



OPEN

Characterization of interactions' persistence in time-varying networks

Francisco Bauzá Mingueza^{1,3}, Mario Floría^{3,4}, Jesús Gómez-Gardeñes^{3,4}, Alex Arenas² & Alessio Cardillo^{2,3,5}✉

Many complex networked systems exhibit volatile dynamic interactions among their vertices, whose order and persistence reverberate on the outcome of dynamical processes taking place on them. To quantify and characterize the similarity of the snapshots of a time-varying network—a proxy for the persistence,—we present a study on the persistence of the interactions based on a descriptor named *temporality*. We use the average value of the temporality, \bar{T} , to assess how “special” is a given time-varying network within the configuration space of ordered sequences of snapshots. We analyse the temporality of several empirical networks and find that empirical sequences are much more similar than their randomized counterparts. We study also the effects on \bar{T} induced by the (time) resolution at which interactions take place.

Over the last decades, complex networks have been used successfully to study a wide range of complex systems: from biological to technological systems, from social to economical ones just to cite a few^{1,2}. Nevertheless, the evolving nature of many complex systems at timescales of specific studies still requires of quantification tools^{3–5}.

The evolution of interactions between agents in complex systems over time does not only affect the structural properties of networked systems⁶, but also the dynamics taking place on them. Indeed, it has been found that time-varying interactions change the behaviour of dynamical processes like: epidemic spreading^{7–9}, diffusion¹⁰, synchronization^{11–13}, pattern formation¹⁴, and evolutionary game theory¹⁵. In particular, the speed of the variation of the interactions plays a pivotal role on the outcome of the dynamics¹⁶. In some cases, the time-scales at which the interactions and the dynamics evolve are distinct¹⁷, allowing the system to be studied under either the *quenched* (i.e., static) approximation^{18,19} or the annealed one^{20,21}. The former approach is more suitable when the dynamics evolves much faster than the network's structure (which can be thought as if it is static), whereas the latter approach (which leverages a well-mixing approach) is more appropriate in the opposite scenario, since it is equivalent to the case where every individual contacts a sufficient number of individuals to have information on the overall system state. More often, the two time-scales are not distinguishable^{22–24}, thus requiring more sophisticated techniques to study first, and understand then, the phenomenology observed.

A key feature of the evolution of interactions is their degree of *persistence* (particularly at a short-range in time). It is known that the persistence has effects on different types of dynamics such as evolutionary game theory^{25,26}, synchronization²⁷, and diffusive processes^{8,28}, as well as on the properties of communication patterns among people²⁹. Beyond that, recent studies have highlighted a clear relationship between the interactions' persistence and the differentiation, in networks, between the main backbone of the interactions and the noise— or spurious interactions,—that simply “switch on and off”^{30,31}.

Thus, the intricate interplay between the evolution of the interactions and the dynamics taking place on a complex network calls for a deeper understanding of time-varying interactions' characteristics, with special attention on their persistence. One way to achieve such a goal would be by answering to the question: *How special/rare is the observed temporal order of the interactions occurring in a time-varying network?* To solve this conundrum, we propose to set: (i) A null hypothesis/model to be used for generating a benchmark interactions' order. (ii) A meaningful property capturing the features of the temporal order of the interactions of a time-varying network.

Concerning the former, we assume that the underlying and invariant feature of a time-varying network is the set of interactions taking place at a given instant i.e., the set of N_s snapshot graphs (see Fig. 1a) constituting

¹Department of Theoretical Physics, University of Zaragoza, 50006 Zaragoza, Spain. ²Department of Computer Science and Mathematics, University Rovira i Virgili, 43007 Tarragona, Spain. ³GOTHAM Lab, Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, 50018 Zaragoza, Spain. ⁴Department of Condensed Matter Physics, University of Zaragoza, 50006 Zaragoza, Spain. ⁵Internet Interdisciplinary Institute (IN3), Open University of Catalonia, 08018 Barcelona, Spain. ✉email: acardillo@uoc.edu

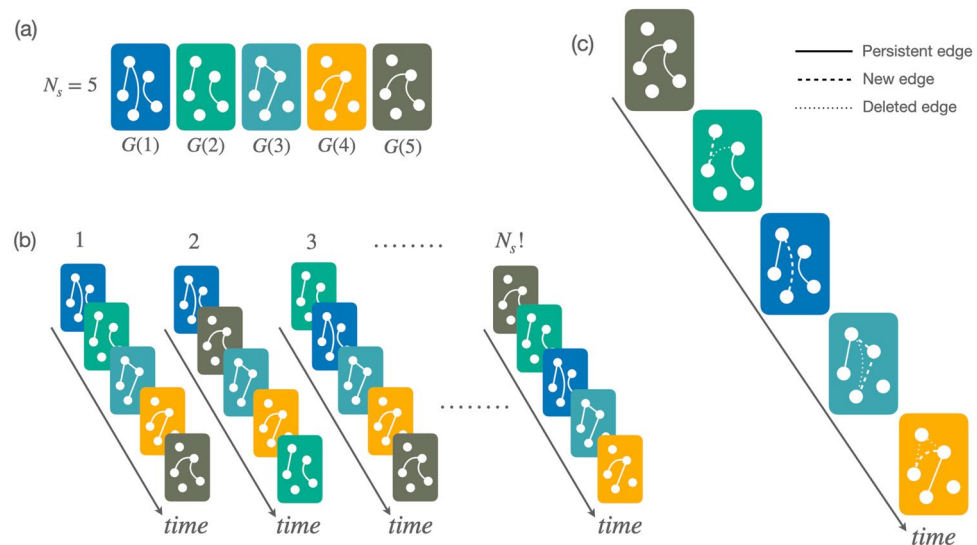


Figure 1. Schematic illustration of the basic features of our framework. **(a)** A simple collection of $N_s = 5$ snapshot graphs constituting the building-blocks of a time-varying network. In **(b)**, the former snapshots are randomly shuffled yielding different instances of the time-varying graphs. The real graph corresponds to one of the $N_s!$ possible orderings. Finally, in **(c)** we analyse a given sequence of snapshots comparing each snapshot with the previous one, highlighting: (i) edges that persist (solid lines), (ii) the new edges (dashed lines), and (iii) removed edges (dotted lines).

the time-varying network. The ordered sequence of snapshots is a well-defined ordination out of all the possible ones. Such a relationship can be thought as the outcome of a shuffling process taking place in the configuration space of all the possible ordinations, with a specific sequence corresponding to one point of such a space^{8,16}. The volume of the configuration space is equal to the number of possible configurations N_{conf} , which—in the absence of additional assumptions,—is the number of permutations of the N_s snapshots, $N_s!$ (see Fig. 1b). Under these premises, assessing how special/rare a time-varying network is, corresponds to nothing else than computing the probability of obtaining an ordination with a given property within the configuration space. This physical-statistical framework—and the idea of a random reshuffling of time snapshots as a generator of the configuration space,—takes advantage of a group of models and processes, known within the literature on the topic with the name of Microcanonical Randomized Reference Model (MRRM) (see³² and references therein).

Regarding the choice of a descriptor encapsulating the features of the snapshots' ordination, an ideal candidate should be able to capture one of the main properties of the latter: how interactions—i.e., the set of edges,—evolve from one snapshot to the next one (see Fig. 1c). Over the years several metrics have been proposed (see^{33,34} and references therein), but we decided to use an indicator, \mathcal{T} , named *temporality*²⁶. Given an ordered sequence of N_s snapshots, $G(t)$, we can associate one value of \mathcal{T} to each of its $N_s - 1$ pairs of adjacent snapshots. Such a sequence (i.e., time series) of values can be thought as a fingerprint of $G(t)$ itself. Hence, we can assume that the properties of the distribution of the values of the temporality series can be used as a proxy to characterize the ordering of $G(t)$. In particular, the average value of the temporality $\overline{\mathcal{T}}$ —computed over the sequence of $N_s - 1$ values,—is enough to grasp the main features of the snapshots' ordination.

