Higher-order shortest paths in hypergraphs

Berné L. Nortier [1,2,*] Simon Dobson [1] and Federico Battiston [2,†]
[1]*Department of Computer Science, University of St. Andrews, St. Andrews KY16, Scotland*
[2]*Department of Network and Data Science, Central European University Vienna, Vienna 1100, Austria*

(Received 11 February 2025; accepted 22 September 2025; published 5 November 2025)

One of the defining features of complex networks is the connectivity properties that we observe emerging from local interactions. Recently, hypergraphs have emerged as a versatile tool to model networks with nondyadic, higher-order interactions. Nevertheless, the connectivity properties of real-world hypergraphs remain largely understudied. In this work we introduce path size as a measure to characterize higher-order connectivity and quantify the relevance of nondyadic ties for efficient shortest paths in a diverse set of empirical networks with and without temporal information. By comparing our results with simple randomized null models, our analysis presents a nuanced picture, suggesting that nondyadic ties are often central and are vital for system connectivity, while dyadic edges remain essential to connect more peripheral nodes, an effect which is particularly pronounced for time-varying systems. Our work contributes to a better understanding of the structural organization of systems with higher-order interactions.

DOI: 10.1103/1mxy-3cnl

I. INTRODUCTION

Networks, collections of nodes and their interactions, are one of the primary tools used to study complex systems, allowing us to describe collective, emergent system properties which can not be captured by looking at the individual units in isolation [1]. A typical case is the emergence of global connectivity from local interactions, with many real-world networks characterized by the emergence of large connected components [2]. Many such systems display low diameter and are hence dubbed "small worlds" [3], displaying a relational structure able to support both efficient system-wide communications [4,5] and making them easy to navigate [6]. The efficient structural connectivity of these systems has a profound impact on their functionality, for both stochastic and deterministic processes. The presence of shortcuts is known to influence the speed of contagion and emergence of cooperation [7], impacting also the ability of a system to synchronize [8]. For all these reasons, efficiently computing shortest paths in graphs is a problem that has attracted enormous interest in the research community for many decades [9], and it continues to do so.

The analysis of connectivity has also produced relevant insights in the case of temporal networks, where edges are not permanent but created and destroyed over time [10]. Structures such as paths [11] and connected components [12] have been formulated in the setting where network structure evolves over time, as well as numerous measures of centrality [13]. The addition of a temporal dimension, in particular the presence of nontrivial temporal correlations [14] and memory [15] among interactions, has stimulated important research on increasingly complex temporal network models, able to reproduce empirical patterns [16–18]. Moreover, since the first observations about epidemic processes [19,20], network temporality was found to have interesting consequences also for the network dynamics, influencing the behavior of individuals across a range of processes [21,22].

Networks have classically modeled interactions through links, describing relations between pairs of entities only, even though in many real-world systems interactions simultaneously involve multiple nodes [23]. Recently, a wide variety of structural descriptors have been extended to account for the presence of such higher-order interactions, including algorithms for motif discovery [24] and analysis [25], community detection [26,27], centrality measures [28], as well as network filtering [29] and reconstruction [30] procedures. Nondyadic interactions generate new dynamical behaviors and collective phenomena [31,32], from contagion [33,34] to synchronization [35–37] and evolutionary games [38,39]. More recently, temporality has also been considered in higher-order networks, from burstiness [40] and temporal-topological correlations [41], to structural models with [42,43] and without [44,45] memory.

Despite these advances, characterizing shortest paths and connectivity in systems with higher-order interactions remains an open problem. Recently, efforts have been devoted to characterize the concepts of distance [46] and walks [47] in networks with nondyadic ties, as well as proposing efficient algorithms to extract shortest paths in hypergraphs [48], and randomize hypergraphs preserving shortest path lengths [49]. However, all these are analyses limited to static systems.

In this work, we provide a systematic investigation of the effect of higher-order interactions on system connectivity across a variety of real-world data sets of social interactions. We quantify the contribution of nondyadic ties for shortest path length, computing length distribution, average path size

*Contact author: bln1@st-andrews.ac.uk
†Contact author: battistonf@ceu.edu

and the fraction of purely dyadic segments in each path, and show that all such features are compatible with a simple null model which preserves the higher-order degree distribution. Next, we consider time-varying interactions, and extend the concept of higher-order path and components to temporal hypergraphs. Our analysis of multiple real-world systems reveals the crucial role of higher-order interactions to ensure efficient connectivity in temporal higher-order networks, characterizes differences between topologically shortest and temporally fastest paths, and shows how the observed nontrivial empirical patterns can not be reproduced with simple randomized null models. Our work provides insights into the connectivity properties and the structural organization of real-world hypergraphs, both for static and time-varying systems.

II. RESULTS

A. Data

To study shortest paths in temporal hypergraphs, we collected 24 publicly available data sets of real-world temporal systems with higher-order interactions. Specifically, our study focuses on hypergraphs that describe social interactions from a broad range of domains such as households and schools, hospitals, workspaces and conferences, as well as political interactions between individuals in Congress and even contact data from baboons.

We provide relevant summary statistics for the static versions of each data set in Table I, alongside the amount and resolution of successive time stamps for the temporal networks. A fuller description of the data sets is provided in the Methods section.

TABLE I. Summary statistics of real-world temporal hypergraphs. V indicates the number of nodes, E_{HO} denotes the number of higher-order interactions and E_{DY} the number of purely dyadic interactions in the static system, and $|e^{\mathrm{max}}_{\mathrm{HO}}|$ indicates the maximum size of a higher-order interaction. T measures the number of total time stamps and dt the interval between successive time stamps in the temporal hypergraphs.

