POLITECNICO DI TORINO
Repository ISTITUZIONALE

Detecting network backbones against time variations in node properties

Original
Detecting network backbones against time variations in node properties / Nadini, M., Bongiorno, C., Rizzo, A., Porfiri, M.. - In: NONLINEAR DYNAMICS. - ISSN 0924-090X. - ELETTRONICO. - 99:1(2020), pp. 855-878. [10.1007/s11071-019-05134-y]

Availability:
This version is available at: 11583/2743532 since: 2020-01-10T12:13:59Z

Publisher:
Springer

Published
DOI:10.1007/s11071-019-05134-y

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Springer postprint/Author's Accepted Manuscript

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s11071-019-05134-y

(Article begins on next page)

09 November 2026

Noname manuscript No.
(will be inserted by the editor)

Detecting network backbones against time-variations in node properties

Matthieu Nadini · Christian Bongiorno · Alessandro Rizzo · Maurizio Porfiri*

Received: date / Accepted: date

Abstract Many real systems can be described through time-varying networks of interactions that encapsulate information sharing between individual units over time. These interactions can be classified as being either reducible or irreducible: reducible interactions pertain to node-specific properties, while irreducible interactions reflect dyadic relationships between nodes that form the network backbone. The process of filtering reducible links to detect the backbone network could allow for identifying family members and friends in social networks or social structures from contact patterns of individuals. A pervasive hypothesis in existing methods of backbone discovery is that the specific properties of the nodes are constant in time, such that reducible links have the same statistical features at any time during the observation. In this work, we release this assumption toward a new method for detecting network backbones against time-variations in node properties. Through analytical insight and numerical evidence on synthetic and real datasets, we demonstrate the viability of the proposed approach to aid in the discovery of network backbones from time series. By critically comparing our approach with existing methods in the technical literature, we show that neglecting time variations in node-specific properties may beget false positives in the inference of the network backbone.

Keywords Activity-driven · backbone network · statistical filtering · time-varying network

Matthieu Nadini
Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Brooklyn, NY, 11201, USA. ORCID 0000-0003-4542-7481

Christian Bongiorno
Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, 10129 Torino, ITALY. ORCID 0000-0002-1730-2895

Alessandro Rizzo
Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, 10129 Torino, ITALY, and Office of Innovation, New York University Tandon School of Engineering, Brooklyn, New York 11201, USA. ORCID 0000-0002-2386-3146

Maurizio Porfiri*
Department of Mechanical and Aerospace Engineering and Department of Biomedical Engineering, New York University Tandon School of Engineering, Brooklyn, NY, 11201, USA. ORCID 0000-0002-1480-3539 E-mail: mporfiri@nyu.edu

1 Introduction

Dealing with real, temporal datasets brings forward several challenges. One of the most ambitious goals is to elucidate the role of temporal interactions in complex systems [1–5]. The presence of temporal interactions questions the very basis of a network approach to complex systems. As articulated in [6], temporal links could be related to intrinsic node properties that do not require the truly dyadic nature of a network. Such temporal links are reducible, whereby they are fully explained by node-specific features. Devising robust methodologies to filter out reducible links for the inference of the irreducible backbone of temporal interactions is an open research topic.

A fruitful approach entails the formulation of null models to explain the reducible part of the temporal interactions and guide the process of filtering, as illustrated in Fig. 1. Filtering is carried out within a statistically-principled approach, where one seeks to detect links that are incompatible with the null hypothesis of links being produced by the null model [7–14]. More concretely, the approach assigns a "strength" to link candidates and filters out weak links, which are statistically unlikely to pertain to the backbone network.

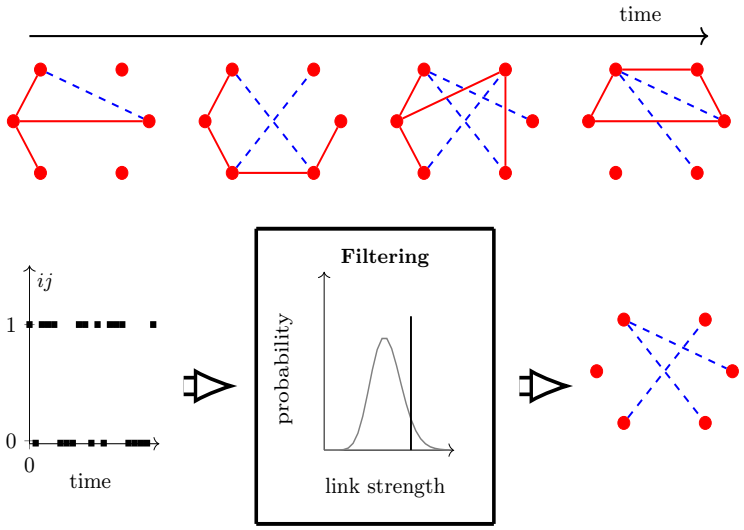

Fig. 1 Top: Time evolution of a complex system, showing node-specific, reducible interactions (solid red links), and the irreducible backbone (dashed blue links). Bottom: empirical observations of temporal interactions between any node pair are used by a filtering algorithm to reconstruct the backbone. The resulting backbone network is composed of a set of aggregated, static links. Retaining a link in the backbone is informed by a statistical comparison that tests the hypothesis of the link being explained by the null model.

Despite significant progress, most research studies assume that nodes have time invariant properties, such that the empirical time series are realizations of a stationary stochastic processes. However, time-varying connections might be affected by several factors, such as individual propensity to generate links over time and physical constraints on the network evolution. In addition, connections may vary non-uniformly in time, exhibiting highly dynamic patterns that could challenge the possibility of network reconstruction. The chief objective of our work is to explore the feasibility of inferring the backbone network, in the presence of richer time-varying connections.

1.1 Background and related studies

A key step toward the inference of the backbone network is the formulation of reliable and comprehensive null models. A recent promising modeling paradigm is offered by activity-driven networks (ADNs) [15]. Within the ADN paradigm, individual propensity of generating links over time is encapsulated by a single, heterogeneously distributed parameter, called activity. In its original formulation, the activities of all the nodes are assumed to be constant in time and the process of network assembly is carried out in a discrete-time setting. A similar approach can be undertaken in continuous time [16].

Because of its analytical tractability, activity-driven models have been extended to comprehend features of real networks, such as memory effects in the link wiring [17], self-exciting mechanisms in individual activities [18], presence of communities [19], and spreading over multiple layers [20, 21]. For example, recent studies have examined adaption of individual activities based on the node's health status [22,23]. Building on this promising line of research, a predictive model of the 2014 Ebola spreading in Liberia has been established [24]. Finally, circadian and weekly patterns have been included in the ADN paradigm in [25].

Upon formulating a valid null model, the next step to backbone inference entails a statistical test to identify irreducible links. The simplest approach is to set a threshold that filters out links with lower strength [26]. However, such an approach could fail to capture the inherent heterogeneity of many complex systems where the formation of reducible links drastically vary in time and throughout the network. The disparity filter [9], which has been recently extended to consider self-exciting mechanisms [14], could address inference of heterogeneous networks. A more established approach is the *statistically validated network* (SVN) [10], first introduced to study bipartite networks, and then applied to temporal networks [11, 27–29].

The SVN works on an aggregated version of the temporal network, that is, a static weighted network formed by retaining all the links occurring at any time instant. Each link has a weight equal to the total number of temporal connections formed over time between two nodes. The SVN approach has helped to elucidate many as-

pects of real systems, such as connections between the backbone network and the network's community structure [30], the influence of time correlations [28,29], and the time evolution of the backbone network [27].

A further improvement on the SVN approach is constituted by the *temporal fitness model* (TFM) [7]. The TFM utilizes an ADN as a null model, in which individual activities are considered to be constant in time. Their values are identified through maximum likelihood estimation. The approach can be extended to study daily patterns and circadian rhythms, within the so-called $\text{TFM}_{\text{rhythm}}$ [7], which utilizes a common function to modulate the overall network evolution. The SVN, TFM, and $\text{TFM}_{\text{rhythm}}$ are summarized in the Appendix.

Overall, these approaches assume that individual properties of the nodes are constant in time. As a result, they cannot be utilized to infer backbone networks in the presence of changes in individual behavior.

1.2 Our contribution

Here, we seek to propose a new methodology to improve the detection of a backbone network in the presence of complex temporal variations of activity patterns. To this end, we introduce an extended version of ADNs, where individual activities are piece-wise constant in time and heterogeneously distributed throughout the network. The null model assumes that all connections are formed uniformly at random, that is, the probability of creating a link at a specific time instant between two nodes is the product of the individual activities of the nodes at that time. In this vein, a very active node is more likely to form connections with another high active node than with a low active node.

Accounting for time-varying activities in the null model calls for two main steps to find irreducible links. First, it is necessary to estimate activity values as piece-wise constant functions of time. Then, links are included in the backbone network if their overall weight is significantly higher than what would be expected from the null model.

While the latter step can be tackled through a statistical test similar to [7,10], dealing with estimation of activity values requires a novel scheme. Specifically, we divide the total observation window of network evolution in independent intervals, containing a uniform total number of connections. Activities are then estimated according to the weighted configuration model [31,32], which has been shown to offer reliable estimates for large networks.

Partitioning the observation window into independent intervals is a crucial step that can be carried out in three ways, depending on the available information of the network evolution. If these intervals are known a priori, they can be used as inputs for the estimation of activity values. If only the number of these intervals is known, then a *supervised* method is necessary, which takes the number of intervals as an input, and returns an interval partition. Finally, if no information is available, an *unsupervised* method is necessary to identify the partition from the available time series.

The simplest supervised method entails choosing the length of the intervals at random, such that their sum equals the length of the total observation window. This naïve approach should set a lower bound for the performance of our approach to the backbone inference. Other effective supervised methods include the parsimonious temporal aggregation [33], piece-wise constant approximation [34], and V-optimal histograms [35]. A freeware software that implements these methods is available in [36].

A convenient unsupervised method is the Bayesian blocks (BB) representation [37]. The BB method employs maximum likelihood and marginal posterior functions to separate statistically significant features from random observational errors. In this way, it relaxes common assumptions regarding the smoothness or shape of the overall temporal evolution, without constraining the process of partitioning the observation window. We refer to our methodology toward backbone inference as *evolving activity-driven model* (EADM), encompassing the null model formulation, the identification of the time-varying activities, and the statistical test.

We acknowledge that partition into intervals is not always necessary. For instance, if a system is stationary, then the number of connections generated at each time step is constant. In this case, the total observation window is contained in only one interval. To investigate such a scenario, we examine a simplified version of the EADM, where only one interval is present so that the EADM reduces to a classical ADN (referred to as $\text{EADM}_{I=1}$).

Beyond comparing our approach with its simplest incarnation that utilizes a single time interval, we further consider three different methods: SVN, TFM, and $\text{TFM}_{\text{rhythm}}$. We consider both an artificial, synthetic network (benchmark) and seven real-world networks (datasets). For each network (artificial or real), we set the maximum computational time of 24 hours, thereby dismissing longer processes.

The synthetic network is useful for validating our model in a controlled setting. In fact, it considers activity values as piece-wise constant functions in time with a ground-truth on the backbone network. We consider three different scenarios. First, we assume knowledge

about the interval partition, thereby fully exploiting the capabilities of our method. Then, we consider the case in which limited information is available about the interval partition. When only the number of intervals is available, we use the naïve supervised method and estimate the length of the intervals at random. When no information about the interval partition is accessible, we utilize the unsupervised BB method. Afterwards, when tackling backbone detection of real systems, we focus on the BB method, as we have no prior knowledge about the interval partition.

1.3 Main results

A critical result of our study is the analytical characterization of the conditions in which one must account for time-varying individual properties to accurately infer backbone networks. Our analysis suggests that considering time-varying properties is necessary when the system is not stationary or when the activation pattern of a node is correlated with the activation pattern of another node.

Further, from the analysis of synthetic networks, we conclude that our methodology outperforms the SVN, TFM, and $\text{TFM}_{\text{rhythm}}$, whereby it leads to more reliable inference of backbone networks in synthetic data, where a ground-truth backbone is known. Interestingly, in both synthetic and real networks, we find that our method reconstructs a backbone with a subset of links found by other methods, thereby diminishing the number false positive links (links wrongly identified as part of the backbone network). Overall, the three methods available in the literature result in equivalent inferences, similar to the special case of $\text{EADM}_{I=1}$, in which we execute our approach without partitioning the observation window.

Assuming that individual activities are constant in time could lead to incorrect classification of irreducible links and parts of the backbone network. Considering individual activities as piece-wise constant functions of time offers improved estimates and more reliable results.

1.4 Paper organization

The paper is organized as follows. In Section 2, we introduce the null model and articulate our procedure to detect significant interactions in time-varying networks. In Section 3, we describe our main findings by comparing the performance of our approach with other methods on synthetic networks, in which the backbone network is known, and on real datasets where different claims can be formulated depending on the method that is pursued. Finally, in Section 4, we draw our main conclusions and outline potential directions for further inquiry.

