

PhD Thesis

**Social Contagion Mechanisms Inference on Temporal
Networks from Local and Global Views**

by

Elsa Andres

Submitted to
Central European University
Department of Network and Data Science

*In partial fulfilment of the requirements for the degree of Doctor of
Philosophy in Network Science*

Supervisor: Prof. Márton Karsai

Vienna, Austria
2024

To my Grandma Simone

Declaration of Authorship

I, the undersigned, Elsa Andres, candidate for the PhD degree in Network Science declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 4th of October, 2024



Signature

Copyright Notice

Copyright © Elsa Andres, 2024. *Social Contagion Mechanisms Inference on Temporal Networks from Local and Global Views* - This work is licensed under [Creative Commons Attribution-NonCommercial-ShareAlike \(CC BY-NC-SA\) 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/) International license.



1

¹Icon by [Font Awesome](https://fontawesome.com/)

Abstract

Networks are powerful tools to model systems composed of interacting entities, like societies where individuals are interconnected. These structures are particularly useful for studying the spread of social behaviours such as fashion trends, product adoption, or ideas themselves, as they propagate through human interactions. As societies are dynamic and constantly evolving, temporal networks, where connections between entities change over time, provide an accurate framework to study spreading phenomena.

The dynamics of temporal networks can affect propagation processes, thus it is essential to understand their evolution. In particular, their dynamics may evolve on multiple time scales characterising periodic activity patterns or structural changes. The detection of these time scales can be challenging from the direct observation of simple dynamical network properties like the activity of nodes or the density of links. In the first part of the thesis I propose two new methods based on static representations of temporal networks, which allow us to define dissimilarity metrics and compute their power spectra from their Fourier transforms. I demonstrate these methods outperform the reference measures using synthetic and real-world data sets. One approach identifies more easily periodic changes in network density, while the other one is better suited to detect periodic changes in the group structure of the network.

After understanding the characteristics of temporal networks, I delve deeper into the study of contagion processes on networks. The adoption of behaviours is largely determined by stimuli from social interactions or external sources. While some individuals may change behaviour after a single peer's influence, others require multiple exposures from their social circles or act independently. Those mechanisms, known as simple, complex and spontaneous contagions, often coexist in real-world social contagion processes. The goal of the second part of the thesis is to understand whether coexisting adoption mechanisms can be distinguished at the egocentric network level, without requiring global network information. I formulate this question as a classification problem, employing likelihood analysis and random forest classifiers in various synthetic and data-driven experiments.

While this last analysis is conducted on static networks, a more realistic scenario would involve temporal networks where individuals can be infected by both the simple and the complex contagions. Each person's behaviour adoption depends on factors like personal characteristics, the propagating behaviour, and the nature of social ties. My objective is to determine which contagion mechanism predominates in social spreading with time-varying interactions. I approach this as a classification problem using a mixed synthetic propagation model on temporal networks. By analysing the simulation curves, I identify three categories

of propagation and, through an analytical study, develop methods to detect transitions between them.

This study as a whole offers a novel perspective on the phenomena occurring at multiple time scales in temporal networks, as well as on the nature of propagation processes. Those insights allow a better understanding of contagion mechanisms from both local and global views, contributing to the broader study of dynamic systems.

List of publications

Andres E., Barrat A., Karsai M.. (2024) *Detecting periodic time scales of changes in temporal networks*. Journal of Complex Networks 12 (2), cnae004.

Andres E., Ódor G., Iacopini I., Karsai M. (2024) *Distinguishing mechanisms of social contagion from local network view*. arXiv preprint arXiv:2406.18519.

Andres E., Pastor-Satorras R., Starnini M., Karsai M. (2024) *Competition between simple and complex contagion at early time behaviour*. arXiv preprint arXiv:2410.22115.