Data set	V	E_{HO}	E_{DY}	$\lvert e^{\mathrm{HO}}_{\mathrm{max}}\rvert$	T	dt
Copenhagen	692	449413	38088	21	8064	300 s
Elem1	339	26543	14184	8	2244	20 s
F&F: 2010-08	44	436	163	9	66	8 h
F&F: 2010-09	58	1155	349	11	90	8 h
F&F: 2010-10	128	4832	1293	13	90	8 h
F&F: 2010-11	126	6094	1504	20	90	8 h
F&F: 2010-12	121	4529	1091	21	90	8 h
F&F: 2011-01	118	4070	1056	11	90	8 h
F&F: 2011-02	117	6340	1268	20	90	8 h
F&F: 2011-03	112	6529	1330	10	90	8 h
F&F: 2011-04	112	6787	1320	16	90	8 h
F&F: 2011-05	97	1257	364	8	18	8 h
HS11	126	3642	151	44	610	20 s
HS12	180	7458	378	46	1321	20 s
InVS13	95	10502	2342	43	20129	20 s
InVS15	219	35764	7846	71	21536	20 s
Kenyan	75	972	272	11	42	1 h
LH10	76	1854	1093	5	7639	20 s
LyonSchool	242	12704	7748	5	3100	20 s
Malawi	86	14051	194	24	1127	20 s
Mid1	591	76062	49400	10	2505	20 s
SFHH	403	10541	8268	9	3509	1 h
Thiers13	327	7818	5498	5	7375	1 h
Congress Bills	1718	82873	13845	25	5305	1 months

B. Static hypergraphs

In simple graphs, relational information is described by nodes and edges, which encode pairwise interactions between pairs of nodes only. A *path* between 2 nodes i and j is an ordered, nonrepeating sequence of pairwise edges. Similarly, a hypergraph is a collection of nodes and hyperedges which encode interactions among an arbitrary number of nodes. Hyperedges are called *pure pairwise* or *dyadic* when they correspond to the classic notion of an edge by connecting only 2 nodes and are distinguished from *truly higher-order* interactions where 3 or more nodes participate in an interaction. The *size* of a hyperedge is defined as the number of nodes contained therein. A *hyperpath* or a *higher-order path* between 2 nodes in a hypergraph is defined as an ordered, nonrepeating sequence of hyperedges, where each subsequent step occurs between nodes within the same hyperedge.

The *length* of a higher-order path between any two nodes i and j, ℓ_{ij}, is defined as the number of edges traversed. If no such path exists between 2 nodes, we set $\ell_{ij} = \infty$. Two nodes are *connected* if a path of finite length exists between them. While a pair of nodes can be connected by multiple distinct paths, the *shortest path* between any two nodes in a static network is defined as a path of minimum length, which might not necessarily be unique. Technical details regarding the exact computation of such hyperpaths are provided in the relevant section of Methods. Clearly, hyperpaths in higher-order networks will traverse hyperedges of varying sizes, where each segment of a hyperpath traverses a hyperedge of a potentially different size. As such, one may attach to each individual segment of a hyperpath its associated size s, defined as the size of the hyperedge that is traversed during that particular step. In some cases, such a step may be contained in multiple hyperedges, in which case we choose the hyperedge of smallest size to define its size. Results concerning alternatives, such as selecting the mean or maximum in the case multiple hyperedges, and the frequency with which they occur are reported in Sec. 3 of the Supplemental Material [50]. Taken together, our analyses show that different strategies with regard to the hyperedge to traverse in case of redundancies do not induce a qualitative difference in our findings. Paths in higher-order networks contain rich information and one may calculate the average size $\langle s \rangle$ of a path as the average size across traversed segments. For example, the average size of a path allows one to determine how much of a path makes use of dyads as opposed to genuinely higher-order ($s \geqslant 3$) interactions. The minimum size of a path is $s = 2$, which occurs when the hyperpath consists of purely pairwise interactions.

To illustrate these concepts, we consider the hypergraph in Fig. 1. A shortest hyperpath between nodes B and K has length $\ell = 3$. One of these is the hyperpath (B, F, H, K), where the segments (B, F), (F, H), and (H, K) have sizes 4, 5, and 2. The average size of such a hyperpath is then $\langle s \rangle = 11/3$. We note

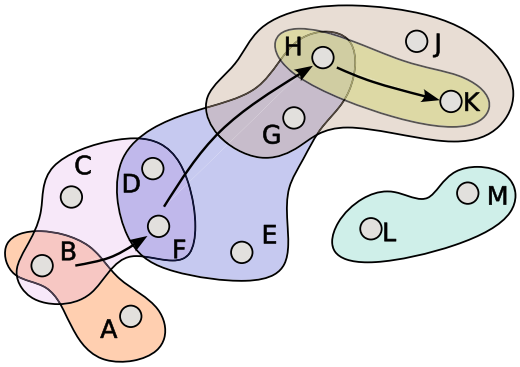

FIG. 1. A shortest hyperpath between nodes B and K has length $\ell = 3$ and average size $\langle s \rangle = (4 + 5 + 2)/3$.

that the segment (H $\rightarrow$ K) is also contained in a hyperedge of size four but that we select the hyperedge of minimum size to assign a size. Importantly, other shortest hyperpaths of the same length can exist. For instance, here the hyperpath (B, D, H, K) also has the same $\langle s \rangle = 11/3$. Another of such shortest hyperpaths is (B, D, G, K) which has a higher average size $\langle s \rangle = 13/3$.

We are interested in studying the higher-order organization of shortest paths and the extent to which group interactions contribute to the forming of these paths in real-world systems.

We begin by considering static hypergraphs where we neglect information about the temporal nature of the interactions (see Methods for details). As an illustrative example, in Fig. 2, we show results for the social hypergraph from the well-known Copenhagen network study, which describes social interactions over 4 weeks between 576 university students on a university campus. Results of analyses on all other data sets are provided in the Supplemental Material, Secs. 2 and 3 [50].

In Fig. 2(a), we plot the distribution of shortest path lengths for higher-order paths in blue. We compare it with the distribution of path lengths of purely dyadic paths, obtained when taking considering only dyadic interactions, shown in gray. When higher-order interactions are not considered, we observe a distribution shift to the right, as such paths take longer to reach the same target as their higher-order counterparts. Interestingly, the same distributions of higher-order and dyadic path lengths are accurately reproduced by randomizing the data with a higher-order configuration model (see Methods section for details), as shown in Fig. 2(b).