2 Significant links

In this Section, we articulate the EADM, our approach to the detection of the irreducible backbone from the time series of each individual link. First, we present our null model, which defines the process of generating temporal interactions from node-specific and piece-wise constant properties. Then, we elucidate the inference procedure of the nodes' activities within the null model from available time series, assuming to be able to access the switching events. Further, we present the statistical test from which we filter reducible links and retain irreducible ones, thus finding the backbone network. Finally, we discuss the computational complexity of our methodology.

2.1 Null model

We consider a time-varying network of N nodes evolving in a observation window of $T \gg 1$ time steps, labeled by time index $t = 1, ..., T$, with a unitary resolution. The same modeling framework is valid for a continuous time evolution.

At each time step t, nodes are connected through a binary, possibly disconnected, undirected network whose adjacency matrix, $A(t)$, stochastically varies in time. Each temporal connection is the realization of a Bernoulli variable, whereby the probability that two distinct nodes i and j are connected at time t is equal to

$$p_{ij}(t) = a_i(t)a_j(t), \tag{1}$$

where $a_i(t)$ and $a_j(t)$ are the activities of nodes i and j at time t, respectively.

Activities vary according to a switching rule, whereby they are kept constant over I disjoint time intervals indexed by $\Delta = 1, \ldots, I$. The generic Δth time interval starts at time $t_{\text{in}}(\Delta)$ and has a duration $\tau(\Delta)$, such that $\sum_{\Delta=1}^{I} \tau(\Delta) = T$. The interval partition might be a priori known or it should be determined from the time series as explained below.

When only the number of intervals I is known, a supervised method should be used to determine the interval partition. A crude possibility is to assume a random partition in I intervals, which strains the use of the null model and sets a lower bound for the EADM performance. On the contrary, if I is unknown, then,

an unsupervised method should be used. Specifically, we use the BB representation [37]. In this case, we analyze the total number of temporal links created at time t

$$\Omega(t) = \sum_{i,j=1;i<j}^{N} A_{ij}(t), \tag{2}$$

where $A_{ij}(t)$ is the ijth entry of the network adjacency matrix at time t. This method returns a set of independent intervals containing a uniform total number of connections.

To characterize the network evolution at the intermediate scale of the switching rule, that is, over successive intervals, we define a weight matrix for each interval, summing the number of occurrences of links between any two nodes. Specifically, in the Δth interval we define the random variable

$$w_{ij}(\Delta) = \sum_{t=t_{\rm in}(\Delta)}^{t_{\rm in}(\Delta)+\tau(\Delta)-1} A_{ij}(t). \tag{3}$$

To count the overall number of temporal connections between nodes i and j forming a link ij along the observation window, it is sufficient to sum the corresponding weights, resulting into the following aggregated random variable:

$$\overline{w}_{ij} = \sum_{\Delta=1}^{I} w_{ij}(\Delta) = \sum_{t=1}^{T} A_{ij}(t). \tag{4}$$

By construction, the weight $w_{ij}(\Delta)$ is a binomial variable and $\overline{w}_{ij}$ the sum of non-identical binomial random variables, described through a Poisson binomial distribution. Since no closed-form expression is available for the Poisson binomial distribution, this is usually approximated by the Poisson distribution [38–40], with expected value

$$\mathrm{E}\left[\overline{w}_{ij}\right] = \sum_{t=1}^{T} p_{ij}(t). \tag{5}$$

From the weight matrix, we define the *strength* of the ith node in the Δth interval as

$$s_i\left(\Delta\right) = \sum_{j=1}^{N} w_{ij}(\Delta). \tag{6}$$

This quantity encapsulates the total number of temporal links generated by the ith node within an interval. The total number of temporal links generated in the whole network in the Δth interval is therefore

$$W(\Delta) = \frac{1}{2}\sum_{i=1}^{N} s_i(\Delta). \tag{7}$$

Both $s_i\left(\Delta\right)$ and $W(\Delta)$ can be approximated by Poisson random variables, being linear combinations of independent non-identical binomial random variables.

2.2 Estimation of the activities from time series

In order to compute the probability that two distinct nodes i and j are connected at time t, as given in Eq. (1), we must estimate the time-varying activities $a_i(t)$ and $a_j(t)$, assumed to be piece-wise constant over known successive intervals. A possible line of approach entails the use of the weighted configuration model [31, 32], which implies that the activity of node i in the Δth time interval $t_{\rm in}(\Delta), \ldots, t_{\rm in}(\Delta)+\tau(\Delta)-1$ can be estimated from the time series of the temporal connections $A^{\rm ts}_{ij}(t)$, where we utilize a superscript "ts" to identify that the realizations from the corresponding random variables are experimental or numerical time series.

Hence, we obtain

$$\begin{aligned} a_i\left(t\right) &= \frac{s_i^{\rm ts}\left(\Delta\right)/\tau(\Delta)}{\sqrt{\left(2W^{\rm ts}(\Delta)-1\right)/\tau(\Delta)}} = \\ &= \frac{s_i^{\rm ts}\left(\Delta\right)}{\sqrt{\left(2W^{\rm ts}(\Delta)-1\right)\tau(\Delta)}}, \end{aligned} \tag{8}$$

where $s_i^{\rm ts}\left(\Delta\right)$ and $W^{\rm ts}(\Delta)$ are estimated from the time series, and $\tau(\Delta)$ is derived from the interval partition. In Eq. (8), the activity $a_i(t)$ in the Δth interval is estimated as the ratio between the average number of temporal links created per time step by node i, $s_i^{\rm ts}\left(\Delta\right)/\tau(\Delta)$, over a measure of the same quantity for the entire network, $\sqrt{\left(2W^{\rm ts}(\Delta)-1\right)/\tau(\Delta)}$. We note that the use of a square root in the denominator is consistency with the weighted configuration model [31]. Further, when only one link is created in the Δth interval, $W^{\rm ts}(\Delta)=1$, such that, the factor $2W^{\rm ts}(\Delta)-1=1$, in agreement with the static configuration model [32]. The accuracy of the estimate relies on the assumption that $W^{\rm ts}(\Delta)\gg 1$ and the network is large, that is, a large number of events is occurring in each interval and a large number of nodes is participating in the system's evolution. In the Appendix, we examine the accuracy of Eq. (8) as a function of the network size.

By replacing Eq. (8) in Eq. (1), we obtain the probability[1] of observing a link ij in the Δth time interval $t_{\rm in}(\Delta), \ldots, t_{\rm in}(\Delta)+\tau(\Delta)-1$

$$p_{ij}(t) = \frac{s_i^{\rm ts}\left(\Delta\right)s_j^{\rm ts}\left(\Delta\right)}{\left(2W^{\rm ts}(\Delta)-1\right)\tau(\Delta)}. \tag{9}$$

[1] According to the weighted configuration model [31, 32], Eq. (9) represents the expected number of links formed between node i and j in the $\tau(\Delta)$ snapshots in the Δth interval. Since most temporal networks are sparse, we can assume that $p_{ij}(t)\in[0,1)$ and refer to it as a probability.

2.3 Statistical analysis

To determine whether a link is a node-specific temporal connection or part of the irreducible backbone, we compute a p-value α_{ij} for each link observed at least once in the evolving network and compare it with a proper significance threshold. If the p-value is below the significance threshold, then the corresponding link appears more often than what the null model would predict and should therefore be associated with the backbone.

Thus, we examine the probability distribution of the generic weight of the ijth link over the entire observation window. As previously stated, the distribution is conveniently described by a Poisson distribution as

$$P\left(x;\mathrm{E}\left[\overline{w}_{ij}\right]\right)=\frac{1}{x!}\mathrm{E}\left[\overline{w}_{ij}\right]^{x}e^{-\mathrm{E}\left[\overline{w}_{ij}\right]}, \tag{10}$$

where x is the realization of the random variable. The distribution in Eq. (10) can be explicitly computed from empirical data, using Eq. (5) and the estimation of $p_{ij}(t)$ in Eq. (9), as

$$P\left(x;\mathrm{E}\left[\overline{w}_{ij}\right]\right)=\frac{1}{x!}\left[\sum_{\Delta=1}^{I}\frac{s_i^{\mathrm{ts}}\left(\Delta\right)s_j^{\mathrm{ts}}\left(\Delta\right)}{2W^{\mathrm{ts}}(\Delta)-1}\right]^{x}\exp\left[-\sum_{\Delta=1}^{I}\frac{s_i^{\mathrm{ts}}\left(\Delta\right)s_j^{\mathrm{ts}}\left(\Delta\right)}{2W^{\mathrm{ts}}(\Delta)-1}\right]. \tag{11}$$

The p-value α_{ij} of the link ij in the *overall* network evolution is then computed according to the cumulative function of the Poisson distribution

$$\alpha_{ij}\equiv 1-\sum_{x=0}^{\overline{w}_{ij}^{\mathrm{ts}}-1}P\left(x;\mathrm{E}\left[\overline{w}_{ij}\right]\right). \tag{12}$$

Upon computing a p-value for every pair of nodes in the network, one should perform a statistical test on all the links observed at least once in the evolving network. Given that multiple hypotheses are tested, a multiple hypothesis test correction is required [41]. We use the Bonferroni correction which modifies the significance threshold to $\beta^{*}=\beta/N_E=0.01/N_E$, where N_E is the number of links observed at least once in the evolving network [42]. This correction ensures that no false positives will be included with probability $1-\beta$. A possible, less restrictive alternative may be a procedure that controls the false discovery rate [43]. Specifically such a procedure ensures that the fraction of false positive is less than β.

2.4 Computational complexity

Here, we examine in detail the computational complexity of our method to detect significant links. For clarity, we discuss separately the three required steps: (i) finding the interval partition; (ii) estimating the individual activities a_i and probability p_{ij} in Eqs. (8) and (9); and (iii) computing the p-values α_{ij} in Eq. (12).

To find the interval partition, we use the BB representation [37]. Given a time series composed of T successive time steps, the total number of links $\Omega(t)$ is computed according to Eq. (2) and used as an input for the BB representation. This method determines whether $\Omega(t)\neq\Omega(t+1)$, $\forall t=1,2,\ldots,T-1$, to identify the number of change points T_{cp} (when the time series changes value). From the knowledge of T_{cp}, the maximum number of possible intervals is computed as $I_{\max}=T_{\mathrm{cp}}+1$. The interval partition is calculated with a computational time that scales as $\mathcal{O}(I_{\max}^2)$, which is affordable even for $I_{\max}\sim 10^6$ [37].

The next step is to estimate the individual activities a_i and probability p_{ij} in Eqs. (8) and (9). These empirical estimations depend on the number of intervals I and the amount of temporal links in each interval. The latter might substantially affect the algorithm's complexity, which ranges from $\mathcal{O}(NI)$ for sparse networks, to $\mathcal{O}(N^2I)$ for dense networks.

Finally, the p-values are computed according with Eq. (12), which has a computational complexity of $\mathcal{O}(N_E)$, where N_E is the number of links observed at least once in the evolving network. For sparse networks, this reduces to $\mathcal{O}(N)$.

The above three steps are independent and their computational costs add up, such that for sparse networks, the bottleneck of our approach is either the detection of the interval partition or the computation of the p-values. If the time series is larger than the number of links observed at least once in the evolving network, then the complexity is $\mathcal{O}(I_{\max}^2)$. In the opposite scenario, our method has a computational cost of $\mathcal{O}(N)$.

3 Results

In this Section, we assess the performance of the EADM in detecting the backbone of temporal networks and we compare such a performance with four models that assume time-invariant activities. We specify our study to the SVN, TFM, $\text{TFM}_{\text{rhythm}}$, and $\text{EADM}_{I=1}$ (a simplified version of our model that uses time-invariant activities). We limit the computational time for each method to 24 hours, on an Intel(R) Xeon(R) CPU E5-2697 v3 @ 2.60GHz, which we consider a reasonable computational burden for the backbone inference.

First, we analytically determine conditions for which the EADM yields equivalent results to the $\text{EADM}_{I=1}$, which allows for speculating when time-varying activities could play a salient role in the backbone detection. This corresponds to cases where the system is not stationary or the activation patterns of the nodes are correlated.

Then, we numerically assess the performance of the EADM, $\text{EADM}_{I=1}$, TFM, and SVN in detecting the backbone of temporal networks generated via an artificial network. Given that the $\text{TFM}_{\text{rhythm}}$ requires the solution of $N+T-1$ equations, its implementation on synthetic data exceeds the computational time limit of 24 hours per simulation. Therefore, its performance is not assessed on synthetic datasets. The key findings of our comparisons are: (i) the EADM offers improved performance with respect to the other methods, thereby reducing the number of false positives in the backbone network; (ii) the $\text{EADM}_{I=1}$, TFM, and SVN have comparable performance for all situations under scrutiny; and (iii) the EADM performs better when using the BB method for time interval partitioning, rather than a naïve interval partition.

Finally, we compare the irreducible backbone extracted from all models under study on several real datasets: Primary school, High school, and Museum contact patterns are from the SocioPatterns project [44]; Message, Email, and Stack overflow datasets are from the SNAP database [45]; and Enron email dataset [46]. For the Primary school, High school, and Museum datasets, we remove the time intervals when no links are recorded. Some simulations of the TFM (which solves N equations) and $\text{TFM}_{\text{rhythm}}$ exceed our computational time limit. In all the seven datasets, the EADM finds less links than other methods, which perform very similar to each other. This observation is in agreement with numerical computations on the synthetic networks, suggesting that assuming individual activities as constant in time leads to an overestimation in the number of links of the backbone network.