Acknowledgment

I would like first to deeply thank my supervisor Márton Karsai, who has made my thesis a superb journey. I value in particular his ability to have great ideas and his creativity. In addition of being an admirable supervisor, he is also a great person on a human level, helping me with any troubles that I have encountered during my PhD. I am also very grateful to the professors with who I have collaborated: Alain Barrat, who taught me a lot about temporal networks since my master thesis, Romualdo Pastor-Satorras who has incredible mathematical skills and who has been very available during my period in Barcelona and Michele Starnini who gave me great advice during the project we did. I was also very lucky to have worked with two postdocs, Iacopo Iacopini and Gergely Ódor, who are not only two very helpful and inspiring persons always available when needed, but also true good friends. Finally, I am very grateful to János Kertész for offering me financial support during an injury.

I would like also to thank the students and postdocs of the department, the atmosphere was constantly shining with all those beautiful people around, who became for some of them really good friends. I am grateful to Adri for everything, since the first day, no words can sum up this journey with her, to Gabri for her warm smile every morning, her immense kindness and for being a huge support in every hard moment, to Leo dG. for his big heart and his questionable jokes, to Sandeep for the sweet moments and to have been my first climbing partner, to Ludo for his constant happiness and his passion for outdoor sports, to Onkar for his powerful high-fives and his hand-waves, to Timur for being the pierre angulaire of the department, to Thomas Robiglio for his French-speaking part and his love-cycling part, to Gale for his billardino games and his giro di Grappa, to Felipe to hike up in the mountains with me, to Doris to sledge down in the mountains with me, to Helcio for washing his radishes every day, to Marti for invoking the ghosts, to the other Marti for her enthusiasm for every sport, to Bojan for feeding us with chewing-gums, to Clara to do so much for the department, to Basti to have been a great mate in Canada, to Leo R. to show me what it is to behave like a proper Italian, to Luka for all the advice and the help, and to Berné for the afternoon walks and the morning swims. I would like to thank also Teo, Yijing, Jan, Sina, Juli, Lorenzo, Onur, Omar, Jun, Qiuya, Piero, Lisette, Bukyang, Max, Thomas L., Rémi, Valentina, Filippo, Andrea, Nelson and Alfonso for sharing the offices and contributing to the good atmosphere around. And a big thanks also to the professors around for spreading their good mood in the corridors, in particular to Fede for his only funny jokes, to Elisa and to Petra. I am also very grateful to my colleagues in Barcelona, in particular Jaume, Andreu, David and Andrea, for the warm welcome and the nice lunch breaks under the Spanish sun.

I am glad to have found in Vienna a place to call home, mainly because I have always found amazing flatmates wherever I have lived. Of course I am thinking about Marion, with who I have spent hours talking late and laughing loud, Mirjam, who was always happy to share and hear personal stories, Kevin, who warmed the flat with his playful laugh and his reflections about the world, Christina, who has a huge heart and great feminist views, Leonie, who comforted me with injury tips, Andrea, who quickly became a climbing partner always ready for adventure, Luca, who brought his sweet personality and his dog adoption project, Gladji, who had a great sense of fun and a great sense of friendship, Cleia who opened me her house in Barcelona, Martina and Sara.

Also, network science is great but mountain sport is even better, and I survived the last four years thanks to the evenings and the weekends I have spent in great company moving around. I particularly thank Jessie B. whose enthusiasm for climbing, cinema and life in general is contagious, Jessie S. for being not only an inspirational person, but also a very warmhearted holiday-lover climber, Fabri to push climbing always harder in Slovenia, Austria and Italy, Gregor for the great advice, Christian who was a great bouldering-partner and beer-drinking-partner, Adria to make me discover the coziness of Klosterneuburg despite the cold of the winter, David for his constant enthusiasm and his big smile, Antoine and Blaz with who I went ski-touring more than once no matter if it was sunny or windy, Thomas B. for always coming back to Vienna motivated to move around. I am also very grateful to the community of climbers of Boulderbar, for the good vibes, the cheering, and the warm smiles, in particular thanks to Andrew, Ciro, Celia and Davide.

A big thank also to the pyjama party girls, Salambo, Ines, Anja, Lena and Núria. Ines once said, "I feel someone is hugging me since I am with you" and it perfectly captures the warmth of these moments together with this group of amazing girls.