Leveraging these assumptions, in this work we characterize the persistence of the interactions of several empirical time-varying networks. In particular, we compare the average value of the temporality with the same quantity computed for a random order of the snapshots. The latter can be computed both analytically from the raw data, as well as numerically via sampling the configuration space. We also estimate the boundaries of the configuration space by identifying the configurations corresponding to the maximum and minimum average temporality. Finally, we perform a coarse-graining of the sequences to study the effects of time resolution on the interactions' persistence.

Our results show that same values of average empirical temporality can stem from different mechanisms of persistence, and that only the comparison with a null model allows to discriminate them. Moreover, we have observed that analysing the system at different (time) resolutions highlights stark differences in the evolution of the persistence at different time scales. Such differences exist even among systems of the same kind, casting doubts on the generally accepted idea that systems of the same type are similar (from our perspective, at least).

Results

We divide the characterization of the persistence of the interactions in time-varying networks in two main parts, as explained in the “Methods”. First, we characterize the “raw” sequence (i.e., without aggregation) by comparing its average temporality with its “randomly shuffled” counterpart. Then we repeat the comparison but, this time, using the aggregated version of the sequences. A brief description of the datasets used in our study is available in the section named “Data” of the “Methods”.

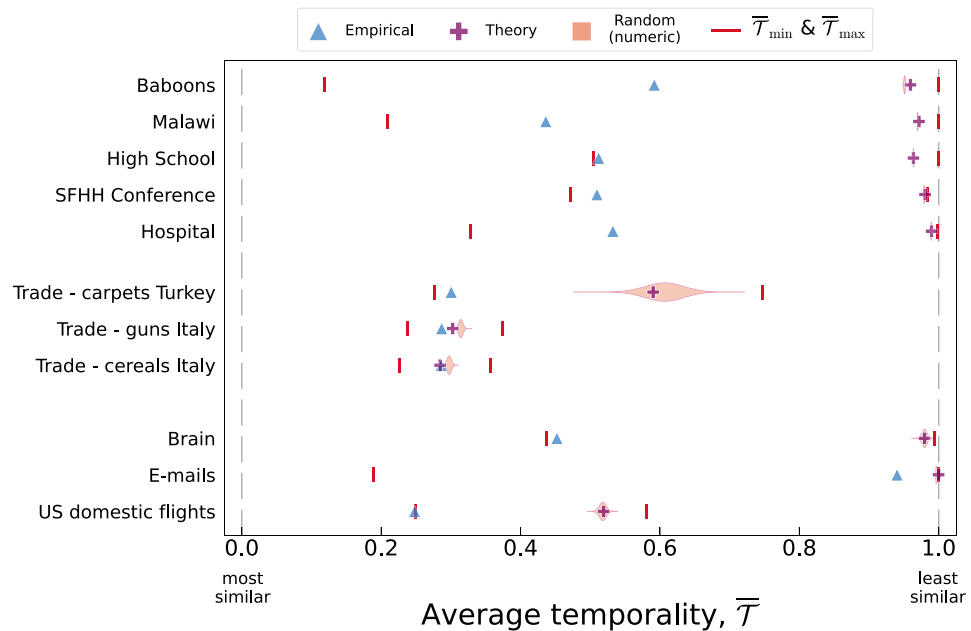


Figure 2. Characterization of the average temporality, \bar{T} , of the empirical datasets considered in our study. For each dataset, we display the empirical value of \bar{T} , its theoretical estimation, the distribution (violin plot) of the temporality of randomly shuffled sequences of snapshots, and its maximum and minimum possible values (\bar{T}_{\max} , \bar{T}_{\min}). The vertical dashed lines at $\bar{T} = 0.0$ and $\bar{T} = 1.0$ highlight the temporality's theoretical boundaries. To compute the violin plots, we generate 5×10^6 randomly shuffled sequences.

Characterization of unaggregated networks. Figure 2 summarises the relationships existing between the values of temporality (see “Methods”) of all the networks considered in our study. More specifically, for each network we report the average value of the temporality computed: (i) numerically from the data (\bar{T} —Empirical); (ii) analytically from the data via Eq. (6) (\bar{T}_{th} —Theory); (iii) sampling numerically sequences extracted from the configuration space (\bar{T}_{rand} ,—Random) (in the text we use \bar{T}_{rand} to refer to the median of the distribution); (iv) extracting the sequences corresponding to the maximum and minimum values (\bar{T}_{\min} , \bar{T}_{\max}) using the optimization algorithm described in the “Methods”.

First of all, we want to remark that for almost all the networks considered in this study, the theoretical estimation \bar{T}_{th} is in good agreement with \bar{T}_{rand} . Such an agreement means that the shuffling process underlying the computation \bar{T}_{rand} fulfils the hypothesis used to derive \bar{T}_{th} and that, more in general, we can use \bar{T}_{th} as a proxy for the estimation of \bar{T}_{rand} .

Also, Fig. 2 shows that most of the values of the empiric \bar{T} fall within the $[0.2, 0.6]$ interval, and that empiric sequences have more persistent interactions than the corresponding theoretical estimation (which also represents the randomized counterparts) i.e., $\bar{T} \leq \bar{T}_{\text{th}}$. The relative position of \bar{T} and \bar{T}_{th} , $d_{\text{th}} = |\bar{T} - \bar{T}_{\text{th}}|$, provides us with a valuable information. Such a quantity, in fact, determines whether the persistence, if any, stems either from the existence of a sort of permanent set/core of interactions (i.e., $d_{\text{th}} \sim 0$) or, alternatively, from the existence of short range correlations between temporal-adjacent snapshots which—in turn,—controls the lifespan of the interactions (i.e., $d_{\text{th}} > 0$). Eyeballing at the diagram, we observe that in all the networks (except for the Italian trade ones) $\bar{T} \ll \bar{T}_{\text{th}}$, implying that the origin of the persistence in the interactions between adjacent snapshots is due to the existence of short-range correlations between temporal-adjacent snapshots.

Another useful indicator is the distance $d_{\min} = |\bar{T} - \bar{T}_{\min}|$. We observe that for the networks of High School, SFHH Conference, Trade-carpets Turkey, Brain, and US domestic flights $d_{\min} \simeq \varepsilon$ with $\varepsilon \rightarrow 0$. Such a trend hints at the existence of some kind of intrinsic optimization behind these networks' organization. For E-mails, instead, the situation is quite the opposite with a system that is almost the least persistent possible. For the Baboons, Malawi, Hospital, and Italian trade networks we observe bigger values of d_{\min} , implying that although the similarities between adjacent snapshots are non negligible (compared to \bar{T}_{rand}), they are stronger with snapshots distant in time (without having information on how far, though).

Finally, we observe that—for most networks,—the theoretical estimation is close to the \bar{T}_{\max} . Such a proximity implies that each snapshot is as different as possible with the majority of the other snapshots, except for those which are adjacent in the original sequence. In the three trade networks and that of US domestic flights the theoretical estimation lays more far away from the maximum. One possible explanation of such a feature could be related with the fact that these networks have the smallest number of snapshots, N_s (see Table 1).