3.1 Analytical derivation

We start by estimating the probability of having the occurence of link ij in the $\text{EADM}_{I=1}$, that is, when individual activities are constant in time. In this case, Eq. (9) reads as

$$p_{ij} = \frac{\overline{s}_i^{\text{ts}}\overline{s}_j^{\text{ts}}}{\left(2\overline{W}^{\text{ts}}-1\right)T}, \tag{13}$$

where we define the total strength in the overall observation window $\overline{s}_i^{\text{ts}} = \sum_{\Delta=1}^{I} s_i^{\text{ts}}(\Delta)$, and the total number of temporal links in the overall observation window $\overline{W}^{\text{ts}} = \sum_{\Delta=1}^{I} W^{\text{ts}}(\Delta)$. Thus, the expected number of links in the $\text{EADM}_{I=1}$ is

$$\text{E}_{I=1}\left[\overline{w}_{ij}\right] = Tp_{ij} = \frac{\overline{s}_i^{\text{ts}}\overline{s}_j^{\text{ts}}}{2\overline{W}^{\text{ts}}-1}, \tag{14}$$

which is equivalent to predictions of the weighted configuration model [31].

In general, $\text{E}_{I=1}\left[\overline{w}_{ij}\right]$ in Eq. (14) is different from $\text{E}\left[\overline{w}_{ij}\right]$ in Eq. (5), thereby begetting different statistical inferences of the backbone. Under the following conditions, we show that the two inferences are similar:

(i) if the system is stationary, $W^{\text{ts}}(\Delta) \approx \overline{W}^{\text{ts}}/T$, and
(ii) if, for any link ij, the activation pattern of node i is independent of the one of node j.

To prove this claim, we compute $\text{E}\left[\overline{w}_{ij}\right]$ and demonstrate that it converges to $\text{E}_{I=1}\left[\overline{w}_{ij}\right]$ for $\overline{W}^{\text{ts}} > W^{\text{ts}}(\Delta) \gg 1$ and large networks. By replacing Eq. (9) into Eq. (5) for $\overline{W}^{\text{ts}} > W^{\text{ts}}(\Delta) \gg 1$, we obtain

$$\text{E}\left[\overline{w}_{ij}\right] \simeq \sum_{\Delta=1}^{I} \frac{s_i^{\text{ts}}(\Delta)\, s_j^{\text{ts}}(\Delta)}{2W^{\text{ts}}(\Delta)}. \tag{15}$$

First, we assume the system as stationary, as in condition (i), so that

$$\text{E}\left[\overline{w}_{ij}\right] \simeq \frac{T^2}{2\overline{W}^{\text{ts}}}\frac{1}{T}\sum_{\Delta=1}^{I} s_i^{\text{ts}}(\Delta)\, s_j^{\text{ts}}(\Delta). \tag{16}$$

Then, we apply condition (ii), which for large networks implies the mean-field approximation $\langle s_i^{\text{ts}}(\Delta)\rangle\langle s_j^{\text{ts}}(\Delta)\rangle \simeq \langle s_i^{\text{ts}}(\Delta)\, s_j^{\text{ts}}(\Delta)\rangle$, obtaining

$$\text{E}\left[\overline{w}_{ij}\right] \simeq \frac{T^2}{2\overline{W}^{\text{ts}}}\left[\frac{1}{T}\sum_{\Delta=1}^{I} s_i^{\text{ts}}(\Delta)\right]\left[\frac{1}{T}\sum_{\Delta'=1}^{I} s_j^{\text{ts}}(\Delta')\right], \tag{17}$$

and, from the time series of $\overline{s}_i^{\rm ts}$ and $\overline{s}_j^{\rm ts}$ in Eqs. (6), we establish

$$\mathrm{E}\left[\overline{w}_{ij}\right] \simeq \frac{\overline{s}_i^{\rm ts}\overline{s}_j^{\rm ts}}{2\overline{W}^{\rm ts}}. \tag{18}$$

Finally, we observe that Eq. (18) corresponds to $\mathrm{E}_{\mathrm{I}=1}\left[\overline{w}_{ij}\right]$ in Eq. (14) under the assumption that $\overline{W}^{\rm ts} \gg 1$, which concludes our proof.

If the system is not stationary or the activation patterns of nodes are correlated, one might expect that the EADM will yield different predictions than the $\mathrm{EADM}_{I=1}$, supporting the need for properly partitioning the observation window toward the successful detection of the backbone network.

3.2 Performance comparison on synthetic data

The considered synthetic data begets a temporal network where reducible links, generated by the EADM, coexist with the irreducible backbone. Reducible links evolve over a observation window T, partitioned into I successive intervals. Nodes have interval-dependent (piece-wise constant) activities $a(t)$ drawn from a power law distribution $F(a) \sim a^{-2.1}$, with $a \in [a_{\min}, 1]$. The value $a_{\min}$ represents the minimum possible value for the individual activity in the system and it is chosen to be greater than zero to avoid divergence in the distribution [15,24,47].

Between two consecutive intervals, $t_1 \in [t_{\rm in}(\Delta - 1), t_{\rm in}(\Delta - 1) + \tau(\Delta - 1) - 1]$ and $t_2 \in [t_{\rm in}(\Delta), t_{\rm in}(\Delta) + \tau(\Delta) - 1]$, the activity values vary according to

$$a_i(t_2) = a_i(t_1)p + y(1-p), \tag{19}$$

where p is an autocorrelation parameter and y a random number extracted from $F(a)$. For $p = 1$, individual activities are time-invariant, while for $p < 1$, they exhibit temporal correlations.

A small fraction δ of all the links observed at least once in the network is arbitrarily assigned to the backbone network. An additional parameter λ is used to measure the preponderance of the backbone during the observation window, such that if $\lambda = 1$, these links are always present, and if $\lambda < 1$, they could not be present at all times. Details about the algorithm to construct synthetic data are presented in the Appendix.

We numerically assess the improvement provided by the EADM in the backbone detection with respect to the TFM, SVN, and $\mathrm{EADM}_{I=1}$. Performing our numerical experiments using the $\mathrm{TFM}_{\rm rhythm}$ exceeds our allotted computational time, such that its performance could not be tested against this artificial network. Performance is otherwise scored using two well-known metrics, precision and recall [48]. The former is computed as the ratio between the number of links detected, which belong to the irreducible backbone (true positives), divided by the total number of detected links (sum of true and false positives). The latter metric is the ratio between the true positives divided by the total number of links in the irreducible backbone (sum of true positives and false negatives).

First, we assume that the partition into intervals is known and we estimate the activity values according to Eq. (8), thereby applying the EADM. Then, we release this assumption toward choosing the length of the intervals at random or we employ the unsupervised BB method to estimate such a partition.

3.2.1 The EADM improves backbone detection

In our comparison, we assess the role of two important parameters: (i) the autocorrelation parameter p, which regulates the variation of individual activities over time, from $p = 0$ (completely uncorrelated individual activities) to $p = 1$ (time-invariant activities), and (ii) the ratio between the average interval length and the total length of the observation window $\langle\tau(\Delta)\rangle/T$, which quantifies the fraction of switches in activity patterns. For $\langle\tau(\Delta)\rangle/T = 1$, individual activities are constant in time, while as $\langle\tau(\Delta)\rangle/T$ approaches zero, individual activities rapidly change over time. We select two values of λ, which lead to different scenarios: a larger value of λ that begets an easily detectable backbone where all irreducible links can be discovered, examined in Fig. 2; and a smaller value of λ that results into a partially hidden backbone where some irreducible links cannot be discovered, considered in Fig. 3.

Figures 2 (a) and (c) support the claim that the EADM is a valuable approach to infer the backbone networks for any choice of the autocorrelation parameter, since precision and recall are always close to one. Figures 3 (a) and (c) confirms that no false positive are detected by the EADM even if the backbone is not preponderant; however, some irreducible links cannot be discovered and the recall is lower than one. On the contrary, the TFM, SVN, and $\mathrm{EADM}_{I=1}$ are successful only when the value of the autocorrelation parameter approaches 1, such that individual activities are practically time-invariant. In this case, we register values of the precision close to 1.

Figures 2 (b) and (d), and Figs. 3 (b) and (d), suggest that the EADM outperforms the other methods for intermediate values of the number of switching intervals in terms of precision. Performance is, on the other

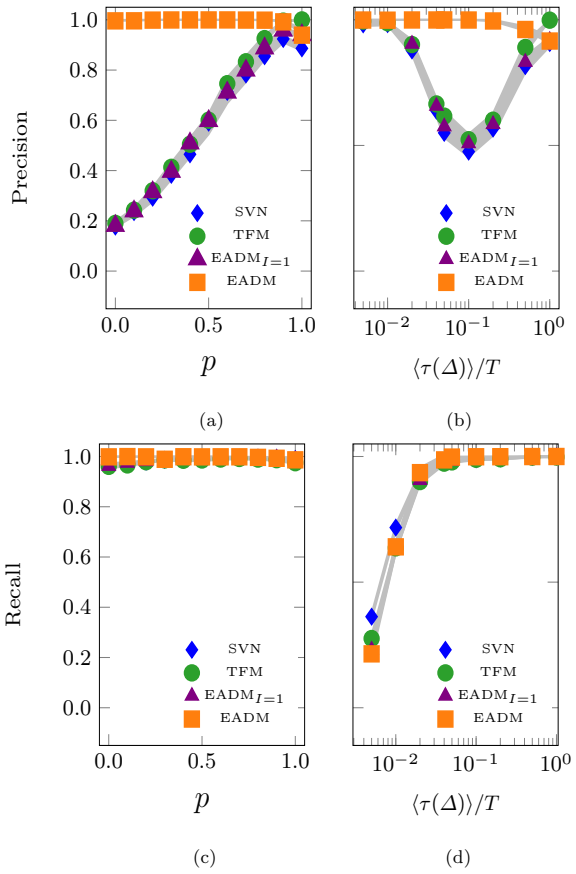

Fig. 2 Performance comparison against the synthetic network, assuming a priori knowledge of the interval partition for the EADM implementation. We assess precision and recall as a function of the autocorrelation parameter p and ratio between the average interval length and the total observation window $\langle\tau(\Delta)\rangle/T$. The horizontal axis in panels (b) and (d) is obtained by fixing $T = 5,000$ and varying I to span different values of $\langle\tau(\Delta)\rangle = T/I$. In (a) and (c), we hold $\langle\tau(\Delta)\rangle$ and I, fixed to 500 and 10, while in (b) and (d) we set $p = 0.4$. Other parameter values are: $N = 100$, $\delta = 0.01$, $\lambda = 0.025$, and $a_{\min} = [\sqrt{\langle\tau(\Delta)\rangle}]^{-1}$. Markers indicate the average of 10^2 independent simulations, 95% confidence interval is displayed in gray.

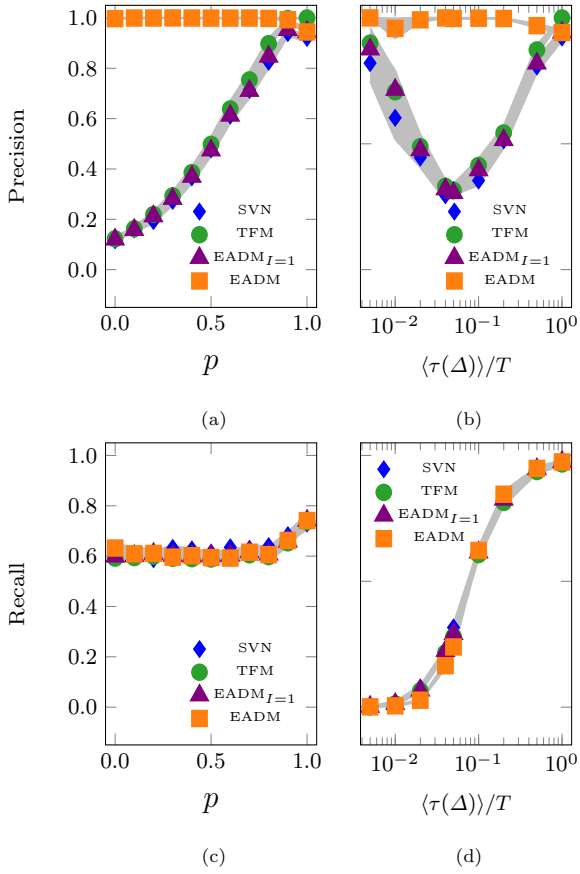

Fig. 3 Performance comparison against the synthetic network, assuming a priori knowledge of the interval partition for the EADM implementation. We assess precision and recall as a function of the autocorrelation parameter p and ratio between the average interval length and the total observation window $\langle\tau(\Delta)\rangle/T$. The horizontal axis in panels (b) and (d) is obtained by fixing $T = 5,000$ and varying I to span different values of $\langle\tau(\Delta)\rangle = T/I$. In (a) and (c), we hold $\langle\tau(\Delta)\rangle$ and I, fixed to 500 and 10, while in (b) and (d) we set $p = 0.4$. Other parameter values are: $N = 100$, $\delta = 0.01$, $\lambda = 0.010$, and $a_{\min} = [\sqrt{\langle\tau(\Delta)\rangle}]^{-1}$. Markers indicate the average of 10^2 independent simulations, 95% confidence interval is displayed in gray.

hand, comparable for the extreme cases of $\langle\tau(\Delta)/T\rangle$ approaching one or zero. While the comparable predictions that we register for the former case $\langle\tau(\Delta)/T\rangle \simeq 1$ can be anticipated due to the limited variability of the activity patterns, the similar performance registered for the latter case $\langle\tau(\Delta)/T\rangle \simeq 0$ deserves some comments. Under fast switching conditions, none of the algorithms leads to large values of the recall, such that only a small fraction of the backbone can be reconstructed, although with high accuracy. Under fast switching conditions, the SVN, TFM, and EADM$_{I=1}$ would practically capture an annealed version of the network that is not representative of the backbone. On the other hand, an algorithm like ours that tracks time-variations requires a large number of realizations for performing the statistical test, which become unfeasible for time series of limited length with several switches. The similar performance registered for the TFM, SVN, and EADM$_{I=1}$ is discussed in the Appendix.