In addition, the period I spent in Barcelona was very special to me thanks to amazing people. I had great moments with Denis who is a great inspiring, devoted climber and a caring person, Virginia with who I drank more than one vermouth, Alice who welcomed me so warmly, Pato with who I could train my Chilean expressions, and Statis for the sunny vibes.

Austria was my country of adoption but I felt so great there thanks to the people supporting me from abroad. I have no words to say how much Lea has been important to me, 10 years of intense friendship and endless support, she is a ray of sunshine. There were also Lucas B. who was always there somehow, sharing intimate stories. François, who supported me at the beginning of this journey, Thomas R. who taught me that running is fun, Marouchka and Sabrina who shared their life stories from Germany to the USA including Norway, and my childhood friends Anto, Solène, Laurine, Justine and Manon who always gave me the will to come back to Gap.

In general, I thank with all my heart every person who was present during my accident and the recovery, it was definitely a hard period and you mean the world to me.

And of course I will definitely not have written this thesis without the unconditional support of my family, who has always shown me solid love to grow and flourish. I am extremely glad to be surrounded by such affectionate and caring people. In particular, many thanks to my grandparents, Simone, Pierre, Jacqueline and René, who taught me so much about life. Many thanks to my brothers Romain and Mathis, even if they have asked me thousands of time to be written in these acknowledgements, I still like them a lot. And many thanks to my parents Mireille and Bernard, to be a constant refuge of love and support.

Contents

1	Introduction	1
1.1	Human dynamics as complex systems	1
1.2	The static network approach	2
1.2.1	Representations	2
1.2.2	Properties	3
1.2.3	Models of network structures	4
1.3	The temporal network approach	8
1.3.1	Representations	8
1.3.2	Properties	9
1.3.3	Representations and models of temporal networks	9
1.4	Dynamical processes	12
1.4.1	Simple contagion process (Sm)	13
1.4.2	Complex contagion process (Cx)	14
1.4.3	Spontaneous adoption process (St)	15
1.4.4	Comparison and distinguishability of the spreading processes	15
1.5	Outline of the thesis	18
2	Detecting periodic time scales of changes in temporal networks	20
2.1	Introduction	20
2.2	Methods	22
2.3	Validation on synthetic data sets	27

2.3.1	Results	29
2.3.2	Parameter dependencies and limitations	30
2.4	Applications on real networks	31
2.4.1	Data sets	32
2.4.2	Results	33
2.4.3	Shuffling of the data	35
2.5	Conclusion	37
3	Distinguishing mechanisms of social contagion from local network view	40
3.1	Introduction	40
3.2	Results	41
3.2.1	Methods of classification	41
3.2.2	Different mechanisms of social contagion	44
3.2.3	Process classification with known parameters	45
3.2.4	Process classification with unknown parameters	54
3.2.5	Case study: adoption mechanisms on Twitter	55
3.3	Discussion	61
4	Competition between simple and complex contagion on temporal networks	65
4.1	Introduction	65
4.2	Model definition	66
4.3	Analytical study	67
4.3.1	Pure complex contagion	68
4.3.2	Mixed simple and complex contagion	69
4.4	Numerical simulations	71
4.5	Conclusion	76
5	Conclusion	79
A	Detecting periodic time scales of changes in temporal networks	81

A.1	Sensitivity analysis: size and length of the temporal network, sliding window parameters	81
B	Distinguishing mechanisms of social contagion from local network view	86
B.1	Distribution of the features of the random forest of Experiment 2	86
B.2	Best subset of features for the random forest on Experiment 2	88
B.3	Accuracies of the classification of spontaneous adoption on Experiment 3 .	91
C	Competition between simple and complex contagion on temporal networks	93
C.1	General case $z > 2$	93
C.2	Simulations on the extended parameter-space (β, p)	94

