τ in which it undergoes a group change. We compute the probability that it changes from a group of size k to a group of size k', $P(n_x^{\tau+1} = k' | n_x^\tau = k)$, as

$$T_{kk'} \equiv P(n_x^{\tau+1} = k' | n_x^\tau = k) = \frac{\sum_{t,i} \delta_{n_i^t k} \delta_{n_i^{t+1} k'} (1 - \delta_{\sigma_i^t \sigma_i^{t+1}})}{k' \sum_{k'',t,i} \delta_{n_i^t k} \delta_{n_i^{t+1} k''} (1 - \delta_{\sigma_i^t \sigma_i^{t+1}})}, \quad (5)$$

where the sum at the numerator takes into account all the transitions of all nodes $i \in \mathcal{V}$ taking place at any time t, from a group of size k to a group of size k', and the sum at the denominator takes into account all such transitions, but to a group of any size k'' ($\delta_{x,y}$ is the Kronecker delta function, i.e., $\delta_{x,y} = 1$ if $x = y$ and zero otherwise). The normalization by the size of the target group in Eq. (5) ensures that changes to groups of different sizes are comparable. Without this, a single node leaving a group of, say, 5 nodes —assuming no further changes to the group— would result in 4 contributions (the remaining nodes) to the transitions from size 5 to size 4.

Computing group aggregation and disaggregation matrices

Studying group aggregation and disaggregation helps us to understand how groups that form/dismantle behave just before/after the event. Each group interaction σ, of size $k \equiv |\sigma|$, is associated to a time of birth t_β and a time of death t_δ defining a temporal span $\tau = t_\delta - t_\beta$ in which all the members of the group stayed continuously together. To study the aggregation and disaggregation phases, we look at how the k members of σ were respectively distributed among groups at $t_\beta - 1$ and $t_\delta + 1$ (if these timestamps are present within the considered context of interaction). In particular, the probability heatmaps shown in Fig. 3 are constructed, for each group size k, from the frequencies of the sizes of the maximal sub-groups of σ right before its birth,

$$\max_{\sigma' \in \mathcal{K}^{t_\beta - 1} : |\sigma'| < |\sigma|} |\sigma \cap \sigma'|, \quad (6)$$

and right after its death,

$$\max_{\sigma' \in \mathcal{K}^{t_\delta + 1} : |\sigma'| < |\sigma|} |\sigma \cap \sigma'|. \quad (7)$$

Notice how we intentionally restrict our attention to the sub-groups of smaller sizes, thus splitting the dynamics into groups that either grow or shrink. Within this dichotomy, for example, a group of size 3 that detaches from a group of size 5 will not contribute to

building the probability distribution associated to the aggregation dynamics for $k = 3$, but only to the disaggregation one for $k = 5$.

Computing group-change probabilities

The group-change probability for each data set and context of interaction is computed by considering for all time steps t and all nodes i the number of times each node, belonging to a group of size k, leaves the group after τ timestamps, over all the possible times. This is defined as:

$$P^{\curvearrowright}(k,\tau) = \frac{\sum_{t,i}\left(1 - \delta_{\sigma_i^{t+1}\sigma_i^t}\right)\prod_{\Delta=0}^{\tau+1}\delta_{\sigma_i^t\sigma_i^{t-\Delta}}\left(1 - \delta_{\sigma_i^t\sigma_i^{t-\tau}}\right)\delta_{n_i^t k}}{\sum_{t,i}\prod_{\Delta=0}^{\tau+1}\delta_{\sigma_i^t\sigma_i^{t-\Delta}}\left(1 - \delta_{\sigma_i^t\sigma_i^{t-\tau}}\right)\delta_{n_i^t k}}. \quad (8)$$

Figure 2 shows the results aggregated over all group sizes, while the full results from Eq. (8) are given in Supplementary Information, Figs. S8 and S9 for the CNS and the DyLNet data, respectively.