Apart from the characterisation of empirical networks, we have computed these metrics for synthetic time-varying networks for which the edges' persistence and the origin of such a persistence can be tuned. In Supplementary Note S3 of the SM, we present different methods for the generation of synthetic datasets in which the persistence of edges is due to the existence of a stable (time-invariant) core of interactions or, alternatively, to

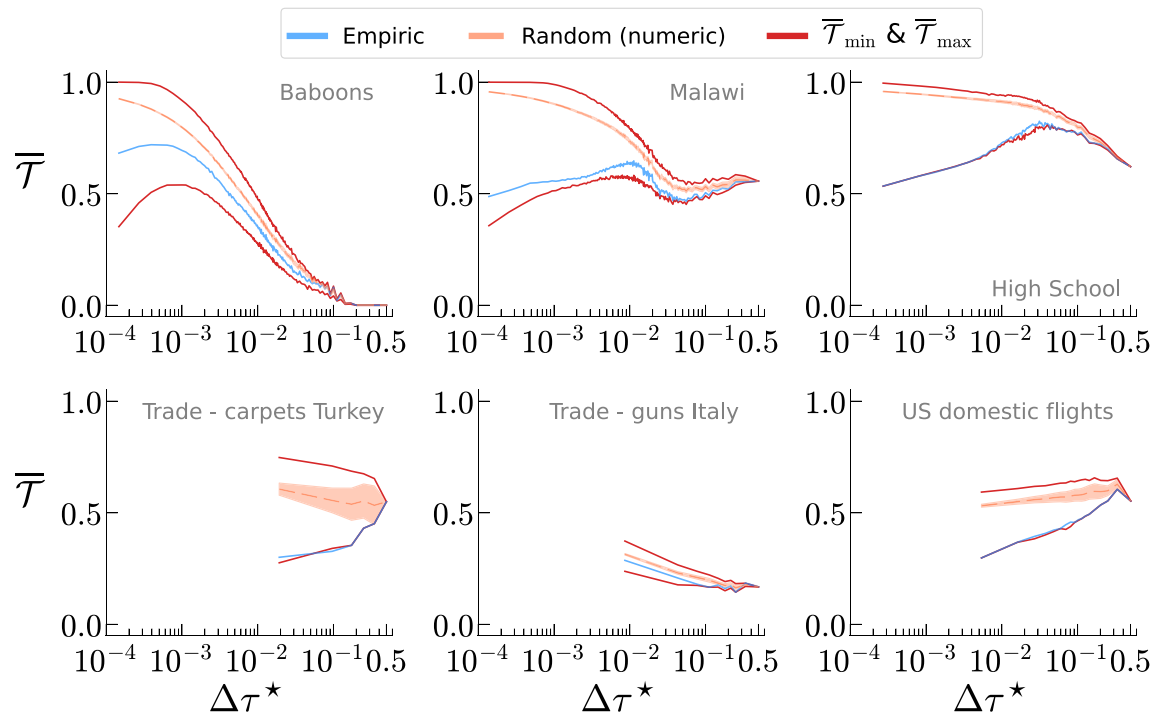


Figure 3. Effects of the coarse graining (i.e., time resolution) on the temporality. We display the values of the empiric temporality, \bar{T} , of the mean and standard deviation of the distribution of temporality of randomly shuffled sequences, \bar{T}_{rand} , and of the maximum and minimum values (\bar{T}_{max} , \bar{T}_{min}) as a function of the rescaled time resolution $\Delta\tau^*$. Each panel refers to a distinct dataset.

the presence of short-range correlations between temporal-adjacent snapshots. The results of our characterisation (see Supplementary Note S3.2 of the SM) provide some insights that argue for the role of different types of persistence on the values of temporality observed in our empirical networks.

Effects of changing the time resolution. After characterizing the persistence of the interactions, and estimating how special its value is compared with some null hypothesis, we study what are the effects of changing the time resolution on the phenomenology observed. For this reason, we perform a coarse-graining (i.e., aggregation as described in the “Methods”) and compute \bar{T} as a function of the size of the aggregation window, $\Delta\tau$. One of the goals of studying the aggregation process is to estimate the level of coarse-graining (i.e., the size of the aggregation window) at which the empirical sequence’s order is statistically equivalent to the random one. Said in other terms, we seek to find the point at which the correlations between adjacent snapshots do not play any role for the persistence of the interactions.

Figure 3 portrays the behaviour of \bar{T} as a function of the aggregation’s level, $\Delta\tau$, for six networks whose behaviours encompass the whole spectrum of the phenomenology observed (see Supplementary Fig. S3 in Supplementary Note S5 for the same picture displaying the whole set of datasets considered in our study). For each network, we display the value of \bar{T} of the empiric sequence, the mean (and standard deviation) of the random sampling of the configuration space, and the maximum and minimum values. We decided to not display the theoretical estimation too, since its estimation’s accuracy depends on N_s which, in turn, decreases with the coarse-graining (at the coarsest level $N_s = 2$). Finally, to enable the comparison between distinct networks at the same aggregation level, we have to replace $\Delta\tau$ with its rescaled counterpart $\Delta\tau^* = \frac{\Delta\tau}{N_s} = \frac{n\tau}{N_s}$.

Overall, we observe the following behaviours: **Baboons:** As we aggregate together the snapshots, we observe a rise and fall of the empiric \bar{T} and \bar{T}_{min} , with a maximum located around $\Delta\tau^* \sim 0.001$ and corresponding to the resolution at which the snapshots are the least similar to each other. As for the non-aggregated data, the empiric temporality is closer to the minimum than the maximum with the random displaying the opposite feature, instead. The relative distance between the random/empiric temporality and the extreme values decreases as we aggregate more, with the four curves merging together around $\Delta\tau^* \sim 0.15$. After that, $\bar{T} = 0$ as the network becomes complete/fully connected. **Malawi:** Aggregating the snapshots produces oscillations on the empiric \bar{T} with the global maximum located around $\Delta\tau^* \sim 0.015$. Moreover, we notice that the empiric and random values of \bar{T} remain always close to \bar{T}_{min} and \bar{T}_{max} , respectively. The behaviour of the empiric and random \bar{T} suggest that long and short range correlations between snapshots are nearly the same. It is worth mentioning that also the E-mails network displays a similar behaviour [see Supplementary Fig. S3 of the Supplementary Materials (SM)]. **High School:** We observe an almost perfect overlap between the empiric temporality and the minimum one, and between the random and the maximum one across the whole $\Delta\tau^*$ ’s range. As for Baboons, the empiric temporality has a maximum for $\Delta\tau^* \sim 0.02$ and the four curves merge together around $\Delta\tau^* \sim 0.3$. The SFHH Conference, Hospital, and Brain data display a similar behaviour (see Supplementary Fig. S3 of

the SM). **Trade-carpets Turkey**: As for the previous case, the empiric and minimum temporality curves overlap almost perfectly across the whole $\Delta\tau^*$'s range. However, the mean value of $\overline{\mathcal{T}}_{\text{rand}}$ lays more or less in the middle between $\overline{\mathcal{T}}_{\text{max}}$ and $\overline{\mathcal{T}}_{\text{min}}$. Also, the value of $\overline{\mathcal{T}}_{\text{rand}}$ is not affected too much by the temporal resolution at which we study the system. Finally, the curves overlap with each other only at $\Delta\tau^* \simeq 0.5$. **Trade-guns Italy**: We observe a monotonous decrease of $\overline{\mathcal{T}}$ for all indicators. Moreover, the empiric and mean value of $\overline{\mathcal{T}}_{\text{rand}}$ are very similar for almost the whole range of $\Delta\tau^*$ values, implying that the phenomenology observed for unaggregated data is not affected by the coarse graining. **US domestic flights**: We observe a behaviour similar to the Turkish trade dataset one, albeit in this case the temporality is monotonically increasing with $\Delta\tau^*$. Such a behaviour denotes that the snapshots get less similar as we aggregate them. Furthermore, it is worth mentioning that for every dataset all the temporality values must coincide for $\Delta\tau^* = 0.5$ (i.e., when the sequence is made only by two snapshots). Such a phenomenon, stems from the fact that there are only two possible snapshots' orders which correspond to the same temporality. Apart from that, one might expect that the value of $\overline{\mathcal{T}}$ becomes smaller for higher values of $\Delta\tau^*$ as the snapshot networks becomes more akin to complete networks. However, this is not the case and $\overline{\mathcal{T}}$ at $\Delta\tau^* = 0.5$ falls within the whole range of possible values with $\overline{\mathcal{T}} \simeq 1$ for SFHH Conference and $\overline{\mathcal{T}} = 0$ for Baboons, instead (see Supplementary Fig. S3 of the SM).