List of Figures

1.1	Example of the representations of a temporal network realised with the Supra-Adjacency and the Event-Graph methods.	10
1.2	Illustration of the three contagion mechanisms: the simple contagion, the complex contagion and the spontaneous adoption.	13
1.3	Speed dependency of the simple and complex contagions in function of the parameters β and ϕ	16
2.1	Methodology pipeline to measure the time scales of changes of a temporal network G_T	22
2.2	Sketch of the method to compute the tensor portraits BD_*^m of the temporal network G_T^m	24
2.3	Schematic representation of three settings simulated with the Activity-Driven temporal network model with periodic changes of parameters.	28
2.4	Periods corresponding to the two first harmonics measured through the SA-method and the EG-method for periodic synthetic temporal networks generated through the <i>Change of activity and grouping</i> setting.	30
2.5	Number of events as a function of time for the four data sets: the US school, the US flight, the Conference and the Resistance game.	32
2.6	Power spectra of dissimilarity and activity functions of four real-world data sets: a US middle-school, a US flight network, a conference, and a resistance game network.	34
2.7	Measures of the time-scales of changes of the US middle school temporal network with the SA-method and the EG-method.	35
2.8	Power spectra for the data sets US school, US flight, Conference and Resistance game networks shuffled using the two shuffling methods $P_p(\Gamma)$ and P_t	36

3.1	Overview of experimental setups.	44
3.2	Classification accuracy values of the likelihood method when it is obtained theoretically and by simulation and of the random forest method.	46
3.3	Frequency of observation of the features used to train the random forest classifier among the top-3 most important ones across the full parameter space for Experiment 2 and 4.	54
3.4	Accuracies obtained by classifying the infection instances from Experiments 2 and 3 on different networks and with different methods.	56
3.5	Parameter distributions and dependencies of Experiment 4 inferred from the #GiletsJaunes Twitter dataset.	59
3.6	Classification of contagion mechanisms of the #GiletsJaunes Twitter dataset as the function of $\hat{\phi}$ and $\hat{\beta}$ parameters.	62
4.1	Network and contagion dynamics.	67
4.2	Fraction of infected nodes ρ , proportion of nodes infected by simple ρ_s and complex ρ_c contagion as a function of time, for $z = 2$	72
4.3	Comparison of the contagion curves and the final ratio Λ when every node has the same activity and when the activities are sampled from a power law.	73
4.4	The $1/\rho^{z-1}$ curve obtained from simulation averages with $(\beta, p) = (0.037, 0.037)$, fitted by a linear function on the period before t_{init} for $z = 2$ and $z = 3$	75
4.5	Areas corresponding to the three categories in the parameter-space (β, p) for $z = 2$ and for $z = 3$	77
A.1	Power spectra of the temporal network for different number of nodes.	82
A.2	Power spectra of the temporal network for different periods.	83
A.3	Power spectra of the temporal network for different number of periods.	83
A.4	Power spectra of the temporal network for different (γ_1, γ_2)	84
A.5	Power spectra of the temporal network for different parameters of the sliding window.	85
B.1	Distribution of the features of the random forest algorithms across the parameter space.	87

B.2	Subsets of features giving the best accuracies in the parameter space (β, ϕ) for a certain length of subset in the classification with the random forest of Experiment 2.	89
B.3	Subsets of features giving the best accuracies in the parameter space (β, ϕ) for a certain length of subset in the classification with the random forest of Experiment 4.	90
B.4	Parametrization and accuracy of the classification of Experiment 4 constructed based on the #GiletsJaunes Twitter dataset.	92
C.1	Fraction of infected nodes, $\rho(t)$, as the function of time for simulated spreading scenarios, together with the proportion of nodes infected by the simple, ρ_s , and the complex, ρ_c , contagions for $z = 2$	95
C.2	Inverse of the proportion of infected neighbours to the power $z-1$, fitted with a linear function on the first part of the propagation for the whole parameter-space.	97

List of Tables

3.1	Average over the whole parameter-space of the accuracies on the classification of the contagion cases from Experiment 2.	52
3.2	Number of instances of contagion mechanisms inferred by the likelihood and random forest methods on the #GiletsJaunes Twitter dataset.	57
B.1	Accuracy of the classification of the spontaneous adoptions on Experiment 4 with the random forest.	91