Null models for group change

When checking for the presence of *social memory* effects in the empirical and in the synthetic data, we also define two null models for comparison. Let us consider the case of a group change performed by a node i that switches from group ω_i^t at time t to a different group ω_i^{t+1} at time $t + 1$. Notice that, despite the splitting of the datasets into different temporal windows (as given by the different contexts of interaction), we do not have problems at the borders as we restrict our attention to group transitions that were actually recorded in the data. The first baseline scenario we consider is the case of a random selection, in which i chooses instead a random group $\omega \in \mathcal{K}^{t+1}$ uniformly at random from the set of available groups at $t + 1$. Notice that there will always be at least one group to choose from, that is the one found in the data. As a second baseline scenario we add to this random selection a constraint on the size, such that the group is chosen uniformly at random from the subset of the groups in $\mathcal{K}^{t+1}$ that have the same size of the target group found in the empirical transition. This second type of null model does not work for the case of an empirical transition towards a group of unitary size (a node that becomes isolated). All these transitions are thus discarded from the computation. Notice however that this does not jeopardise the comparison as the computation of the density of known nodes in this transition always leads to a 0 —ultimately reducing the differences with the null models.

Model parametrization and fitting

The model is fitted by selecting the best-performing run among different combinations of parameters and with respect to two target observables. In particular, we perform different realisations of the model for different combinations of the parameters $\theta = \{\epsilon, \alpha, n_0\}$ that take values in the intervals $\epsilon \in [1, 30]$, $\alpha \in [0.05, 0.95]$, $n_0 \in [3, 14]$, while keeping constant $N = 700$ (as the number of students at the university), $\beta = 0.8$ (as measured, see Fig. 4), and for a number of time steps equals to $T = 2000$ (notice that each time step involves the activation of every node in a random order). The optimal set of parameters θ^* is selected based on a joint minimisation of the Kullback-Leibler (KL) divergence $D_{\mathrm{KL}}(\cdot \| \cdot)$ with respect to the logarithm of the empirical group-size distribution $\hat{P}(k)$ and the node transition matrix $\hat{T}_{kk'}$:

$$\theta^* = \arg\min_{\theta}\left[\mu D_{\mathrm{KL}}\left(\hat{P}(k)\|P(k|\theta)\right) + [1 - \mu]D_{\mathrm{KL}}\left(\hat{T}_{kk'}\|T(kk'|\theta)\right)\right], \quad (9)$$

with $\mu = 1/2$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The Copenhagen Network Study data are available from the original source[43] at https://doi.org/10.6084/m9.figshare.7267433. The DyLNet data are available from the original source[44] at https://doi.org/10.7303/syn26560886. The GESIS data are available upon request from the original source[61] at https://doi.org/10.7802/2351.

Code availability

The entire analysis was conducted using Python. In particular, the model was coded using the XGI Python library for *compleX Group Interactions*[86]. All scripts and notebooks that support the findings of this study can be found at the Github repository https://github.com/iaciac/temporal-group-interactions[87].

References

1. Wasserman, S. & Faust, K.*Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences (Cambridge University Press, 1994).
2. Lehmann, J., Korstjens, A. H. & Dunbar, R. I. Group size, grooming and social cohesion in primates. *Anim. Behav.* **74**, 1617–1629 (2007).
3. Dunbar, R. I. The anatomy of friendship. *Trends Cogn. Sci.* **22**, 32–51 (2018).
4. Dunbar, R. Structure and function in human and primate social networks: Implications for diffusion, network stability and health. *Proc. R. Soc. A* **476**, 20200446 (2020).
5. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47 (2002).
6. Newman, M. E. The structure and function of complex networks. *SIAM review* **45**, 167–256 (2003).
7. Barrat, A., Barthélemy, M. & Vespignani, A.*Dynamical Processes on Complex Networks* (Cambridge University Press, 2008).
8. Latora, V., Nicosia, V. & Russo, G.*Complex Networks: Principles, Methods and Applications*. Complex Networks: Principles, Methods and Applications (Cambridge University Press, 2017).
9. Vespignani, A. Twenty years of network science (2018).
10. Holme, P. & Saramäki, J. Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).
11. Holme, P. Modern temporal network theory: a colloquium. *Eur. Phys. J. B* **88**, 1–30 (2015).
12. Battiston, F. et al. Networks beyond pairwise interactions: Structure and dynamics. *Phys. Rep.* **874**, 1–92 (2020).
13. Battiston, F. et al. The physics of higher-order interactions in complex systems. *Nat. Phys.* **17**, 1093–1098 (2021).
14. Torres, L., Blevins, A. S., Bassett, D. & Eliassi-Rad, T. The why, how, and when of representations for complex systems. *SIAM Rev.* **63**, 435–485 (2021).
15. Bianconi, G.*Higher-Order Networks*. Elements in Structure and Dynamics of Complex Networks (Cambridge University Press, 2021).
16. Milojević, S. Principles of scientific research team formation and evolution. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3984–3989 (2014).
17. Juul, J. L., Benson, A. R. & Kleinberg, J. Hypergraph patterns and collaboration structure. *Front. Phys.* **11**, 1301994 (2024).
18. Katz, Y., Tunstrøm, K., Ioannou, C. C., Huepe, C. & Couzin, I. D. Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18720–18725 (2011).
19. McGrath, J.*Groups: Interaction and Performance* (Prentice-Hall, 1984).
20. Patania, A., Petri, G. & Vaccarino, F. The shape of collaborations. *EPJ Data Sci.* **6**, 1–16 (2017).
21. Benson, A. R., Abebe, R., Schaub, M. T., Jadbabaie, A. & Kleinberg, J. Simplicial closure and higher-order link prediction. *Proc. Natl. Acad. Sci. USA* **115**, E11221–E11230 (2018).