Next, in Fig. 2(c), we study the average size $\langle s \rangle$ of shortest paths between connected nodes in the higher-order network as a function of their length ℓ. We observe that shortest paths which are longer also tend to be those with a lower average size. This suggests that higher-order interactions are often central in the system, in agreement with recent work on hypercores [51] while dyadic edges remain crucial to connect more peripheral nodes. Due to the small sample (2) of maximum-length paths, the randomization does not reflect the downward trend at all lengths. To further corroborate our intuition, in Fig. 2(d) we plot the fraction of a shortest path that consists of purely dyadic segments, grouped by the length of the path itself. As expected, the curve is increasing, and such a behavior is again well reproduced in the randomized system. In Fig. 3 we extend the analysis of higher-order shortest paths to all our data sets. We observe that the average size $\langle s \rangle$ as a function of ℓ decreases in all systems, except for a data set about social interactions in an office. We hypothesize that finite-size effects play a role for this workplace data, as the number of paths of length 3 are three orders of magnitude less than the other path length counts.

The inset of Fig. 3 shows for all data sets the difference in average path length for higher-order and purely dyadic paths. As expected, those systems which have $\langle s \rangle \approx 2$ in the main figure for all lengths are those which fall closest to the line $x = y$ in the inset. Finally, we observe that the lengths of shortest paths in all data sets are very short, never having a length of more than 5, in agreement with the small-world nature of complex networks [3].

In summary, while higher-order interactions are important to ensure connectivity in static hypergraphs, in such a simple scenario most shortest path features can be reproduced by preserving the higher-order degree distribution. As we will see in the next section, a much richer picture emerges once we move beyond static systems and expand our analysis to consider the role of time.

C. Temporal hypergraphs

We model the temporal evolution of a networked system with a temporal network. A temporal network is a sequence of consecutive time-stamped static graphs, or alternatively a collection of nodes and a collection of time-stamped edges. At any time t hyperlinks can be either active or inactive.

A *temporal path* is an ordered sequence of successive (edge, activation time) pairs, where consecutive pairs' time stamps must be increasing. An important consequence of temporality is that paths are not symmetric (even in the case of undirected graphs). In other words, the existence of a path from i to j does not imply the existence of a path from j to i and, in general, even when both exist, they might be different and have different lengths. In this work, we assume that a message cannot be propagated further than its direct neighbors within any single snapshot (although the alternative makes sense when time stamps are at longer intervals), and that a node can "retain" a message across multiple snapshots, following the approach of Ref. [13].

The concept of length in the temporal setting may be defined in two complementary ways. The *temporal duration*, denoted d, of a temporal path from node i to node j is defined as the time elapsed between the time stamp at which node i is first active and the time stamp at which node j is active. By contrast, the *length* or step count, denoted ℓ, of a temporal path from nodes i to node j is defined as the total number of temporal edges traversed.

Shortest temporal paths can therefore be either those that have minimum duration ($d_{\min}$), or minimum length ($\ell_{\min}$). The *fastest path* (of shortest duration) between nodes i and j is the temporal path of minimum duration, and is set to infinity if such a path does not exist. The *shortest path* (with the fewest steps) counts the number of steps, and is hence analogous to the static case.

When a system describes time-varying interactions between groups of nodes, we make use of temporal hypergraphs.

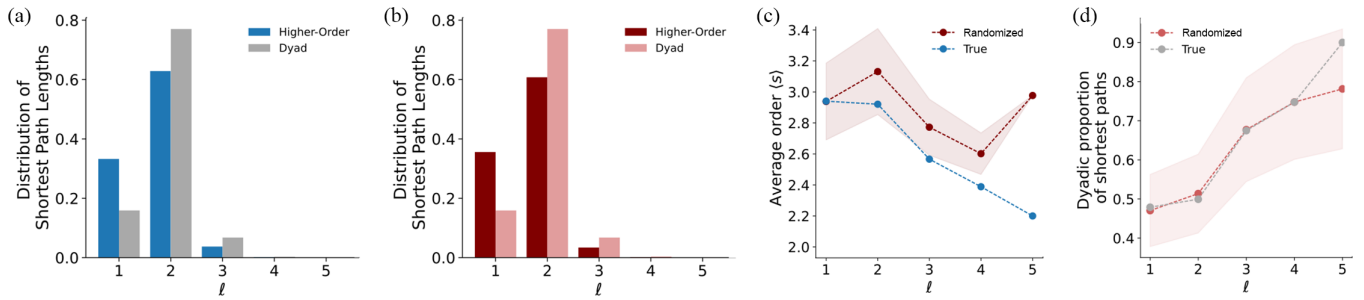

FIG. 2. Shortest paths in the static hypergraph of the Copenhagen study. Path-length distributions of higher-order and purely dyadic paths for hypergraphs of the Copenhagen data set for (a) the empirical data and (b) a randomized null model. (c) Average interaction size $\langle s \rangle$ of the higher-order path as a function of the path length for empirical (blue) and randomized (dark red) data. (d) Average fraction of each path which is purely dyadic as a function of the path length for empirical (gray) and randomized (light red) data. Shaded areas represent a 97.5% confidence interval obtained over 100 randomizations.

A *temporal hypergraph* is a sequence of static hypergraphs, indexed by time. The two notions of temporal length (duration and step count) can be defined as in the pairwise case. In Fig. 4, a simple temporal hypergraph illustrates how such measures can differ. A temporal path consisting of time-stamped links $(A \to C, t_1)$, $(C \to D, t_2)$, $(D \to E, t_4)$, $(E \to G, t_5)$ has a temporal length of 5 units, but a topological length of 4 links. In our temporal analysis, we mainly focus on fastest paths, although we present some results for both fastest and shortest in Sec. 4 of the Supplemental Material [50].

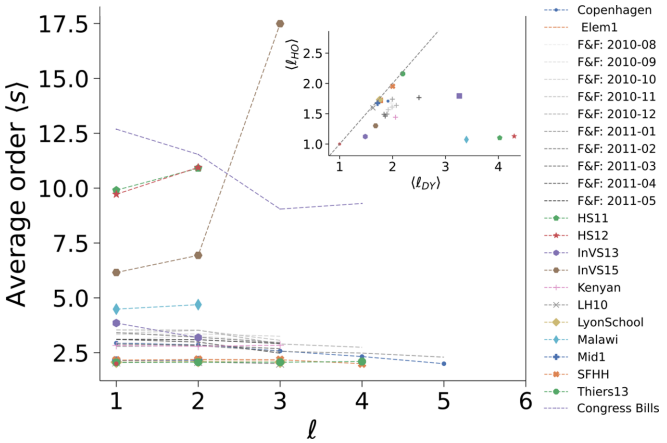

FIG. 3. Summary statistics for shortest paths in static hypergraphs. Average size $\langle s \rangle$ as a function of the path length across 24 data sets describing social interactions. (Inset) Average shortest higher-order path length $\langle \ell_{\mathrm{HO}} \rangle$ against the average shortest dyadic path length $\langle \ell_{\mathrm{DY}} \rangle$ for all data sets.