Taken together, the higher precision of the EADM and its comparable recall to other methods, suggest that the EADM is successful in reducing the number of false positives. These advantages will be explored and further detailed when we examine real networks.

	SVN	TFM	$\text{EADM}_{I=1}$	EADM+R	EADM+BB
Precision					
$\lambda = 0.025$	0.483 (0.461, 0.504)	0.510 (0.488, 0.532)	0.490 (0.468,0.511)	0.995 (0.992, 0.998)	0.991 (0.985, 0.996)
$\lambda = 0.010$	0.365 (0.345, 0.385)	0.379 (0.355, 0.403)	0.361 (0.339, 0.384)	0.992 (0.988, 0.996)	0.987 (0.980, 0.995)
Recall					
$\lambda = 0.025$	0.991 (0.988, 0.994)	0.984 (0.980, 0.988)	0.986 (0.983, 0.990)	0.999 (0.999, 1.000)	0.999 (0.998, 1.000)
$\lambda = 0.010$	0.617 (0.603, 0.631)	0.584 (0.569, 0.599)	0.594 (0.580, 0.609)	0.593 (0.577, 0.609)	0.602 (0.587, 0.616)

Table 1 Performance comparison in the synthetic network, assuming limited information of the interval partition for the EADM implementation. For the EADM+R, the number of intervals I is known. For the EADM+BB, no a priori information of the interval partition is assumed. We study two values of λ, which exemplify two levels of preponderance of the backbone. Parameter values are: $N = 100$, $T = 5,000$, $I = 10$, $\langle \tau(\Delta) \rangle = 500$, $\delta = 0.01$, and $a_{\min} = [\sqrt{\langle \tau(\Delta) \rangle}]^{-1}$. Tabulated values are the average of 10^2 independent simulations, the 95% confidence interval is displayed in brackets.

3.2.2 The backbone inference does not require knowledge about activity patterns

Thus far, we have assumed complete knowledge about the interval partition, which is used as an input parameter in the EADM. However, this situation is rarely met in reality, where only limited information about the interval partition may be available. To improve the degree of realism of the analysis, we consider two different scenarios. In the first one, we assume knowledge about the number of intervals and choose their length at random. This naïve approach sets a lower bound for the EADM performance. We identify this setting as EADM+R, where "R" stands for random. In the second scenario, we assume no a priori knowledge about the interval partition, and we resort to the unsupervised BB method. We identify this situation as EADM+BB.

In Table 1, we study precision and recall of the five methods for two choices of the parameter values, considered in Figs. 2 and 3. The two cases pertain to two different choices of λ, where we were fully successful in reconstructing the backbone or registered a recall less than one with full knowledge about the interval partitions.

Results in Table 1 indicate that all the five methods lead to a comparable recall, which is equivalent to results in Figs. 2 and 3. However, we document a remarkable improvement in precision for the EADM+R and EADM+BB, when compared to the other three methods that do not account for time-variations of activity patterns. Given that the EADM+BB does not require any knowledge about the intervals, it should be the approach of choice in backbone inference. In the Appendix, we report further insight on the comparison between the EADM+R and EADM+BB, which indicate that the EADM+R might lead to inadequate inferences if the number of intervals is not exactly known. This is the case of real networks, which motivates the systematic use of the EADM+BB in the discovery process.

3.3 Application to real networks

Based on our previous assessment on synthetic data, we turn to real networks where we compare predictions of the EADM+BB with other existing methods.

The comparison is carried out using three different metrics: (i) the number of significant links; (ii) the Jaccard index [49]; and (iii) the overlap coefficient [50]. We denote the set of irreducible links detected by our method as $L_{\text{EADM+BB}}$, and the others as L_x, where $x = \text{EADM}_{I=1}$, TFM, $\text{TFM}_{\text{rhythm}}$, or SVN. The Jaccard coefficient is defined as

$$J(L_{\text{EADM+BB}}, L_x) = \frac{|L_{\text{EADM+BB}} \cap L_x|}{|L_{\text{EADM+BB}} \cup L_x|}, \tag{20}$$

where $|\cdot|$ indicates the set cardinality. The overlap coefficient is defined as

$$O(L_{\text{EADM+BB}}, L_x) = \frac{|L_{\text{EADM+BB}} \cup L_x|}{\min\left(|L_{\text{EADM+BB}}|, |L_x|\right)}. \tag{21}$$

The Jaccard coefficient yields the fraction of common links between the EADM+BB and each of the other methods, while the overlap coefficient quantifies the extent of the overlap between the two detected backbones.

Each real dataset is examined at four different time resolutions obtained by counting, without repetitions, all the links that occur at the nominal frequency of acquisition of the experimental observation. Table 2 summarizes the seven datasets considered in this work. For ease of illustration, in this main document, we focus on the Primary school and the Museum datasets; the Appendix contains the analysis of all datasets. Similar to the study of synthetic data, simulations are terminated after 24 hours of computational time.

Data	# nodes	# temporal links	# aggregated links	Time span	Resolution (r_1, r_2, r_3, r_4)
Primary school	242	125,773	8,317	2 days	(20 sec., 1 min., 5 min., 15 min.)
High school	126	28,561	1,708	4 days	(20 sec., 1 min., 5 min., 15 min.)
Enron	182	125,235	2,097	1,313 days	(15 min., 1 hour, 1 day, 1 week)
Email	986	329,910	16,025	526 days	(15 min., 1 hour, 1 day, 1 week)
Message	1,899	59,835	13,838	194 days	(15 min., 1 hour, 1 day, 1 week)
Stack overflow	24,759	506,550	187,986	2,351 days	(15 min., 1 hour, 1 day, 1 week)
Museum	10,852	411,490	44,120	81 days	(1 min., 5 min., 15 min., 30 min.)

Table 2 Data summary of the seven real datasets under consideration. The "# temporal links" column indicates the total number of temporal links in the dataset. The "Resolution" column lists four different time resolutions for conducting the inference. For brevity, in the manuscript we use symbols from $r_1, \ldots, r_4$ to refer to the different resolutions, ordered from the smallest to the largest. For the Primary school, High school, and Museum datasets, we remove the time intervals when no links are recorded.

In Fig. 4, we summarize our comparison. In panels (a) and (d), we show a sample of the time series of the total number of temporal links, $\Omega^{\rm ts}(t)$, and the interval partition identified by the BB method. For both datasets, $\Omega^{\rm ts}(t)$ is not stationary, reflecting the complexity of the time evolution where each student or teacher in the Primary school dataset, or museum visitor in the Museum dataset will come irregularly into contact with others. In panels (b) and (e), we compare the number of significant links detected by the five methods considered in this work. In agreement with evidence from Figs. 2 and 3 and Table 1 on synthetic data, the EADM+BB identifies a smaller number of links than other methods, whose predictions are equivalent.

We also observe that improving on the resolution of the data, by lowering the time step, increases the number of significant links detected by all the methods. This is related to the decrease of the number of temporal links $\overline{W}^{\rm ts}$ due to the deletion of the repeated temporal links (no multiedges are allowed in a single time step). Such a deletion affects mostly the nodes with highest activity, which generate many links over time. In this way, the heterogeneity of the system is reduced, reflecting in a lower number of detected significant links. Although all the methods are affected by the time resolution of the dataset, the EADM+BB is the one that shows the strongest tendency, as it requires the identification of switches in the activity patterns, which could be masked by node-specific links in poorly resolved datasets.

In Fig. 5, we compare the detected backbone networks using the Jaccard index and the overlap coefficient. The Jaccard index suggests a strong similarity in the case of the Primary school dataset and a weak similarity in the case of the Museum dataset. On the other hand, the overlap coefficient suggests that in both cases our method identifies a subset of links within those detected by other methods.

Individual activities have different temporal features in the two datasets. In the Primary school dataset, most students and teachers are recorded for the entire observation window, and can recurrently interact with each other. As a result, the impact of explicitly considering time-varying activities is limited, and a time-averaged representation of the phenomenon constitutes an acceptable approximation. On the other hand, in the Museum dataset, visitors spend only a few hours in the museum, which comprises a small fraction of the observation window of 81 days. In this case, approximating individual activities with constant quantities along the whole observation window is an oversimplification of the problem that could lead to several false positives in the backbone detection.

In Fig. 6, we assess the accuracy of the methods in estimating the overall network connectivity, measured in terms of the total number of links in the observation window. We compare the expected number of temporal links, $\mathrm{E}\big[\overline{W}\big]$, with observations in the time series, $\overline{W}^{\rm ts}$. We specifically compute the relative error, $|\mathrm{E}\big[\overline{W}\big] - \overline{W}^{\rm ts}|/\overline{W}^{\rm ts}$, where we use $\mathrm{E}\big[\overline{W}\big] = \sum_{i,j=1;i<j}^{N} \sum_{t=1}^{T} p_{ij}(t)$ for the EADM+BB; $\mathrm{E}_{I=1}\big[\overline{W}\big] = \sum_{i,j=1;i<j}^{N} T p_{ij}$ for the EADM$_{I=1}$; Eq. (25) in the Appendix for the TFM; and Eq. (29) in the Appendix for the TFM$_{\rm rhythm}$. The SVN is excluded from this analysis as it takes $\overline{W}^{\rm ts}$ as an input parameter. For all the considered datasets and all backbone detection methods, relative error is at most 5%, thereby indicating that all the methods are accurate in capturing the evolution of the network connectivity. In agreement with our expectation, the relative error for the TFM and the TFM$_{\rm rhythm}$ (when available) is lower than those for the EADM$_{I=1}$ and the EADM+BB. In fact, as previously discussed, the TFM and TFM$_{\rm rhythm}$ refine the estimation of individual activities through a maximum likelihood approach.

While all the methods work with approximately the same number of links throughout the temporal evolu-

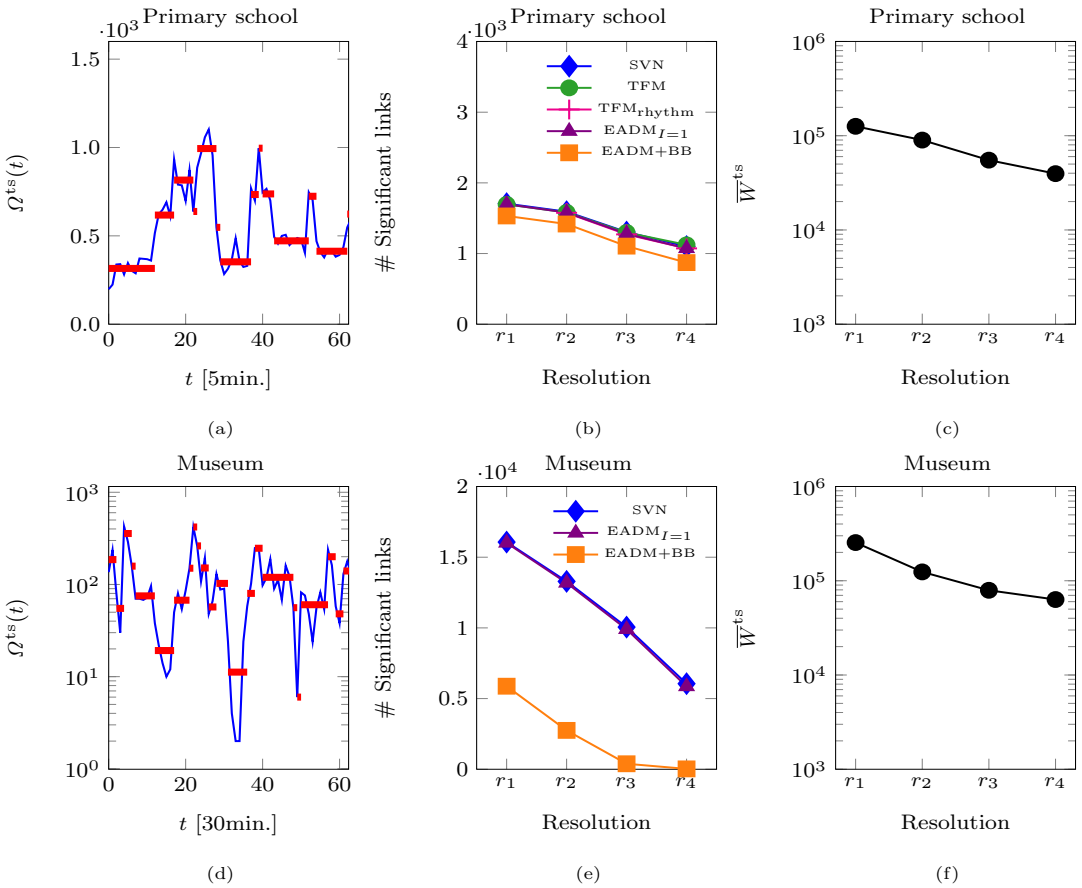

Fig. 4 Influence of temporal patterns on backbone detection. In panels (a) and (d), we show the total number of temporal links created over time, $\Omega^{\mathrm{ts}}(t)$, for one chosen resolution (indicated in square brackets) of the Primary school and Museum datasets, respectively. For visualization purposes, we select the first 60 time steps. Partition into intervals is performed by applying the Bayesian blocks (BB) method to the time series. Horizontal red segments represent the average number of temporal links in a specific interval. In panels (b) and (e), we compare the number of significant links found by the methods under scrutiny for the same two datasets. Inferences not reported correspond to simulations that exceed our time limit of 24 hours. In panels (c) and (f), we display the number of temporal links, $\overline{W}^{\mathrm{ts}}$, as a function of the resolution for the same two datasets. The exact values of the resolution are found in Table 2.

tion, as shown in Fig. 6, they yield different predictions for the underlying backbone network as shown in Figs. 4 and 5. The most remarkable difference depends on whether one is accounting or not for time-varying activities.