22. Cencetti, G., Battiston, F., Lepri, B. & Karsai, M. Temporal properties of higher-order interactions in social networks. *Sci. Rep.* **11**, 1–10 (2021).

23. Iacopini, I., Petri, G., Baronchelli, A. & Barrat, A. Group interactions modulate critical mass dynamics in social convention. *Commun., Phys.* **5**, 1–10 (2022).

24. Korbel, J., Lindner, S. D., Pham, T. M., Hanel, R. & Thurner, S. Homophily-based social group formation in a spin glass self-assembly framework. *Phys. Rev. Lett.* **130**, 057401 (2023).

25. Forsyth, D. R.*Group dynamics* (Cengage Learning, 2018).

26. Geard, N. & Bullock, S. Competition and the dynamics of group affiliation. *Adv. Complex Syst.* **13**, 501–517 (2010).

27. Lotito, Q. F., Musciotto, F., Montresor, A. & Battiston, F. Higher-order motif analysis in hypergraphs. *Commun. Phys.* **5**, 1–8 (2022).

28. Mancastroppa, M., Iacopini, I., Petri, G. & Barrat, A. Hyper-cores promote localization and efficient seeding in higher-order processes. *Nat. Commun.* **14**, 6223 (2023).

29. Zhao, K., Stehlé, J., Bianconi, G. & Barrat, A. Social network dynamics of face-to-face interactions. *Phys. Rev. E* **83**, 056109 (2011).

30. Ceria, A. & Wang, H. Temporal-topological properties of higher-order evolving networks. *Sci. Rep.***13**, https://doi.org/10.1038/s41598-023-32253-9 (2023).

31. Gallo, L., Lacasa, L., Latora, V. & Battiston, F. Higher-order correlations reveal complex memory in temporal hypergraphs. *Nat. Commun.* **15**, 4754 (2024).

32. Sekara, V., Stopczynski, A. & Lehmann, S. Fundamental structures of dynamic social networks. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9977–9982 (2016).

33. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591 (2009).

34. Vespignani, A. Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.* **8**, 32–39 (2012).

35. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).

36. Vestergaard, C. L., Génois, M. & Barrat, A. How memory generates heterogeneous dynamics in temporal networks. *Phys. Rev. E* **90**, 042805 (2014).

37. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 2485 (2019).

38. St-Onge, G. et al. Influential groups for seeding and sustaining nonlinear contagion in heterogeneous hypergraphs. *Commun. Phys.* **5**, 1–16 (2022).

39. Papanikolaou, N., Vaccario, G., Hormann, E., Lambiotte, R. & Schweitzer, F. Consensus from group interactions: An adaptive voter model on hypergraphs. *Phys. Rev. E* **105**, 054307 (2022).

40. Sheng, A., Su, Q., Wang, L. & Plotkin, J. B. Strategy evolution on higher-order networks. *Nat. Comput. Sci.* 1–11, https://doi.org/10.1038/s43588-024-00621-8 (2024).

41. Civilini, A., Sadekar, O., Battiston, F., Gómez-Gardeñes, J. & Latora, V. Explosive cooperation in social dilemmas on higher-order networks. *Phys. Rev. Lett.* **132**, 167401 (2024).

42. Chowdhary, S., Kumar, A., Cencetti, G., Iacopini, I. & Battiston, F. Simplicial contagion in temporal higher-order networks. *J. Phys. Complexity* **2**, 035019 (2021).

43. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the copenhagen networks study. *Sci. Data* **6**, 1–10 (2019).

44. Dai, S. et al. Longitudinal data collection to follow social network and language development dynamics at preschool. *Sci. Data* **9**, 1–17 (2022).

45. Dai, S. et al. Temporal social network reconstruction using wireless proximity sensors: model selection and consequences. *EPJ Data Sci.* **9**, 19 (2020).

46. Braha, D. & Bar-Yam, Y. From centrality to temporary fame: Dynamic centrality in complex networks. *Complexity* **12**, 59–63 (2006).

47. Braha, D. & Bar-Yam, Y. Time-dependent complex networks: Dynamic centrality, dynamic motifs, and cycles of social interactions. In *Adaptive networks: Theory, models and applications*, 39–50, https://doi.org/10.1007/978-3-642-01284-6_3 (Springer, 2009).

48. Pedreschi, N. et al. Dynamic core-periphery structure of information sharing networks in entorhinal cortex and hippocampus. *Netw. Neurosci.* **4**, 946–975 (2020).

49. Cattuto, C. et al. Dynamics of person-to-person interactions from distributed rfid sensor networks. *PLoS One* **5**, e11596 (2010).

50. Barrat, A., Cattuto, C., Tozzi, A. E., Vanhems, P. & Voirin, N. Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases. *Clin. Microbiol. Infect.* **20**, 10–16 (2014).

51. Perra, N., Gonçalves, B., Pastor-Satorras, R. & Vespignani, A. Activity driven modeling of time varying networks. *Sci. Rep.* **2**, 1–7 (2012).

52. Starnini, M., Baronchelli, A. & Pastor-Satorras, R. Modeling human dynamics of face-to-face interaction networks. *Phys. Rev. Lett.* **110**, 168701 (2013).

53. Karsai, M., Perra, N. & Vespignani, A. Time varying networks and the weakness of strong ties. *Sci. Rep.* **4**, 1–7 (2014).

54. Nadini, M. et al. Epidemic spreading in modular time-varying networks. *Sci. Rep.* **8**, 1–11 (2018).

55. Le Bail, D., Génois, M. & Barrat, A. Modeling framework unifying contact and social networks. *Phys. Rev. E* **107**, 024301 (2023).

56. Stehlé, J., Barrat, A. & Bianconi, G. Dynamical and bursty interactions in social networks. *Phys. Rev. E* **81**, 035101 (2010).

57. Petri, G. & Barrat, A. Simplicial activity driven model. *Phys. Rev. Lett.* **121**, 228301 (2018).

58. Hatcher, A., Press, C. U. & of Mathematics, C. U. D.*Algebraic Topology*. Algebraic Topology (Cambridge University Press, 2002).

59. Simmel, G. The number of members as determining the sociological form of the group. *Am. J. Sociol.* **8**, 1–46 (1902).

60. Egbert, M. M. Schisming: The collaborative transformation from a single conversation to multiple conversations. *Res. Lang. Soc.* **30**, 1–51 (1997).

61. Génois, M. et al. Combining sensors and surveys to study social interactions: A case of four science conferences. *Pers. Sci.* **4**, 1–24 (2023).

62. Shang, Y. Non-linear consensus dynamics on temporal hypergraphs with random noisy higher-order interactions. *J. Complex Netw.* **11**, cnad009 (2023).

63. Barrat, A., Ferraz de Arruda, G., Iacopini, I. & Moreno, Y. Social contagion on higher-order structures. In *Higher-Order Systems*, 329–346, https://doi.org/10.1007/978-3-030-91374-8_13 (Springer, 2022).

64. Neuhäuser, L., Lambiotte, R. & Schaub, M. T. Consensus dynamics and opinion formation on hypergraphs. In *Higher-Order Systems*, 347–376, https://doi.org/10.1007/978-3-030-91374-8_14 (Springer, 2022).

65. Skardal, P. S. & Arenas, A. Explosive synchronization and multistability in large systems of kuramoto oscillators with higher-order interactions. In *Higher-Order Systems*, 217–232, https://doi.org/10.1007/978-3-030-91374-8_8 (Springer, 2022).