We investigate temporal connectivity and fastest paths in various real-world systems to discover how much of the system's connectivity is due to temporal higher-order interactions. Similarly to the static case, in Fig. 5, we begin our exploration of the temporal connectivity properties of real-world systems by focusing on the Copenhagen data set. As before, all analyses for the remaining data sets are available in Secs. 2 and 3 of the Supplemental Material [50], and we mention here only that similar trends are observable across the other data sets. In Fig. 5(a) we plot side by side the distribution of durations of fastest paths across all temporal interaction sizes in blue, and for only purely pairwise temporal interactions in gray. We note that the duration of such paths can be very long, even well above 100 time units. When only pairwise interactions are considered, there is a distinct distribution shift to the right of temporal path lengths. Such an effect of higher-order interactions is much greater than what we observed for the path length in static hypergraphs, making explicit the very relevant role of nondyadic connections in determining the connectivity properties of temporal systems.

In Fig. 5(b) we plot the same distribution of durations for a randomization of the data where hyperedges' time information is shuffled but the underlying static graph is kept constant. As shown, such a randomization fails to reproduce the distribution of durations of the actual system, as washing out temporal correlations among hyperedges makes the typical duration of fastest paths in the null model much shorter and narrower than in the real-world data.

Next, in Fig. 5(c) we plot the average size $\langle s \rangle$ of fastest paths between connected nodes in the higher-order network, grouped by their duration, for both real and randomized data.

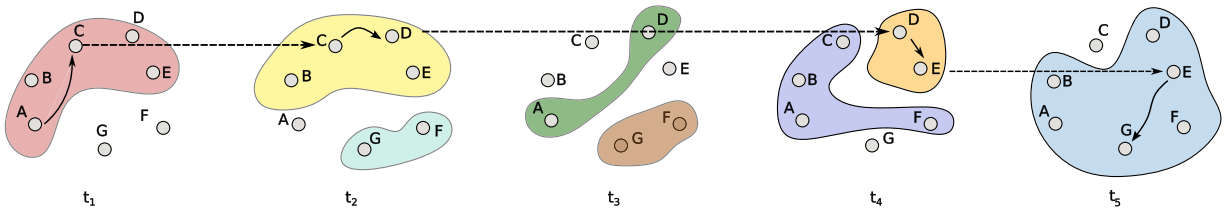

FIG. 4. Example of a temporal path traveling from node A at t_1 to node G at t_5. The path has time duration $d = 5$ and path topological length $\ell = 4$.

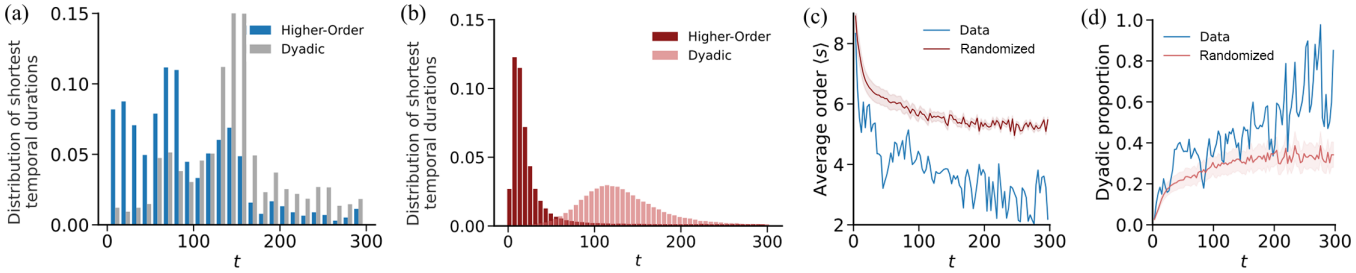

FIG. 5. Fastest paths in the temporal hypergraph of the Copenhagen study. Path-duration distributions of higher-order and purely dyadic paths for hypergraphs of the Copenhagen data set for (a) the empirical data and (b) a randomized null model. (c) Average temporal interaction size $\langle s \rangle$ of the higher-order path as a function of the path duration for empirical (blue) and randomized (dark red) data. (d) Average fraction of each path which is purely dyadic as a function of the path duration for empirical (gray) and randomized (light red) data. Shaded areas represent a 97.5% confidence interval obtained over 100 randomizations.

As in the static case, we observe that paths that last longer also tend to have a lower average size in both the true and randomized networks. Faster paths traverse more higher-order temporal interactions. In contrast, paths with longer duration use more dyadic edges, as illustrated in Fig. 5(d), which shows the average fraction of fastest paths composed of pure dyadic temporal interactions, grouped by path duration. While the null model qualitatively reproduces the observed trends, we observe that fastest paths in the randomized data systematically underestimate the usage of dyadic edges in temporal paths.

In Fig. 6(a) we broaden our investigation to consider fastest paths for all data sets in our collection and plot $\langle s \rangle$ against t, the number of time steps across all data sets. Again, we observe again the same downward trend of decreasing average size with increasing path duration which is even more pronounced when plotting the topological length of fastest paths, and an increasing dyadic proportion for longer path lengths (last two not shown here). In Fig. 6(b), we calculate the proportion of node pairs in each data set that remain connected when dyadic edges are removed, i.e., when they are only allowed to traverse temporal interactions involving groups. While there is variability across real-world systems, we observe that in many cases a high percentage of paths disappear when only pure dyadic segments are considered, highlighting the importance of higher-order interactions for the global architecture of the system. For those nodes that remained connected even when only dyadic temporal interactions were allowed, in Fig. 6(c) we visualize how much longer on average a dyadic temporal path will tend to be than its higher-order counterpart. We note that in a some cases such differences can be very large, such as for the InVS13 and the Baboons data sets.