Based on the study of the synthetic datasets in Figs. 2 and 3, we propose that the discovery process of the backbone network should be formulated by assuming, in general, that activity patterns are time-varying.

4 Discussion

In this paper, we have introduced the *evolving activity-driven model*, a novel approach to detect the backbone network against time variations of node-specific properties, encapsulated by the activity. The activity of a node represents its propensity to generate links over time, which, in real systems, is seldom constant [51]. Should one look at temporal networks formed by humans, the individual activity might be low during sleeping hours and breaks, while it should be high during working hours. Whether differences in individual behavior modify the backbone network is the topic of our study.

To this end, we analytically identify conditions in which temporal patterns of the activity will have a secondary role on the detection of the backbone. These conditions correspond to the system being stationary and the activation patterns of the nodes not correlated. Based on these claims, we speculate that straining either of these conditions will lead to a salient role of temporal variations of the activity patterns on the backbone detection. Afterwards, we compare the backbone networks detected by our methodology with inferences

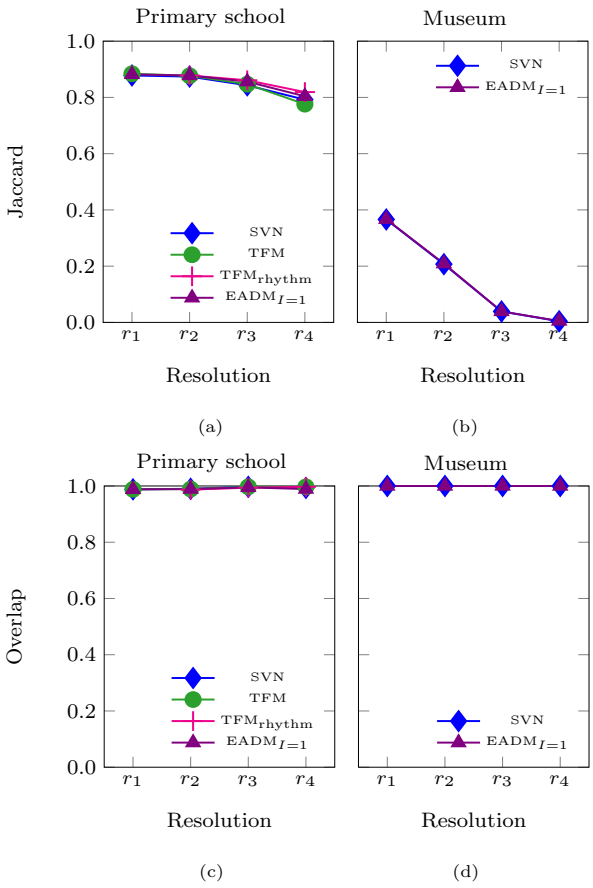

Fig. 5 Differences and similarities in the backbone networks detected by the EADM+BB and the other methods (indicated in the legends). In panels (a) and (c), we show the Jaccard index for the Primary school and Museum datasets, respectively. In panels (b) and (d), we display the overlap coefficient for the same two datasets. Inferences not reported correspond to simulations that exceed our time limit of 24 hours.

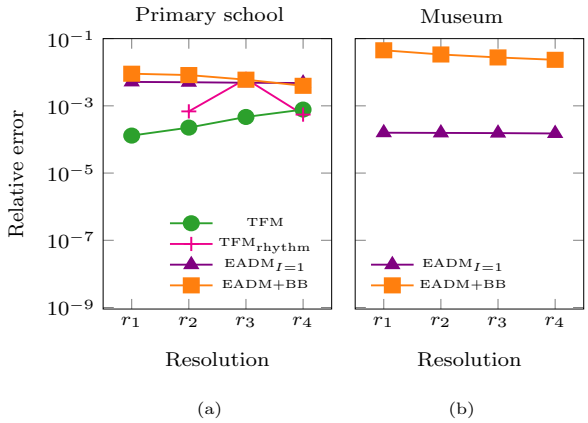

Fig. 6 Relative error between the total number of temporal links found in the time series, $\overline{W}^{\text{ts}}$, and the number of total temporal links estimated from the backbone detection algorithms under consideration. The SVN is discarded from this analysis since it uses $\overline{W}^{\text{ts}}$ as an input for filtering reducible links. Inferences not reported correspond to simulations that exceed our time limit of 24 hours.

supported by four other approaches all of which assume that individual activities are constant in time. Specifically, we focus on a modification of the evolving activity-driven model with constant activities; the *statistically validated network* [10]; and two versions of the *temporal fitness model* [7]. In the first version of the temporal fitness model, activities are kept constant in time and their estimates are refined through a maximum likelihood approach; whilst, in the second one, a time-varying parameter is utilized to encapsulate circadian and weekly patterns.

For both synthetic and real datasets, our model identifies a subset of the links determined by the other methods. By utilizing a ground-truth backbone network from the synthetic data, we conclude that our approach reduces the number of links that are incorrectly classified as part of the backbone network (false positives) and improves the precision of the detection process. These results suggest that accounting for temporal variations in the activity plays an important role in backbone detection, potentially leading to the discovery of a different backbone network. The most remarkable differences are noted when nodes display activity patterns that intensively vary in time, without a recurrent behavior. For instance, in the Museum dataset, visitors spend only a few hours in the museum, which is a small fraction of the total observation window of 81 days. In contrast with other methods that all yield equivalent predictions, our approach discovers a small backbone network, representative of people visiting museums in small groups that constitute a backbone network. We expect a similar behavior when analyzing airports, restaurants, hotels, websites, and chat rooms, where people access alone or in small groups and only for a limited time.

The size of the backbone network discovered by our approach is influenced by the time resolution of the dataset. Working with poorly resolved data will challenge the feasibility of network inference, which is evident when dealing with visitors in a museum, and calls for the careful selection of a time resolution, which could be a confounding factor in detecting the backbone network of a system. This claim is in line with [52], which focused on random walks over temporal networks.

The main advantages of the proposed evolving activity-driven methodology are three: (i) its limited computational time, whereby it allows for fast network discovery even when dealing with long time series and large networks (simulations presented in this paper are only

a few minutes long); (ii) its ability to cogently model temporal activity patterns, which cannot be addressed by the current state-of-the-art approaches; and (iii) its consistency with the literature, whereby it yields equivalent predictions to existing methods when dealing with time-invariant activity patterns.

Our approach can find applications across several domains of science and engineering, beyond the exemplary social networks examined herein. For example, it could be implemented in the study of functional networks in the brain, which primarily relies on simple thresholding [26], or in the analysis of the World wide web, power grids, chemical reaction networks, where topology identification methods [53–55] can benefit from a statistically-principled approach to discard reducible links.

However, our approach is not free of limitations. We detect switches in the individual activities over successive disjoint intervals by considering the overall system evolution, rather than the individual time series. In principle, we cannot exclude the possibility that individual activities could vary in time in such a way that the overall system evolution remains stationary. In this case, our approach would not be able to detect time variations in individual activities. In principle, we could attempt at working with individual time series, but this would challenge the use of the Bayesian block representation [37] that relies on nodes to activate multiple times – a condition that is not satisfied by the sparse datasets considered in our study. In addition, the overall computational cost would depend also on the size of the system, thereby hindering implementation for large networks. At the same time, we acknowledge that our approach is not applicable to small networks, composed of only a few tens of nodes, because we conduct the estimates of the individual activities using a weighted configuration model that requires large networks [31].

Future research will involve the formulation of algorithms for the optimal selection of the resolution which are needed for enhancing the performance of our methodology and the one proposed in [7]. The evolving activity-driven model might be further extended through the detection of individual interval partitions, one for each node in the network, overcoming the assumptions that the interval partition is unique and that all of the activities switch synchronously. More long-term, fruitful lines of research should aim at unraveling the intricate interplay between individual features and the formation of temporal interaction patterns.

Code availability

Python 2.7 codes are freely available here [56].

Funding

The authors acknowledges financial support from the National Science Foundation under grant No. CMMI-1561134. A.R. acknowledges financial support from Compagnia di San Paolo, Italy, and the Italian Ministry of Foreign Affairs and International Cooperation, within the project, “Macro to Micro: uncovering the hidden mechanisms driving network dynamics”.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. P. Holme, J. Saramäki, Physics Reports **519**(3), 97 (2012)
2. P. Holme, The European Physical Journal B **88**, 1 (2015)
3. N. Masuda, R. Lambiotte, World Scientific **4** (2016)
4. D. Lazer, A. Pentland, L. Adamic, S. Aral, A.L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al., Science **323**(5915), 721 (2009)
5. T. Ivancevic, L. Jain, J. Pattison, A. Hariz, Nonlinear Dynamics **56**(1-2), 23 (2009)
6. S. Battiston, J.D. Farmer, A. Flache, D. Garlaschelli, A.G. Haldane, H. Heesterbeek, C. Hommes, C. Jaeger, R. May, M. Scheffer, Science **351**(6275), 818 (2016)
7. T. Kobayashi, T. Takaguchi, A. Barrat, Nature Communications **10**(1), 220 (2019)
8. Z. Wu, L.A. Braunstein, S. Havlin, H.E. Stanley, Physical Review Letters **96**(14), 148702 (2006)
9. M.Á. Serrano, M. Boguná, A. Vespignani, Proceedings of the National Academy of Sciences **106**(16), 6483 (2009)
10. M. Tumminello, S. Micciche, F. Lillo, J. Piilo, R.N. Mantegna, PloS One **6**(3), e17994 (2011)
11. M.X. Li, V. Palchykov, Z.Q. Jiang, K. Kaski, J. Kertész, S. Miccichè, M. Tumminello, W.X. Zhou, R.N. Mantegna, New Journal of Physics **16**(8), 083038 (2014)
12. V. Gemmetto, A. Cardillo, D. Garlaschelli, arXiv preprint arXiv:1706.00230 (2017)
13. G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, G. Caldarelli, Nature Reviews Physics **1**(1), 58 (2019)
14. R. Marcaccioli, G. Livan, Nature Communications **10**(1), 745 (2019)
15. N. Perra, B. Gonçalves, R. Pastor-Satorras, A. Vespignani, Scientific Reports **2**, 469 (2012)
16. L. Zino, A. Rizzo, M. Porfiri, Journal of Complex Networks **5**(6), 924 (2017)
17. K. Sun, A. Baronchelli, N. Perra, The European Physical Journal B **88**(12), 326 (2015)
18. L. Zino, A. Rizzo, M. Porfiri, SIAM Journal on Applied Dynamical Systems **17**(4), 2830 (2018)
19. M. Nadini, K. Sun, E. Ubaldi, M. Starnini, A. Rizzo, N. Perra, Scientific Reports **8**(1), 2352 (2018)
20. Q.H. Liu, X. Xiong, Q. Zhang, N. Perra, Physical Review E **98**(6), 062303 (2018)
21. Y. Lei, X. Jiang, Q. Guo, Y. Ma, M. Li, Z. Zheng, Physical Review E **93**(3), 032308 (2016)