66. Millán, A. P., Restrepo, J. G., Torres, J. J. & Bianconi, G. Geometry, topology and simplicial synchronization. In *Higher-Order Systems*, 269–299, https://doi.org/10.1007/978-3-030-91374-8_11 (Springer, 2022).

67. Traulsen, A. & Nowak, M. A. Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10952–10955 (2006).

68. Perc, M., Gómez-Gardenes, J., Szolnoki, A., Floría, L. M. & Moreno, Y. Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**, 20120997 (2013).

69. Schweitzer, F. & Andres, G. Social nucleation: Group formation as a phase transition. *Phys. Rev. E* **105**, 044301 (2022).

70. Rosenthal, S. B., Twomey, C. R., Hartnett, A. T., Wu, H. S. & Couzin, I. D. Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4690–4695 (2015).

71. Iacopini, I., Foote, J. R., Fefferman, N. H., Derryberry, E. P. & Silk, M. J. Not your private tête-à-tête: leveraging the power of higher-order networks to study animal communication. *Philos. Trans. R. Soc. B: Biol. Sci.* **379**, 20230190 (2024).

72. Gelardi, V., Godard, J., Paleressompoulle, D., Claidière, N. & Barrat, A. Measuring social networks in primates: wearable sensors versus direct observations. *Proc. R. Soc A* **476**, 20190737 (2020).

73. Flierl, G., Grünbaum, D., Levin, S. & Olson, D. From individuals to aggregations: the interplay between behavior and physics. *J. Theor. Biol.* **196**, 397–454 (1999).

74. Conradt, L. & Roper, T. J. Activity synchrony and social cohesion: a fission-fusion model. *Proc. Royal Soc. B* **267**, 2213–2218 (2000).

75. Wittemyer, G., Douglas-Hamilton, I. & Getz, W. M. The socioecology of elephants: analysis of the processes creating multitiered social structures. *Anim. Behav.* **69**, 1357–1371 (2005).

76. Archie, E. A., Moss, C. J. & Alberts, S. C. The ties that bind: genetic relatedness predicts the fission and fusion of social groups in wild african elephants. *Proc. Royal Soc. B* **273**, 513–522 (2006).

77. Gavrilets, S. Collective action problem in heterogeneous groups. *Philos. Trans. R. Soc. B* **370**, 20150016 (2015).

78. Alvarez-Rodriguez, U. et al. Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* **5**, 586–595 (2021).

79. Young, J.-G., Petri, G. & Peixoto, T. P. Hypergraph reconstruction from network data. *Commun. Phys.* **4**, 135 (2021).

80. Musciotto, F., Battiston, F. & Mantegna, R. N. Detecting informative higher-order interactions in statistically validated hypergraphs. *Commun. Phys.* **4**, 218 (2021).

81. Alvarez-Rodriguez, U., Petrović, L. V. & Scholtes, I. Inference of time-ordered multibody interactions. *Phys. Rev. E* **108**, 034312 (2023).

82. of Denmark, T. U. Course base. https://www.dtu.dk/english/education/course-base (2022).

83. Sekara, V. & Lehmann, S. The strength of friendship ties in proximity sensor data. *PLoS One* **9**, e100915 (2014).

84. Hall, E. *The Hidden Dimension* (Anchor Books, 1990).

85. Reis, H. T. & Wheeler, L. Studying social interaction with the rochester interaction record. *Adv. Exp. Soc. Psychol.* **24**, 269–318 (1991).

86. Landry, N. W. et al. Xgi: A python package for higher-order interaction networks. *J. Open Source Softw.* **8**, 5162 (2023).

87. Iacopini, I., Karsai, M. & Barrat, A. The temporal dynamics of group interactions in higher-order social networks. *iaciac/temporal-group-interactions* https://doi.org/10.5281/zenodo.12698353 (2024).

88. Alstott, J., Bullmore, E. & Plenz, D. powerlaw: a python package for analysis of heavy-tailed distributions. *PLoS One* **9**, e85777 (2014).

Author contributions

I.I., M.K., A.B. designed and conceived the study. I.I. performed the data analysis, the implementation of the model, and the numerical simulations. I.I., M.K., A.B. analyzed and discussed the results. I.I., M.K., A.B. wrote the manuscript.

Competing interests

The authors declare that they have no competing interests. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-50918-5.

Correspondence and requests for materials should be addressed to Iacopo Iacopini.

Peer review information *Nature Communications* thanks Jesus Gómez-Gardeñes, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.