Finally, one feature observed in all datasets is that the distance between $\overline{\mathcal{T}}$ and $\overline{\mathcal{T}}_{\text{rand}}$ decreases with $\Delta\tau^*$ and, eventually, goes to zero for $\Delta\tau^* = 0.5$. This phenomenon indicates, according to the analysis performed on the characterisation of non-aggregated sequences, that as the aggregation window increases, a core of fixed interactions emerges.

Discussion and conclusion

The static network paradigm is short to fully mimic the rich phenomenology displayed by dynamics taking place in those systems whose interactions evolve in time. Using a time-varying network paradigm allows to overcome such limitations and, in turns, to attain a better description of the interplay between dynamics and the persistence of evolving interactions^{8,16}.

In this work, we used a metric to gauge the interactions' persistence named temporality²⁶, and proposed an approach based on statistical physics to characterise the features of several empirical time-varying networks. By comparing the values of the average temporality, $\overline{\mathcal{T}}$, with its counterpart obtained from a random sampling of the configuration space, we assess how "special" the empirical order of snapshots is. Remarkably, we have found that the empiric interactions tend to be more persistent than in a random sequence of snapshots, and that some systems follow some kind of optimisation principle behind the arrangement of their snapshots. Besides, we have studied also the effects of time-resolution—i.e., coarse-graining— $\Delta\tau$ on the temporality and the evolution of the hierarchy between its values computed for empiric, random, and limit ($\overline{\mathcal{T}}_{\text{max}}$, $\overline{\mathcal{T}}_{\text{min}}$) sequences. We have observed how systems belonging to the same category (e.g., face-to-face interactions) display distinct behaviours as we reduce the time resolution. According to the formalism of statistical physics, such an approach could be extended from the canonical formulation—in which the temporality (akin to the energy) is not fixed, but the number of pairs of snapshots (a proxy for the number of particles) is,—to the grand canonical formulation; thus allowing for the characterisation of the datasets with respect to a configurations' space in which the number of snapshots, N_s , varies³⁵.

The comparison of the empiric value of $\overline{\mathcal{T}}$ and the mean of the distribution of temporality obtained from the random sampling of the configuration space, $\overline{\mathcal{T}}_{\text{rand}}$, provides us with a valuable information about the origin of the interactions' persistence. In particular, we can ask ourselves whether the latter is due either to the existence of intrinsic (temporal) correlations between adjacent snapshots (i.e., memory)³⁶, or to the presence of a set of persistent interactions whose existence is not affected by the temporal order of the snapshots. Two examples of these extreme configurations are the Italian trade datasets and the Brain one. Both exhibit a low value of $\overline{\mathcal{T}}$ implying the persistence of the interactions between adjacent snapshots. However, trade datasets display also low values of $\overline{\mathcal{T}}_{\text{rand}}$, whereas Brain has $\overline{\mathcal{T}}_{\text{rand}} \sim 1$. Such a difference highlights the existence of a stable, persistent, set of interactions for the former networks, whereas in the latter network short-range memory prevails.

Finally, the methodology presented in this work has some technical—and conceptual,—limitations which could become the subject of further studies. For instance, one could generalise the estimation of $\overline{\mathcal{T}}$ to account for the existence of correlations in the probability that the same edge exists in adjacent snapshots (which is one of the hallmarks of real time-varying networks)^{3,36,37}. Another possibility is to leverage the information carried by the set of each link's local values of temporality. Among the potential extensions, one is to explore the interplay between the edges belonging to the time-varying backbone^{30,38} and those contributing to the temporality observed in empirical systems. Studying the evolution of topological descriptors could help to grasp the behaviour of the temporality when one changes the size of the aggregation window (e.g., to find a "characteristic" time-scale). Finally, the approach presented in this work could be used as a basis for the design of a method (akin to a configuration model^{39,40}) to generate time-varying networks with a given value of average temporality, and use it to test the role of persistent interactions on dynamical processes in a more controlled way.

Methods

A time-varying network $G(\mathcal{N}, \mathcal{E}(t))$ with $N \equiv |\mathcal{N}|$ vertices (or nodes) is defined as a set of interaction's triples $e \equiv (i, j, t) \in \mathcal{E}(t)$ where $i, j \in \{1, \dots, N\}$ are the indices of the interacting vertices, and t denotes the time at which such an interaction occurs (we are assuming that the set of vertices, \mathcal{N} , does not change over time). Time-varying networks can be thought also as a sequence of N_s graphs (snapshots), $G(t) \equiv G(\mathcal{N}, \mathcal{E}_t)$, each made only by those interactions occurring at the same, discrete, timestep $t^{3,5}$. In this section, first we introduce the concept of temporality \mathcal{T} and how to estimate its average value, $\overline{\mathcal{T}}$, over the sequence $G(t)$. Then, we describe how we

merge together adjacent snapshots, shuffle the elements of $G(t)$, as well as find the sequences $G^{\min}(t)$ and $G^{\max}(t)$ corresponding to the maximum and minimum values of \overline{T} within the configuration space.

Temporality. Given a pair of snapshots graphs $G_m, G_n \in G(t)$ (with $m, n \in \{1, \dots, N_s\}$), we define its *temporality*, $\mathcal{T}_{m,n}$, as:

$$\mathcal{T}_{m,n} = \frac{\sum_{i,j=1}^N |a_{i,j}(m) - a_{i,j}(n)|}{\sum_{i,j} \max\{a_{i,j}(m), a_{i,j}(n)\}}, \quad (1)$$

where $a_{i,j}(m)$ is an element of the adjacency matrix \mathcal{A}_m of the graph G_m ⁴¹. In particular, $a_{i,j}(m)$ is equal to one if vertices i and j are connected in graph G_m , and is equal to zero otherwise. Equation (1) quantifies nothing else than the ratio between the number of distinct edges of G_m and G_n (i.e., those existing in one snapshot but not in the other), divided by the number of common and non-common edges. We can rewrite Eq. (1) in terms of the set of edges \mathcal{E} as:

$$\mathcal{T}_{m,n} = \frac{|\bigcup_{m,n} \mathcal{E}| - |\bigcap_{m,n} \mathcal{E}|}{|\bigcup_{m,n} \mathcal{E}|} = 1 - \frac{|\bigcap_{m,n} \mathcal{E}|}{|\bigcup_{m,n} \mathcal{E}|}, \quad (2)$$

where $|\bigcup_{m,n} \mathcal{E}|$ is the size of the union of the edges' sets of graphs G_m and G_n (i.e., $\bigcup_{m,n} \mathcal{E} \equiv \mathcal{E}_m \cup \mathcal{E}_n$), whereas $|\bigcap_{m,n} \mathcal{E}|$ is the size of the intersection of those sets, (i.e., $\bigcap_{m,n} \mathcal{E} \equiv \mathcal{E}_m \cap \mathcal{E}_n$).