22. A. Rizzo, M. Frasca, M. Porfiri, Physical Review E **90**(4), 042801 (2014)
23. M. Nadini, A. Rizzo, M. Porfiri, IEEE Transactions on Network Science and Engineering (2018)
24. A. Rizzo, B. Pedalino, M. Porfiri, Journal of Theoretical Biology **394**, 212 (2016)
25. A. Moinet, M. Starnini, R. Pastor-Satorras, Physical Review Letters **114**(10), 108701 (2015)
26. V.M. Eguiluz, D.R. Chialvo, G.A. Cecchi, M. Baliki, A.V. Apkarian, Physical Review Letters **94**(1), 018102 (2005)
27. F. Musciotto, L. Marotta, J. Piilo, R.N. Mantegna, Palgrave Communications **4**(1), 92 (2018)
28. C. Curme, M. Tumminello, R.N. Mantegna, H.E. Stanley, D.Y. Kenett, Quantitative Finance **15**(8), 1375 (2015)
29. D. Challet, R. Chicheportiche, M. Lallouache, S. Kassibrakis, Advances in Complex Systems p. 1850019 (2018)
30. C. Bongiorno, A. London, S. Micciché, R.N. Mantegna, Physical Review E **96**(2), 022321 (2017)
31. M.Á. Serrano, M. Boguñá, AIP Conference Proceedings **776**(1), 101 (2005)
32. M.E.J. Newman, Oxford university press (2010)
33. J. Gordevičius, J. Gamper, M. Böhlen, The VLDB Journal **21**(3), 309 (2012)
34. H. Konno, T. Kuno, Operations Research Letters **7**(4), 205 (1988)
35. H.V. Jagadish, N. Koudas, S. Muthukrishnan, V. Poosala, K.C. Sevcik, T. Suel, in *VLDB*, vol. 98 (1998), vol. 98, pp. 24–27
36. G. Mahlknecht, M.H. Bohlen, A. Dignös, J. Gamper, Proceedings of the 29th International Conference on Scientific and Statistical Database Management p. 40 (2017)
37. J.D. Scargle, J.P. Norris, B. Jackson, J. Chiang, The Astrophysical Journal **764**(2), 167 (2013)
38. A. Barbour, G. Eagleson, Advances in Applied Probability **15**(3), 585 (1983)
39. J.M. Steele, The American Mathematical Monthly **101**(1), 48 (1994)
40. L. Le Cam, et al., Pacific Journal of Mathematics **10**(4), 1181 (1960)
41. J.P. Shaffer, Annual Review of Psychology **46**(1), 561 (1995)
42. Y. Hochberg, A. Tamhane, Wiley, New York (2009)
43. Y. Benjamini, Y. Hochberg, Journal of the Royal Statistical Society. Series B (Methodological) pp. 289–300 (1995)
44. www.sociopatterns.org
45. https://snap.stanford.edu/data/index.html
46. https://www.cs.cmu.edu/~./enron/
47. N. Perra, D. Balcan, B. Gonçalves, A. Vespignani, PloS One **6**(8), e23084 (2011)
48. C.D. Manning, P. Raghavan, H. Schütze, Cambridge University Press (2008)
49. P. Jaccard, Bulletin de la Société Vaudoise des Sciences Naturelles **37**, 547 (1901)
50. M. Vijaymeena, K. Kavitha, Machine Learning and Applications: An International Journal **3**, 19 (2016)
51. G. Kossinets, D.J. Watts, Science **311**(5757), 88 (2006)
52. B. Ribeiro, N. Perra, A. Baronchelli, Scientific Reports **3**, 3006 (2013)
53. D.D. Zhou, B. Hu, Z.H. Guan, R.Q. Liao, J.W. Xiao, Nonlinear Dynamics **91**(2), 785 (2018)
54. J. Chen, J.a. Lu, J. Zhou, Nonlinear Dynamics **75**(4), 761 (2014)
55. Y. Xu, W. Zhou, J. Fang, Nonlinear Dynamics **68**(1-2), 195 (2012)
56. https://github.com/matnado/Backbone-Detection

5 Appendix

5.1 Backbone detection methods

Here, we succinctly summarize the *temporal fitness model* (TFM) [7], the *temporal fitness model with rhythm* ($\text{TFM}_{\text{rhythm}}$) [7], and the *statistically validated network* (SVN) [10].

5.1.1 Temporal fitness model

The TFM considers a temporal network formed by N nodes evolving over T discrete time steps. All multiple links occurring within the same time step are removed, so that the total number of temporal links between node i and j is bounded by T. First, individual activities are computed according to

$$a_i = \frac{\overline{s}_i^{\text{ts}}}{\sqrt{2\overline{W}^{ts}T}}. \tag{22}$$

Then, their values are refined through a maximum likelihood approach, which requires the solution of N equations

$$\sum_{j=1;j\neq i}^{N} \frac{\overline{w}_{ij}^{\text{ts}} - Ta_i^* a_j^*}{1 - a_i^* a_j^*} = 0, i = 1, \ldots, N, \tag{23}$$

where $\boldsymbol{a^*} = (a_1^*, \ldots, a_N^*)$ contains the optimal values for the individual activities. Finally, the p-value α_{ij} for the link generated between node i and j is computed from the cumulative function of the Binomial distribution as

$$\alpha_{ij} \equiv 1 - \sum_{x=0}^{\overline{w}_{ij}^{\text{ts}}-1} B\left(x; T, a_i^* a_j^*\right). \tag{24}$$

All p-values, one for each link in the network, are compared with a threshold value β, properly corrected by using a multiple hypotheses correction [42,43], and any value lower than β adds a link to the backbone network.

For our purposes, we also compute the expected total number of temporal links in the overall temporal evolution

$$\text{E}\left[\overline{W}\right] = T \sum_{i,j=1;i<j}^{N} a_i^* a_j^*. \tag{25}$$

5.1.2 Temporal fitness model with rhythm

The $\text{TFM}_{\text{rhythm}}$ adds to the TFM T time-varying coefficients, one for each time step, $\boldsymbol{\xi} = (\xi(1), \ldots, \xi(T))$. First, every element in the time-varying vector is manually set to 0.999, with the exception of $\xi(1)$ which is set equal to one. Individual activities are estimated according to Eq. (22). To determine the optimal values $(\boldsymbol{a^*}, \boldsymbol{\xi^*})$ in the maximum likelihood sense, we solve the system of $N + T - 1$ equations

$$\begin{aligned} \sum_{t=1}^{T}\sum_{j=1;j\neq i}^{N} \frac{A_{ij}^{\text{ts}}(t) - a_i^* a_j^* \xi^*(t)}{1 - a_i^* a_j^* \xi^*(t)} &= 0, \quad i = 1, \ldots, N, \\ \sum_{i,j=1;j\neq i}^{N} \frac{A_{ij}^{\text{ts}}(t) - a_i^* a_j^* \xi^*(t)}{1 - a_i^* a_j^* \xi^*(t)} &= 0, \qquad t = 2, \ldots, T, \end{aligned} \tag{26}$$

where $A_{ij}^{\text{ts}}(t)$ is the adjacency matrix at time t estimated from the time series. The expected number of links is computed as

$$\text{E}\left[\overline{w}_{ij}\right] = \sum_{t=1}^{T} a_i^* a_j^* \xi^*(t). \tag{27}$$

Finally, the p-value α_{ij} for the link generated between node i and j is computed from the cumulative function of the Poisson distribution as

$$\alpha_{ij} \equiv 1 - \sum_{x=0}^{\overline{w}_{ij}^{\text{ts}}-1} P\left(x; \text{E}\left[\overline{w}_{ij}\right]\right). \tag{28}$$

All the p-values, one for each link in the network, are compared to a threshold value β, properly corrected by using a multiple hypotheses correction [42,43]. Any value lower than β leads to a link in the backbone network.

For our purposes, we also compute the expected total number of temporal links in the overall temporal evolution

$$\text{E}\left[\overline{W}\right] = \sum_{i,j=1;i<j}^{N}\sum_{t=1}^{T} a_i^* a_j^* \xi^*(t). \tag{29}$$

5.1.3 Statistically validated network

The SVN considers a temporal network of N nodes evolving over an observation time window that can be either discrete or continuous in time. Temporal links are aggregated to form a weighted static network. The p-value α_{ij} for the link generated between node i and j is computed from the cumulative function of the Hypergeometric distribution as

$$\alpha_{ij} \equiv 1 - \sum_{x=0}^{\overline{w}_{ij}^{\text{ts}}-1} H\left(\overline{w}_{ij} \middle| 2\overline{W}^{\text{ts}}, \overline{s}_i^{\text{ts}}, \overline{s}_j^{\text{ts}}\right). \tag{30}$$

The p-values are compared with a threshold value β, properly corrected by using a multiple hypotheses correction [42, 43], and a link is added to the backbone network of the p-value is less than β.

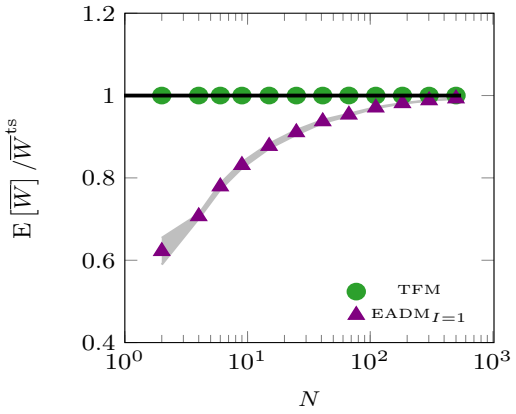

Fig. 7 Accuracy of the EADM$_{I=1}$ and the TFM in estimating the total number of temporal links in the overall time series. A perfect identification should yield a ratio between $\mathrm{E}[\overline{W}]$ and $\overline{W}^{\mathrm{ts}}$ of one (black solid line). In these simulations, we use our artificial network where no backbone is present ($\delta = \lambda = 0$) and activities are constant in time ($T = 5,000$, $I = 1$, $\langle\tau(\Delta)\rangle = T/I = 5000$, $a_{\min} = [\sqrt{\langle\tau(\Delta)\rangle}]^{-1}$, and $p = 0$). Markers indicates the average of 10^2 independent simulations, 95% confidence interval is displayed in gray.

5.2 On the similarity among the EADM$_{I=1}$, SVN, and TFM

Here, we discuss why these three methods yield similar results for both synthetic and real datasets. First, we show that the EADM$_{I=1}$ is a valid approximation of the TFM for large networks (hundreds of nodes or more). Then, we analytically examine the convergence of the SVN to the EADM$_{I=1}$.

5.2.1 On the similarity between the TFM and EADM$_{I=1}$

We consider a long observation window T, for which the Binomial distribution in Eq. (24) converges to a Poisson distribution used in our method in Eq. (12). While in the EADM$_{I=1}$ activities are estimated from the dataset using Eq. (8), in the TFM they are identified in a maximum likelihood sense [7]

In Fig. 7, we assess the ability of the EADM$_{I=1}$ and the TFM to estimate the total number of temporal links. We compute the expected values of the number of links for the EADM$_{I=1}$ as $\mathrm{E}[\overline{W}] = \sum_{i,j=1;i<j}^{N} T p_{ij}$, while we use Eq. (25) for the TFM. These values are compared with the total number of temporal links observed in the time series $\overline{W}^{\mathrm{ts}}$. As expected, the TFM works well for any network size, due to the use of the maximum likelihood. Nevertheless, the maximum likelihood approach becomes computational demanding for networks of around 1,000 nodes and beyond, thereby becoming useless for very large networks. On the other hand, the EADM$_{I=1}$ shows poor performance for small networks, while reaching the TFM for networks of 100 nodes. This improvement in performance of the EADM$_{I=1}$ is explained in [31], where it is shown that Eq. (14) is in excellent agreement with numerical simulations for large networks.

5.2.2 On the similarity between the SVN and EADM$_{I=1}$

When $\overline{W}^{\mathrm{ts}} \gg 1$, the Hypergeometric distribution in Eq. (30) converges to a Poisson distribution and its p-value becomes equivalent to the p-value for the EADM$_{I=1}$

$$\alpha_{ij} = 1 - \sum_{x=0}^{\overline{w}_{ij}^{\mathrm{ts}}-1} P\left(x; \frac{\overline{s}_i^{\mathrm{ts}}\overline{s}_j^{\mathrm{ts}}}{2\overline{W}^{\mathrm{ts}}}\right). \tag{31}$$

In all the synthetic and real data studied herein $\overline{W}^{\mathrm{ts}}$ is very large, so that Eq. (30) converges to Eq. (31).

5.3 Generation of synthetic temporal networks

To examine the precision and recall of irreducible links, we generate synthetic networks. The procedure of network generation is given as follows.