Path temporality also has consequences for the concept of connected components. Following Ref. [12], we define two nodes as *strongly temporally connected* if there is a temporal path from i to j and also vice versa. A *temporal strongly connected component* is the set of vertices for which any participating node can both reach and be reached by all other nodes in the set. To quantify the reachability of nodes in the systems, we study the temporal connectivity of the data sets in Table II by calculating the number of the strongly connected temporal components and the size of the largest. We compare this with the number and largest size of connected components in the static network. The temporal calculation is known to be an NP-complete problem [13] which is only computationally feasible for smaller graphs, so we present here results only for those networks that we could compute. We observe that the static hypergraph for nearly all data sets has a single giant connected component containing all nodes. In contrast, the temporal systems will be much more fractured, with components that are more numerous and smaller in size.

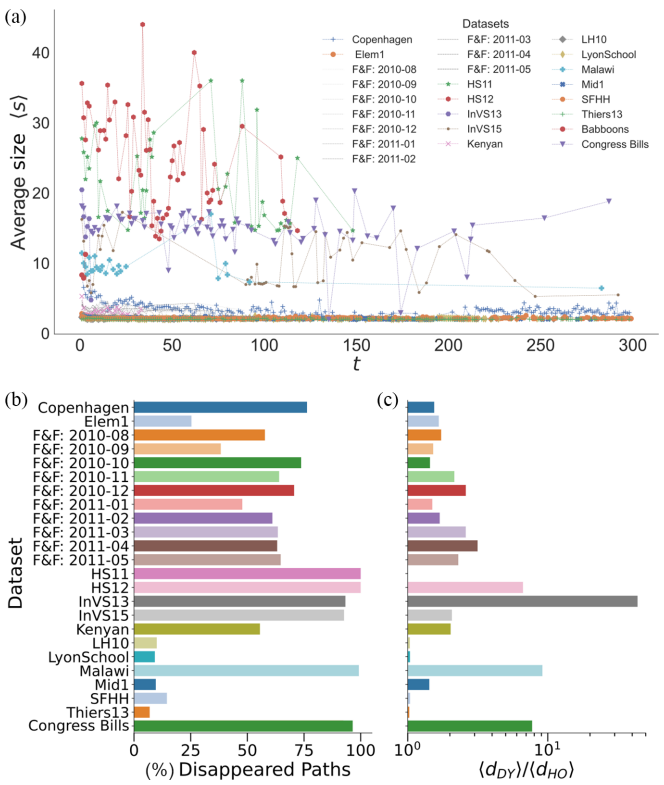

FIG. 6. Features of fastest paths in temporal hypergraphs. Average size $\langle s \rangle$ as a function of the path length across all data sets (a). For each data set, we also show the percentage of fastest paths that only exist as higher-order paths (b), and how much longer in duration the dyadic paths are than the original higher-order paths on average (c).

TABLE II. Connected components in static and temporal hypergraphs. The first column presents the number of connected and temporally strongly connected components (NCC) for static and temporal versions of the dataset. The second column gives the relative size of the largest connected and strongly temporally connected components (|LCC|) when counting included nodes.

	NCC		|LCC|	
Data sets	Static	Temporal	Static	Temporal
Malawi	1	30720	1.00	0.35
LH10	1	4	1.00	0.13
F&F 2010-10	1	96	1.00	0.27
F&F 2010-08	2	992	0.95	0.39
F&F 2010-09	1	128	1.00	0.31
Kenyan	3	692	0.41	0.17
Thiers13	1	50	1.00	0.01

III. CONCLUSION

In this work we have unveiled the higher-order organization of shortest paths in systems with nondyadic interactions and have shown the importance of nondyadic ties for shortest paths. By investigating average path size in many real-world hypergraphs, we have seen that higher-order interactions are crucial to introduce shortcuts into the system, enabling efficient communications beyond system units that would otherwise be disconnected if only pairwise segments were considered. We extended our analysis to the case of temporal hypergraphs, investigating both temporally fastest and topologically shortest paths showing how real-world connectivity cannot be reproduced by simple randomized null models.

In the future it would be interesting to explore how the location and properties of hyperedges impact their contribution to shortest paths, which could be done using the notion of hypercoreness [51] or using more sophisticated reference models such as those that are known to preserve static shortest path lengths across randomization [49]. Another idea to investigate whether truly higher-order links are positioned better than pairwise links could involve counting segment duplication, extending the analysis in S4 of our Supplemental Material [50].

Higher-order connectivity can be defined more generally using, for example, hyperedge overlap [47] or by extending notions of flow to the higher-order case [5]. A future avenue of exploration would be the analysis of our path size metric for networks where connectivity is defined more generally and an investigation of the changes that such definitions would induce [52].

Analyzing large higher-order temporal networks by calculating all shortest and fastest paths has a very high computational cost, as it requires global information about the system. This makes the analysis particularly hard for very dense networks, such as the Copenhagen Network Study, or very large ones. In order to analyze larger systems, it would be important to develop more efficient methods, for instance, based on the development of heuristics for higher-order networks that would allow the calculation of shortest path lengths in an approximate but more efficient way, similar to what was recently done for the problem of higher-order motifs [53]. This would permit a clearer comparison also of the differences between shortest and fastest paths in temporal networks, an analysis that is currently limited by the computational infeasibility of calculating all temporal hyperpaths between 2 nodes. Future work could explore this avenue, considering for example the comparison of subsets of paths between nodes and future versions of themselves instead of between all nodes or other sampling schemes.

Another possible extension of the work would be to investigate the effect of the size of the time window considered in our calculations on the emerging connectivity, as previously done for the case of burstiness [19,40]. Finally, a study of the contribution of hyperlinks of different orders to the shortest paths might be used to develop alternative attack strategies to dismantle the system by affecting shortest paths.

Our work investigates the connectivity properties of static and time-varying hypergraphs, contributing to a better understanding of the structural organization of systems with higher-order interactions.

IV. METHODS

A. Data sets

This section provides more detailed descriptions of the 24 data sets used in the above analyses.