We can use the definition of temporality to gauge the volatility of the interactions of a time-varying network. To this aim, we define the average value of the temporality, \overline{T} , for the whole time-varying network, $G(t)$ (i.e., the whole set of snapshots), as follows:

$$\overline{T} = \frac{1}{N_s - 1} \sum_{m=1}^{N_s-1} 1 - \frac{|\bigcap_{m,m+1} \mathcal{E}|}{|\bigcup_{m,m+1} \mathcal{E}|} = 1 - \frac{1}{N_s - 1} \sum_{m=1}^{N_s-1} \frac{|\bigcap_{m,m+1} \mathcal{E}|}{|\bigcup_{m,m+1} \mathcal{E}|}, \quad (3)$$

where $m, m+1 \in \{1, \dots, N_s - 1\}$ denote the indices of two temporal-adjacent snapshots of the time-varying network $G(t)$. According to the above definition, the temporality values span from 0 for networks having always the same edges, to 1 for completely different networks.

It is possible to estimate analytically \overline{T} by assuming that (i) the microscopic process governing the existence of an interaction between two vertices at time t is independent on both the existence of other interactions at the same time, as well as on the occurrence of such an interaction in the past. (ii) The microscopic process governing the existence of an interaction does not change over time. (iii) The network has a constant size (i.e., N does not vary over time).

Following these assumptions, the average number of interactions (edges) in a snapshot of $G(t)$, $\langle \mathcal{L} \rangle$, can be expressed as:

$$\langle \mathcal{L} \rangle = \frac{N(N-1)}{2} \langle x_{ij} \rangle, \quad (4)$$

where x_{ij} is the probability that an edge between nodes i and j exists in any of the snapshots, and $\langle x_{ij} \rangle$ is the average of such a probability over all the $\frac{N(N-1)}{2}$ possible edges. By leveraging the independence of interactions taking place at different snapshots, we can express the average number of edges belonging to two time adjacent snapshots (i.e., to the intersection of their edges' sets), $\langle |\cap| \rangle$, as:

$$\langle |\cap| \rangle = \frac{N(N-1)}{2} \langle x_{ij \in \cap} \rangle = \frac{N(N-1)}{2} \langle x_{ij}^2 \rangle, \quad (5)$$

where $x_{ij \in \cap}$ is the probability that the edge $e(i, j)$ belongs to the intersection of the edges' sets of all the snapshots which, according to the independence of the probabilities postulated above, is equal to the product of the individual probabilities x_{ij} .

Finally, as the microscopic process governing the existence of the interactions does not change over time (i.e., it is independent of t), we can rewrite Eq. (3) in terms of Eqs. (4) and (5), yielding:

$$\overline{T}_{\text{th}} = 1 - \frac{1}{N_s - 1} \sum_{m=1}^{N_s-1} \frac{|\bigcap_{m,m+1} \mathcal{E}|}{|\bigcup_{m,m+1} \mathcal{E}|} = 1 - \frac{\langle |\cap| \rangle}{\langle |\cup| \rangle} = 1 - \frac{\langle |\cap| \rangle}{2\langle \mathcal{L} \rangle - \langle |\cap| \rangle} \approx 1 - \frac{\langle |\cap| \rangle}{2\langle \mathcal{L} \rangle - \langle |\cap| \rangle} = 1 - \frac{\langle x_{ij}^2 \rangle}{2\langle x_{ij} \rangle - \langle x_{ij}^2 \rangle}. \quad (6)$$

Hence, the theoretical estimation of \overline{T} can be written just in terms of x_{ij} . However, we want to stress that the approximation is valid only if both the variance of \mathcal{L} and $|\cap|$, as well as the covariance between the same quantities (and its higher order moments), are both negligible. For more details, see Supplementary Note S1 of the SM. Hence, to estimate \overline{T} of an empirical network we just have to estimate the probability of appearance of each possible edge. Such a probability can be computed directly from the data as:

$$x_{ij} = \frac{N_s(i, j)}{N_s}, \quad (7)$$

where $N_s(i, j)$ is the number of snapshots in which the edge $e(i, j)$ exists. Ensuring a correct interpretation of these results entails some caveats. For instance, the assumption of the interactions' statistical independence between

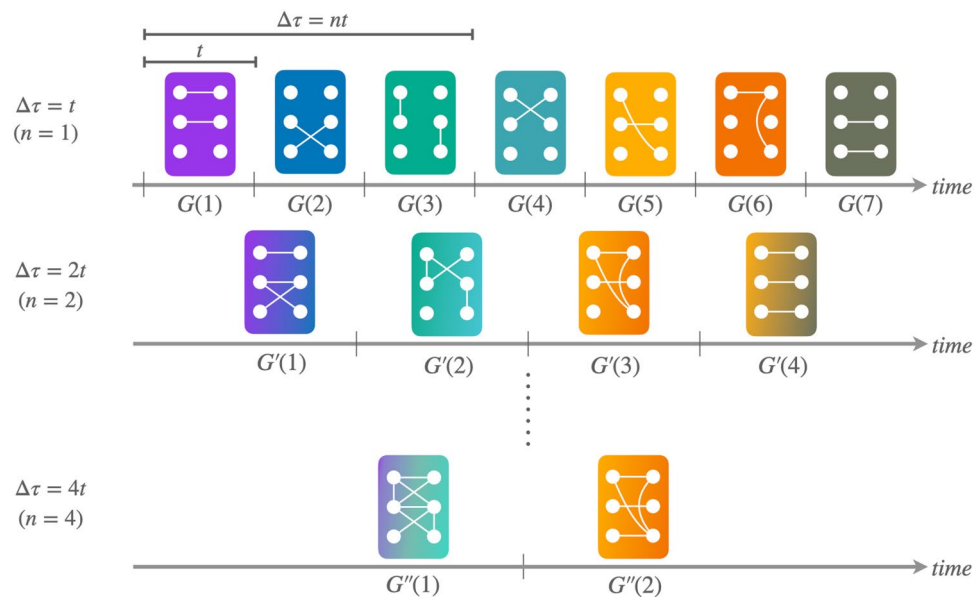


Figure 4. Schematic representation of the aggregation process. Given a sequence of snapshots, $G(t)$, we aggregate it by merging together groups of snapshots of size $\Delta\tau = nt$ with $n \in \mathbb{N}$. Each row accounts for a different level of aggregation: $n = 1$ (top), $n = 2$ (middle), and $n = 4$ (bottom).

snapshots is, in general, not true for empirical systems. Such an issue exists when the system possesses some sort of memory (even just at short range)³⁶. Nonetheless, as we will see, for uncorrelated sequences our theoretical framework estimates quite well the values of \bar{T} . One potential extension of our framework (see Supplementary Note S2 of the SM) is to compute x_{ij} using the so-called *activity driven model*: a popular model used to generate time-varying networks whose structure resembles those of the empirical ones⁴².

Aggregation, shuffling, and finding extreme values of temporality. In the following, we describe how given a sequence of snapshots, $G(t)$, we aggregate, shuffle, and re-order it. Before describing these processes, it is worth mentioning that we perform some pre-processing on the empirical sequences. Specifically, for each dataset we remove those snapshots corresponding to empty graphs (e.g., night recordings of face-to-face interactions). Such operation is justified by the fact that empty graphs do not contribute to the temporality. Pruning inactivity periods from the data leads to a re-definition of the concept of time itself, converting the variable t from a time into a descriptor of the position of the snapshot along the sequence.

Aggregation. One important aspect of the characterization of time-varying networks is the *resolution* at which we study them⁴³. Given a sequence of snapshots $G(t)$, we can convert it into a new sequence $G'(\tilde{t})$ by aggregating its elements into groups of size $\Delta\tau = nt$ with $n \in \mathbb{N}$ (i.e., we create its time-wise coarse graining). There are several ways of generating $G'(\tilde{t})$ ⁵. Here, we consider an aggregation approach analogous to that used by Tang et al.⁴³ consisting in creating a projection network corresponding to the union of the edges' sets of the snapshots' group of size $\Delta\tau$ (see Fig. 4).