1. We consider a temporal network evolving in an observation window of length T, divided into I different intervals. We randomly select without replacement $I-1$ time steps in $\{1, ..., T\}$, which we sort as $t_{\mathrm{in}}(2) \ldots t_{\mathrm{in}}(I)$, and we set $t_{\mathrm{in}}(1) = 1$. Each interval Δ has different length $\tau(\Delta)$, so that, in general, the average length of the interval is $\langle\tau(\Delta)\rangle = T/I$.
2. The N nodes in the network have a time-varying, piece-wise constant, individual activity. We extract activity values from a power law distribution, $F(a) \sim a^{-2.1}$, with $a \in [a_{\min}, 1]$. The time-varying activity $a_i(t)$ is selected according to the following procedure:
   - When $\Delta = 1$, N activity values, one for each node in the network, are randomly extracted from $F(a)$, and held constant within $[t_{\mathrm{in}}(1), t_{\mathrm{in}}(1) + \tau(1) - 1]$.
   - When $2 \leq \Delta \leq I$ activities might be correlated between two successive intervals, $t_1 \in [t_{\mathrm{in}}(\Delta - 1), t_{\mathrm{in}}(\Delta-1)+\tau(\Delta-1)-1]$ and $t_2 \in [t_{\mathrm{in}}(\Delta), t_{\mathrm{in}}(\Delta)+\tau(\Delta)-1]$ according to Eq. (19) in the main text.
3. We generate a temporal network in the observation window $[1, T]$. Each pair of nodes ij within an interval Δ is connected with probability $a_i(\Delta)a_j(\Delta)$. As a result, we obtain a sequence of T undirected

and unweighted networks, with adjacency matrices $\hat{A}(1), \ldots, \hat{A}(T)$. These networks are generated only as a function of the individual activities.

4. Based on the node pairs that are connected at least once over T time steps of the observation window, we define the synthetic backbone. Specifically, we randomly assign a fraction δ of these node pairs to the backbone.
5. We construct T new networks $A(1), A(2), \ldots, A(T)$ from $\hat{A}(1), \hat{A}(2), \ldots, \hat{A}(T)$ by accounting for the synthetic backbone above. First, we set $A_{ij}(t) = \hat{A}_{ij}(t)$ for $t = 1, \ldots T$ for all the pairs that do not belong to the backbone. Then for the generic link ij in the backbone, we initialize $A_{ij}(1) = \hat{A}_{ij}(1)$ and we iterate the following steps for $t = 2, \ldots, T$:
   - if $\hat{A}_{ij}(t) = 1$, we maintain $A_{ij}(t) = 1$;
   - if $\hat{A}_{ij}(t) = 0$, we set $A_{ij}(t) = 1$ with probability λ and $A_{ij}(t) = 0$ with probability $1 - \lambda$.

   The parameter λ measures the preponderance of links associated with the backbone during the observation window.

5.4 Insights on the interval estimation

The EADM+R requires that the number of intervals is known a priori. Nevertheless, when dealing with real networks, our knowledge, I_e, might differ from the true value, I. This mismatch might diminish the accuracy of the backbone inference, as examined below for synthetic data. We focus on two set of parameters, which represents two possible scenarios. In the first case, $a_{\min} = [\sqrt{\langle \tau(\Delta) \rangle}]^{-1}$ and $\lambda = 0.025$, which correspond to a "dense" ADNs with an easily detectable backbone. In the second case, $a_{\min} = [\langle \tau(\Delta) \rangle]^{-1}$ and $\lambda = 0.002$, which represent a "sparse" ADNs with a partially hidden backbone.

In Fig. 8 (a) and (c), we show that if the number of estimated intervals, I_e, is greater or equal to the true value, I, precision and recall are close to one. On the contrary, in Fig. 8 (b) and (d), we observe a more dramatic scenario, in which increasing I_e hinders the performance of the method, leading to filtering out most of the links, that belong to the backbone network.

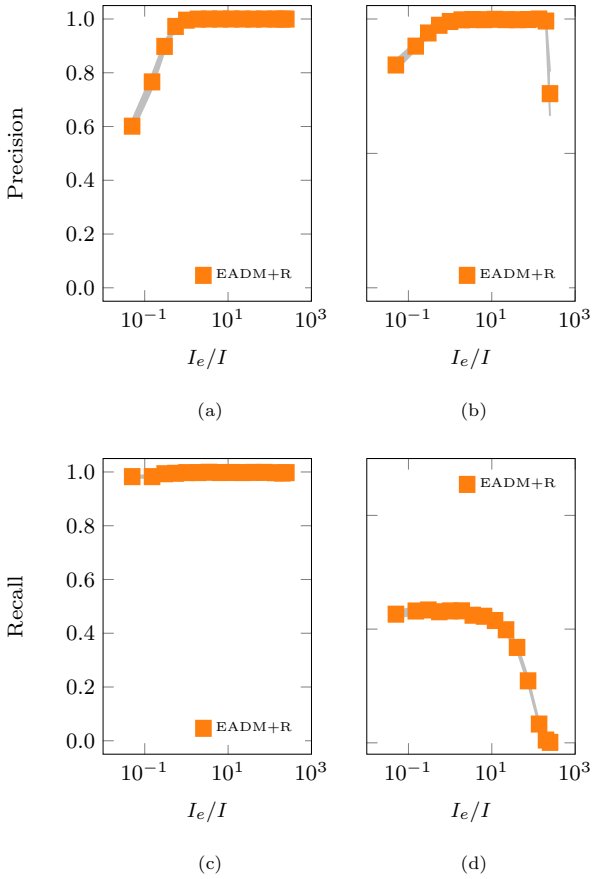

Fig. 8 Sensitivity analysis of the EADM+R to the number of estimated intervals, I_e, from $I_e = 1$ to $I_e = T - 1$. In panels (a) and (c), we set $a_{\min} = [\sqrt{\langle \tau(\Delta) \rangle}]^{-1}$ and $\lambda = 0.025$, to attain a dense ADNs and an easy-to-discover backbone. On the contrary, in panels (b) and (d), we set $a_{\min} = [\langle \tau(\Delta) \rangle]^{-1}$ and $\lambda = 0.002$, to attain sparse ADNs and a partially hidden backbone. Other parameter values are: $N = 100$, $T = 5,000$, $I = 20$, $\langle \tau(\Delta) \rangle = T/I = 250$, $\delta = 0.01$, and $p = 0.4$. Markers indicate the average of 10^2 independent simulations, 95% confidence interval is displayed in gray.

5.5 Analysis of all available real datasets

5.5.1 Significant links

We compare the backbone networks from seven real-world datasets inferred by the five methods under consideration in terms of the number of significant links. The EADM+BB always finds less links than any other methods.

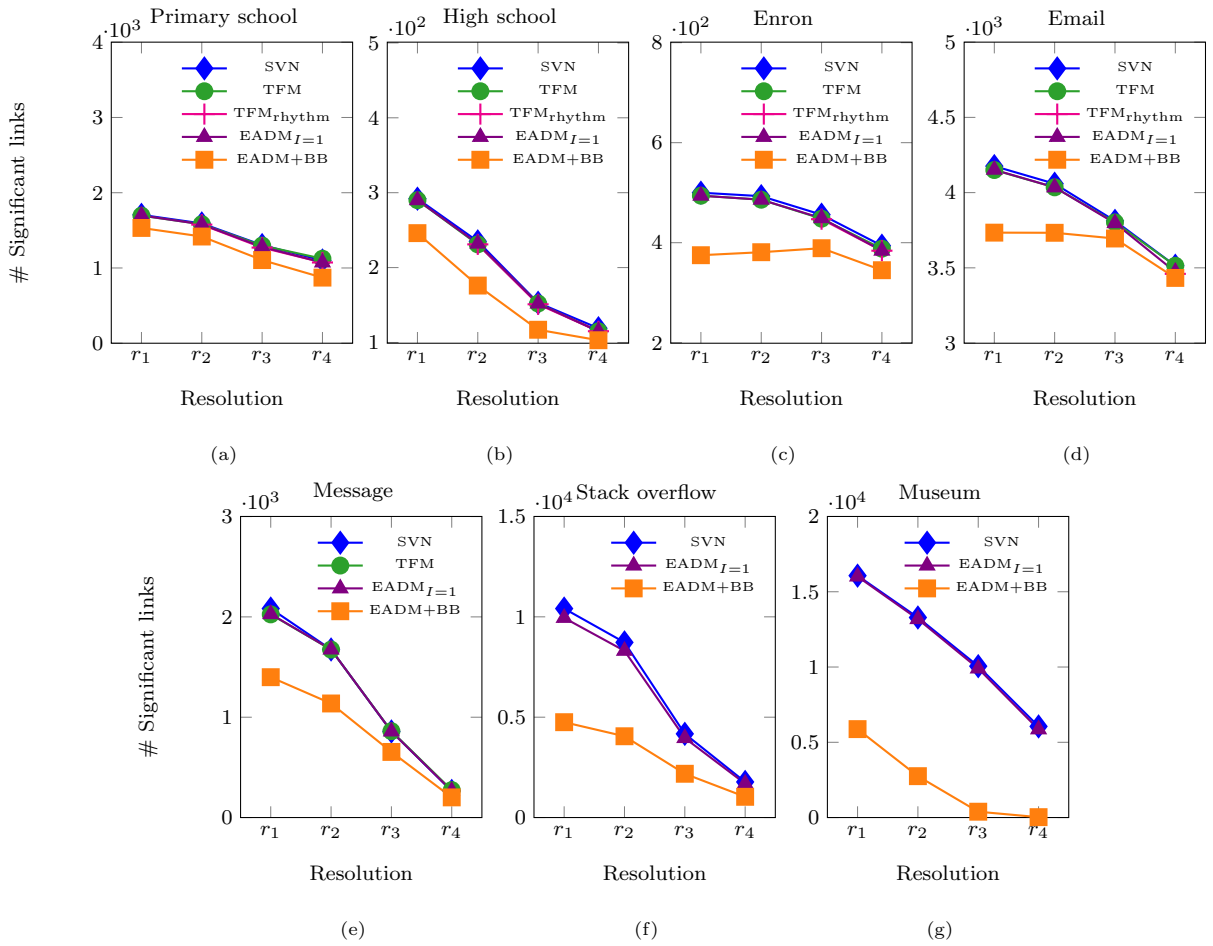

Fig. 9 Number of significant links as a function of the resolution for all real datasets under consideration. Inferences not reported correspond to simulations that exceed our time limit of 24 hours.

5.5.2 Jaccard index

In Fig. 10, we assess differences in the backbone networks detected by the EADM+BB and four methods on seven real-world datasets, in terms of the Jaccard index. We observe that the EADM+BB finds backbones different from the $\mathrm{EADM}_{I=1}$, SVN, TFM, and $\mathrm{TFM}_{\mathrm{rhythm}}$, which are equivalent.

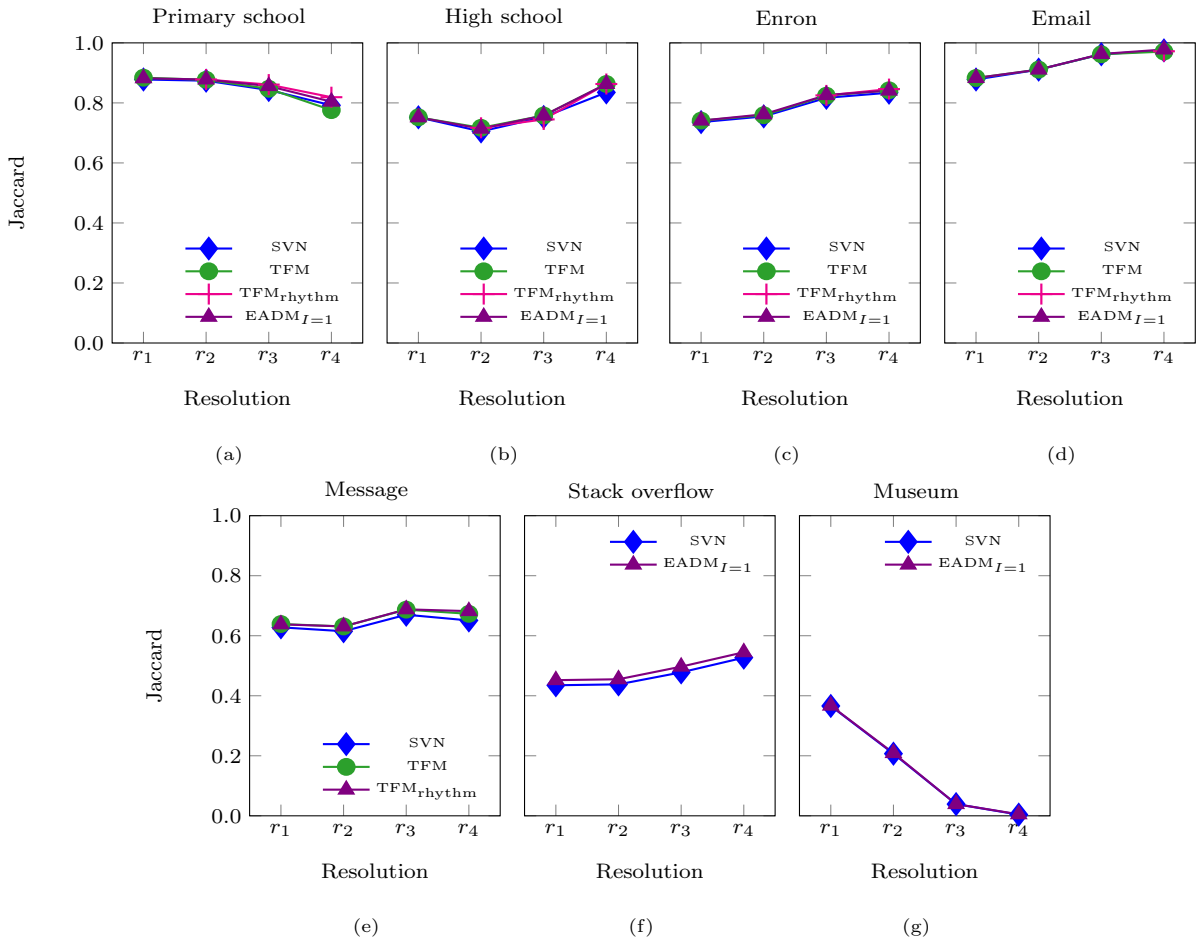

Fig. 10 Jaccard index between EADM+BB and all the other methods as a function of the resolution for all datasets under consideration. Inferences not reported correspond to simulations that exceed our time limit of 24 hours.