Thirteen of the data sets [54–57] [*Copenhagen, Friends & Family (monthly between August 2010–May 2011), Kenyan, Malawi*] describe generic face-to-face interactions between individuals. The Friends & Family data set is particularly rich as it spans an extended time period while simultaneously storing information at a very high time resolution. We therefore opt to split it into 10 smaller data sets, each tracking the temporal evolution over one calendar month. Additionally, following the procedure for intermediate aggregation as described below, we group time frames into blocks of 8 h and create a static hypergraph for each interval.

Six of the data sets [58–60] (*LyonSchool and Thiers13, HS11 and HS12, Elem1 and Mid1*) describe interactions between pupils at school. Five data sets [58,61,62] describe interaction patterns within diverse work environments, namely, a hospital (*LH10, DAWN*), a workspace (*InVS13 and InVS15*), and a conference (*SFHH*). For all of these data sets, nodes correspond to individuals and hyperedges correspond to group interactions.

The final data set (*Congress bills*) describes political interactions between the U.S. House of Representatives and Senate [62]. Nodes here represent congresspersons and each hyperedge contains all sponsors and cosponsors of a specific legislative bill.

In Sec. 5 of the Supplemental Material [50], we additionally perform the same analyses on data describing coauthorship within geological journals across time, which is not inherently social but still a classic example of a setting in which higher order is natural.

B. Higher-order reconstruction

We use 24 freely available data sets for our analyses describing group interactions. Even though interactions involve

groups, the majority of the data sets store interactions between individuals as (time, node, node)-tuples as a computational simplification even when interactions involve groups. We are here interested in investigating the different contributions of higher-order and dyadic edges to the higher-order organization of networks. To this end, we assume that if k individuals are all pairwise connected at time t (i.e., they form a clique), then together they form a group of size k and are promoted to a hyperedge of size k in the temporal hypergraph.

Throughout our analyses, we distinguish between "pure" dyadic interactions and those which are an artifact of the data structures chosen to store the data digitally. We consider an edge to be a "pure pairwise" interaction when it does not form part of any larger clique, i.e., when it cannot be promoted to a group.

We create the static hypergraph by omitting temporal information and generating a hypergraph with all nodes and interactions reported. In effect, this collates all snapshots into a single network. This allows for potentially nested edges to exist in cases where a group of individuals at one time lose or gain members at an earlier or later time, even though in general we assume nested edges do not exist due to the way in which we create hyperedges.

C. Temporal aggregation in higher-order networks

To investigate the temporal evolution of social systems, we construct a temporal network consisting of a sequence of static hypergraphs, each associated with a specific time stamp. In some data sets we have very fine-grained temporal information with data logged at intervals of between 20 s and 5 min and which implies that many of the static hypergraphs are often extremely sparse. As a result, the temporal network will also be much more disconnected as nodes are active much less frequently. Following a procedure standard in the literature [10,63], we address this by performing a preprocessing step of coarse graining where we aggregate multiple smaller snapshots into a single bigger window (e.g., grouping together snapshots of resolution 30 s into groups of 5 min). The choice of aggregation window is determined in a data-driven manner by plotting the size of the largest connected component as a function of the size of the aggregation window and taking note of when its size saturates. We select the smallest window for which the largest connected component no longer grows.

The duration of a temporal path from node i at some time t_1 to node j is defined as the time elapsed between time t_1 and the first time t_n that node j appears in a snapshot (i.e., becomes active) after t_1.

In the case that multiple shortest and fastest paths exist between a given node pair, we select a single path uniformly and at random. Whenever multiple hyperedges connect a single segment of a path, we select the hyperedge of minimum size and assign to the segment its size motivated by the intuition that smaller groups will more accurately transmit information as there are fewer places for information to flow towards [5]. We more fully study the impact of choosing the minimum instead of alternatives in Sec. 3 of the Supplemental Material [50].

D. Calculating shortest paths in static hypergraphs

Shortest paths in hypergraphs may be calculated by relating them to the graph shortest path problem via a clique expansion of the hypergraph.

Specifically, a shortest path between any node pair in a hypergraph is obtained by first transforming a static hypergraph to an undirected graph by placing a pairwise edge between 2 nodes whenever they share the same hyperedge (in effect, the clique expansion of the hypergraph). We then apply Dijkstra's shortest path algorithm [64] to find the shortest paths in the pairwise graph. Finally, we assign to each pairwise link (a, b) in the shortest path the associated hyperedge that contains the pair.

E. Calculating minimum-length paths

Calculating fastest paths in temporal networks is a nontrivial task. To enable us to calculate time-respecting paths, we first map a temporal higher-order network to a static digraph representation whose nodes are now tuples of (time, node) pairs from the original temporal hypergraph [65,66]. We may then easily extract temporal paths from this new digraph. To calculate minimum-length topological paths, we assign a weight of 1 to all edges and employ Dijkstra's shortest path algorithm [64]. To calculate minimum-duration temporal paths, we weight edges by the time interval traversed and again use Dijkstra's algorithm. As before, in the presence of multiple paths we resolve ambiguities by picking one uniformly and at random. If a node "stores" a message across t time steps of length dt, this node is considered to contribute $t\,dt$ units to the temporal path length.

Moreover, calculating shortest paths for large networks is computationally expensive, so we opt to pick a starting time t_0 and window length w and label V_{t_0} as the starting node set. In our case, all data sets have a window size of 300 time steps, except for the Friends & Family data set, which spans the entire calendar month. We then determine the existence and length of fastest and shortest paths from this starting set V_{t_0} to all other nodes in the set V within a timeframe of w consecutive time snapshots. As before, nonexistent paths are set to have a path length of ∞.

F. Randomizations

We consider randomizations of the data in both the static and temporal settings to compare our results. For both cases we generate 100 realizations of the particular null model and plot the resulting averages and 97.5% confidence intervals in the figures. For the static hypergraph, we use the higher-order configuration model, an extension of the well-known configuration model that keeps the degree distribution fixed across all orders and generates a new edge set. For temporal hypernetworks, we use a randomization that shuffles the time stamps of instantaneous events inside individual timelines to create a rearranged version while still preserving both the underlying static hypergraph and the total number of events, following the methodology of [67].