According to such a projection method, given a sequence of snapshots $G(t)$ and a temporal resolution $\Delta\tau$, the resulting aggregated sequence $G'(\tilde{t})$ will be made of $\frac{N_s}{\Delta\tau}$ snapshots (chunks). If N_s is not an integer multiple of $\Delta\tau$, we apply periodic boundary conditions to the sequence and “complete” the last chunk of snapshots by adding those located at the beginning of the sequence (see middle row of Fig. 4). That said, it is worth stressing that distinct values of n (or, equivalently, $\Delta\tau$) may lead to aggregated sequences with the same number of snapshots (chunks). In such a case, we keep only the values of n minimizing the number of snapshots needed to complete the last chunk.

Shuffling. One way to get rid of the correlations of a time series is to shuffle its elements. Similarly, we can get rid of the correlations existing between adjacent snapshots of the sequence $G(t)$ by randomly shuffling their positions, giving rise to a new sequence $G'(t)$. The snapshots' random arrangement is nothing else than the generation process laying behind our configuration space, as explained in the “Introduction”. Given a sequence of graphs $G(t)$ with N_s snapshots, it is—in general—computationally unfeasible to estimate exactly descriptors like the average or the standard deviation of an indicator (e.g., of \mathcal{T}) via exploring the whole configuration space, as its size scales with $N_s!$. For this reason, we compute these descriptors by sampling the configuration space generating a computationally-feasible number of realizations (5×10^6) of the random reshuffling process.

Finally, when studying the effects of time resolution (i.e., aggregation), we aggregate the sequence first and perform the reshuffling then, thus exploring the configuration space of the aggregated sequences.

Dataset	N	N_s	Δt	K_{TOT}	\bar{T}	$\langle \rho \rangle (\times 10^{-4})$	Source
Face-to-face							
Baboons	13	40,845*	20 s	78	0.592	287.47	48
Malawi	86	43,437*		347	0.436	8.51	49
High School	327	7374		5818	0.512	4.80	50
SFHH Conference	403	3508		9565	0.510	2.47	51
Hospital	75	9452		1139	0.532	12.36	52
Trade							
Trade-carpets Turkey	207	52	1 year	206	0.301	34.29	53,57
Trade-guns Italy	156	116	1 month	155	0.287	59.92	
Trade-cereals Italy	157	108		156	0.286	60.41	
Other							
Brain	16	396	$\frac{1}{200}$ s	120	0.452	395.99	54
E-mails	1890	19,380	1 s	4383	0.940	0.01	55,56
US domestic flights	1677	371	1 month	25,890	0.248	24.16	58,59

Table 1. Main characteristics of the datasets considered. For each network, we report the number of nodes, N , the number of snapshots N_s , the temporal resolution, Δt , the number of distinct interactions in the aggregated network, K_{TOT} , the average temporality, \bar{T} , and the edges' density average over the snapshots' set, $\langle \rho \rangle$. Finally, we report the data's bibliographic source.

Optimization algorithm. In the section entitled Temporality, we have mentioned that $\bar{T} \in [0, 1]$. However, such boundaries are just theoretical and correspond to very peculiar configurations. To gauge the effective boundaries of the temporality, we need to identify the arrangements (i.e., sequences) corresponding the maximum and minimum values of \bar{T} . This optimization problem shares many traits with the so-called *Travelling Salesman Problem*⁴⁴. Specifically, each of the N_s snapshots of the sequence $G(t)$ can be thought as the vertex of a graph, and the temporality computed between each of the $\frac{N_s(N_s-1)}{2}$ snapshots' pairs can be thought as the distance between vertices. Therefore, the aforementioned optimization problem consists in finding the open chain (i.e., snapshots' sequence) maximizing—or minimizing,—its total length (a proxy for the average temporality). However, the solution of such an optimization problem is NP hard. To overcome such a limitation, we use a heuristic algorithm based on the *Kruskal's algorithm* (used to extract the minimum spanning tree of a graph)⁴⁵, allowing us to obtain an approximate solution for our problem.

Given the set of snapshots, $\mathbb{G} \equiv \{G_i\}$ with $G_i \in G(t) \forall i = \{1, \dots, N_s\}$, our heuristic algorithm works as follows:

1. Compute using Eq. (2) the elements of the set $\mathbb{T} \equiv \{T_{m,n}\}_{m,n=1}^{N_s} \forall m, n \in \{1, \dots, N_s\}, m \neq n$; and store its values. Then, sort the elements of set \mathbb{T} in ascending (or descending) order.
2. For each $T_{m,n} \in \mathbb{T}$, we add the corresponding edge (G_m, G_n) if and only if G_m and G_n are either both isolated vertices (graphs), or members of distinct chains and each vertex is connected with at most another vertex.
3. Repeat the above operation until the system is made by a single chain with $N_s - 1$ edges.

Inverting the sorting order allows to determine either the snapshots' sequence $G^{\min}(t)$ corresponding to \bar{T}_{\min} , or the sequence $G^{\max}(t)$ corresponding to \bar{T}_{\max} . As mentioned previously, the non-Markovian nature of the algorithm leaves room to the chance that different choices, locally not optimal, could lead to a better final outcome (i.e., the solution found is a local optimum, but not necessarily a global one). One possible workaround is to add the edge (G_m, G_n) with a probability computed using a simulated annealing technique⁴⁶. Finally, we would like to mention that the above algorithm could be modified to reduce the computation time.

Data. We consider eleven time-varying networks grouped in the following categories:

Face-to-face Five networks obtained from the Sociopatterns' repository⁴⁷ Specifically:

Baboons The interactions occurring among a group of Guinea baboons living in an enclosure of a Primate Center in France recorded between June 13th and July 10th of 2019⁴⁸. Note: the great number of snapshots ($\sim 40,000$) of the dataset make unfeasible the sampling and the optimization over the configuration space. For this reason, we need to make a 2-snapshots aggregation even for the unaggregated characterization.

Malawi Observational contact data collected for 86 individuals living in a village in rural Malawi⁴⁹. Note: the great number of snapshots ($\sim 40,000$) of the datasets make infeasible the sampling and the optimization over the configuration space. For this reason, we need to make a 2-snapshots aggregation even for the unaggregated characterization.

High School Contacts and friendship relations between students in a high school in Marseilles (France) recorded during December 2013⁵⁰.

SFHH Conference Interactions among the participants to the SFHH conference in Nice (France) which took place between June 4th and 5th of 2009⁵¹.

Hospital Contacts between patients, patients and health-care workers (HCWs), and between HCWs in a hospital ward in Lyon (France) recorded between December 6th and December 10th of 2010⁵².

Trade Three star networks describing the export relationships over a specific good (commodity) occurring between one country and all the other countries in the world, extracted from the UN-COMMTRADE database⁵³ (see Supplementary Note S4 of the SM for the details). Specifically:

Trade-carpets Turkey Exports of “carpets, carpeting and rugs, knotted” taking place between years 1962 and 2020.

Trade-guns Italy Exports of “Arms and ammunition; parts and accessories thereof” taking place between January 2010 and December 2020.

Trade-cereals Italy Exports of “cereals” taking place between January 2010 and December 2020.

Other Three networks of various types. Specifically:

Brain The functional brain network extracted from the EEG β band activity recorded in several Regions of Interest (ROIs) during a motor task⁵⁴.

E-mails The network of e-mail exchanges between members of the US Democratic Party during the 2016 Democratic National Committee^{55,56}.