5.5.3 Overlap coefficient

Similar to Fig. 10, we examine the overlap coefficient of backbone networks determined by our method and the other four in Fig. 4, confirming that the EADM+BB tends to detect a subset of the links predicted by other methods – which are thus prone to false positives.

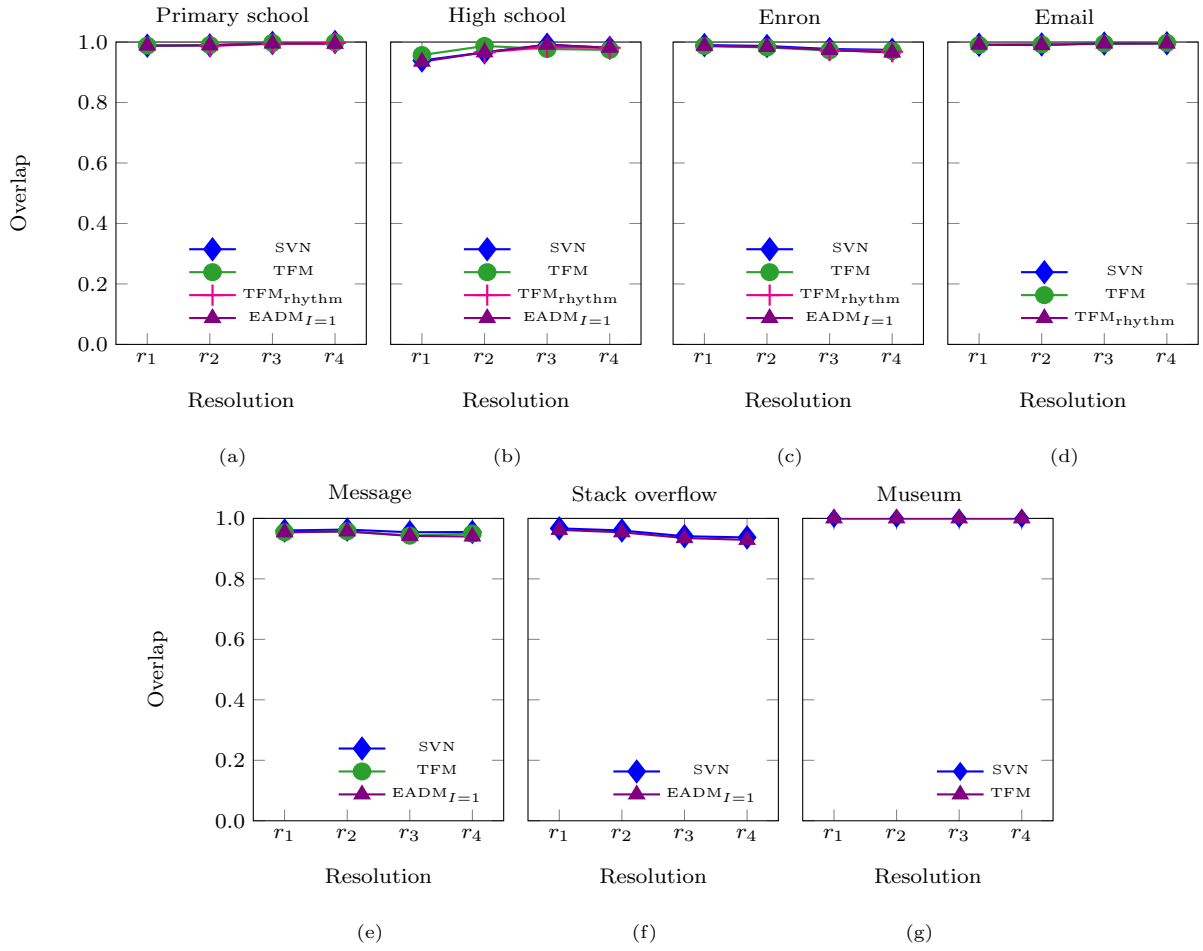

Fig. 11 Overlap coefficient between EADM+BB and all the other models as a function of the resolution for all datasets under consideration. Inferences not reported correspond to simulations that exceed our time limit of 24 hours.

5.5.4 Temporal links

In Fig. 12, we display the total number of temporal links estimated in the time series, $\overline{W}^{\mathrm{ts}}$, for all the considered methods on all the seven real-world datasets. We confirm that the number of links decreases as we increase the time resolution of the dataset.

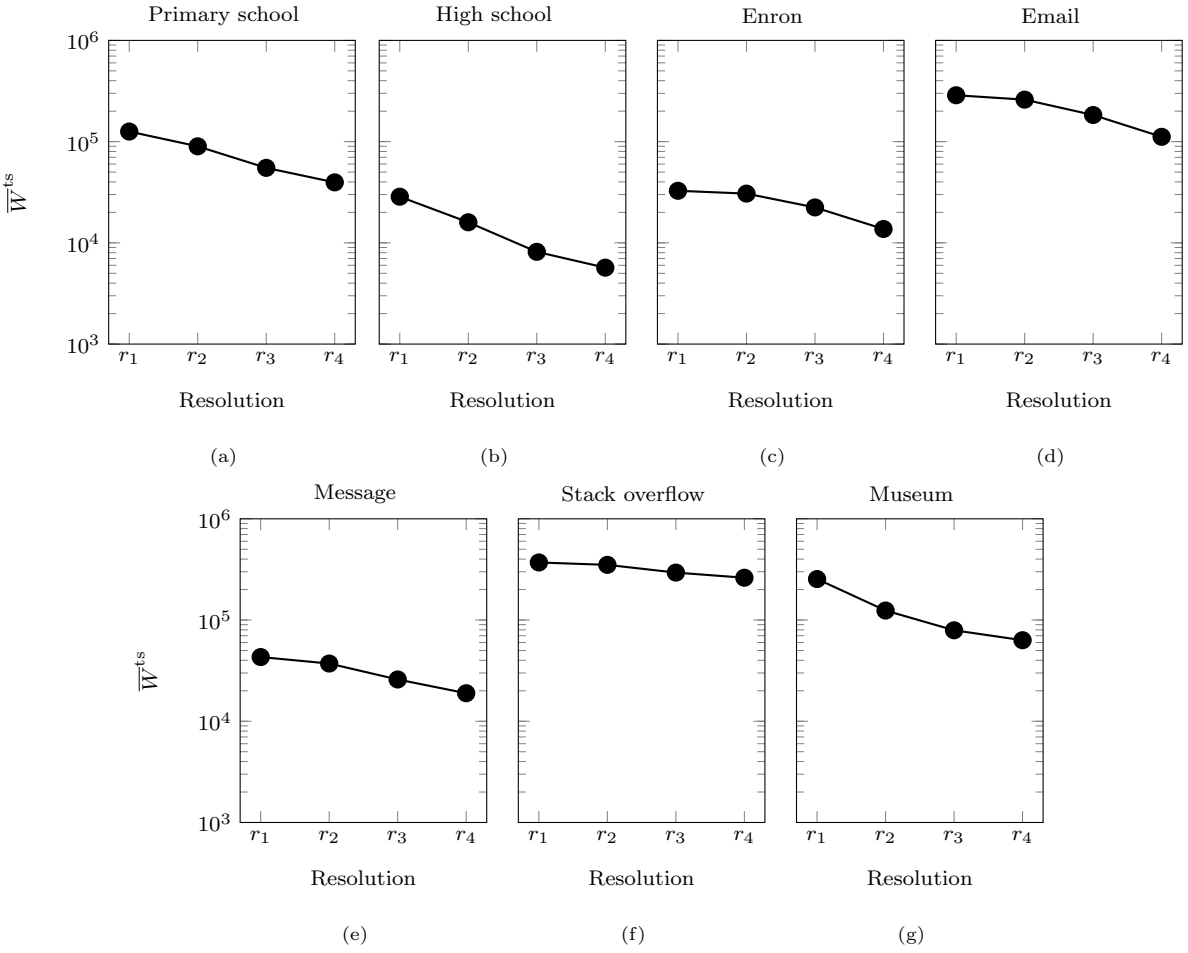

Fig. 12 Total number of temporal links estimated in the time series W^{ts} as a function of the resolution for all datasets under consideration.

5.5.5 Relative error

We analyze the accuracy of the methods in describing the overall system evolution. We compare the expected number of the total temporal links generated in, $\mathrm{E}[\overline{W}]$, with $\overline{W}^{\mathrm{ts}}$. All methods are accurate for the datasets studied herein, with a relative error up to 5%.

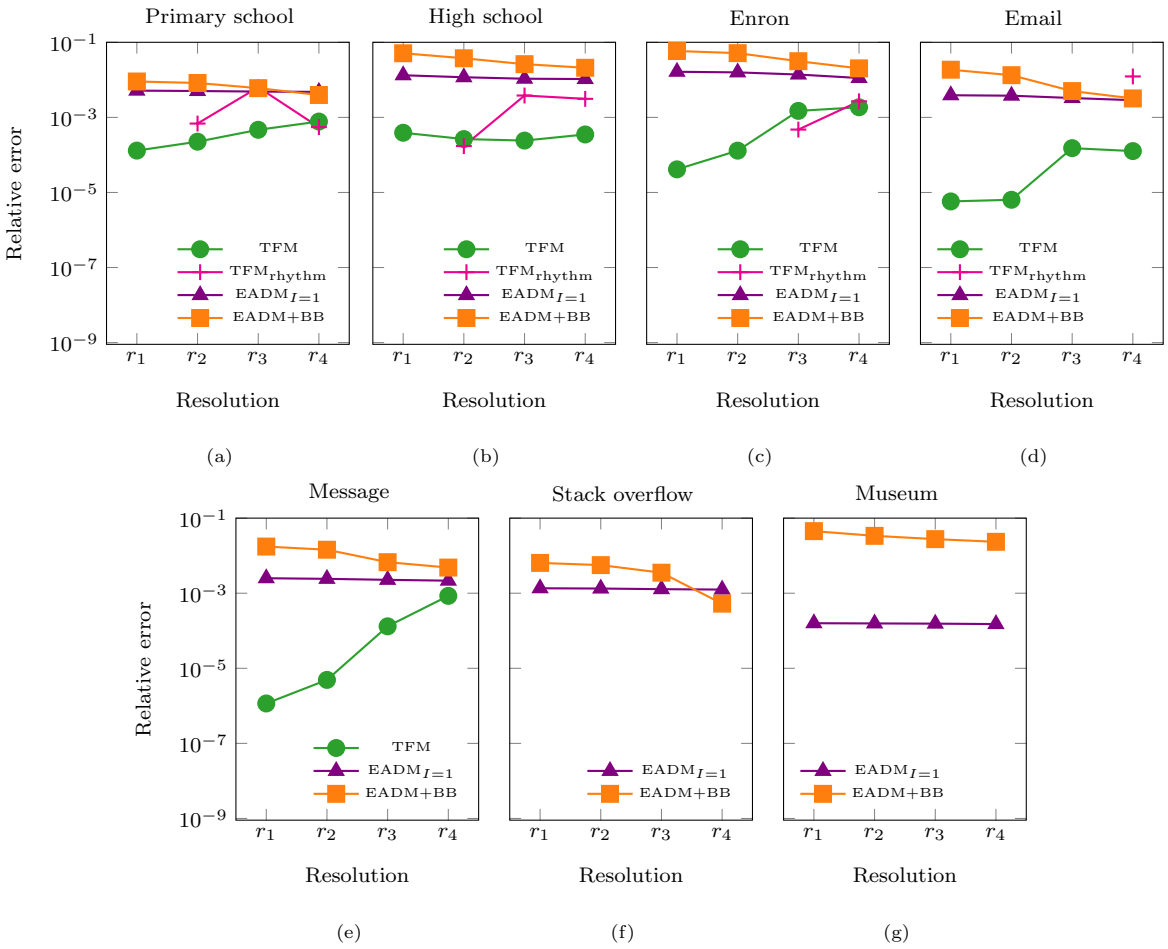

Fig. 13 Relative error between the total number of temporal links, $\overline{W}^{\mathrm{ts}}$, and the number of total temporal links estimated from the backbone detection algorithms under consideration. The SVN is discarded from this analysis because it uses $\overline{W}^{\mathrm{ts}}$ as an input. Inferences not reported correspond to simulations that exceed our time limit of 24 hours.