ACKNOWLEDGMENTS

F.B. acknowledges support from the Austrian Science Fund (FWF) through project 10.55776/PAT1052824 and project 10.55776/PAT1652425. B.L.N. gratefully acknowledges the hospitality of the Central European University during the period in which this manuscript was completed. The computational results have been achieved in part using the Austrian Scientific Computing (ASC) infrastructure.

DATA AVAILABILITY

The data that support findings of this article are openly available [68]. The code to perform the analyses described will soon be publicly available as implemented functions within HypergraphX [69], an open-source Python library for higher-order network analysis.

[1] M. Newman, *Networks* (Oxford University Press, Oxford, 2018), p. 07.

[2] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, Rand. Struct. Alg. **6**, 161, (1995).

[3] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world'networks, Nature (London) **393**, 440 (1998).

[4] V. Latora and M. Marchiori, Efficient behavior of small-world networks, Phys. Rev. Lett. **87**, 198701 (2001).

[5] E. Estrada and N. Hatano, Communicability in complex networks, Phys. Rev. E **77**, 036111 (2008).

[6] J. M. Kleinberg, Navigation in a small world, Nature (London) **406**, 845 (2000).

[7] F. C. Santos, J. F. Rodrigues, and J. M. Pacheco, Epidemic spreading and cooperation dynamics on homogeneous small-world networks, Phys. Rev. E **72**, 056128 (2005).

[8] M. Barahona and L. M. Pecora, Synchronization in small-world systems, Phys. Rev. Lett. **89**, 054101 (2002).

[9] E. W. Dijkstra, A note on two problems in connexion with graphs, *Edsger Wybe Dijkstra: His Life, Work, and Legacy* (Association for Computing Machinery, New York, 2022), pp. 287–290.

[10] P. Holme and J. Saramäki, Temporal networks, Phys. Rep. **519**, 97 (2012).

[11] H. Wu, J. Cheng, S. Huang, Y. Ke, Y. Lu, and Y. Xu, Path problems in temporal graphs, Proc. VLDB Endow. **7**, 721 (2014).

[12] V. Nicosia, J. Tang, M. Musolesi, G. Russo, C. Mascolo, and V. Latora, Components in time-varying graphs, Chaos **22**, 023101 (2012).

[13] V. Nicosia, J. Tang, C. Mascolo, M. Musolesi, G. Russo, and V. Latora, Graph metrics for temporal networks, *Temporal Networks* (Springer, Berlin, 2013), pp. 15–40.

[14] M. Karsai, K. Kaski, A.-L. Barabási, and J. Kertész, Universal features of correlated bursty, Sci. Rep. **2**, 397 (2012).

[15] O. E. Williams, L. Lacasa, A. P. Millán, and V. Latora, The shape of memory in temporal networks, Nat. Commun. **13**, 499 (2022).

[16] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, Activity driven modeling of time varying networks, Sci. Rep. **2**, 469 (2012).

[17] L. Lacasa, J. P. Rodriguez, and V. M. Eguiluz, Correlations of network trajectories, Phys. Rev. Res. **4**, L042008 (2022).

[18] O. E. Williams, P. Mazzarisi, F. Lillo, and V. Latora, Non-Markovian temporal networks with auto-and cross-correlated link dynamics, Phys. Rev. E **105**, 034301 (2022).

[19] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, Small but slow world: How network topology and burstiness slow down spreading, Phys. Rev. E **83**, 025102(R) (2011).

[20] L. E. C. Rocha, F. Liljeros, and P. Holme, Simulated epidemics in an empirical spatiotemporal network of 50, 185 sexual contacts, PLoS Comput. Biol. **7**, e1001109 (2011).

[21] I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C. J. Tessone, and F. Schweitzer, Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks, Nat. Commun. **5**, 5024 (2014).

[22] M. Starnini, J. P. Gleeson, and M. Boguñá, Equivalence between non-Markovian and Markovian dynamics in epidemic spreading processes, Phys. Rev. Lett. **118**, 128301 (2017).

[23] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Networks beyond pairwise interactions: Structure and dynamics, Phys. Rep. **874**, 1 (2020).

[24] Q. F. Lotito, F. Musciotto, A. Montresor, and F. Battiston, Higher-order motif analysis in hypergraphs, Commun. Phys. **5**, 79 (2022).

[25] P. Mann, V. A. Smith, J. B. O. Mitchell, and S. Dobson, Random graphs with arbitrary clustering and their applications, Phys. Rev. E **103**, 012309 (2021).

[26] A. Eriksson, T. Carletti, R. Lambiotte, A. Rojas, and M. Rosvall, Flow-based community detection in hypergraphs, *Higher-Order Systems* (Springer, Berlin, 2022), pp. 141–161.

[27] N. Ruggeri, M. Contisciani, F. Battiston, and C. De Bacco, Community detection in large hypergraphs, Sci. Adv. **9**, eadg9159 (2023).

[28] A. R. Benson, Three hypergraph eigenvector centralities, SIAM J. Math. Data Sci. **1**, 293 (2019).

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

[30] J.-G. Young, G. Petri, and T. P. Peixoto, Hypergraph reconstruction from network data, Commun. Phys. **4**, 135 (2021).

[31] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. F. de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno *et al.*, The physics of higher-order interactions in complex systems, Nat. Phys. **17**, 1093 (2021).

[32] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, What are higher-order networks? SIAM Rev. **65**, 686 (2023).

[33] I. Iacopini, G. Petri, A. Barrat, and V. Latora, Simplicial models of social contagion, Nat. Commun. **10**, 2485 (2019).

[34] G. F. de Arruda, G. Petri, and Y. Moreno, Social contagion models on hypergraphs, Phys. Rev. Res. **2**, 023032 (2020).

[35] P. S. Skardal and A. Arenas, Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching, Commun. Phys. **3**, 218 (2020).

[36] A. P. Millán, J. J. Torres, and G. Bianconi, Explosive higher-order Kuramoto dynamics on simplicial complexes, Phys. Rev. Lett. **124**, 218301 (2020).