US domestic flights The network of domestic flights operated within the USA taking place between January 1990 and December 2021 (see Supplementary Note S4 of the SM for the details).

Table 1 presents a summary of the main properties of the above networks.

Data availability

The data and code used to generate the US domestic flights and COMMTRADE networks are available at: <https://cardillo.web.bifi.es/data.html#flights> and <https://cardillo.web.bifi.es/data.html#trade>. All the other data used in this study are publicly available (see Table 1 for the bibliographic sources).

Received: 5 July 2022; Accepted: 6 December 2022

Published online: 14 January 2023

References

- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D. Complex networks: Structure and dynamics. *Phys. Rep.* **424**, 175–308. <https://doi.org/10.1016/j.physrep.2005.10.009> (2006).
- Barabási, A.-L. The network takeover. *Nat. Phys.* **8**, 14–16. <https://doi.org/10.1038/nphys2188> (2012).
- Masuda, N. & Lambiotte, R. *A Guide to Temporal Networks* 2nd edn. (World Scientific, 2020).
- Zhang, X., Moore, C. & Newman, M. E. J. Random graph models for dynamic networks. *Eur. Phys. J. B* **90**, 200. <https://doi.org/10.1140/epjb/e2017-80122-8> (2017).
- Holme, P. & Saramäki, J. Temporal networks. *Phys. Rep.* **519**, 97–125. <https://doi.org/10.1016/j.physrep.2012.03.001> (2012).
- Granell, C., Darst, R. K., Arenas, A., Fortunato, S. & Gómez, S. Benchmark model to assess community structure in evolving networks. *Phys. Rev. E* **92**, 012805. <https://doi.org/10.1103/PhysRevE.92.012805> (2015).
- Gross, T., D’Lima, C. J. D. & Blasius, B. Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.* **96**, 208701. <https://doi.org/10.1103/PhysRevLett.96.208701> (2006).
- Masuda, N., Klemm, K. & Eguiluz, V. M. Temporal networks: Slowing down diffusion by long lasting interactions. *Phys. Rev. Lett.* **111**, 188701. <https://doi.org/10.1103/PhysRevLett.111.188701> (2013).
- Liu, S., Perra, N., Karsai, M. & Vespignani, A. Controlling contagion processes in activity driven networks. *Phys. Rev. Lett.* **112**, 118702. <https://doi.org/10.1103/PhysRevLett.112.118702> (2014).
- Perra, N. *et al.* Random walks and search in time-varying networks. *Phys. Rev. Lett.* **109**, 238701. <https://doi.org/10.1103/PhysRevLett.109.238701> (2012).
- Lucas, M., Fanelli, D., Carletti, T. & Petit, J. Desynchronization induced by time-varying network. *Europhys. Lett.* **121**, 50008. <https://doi.org/10.1209/0295-5075/121/50008> (2018).
- Kohar, V., Ji, P., Choudhary, A., Sinha, S. & Kurths, J. Synchronization in time-varying networks. *Phys. Rev. E* **90**, 022812. <https://doi.org/10.1103/PhysRevE.90.022812> (2014).
- Frasca, M., Buscarino, A., Rizzo, A., Fortuna, L. & Boccaletti, S. Synchronization of moving chaotic agents. *Phys. Rev. Lett.* **100**, 044102. <https://doi.org/10.1103/PhysRevLett.100.044102> (2008).
- Petit, J., Lauwens, B., Fanelli, D. & Carletti, T. Theory of Turing patterns on time varying networks. *Phys. Rev. Lett.* **119**, 148301. <https://doi.org/10.1103/PhysRevLett.119.148301> (2017).
- Cardillo, A. *et al.* Evolutionary dynamics of time-resolved social interactions. *Phys. Rev. E* **90**, 052825. <https://doi.org/10.1103/PhysRevE.90.052825> (2014).
- Masuda, N. Accelerating coordination in temporal networks by engineering the link order. *Sci. Rep.* **6**, 22105. <https://doi.org/10.1038/srep22105> (2016).
- Darst, R. K. *et al.* Detection of timescales in evolving complex systems. *Sci. Rep.* **6**, 39713. <https://doi.org/10.1038/srep39713> (2016).
- Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M. & Brunham, R. C. Network theory and SARS: Predicting outbreak diversity. *J. Theor. Biol.* **232**, 71–81. <https://doi.org/10.1016/j.jtbi.2004.07.026> (2005).
- Kao, R. R., Green, D. M., Johnson, J. & Kiss, I. Z. Disease dynamics over very different time-scales: Foot-and-mouth disease and scrapie on the network of livestock movements in the UK. *J. R. Soc. Interface* **4**, 907–916. <https://doi.org/10.1098/rsif.2007.1129> (2007).

20. Boguñá, M., Castellano, C. & Pastor-Satorras, R. Langevin approach for the dynamics of the contact process on annealed scale-free networks. *Phys. Rev. E* **79**, 036110. <https://doi.org/10.1103/PhysRevE.79.036110> (2009).
21. Guerra, B. & Gómez-Gardeñes, J. Annealed and mean-field formulations of disease dynamics on static and adaptive networks. *Phys. Rev. E* **82**, 035101. <https://doi.org/10.1103/PhysRevE.82.035101> (2010).
22. Moody, J. The importance of relationship timing for diffusion. *Soc. Forces* **81**, 25–56. <https://doi.org/10.1353/sof.2002.0056> (2002).
23. Bagrow, J. P., Wang, D. & Barabási, A.-L. Collective response of human populations to large-scale emergencies. *PLoS ONE* **6**, 1–8. <https://doi.org/10.1371/journal.pone.0017680> (2011).
24. Wang, P., González, M. C., Hidalgo, C. A. & Barabási, A.-L. Understanding the spreading patterns of mobile phone viruses. *Science* **324**, 1071–1076. <https://doi.org/10.1126/science.1167053> (2009).
25. Cardillo, A., Meloni, S., Gómez-Gardeñes, J. & Moreno, Y. Velocity-enhanced cooperation of moving agents playing public goods games. *Phys. Rev. E* **85**, 067101. <https://doi.org/10.1103/PhysRevE.85.067101> (2012).
26. Li, A. *et al.* Evolution of cooperation on temporal networks. *Nat. Commun.* **11**, 2259. <https://doi.org/10.1038/s41467-020-16088-w> (2020).
27. Fujiwara, N., Kurths, J. & Díaz-Guilera, A. Synchronization in networks of mobile oscillators. *Phys. Rev. E* **83**, 025101. <https://doi.org/10.1103/PhysRevE.83.025101> (2011).
28. Starnini, M., Baronchelli, A., Barrat, A. & Pastor-Satorras, R. Random walks on temporal networks. *Phys. Rev. E* **85**, 056115. <https://doi.org/10.1103/PhysRevE.85.056115> (2012).
29. Godoy-Lorite, A., Guimerá, R. & Sales-Pardo, M. Long-term evolution of email networks: Statistical regularities, predictability and stability of social behaviors. *PLoS ONE* **11**, e0146113. <https://doi.org/10.1371/journal.pone.0146113> (2016).
30. Kobayashi, T., Takaguchi, T. & Barrat, A. The structured backbone of temporal social ties. *Nat. Commun.* **10**, 220. <https://doi.org/10.1038/s41467-018-08160-3> (2019).
31. Presigny, C., Holme, P. & Barrat, A. Building surrogate temporal network data from observed backbones. *Phys. Rev. E* **103**, 052304. <https://doi.org/10.1103/PhysRevE.103.052304> (2021).
32. Gauvin, L. *et al.* Randomized reference models for temporal networks. *SIAM Rev.* **64**, 763–830. <https://doi.org/10.1137/19M1242252> (2022).
33. Masuda, N. & Holme, P. Detecting sequences of system states in temporal networks. *Sci. Rep.* **9**, 795. <https://doi.org/10.1038/s41598-018-37534-2> (2019).
34. Zhan, X.-X. *et al.* Measuring and utilizing temporal network dissimilarity. <https://doi.org/10.48550/arXiv.2111.01334> (2021). 2111.01334.
35. Cimini, G. *et al.* The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58–71. <https://doi.org/10.1038/s42254-018-0002-6> (2019).
36. Williams, O. E., Lacasa, L., Millán, A. P. & Latora, V. The shape of memory in temporal networks. *Nat. Commun.* **13**, 499. <https://doi.org/10.1038/s41467-022-28123-z> (2022).
37. Ferguson, S. T. & Kobayashi, T. Identifying the temporal dynamics of densification and sparsification in human contact networks. *EPJ Data Sci.* **11**, 52. <https://doi.org/10.1140/epjds/s13688-022-00365-3> (2022).
38. Gemmetto, V., Cardillo, A. & Garlaschelli, D. Irreducible network backbones: Unbiased graph filtering via maximum entropy. <https://doi.org/10.48550/arXiv.1706.00230> (2017).
39. Bender, E. A. & Canfield, E. R. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory Ser. A* **24**, 296–307. [https://doi.org/10.1016/0097-3165\(78\)90059-6](https://doi.org/10.1016/0097-3165(78)90059-6) (1978).
40. Fosdick, B. K., Larremore, D. B., Nishimura, J. & Ugander, J. Configuring random graph models with fixed degree sequences. *SIAM Rev.* **60**, 315–355. <https://doi.org/10.1137/16M1087175> (2018).
41. Latora, V., Nicosia, V. & Russo, G. *Complex Networks* (Cambridge University Press, 2017).
42. Perra, N., Gonçalves, B., Pastor-Satorras, R. & Vespignani, A. Activity driven modeling of time varying networks. *Sci. Rep.* **2**, 469. <https://doi.org/10.1038/srep00469> (2012).
43. Tang, J., Scellato, S., Musolesi, M., Mascolo, C. & Latora, V. Small-world behavior in time-varying graphs. *Phys. Rev. E* **81**, 055101. <https://doi.org/10.1103/PhysRevE.81.055101> (2010).
44. Jünger, M., Reinelt, G. & Rinaldi, G. Chapter 4 the traveling salesman problem. In *Network Models*, vol. 7 of *Handbooks in Operations Research and Management Science*, 225–330. [https://doi.org/10.1016/S0927-0507\(05\)80121-5](https://doi.org/10.1016/S0927-0507(05)80121-5) (Elsevier, 1995).
45. Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. *Introduction to Algorithms*. 2nd ed. The MIT Electrical Engineering and computer Science Series (MIT Press, 2001).
46. Newman, M. E. J. *Computational Physics* (CreateSpace Independent Publishing Platform, 2012).
47. Sociopatterns data repository. <http://www.sociopatterns.org/datasets/> (Accessed 01 July 2021).
48. Gelardi, V., Godard, J., Paleressompoulle, D., Claidiere, N. & Barrat, A. Measuring social networks in primates: Wearable sensors versus direct observations. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **476**, 20190737. <https://doi.org/10.1098/rspa.2019.0737> (2020).
49. Ozella, L. *et al.* Using wearable proximity sensors to characterize social contact patterns in a village of rural Malawi. *EPJ Data Sci.* **10**, 46. <https://doi.org/10.1140/epjds/s13688-021-00302-w> (2021).
50. Fournet, J. & Barrat, A. Contact patterns among high school students. *PLoS ONE* **9**, e107878. <https://doi.org/10.1371/journal.pone.0107878> (2014).
51. Génois, M. & Barrat, A. Can co-location be used as a proxy for face-to-face contacts?. *EPJ Data Sci.* **7**, 11. <https://doi.org/10.1140/epjds/s13688-018-0140-1> (2018).
52. Vanhems, P. *et al.* Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS ONE* **8**, e73970. <https://doi.org/10.1371/journal.pone.0073970> (2013).
53. United Nations COMTRADE database. <https://comtrade.un.org/> (Accessed 01 July 2021).
54. De Vico Fallani, F. *et al.* Persistent patterns of interconnection in time-varying cortical networks estimated from high-resolution EEG recordings in humans during a simple motor act. *J. Phys. A Math. Theor.* **41**, 224014. <https://doi.org/10.1088/1751-8113/41/22/224014> (2008).
55. Kunegis, J. KONECT: The Koblenz network collection. In *Proceedings of the 22nd International Conference on World Wide Web, WWW '13 Companion*, 1343–1350. <https://doi.org/10.1145/2487788.2488173> (Association for Computing Machinery, 2013).
56. Pickhardt, R. Extracting 2 social network graphs from the Democratic National Committee Email Corpus on Wikileaks. <https://www.rene-pickhardt.de/index.html%3Fp=1989.html> (2018).
57. Cardillo, A. United Nations COMTRADE datasets. <https://cardillo.web.bifi.es/data.html#trade> (2021).
58. Cardillo, A. US domestic flights datasets. <https://cardillo.web.bifi.es/data.html#flights> (2021).
59. US Bureau of Transportation Statistics—TranStats. <https://www.transtats.bts.gov/> (Accessed 01 July 2021).
60. Oliphant, T. *Guide to NumPy* (Trelgol Publishing, 2006).
61. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22–30. <https://doi.org/10.1109/MCSE.2011.37> (2011).
62. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference*, (eds Varoquaux, G. *et al.*) 11–15 (2008).
63. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95. <https://doi.org/10.1109/MCSE.2007.55> (2007).

Acknowledgements

The authors thank N. Masuda for his comments on a preliminary version of the manuscript, N. Perra for his help with the activity driven model, and F. De Vico Fallani for providing the brain data. AC acknowledges the support of the Spanish Ministerio de Ciencia e Innovación (MICINN) through Grant IJCI-2017-34300. AC acknowledges the support of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 803860). AC, AA, and FB acknowledge the support of the DEIM Department of the University Rovira i Virgili via the "Investigador Activo" funds. AA acknowledges financial support from Spanish MINECO (Grant No. PGC2018-094754-B-C2), from Generalitat de Catalunya (grant No. 2017SGR-896 and 2020PANDE00098), Universitat Rovira i Virgili (grant No. 2019PFR-URV-B2-41), Generalitat de Catalunya ICREA Academia, and the James S. McDonnell Foundation (grant #220020325). LMF and JGG acknowledge financial support from the Departamento de Industria e Innovación del Gobierno de Aragón y Fondo Social Europeo (FENOL group E36_20R), and from grant PID2020-113582GB-I00 funded by MCIN/AEI/10.13039/501100011033. FB acknowledges the support from Departamento de Industria e Innovación del Gobierno de Aragón y Fondo Social Europeo through projects No. E30_17R (COMPHYS group) and doctoral fellowship, and the financial support from the Spanish Ministerio de Ciencia e Innovación (MICINN) through grant PGC2018-094684-B-C22. Numerical analysis has been carried out using the NumPy and NetworkX Python packages^{60–62}. Graphics have been prepared using the Matplotlib Python package⁶³.

Author contributions

A.C. and A.A. secured funding; A.C., A.A., and J.G.G. designed the study; F.B. and M.F. performed the modeling; A.C. contributed with the data; F.B. and A.C. performed the analysis; All authors analysed the results; F.B. and A.C. wrote the paper; A.C. and J.G.G. prepared the graphics. All authors read, reviewed, and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25907-7>.

Correspondence and requests for materials should be addressed to A.C.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023