[37] Y. Zhang, M. Lucas, and F. Battiston, Higher-order interactions shape collective dynamics differently in hypergraphs and simplicial complexes, Nat. Commun. **14**, 1605 (2023).
[38] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, and V. Latoram, Evolutionary dynamics of higher-order interactions in social networks, Nat. Human Behav. **5**, 586 (2021).
[39] A. Civilini, O. Sadekar, F. Battiston, J. Gómez-Gardeñes, and V. Latora, Explosive cooperation in social dilemmas on higher-order networks, Phys. Rev. Lett. **132**, 167401 (2024).
[40] G. Cencetti, F. Battiston, B. Lepri, and M. Karsai, Temporal properties of higher-order interactions in social networks, Sci. Rep. **11**, 7028 (2021).
[41] A. Ceria and H. Wang, Temporal-topological properties of higher-order evolving networks, Sci. Rep. **13**, 5885 (2023).
[42] L. Gallo, L. Lacasa, V. Latora, and F. Battiston, Higher-order correlations reveal complex memory in temporal hypergraphs, Nat. Commun. **15**, 4754 (2024).
[43] I. Iacopini, M. Karsai, and A. Barrat, The temporal dynamics of group interactions in higher-order social networks, Nat. Commun. **15**, 7391 (2024).
[44] G. Petri and A. Barrat, Simplicial activity driven model, Phys. Rev. Lett. **121**, 228301 (2018).
[45] L. Di Gaetano, F. Battiston, and M. Starnini, Percolation and topological properties of temporal higher-order networks, Phys. Rev. Lett. **132**, 037401 (2024).
[46] E. Vasilyeva, M. Romance, I. Samoylenko, K. Kovalenko, D. Musatov, A. M. Raigorodskii, and S. Boccaletti, Distances in higher-order networks and the metric structure of hypergraphs, Entropy **25**, 923 (2023).
[47] S. G. Aksoy, C. Joslyn, C. O. Marrero, B. Praggastis, and E. Purvine, Hypernetwork science via high-order hypergraph walks, EPJ Data Sci. **9**, 16 (2020).
[48] J. Gao, Q. Zhao, W. Ren, A. Swami, R. Ramanathan, and A. Bar-Noy, Dynamic shortest path algorithms for hypergraphs, IEEE/ACM Trans. Netw. **23**, 1805 (2015).
[49] K. Nakajima, K. Shudo and N. Masuda, Randomizing hypergraphs preserving degree correlation and local clustering, IEEE Trans. Netw. Sci. Eng. **9**, 1139 (2022).
[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/1mxy-3cnl for the additional analyses and results.
[51] M. Mancastroppa, I. Iacopini, G. Petri, and A. Barrat, The structural evolution of temporal hypergraphs through the lens of hyper-cores, EPJ Data Sci. **13**, 50 (2024).
[52] P. Van Mieghem and H. Wang, The observable part of a network, IEEE/ACM Trans. Netw. **17**, 93 (2008).
[53] Q. F. Lotito, F. Musciotto, F. Battiston, and A. Montresor, Exact and sampling methods for mining higher-order motifs in large hypergraphs, Computing **106**, 475 (2024).
[54] P. Sapiezynski, A. Stopczynski, D. D. Lassen, and S. Lehmann, Interaction data from the copenhagen networks study, Sci. Data **6**, 315 (2019).
[55] N. Aharony, W. Pan, C. Ip, I. Khayal, and A. Pentland, Social fmri: Investigating and shaping social mechanisms in the real world, Pervasive Mob. Comput. **7**, 643 (2011).
[56] M. C. Kiti, M. Tizzoni, T. M. Kinyanjui, D. C. Koech, P. K. Munywoki, M. Meriac, L. Cappa, A. Panisson, A. Barrat, C. Cattuto *et al.*, Quantifying social contacts in a household setting of rural kenya using wearable proximity sensors, EPJ Data Sci. **5**, 21 (2016).
[57] L. Ozella, D. Paolotti, G. Lichand, J. P. Rodríguez, S. Haenni, J. Phuka, O. B. Leal-Neto, and C. Cattuto, Using wearable proximity sensors to characterize social contact patterns in a village of rural malawi, EPJ Data Sci. **10**, 46 (2021).
[58] M. Génois and A. Barrat, Can co-location be used as a proxy for face-to-face contacts? EPJ Data Sci. **7**, 11 (2018).
[59] J. Fournet and A. Barrat, Contact patterns among high school students, PLoS One **9**, e107878 (2014).
[60] D. J. A. Toth, M. Leecaster, W. B. P. Pettey, A. V. Gundlapalli, H. Gao, J. J. Rainey, A. Uzicanin, and M. H. Samore, The role of heterogeneity in contact timing and duration in network models of influenza spread in schools, J. R. Soc. Interface **12**, 20150279 (2015).
[61] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, Simplicial closure and higher-order link prediction, Proc. Natl. Acad. Sci. USA **115**, E11221 (2018).
[62] J. H. Fowler, Legislative cosponsorship networks in the us house and senate, Comput. Social Netw. **28**, 454 (2006).
[63] N. Masuda and R. Lambiotte, *A Guide to Temporal Netw.* (World Scientific, Singapore, 2016).
[64] T. H. Cormen, Dijkstra's Algorithm, *Introduction to Algorithms*, Sec. 24.3 (MIT Press, Cambridge, MA, 2001), p. 595.
[65] E. Valdano, L. Ferreri, C. Poletto, and V. Colizza, Analytical computation of the epidemic threshold on temporal networks, Phys. Rev. X **5**, 021005 (2015).
[66] K. Sato, M. Oka, A. Barrat, and C. Cattuto, Predicting partially observed processes on temporal networks by dynamics-aware node embeddings (DyANE), EPJ Data Sci. **10**, 22 (2021).
[67] L. Gauvin, M. Génois, M. Karsai, M. Kivela, T. Takaguchi, E. Valdano, and C. L. Vestergaard, Randomized reference models for temporal networks, SIAM Rev. **64**, 763 (2022).
[68] See https://github.com/joanne-b-nortier/higher-order-shortest-paths for data access.
[69] Q. F. Lotito, M. Contisciani, C. De Bacco, L. Di Gaetano, L. Gallo, A. Montresor, F. Musciotto, N. Ruggeri, and F. Battiston, Hypergraphx: A library for higher-order network analysis, J. Comp. Netw. **11**, cnad019 (2023).