Chapter 1

Introduction

1.1 Human dynamics as complex systems

We, as humans, are much less deterministic than atoms and molecules: we are driven by biological and emotional mechanisms which remain partially unknown and make the analysis of our behaviours challenging. Our study is even more complex, thinking of us as social beings, sharing the space and living together in our societies. Indeed, in addition to being complex individuals on our own, we constantly interact with our peers, creating connections with our social circle: family, friends, colleagues or acquaintances. Through those exchanges, various processes emerge and characterise our behaviour like cultural dynamics [1], language evolution [2], crowd behaviour [3] and the formation of hierarchies [4], contributing to the complexity and the ever-changing nature of our societies over time. Namely, by influencing our peers through our conduct and interactions, we thereby impact their decisions to follow behavioural patterns similar to ours. Such patterns, mediated by social influence may be interpreted as a spreading process leading to macroscopic phenomena of mass adoption of products, ideas, beliefs, or information cascades [5–8].

To explore such phenomena, physicists often use tools from their domain even if social propagation seems far from their original field [9]. This approach is justified by the observation that societies, like systems of statistical physics, can be understood as objects made of interacting entities governed by their own rules. This parallel between social systems and physical systems was first explored in the 1960s [10], marking the beginning of a cross-disciplinary approach which continues to evolve today [5, 11]. Those studies have been extended with the recent access to large data bases, allowing to explore further the dynamics of social phenomenon, principally due to the advent of the digital data revolution. In the context of behavioural propagation through social influence, the online social networks and the messaging platforms provide great resources to understand communication patterns

and behaviour adoption on a vast scale. Those new databases significantly surpass the old traditional surveys conducted manually, which reach only a limited number of individuals. Researchers use those new data sources to better design and parametrise the models, giving more accurate forecasts, which helps to control the spreading dynamics of social norms, memes or behavioural patterns.

Those spreading processes are arguably explained by simple decision mechanisms on well-mixed populations [10–13]. However, this type of representation does not effectively encode the underlying structure along which social influence travels, as in this description every individual can have a potential contact with any other. To better capture the complexity of real-world interactions, social networks provide a more realistic framework, as they restrict connections to specific patterns of interactions. In that regard, the importance of social networks has been recognised [14–16] since their structure could critically influence the global outcome of social spreading phenomena unfolding on top of them [6, 17]. Social networks inherently evolve over time, but they can be represented as either static or dynamic structures, depending on data availability or the level of detail required for analysis.

1.2 The static network approach

1.2.1 Representations

Among the different existing network structures, I first introduce the static networks approach, which assumes that connections between individuals remain the same over time. Static networks [18] are represented as graphs constituted of a set of entities V , called nodes, which interact through connections called links or edges, defined by the set E , thus $G = (V, E)$. The number of vertices is typically indicated by N and the number of edges by L . A way to represent networks is to use their adjacency matrices A , for which the element $A[i,j]$ is 1 if there is an edge between nodes i and j and 0 otherwise. In terms of structure, networks can be either undirected, for which the edges have no specific direction, or directed, when the interactions are oriented from one node to another. Moreover, they can be analysed at different scales, ranging from the microscopic level, focusing on individual nodes and their connections, to the macroscopic level, which examines the overall structure and patterns of the network. At the microscopic scale, an important node quantity is the degree k , noted as k_i for node i , measuring the number of nodes with whom it has a connection, known as neighbours. The set of neighbours of node i is indicated as Γ_i , and its ego-network is composed of the set Γ_i and the set of edges connecting the neighbours to node i . Another crucial metric at the node scale is the clustering coefficient, which measures the probability that two neighbours of a node are also connected to each other. The value of the clustering coefficient at the node level is calculated as:

$$C_i = \frac{2e_i}{k_i(k_i - 1)} \quad (1.1)$$

where e_i is the number of edges between the neighbours of i . In this way, the local clustering coefficient is expressed as the ratio between the number of actual edges between the nodes from Γ_i and the number of possible edges between them $k_i(k_i - 1)/2$. This quantity can also be defined at the global level providing an overall indication of the clustering in the entire network as:

$$C = \frac{3 \times \text{number of triangles}}{\text{number of connected triplets}} \quad (1.2)$$

where a triangle is a set of three nodes all connected to each other, and a triplet is a set of three nodes connected by at least two edges. Additionally, at the mesoscopic scale, we can identify communities in the network, which are ensembles of nodes that are more densely connected to each other than to the rest of the network. There also exists at that level connected components, which are groups of nodes in a network where every node can be reached from any other node in the group through a series of edges, without connections to nodes outside the group. Finally, at the global level, the shortest path length between two nodes i and j , indicated as $d(i, j)$, represents the minimum number of edges required to go from node i to node j . Building on this, the diameter of the network is the longest shortest path length, representing the maximum distance between any pair of nodes in the network.

$$D = \max_{i, j \in V} d(i, j) \quad (1.3)$$

In a similar way, the average shortest path length $\langle d \rangle$, calculated over all pairs of nodes, is a widely used measure to quantify how easily it is to navigate in the network.

1.2.2 Properties

Even if every real network is different, we find common properties that are present in almost all of them [18]. First, the degree distribution p_k follows in general a broad distribution, implying that while the majority of the nodes have a low degree, a small fraction of them possess a very high number of connections. For many of these networks, this distribution is a power-law distribution, with an exponent typically between 2 and 3. This contributes to making the network more resilient against random failure, as the high degree nodes are more probable to not be affected by a random removal of the nodes, and thus reduces the risk of disconnecting the network. Another property of real-world networks is a high clustering coefficient. Indeed, this phenomenon is largely observed in social networks in which, for example, the concept of “my friends are also friends between each other” is often true, but can also be present in other types of networks. Finally, most networks tend to have a short

average path length, which enables easy reachability across the network. The combination of high clustering and short path lengths is known as small-worldness. This concept was first highlighted by Stanley Milgram in the 1960s with his Six Degrees of Separation experiment [19]. In his study, participants were asked to send a letter to a target person through acquaintances, only using people they knew by their first name. Milgram found that, on average, it took about six intermediaries, or degrees, for the letter to reach the target, suggesting that any pair of individuals in the world are connected by a surprisingly short chain of social connections. This observation has been repeated recently using emails [20] and online social networks [21] verifying the earlier observed phenomena.

1.2.3 Models of network structures

As real-world networks exhibit specific characteristics, network scientists have developed models to generate random networks that replicate their key properties, namely a power-law degree distribution, a high clustering coefficient and a short average path length. These models are largely used to study network science phenomena within a controlled environment where the properties and the parameters are known. This section presents the main approaches used to construct synthetic static networks, that I am going to use in the upcoming sections describing my research results.

Erdős–Rényi Model

Description

The Erdős–Rényi (ER) network model, known for its simplicity, stands as the fundamental reference model in network theory. In the late 1950s, Erdős and Rényi introduced two versions of this model [22], both beginning with a network of N vertices. In the first version, the L edges of the network are selected at random among the $L_{max} = N(N-1)/2$ possible edges. However, this approach introduces a subtle interdependence among edges as each link can be chosen only once, thus this method is less commonly used. The second version, which is more popular and the one employed in this work, avoids this effect as each of the L_{max} edges is independently present in the network with probability p . Thus the average number of links $\langle L \rangle$ for networks with parameters (N, p) , denoted $G_{N,p}$, is given by $\langle L \rangle = pL_{max}$. Furthermore the correspondence between the two versions, parameterised respectively by (N, L) and (N, p) can be expressed by the probability of constructing a network with exactly L links:

$$P_L(G_{N,p}) = \binom{L_{max}}{L} p^L (1-p)^{L_{max}-L} \quad (1.4)$$

Degree distribution

The degree distribution in Erdős-Rényi model networks follows a binomial distribution given by

$$p_k = \binom{N-1}{k} p^k (1-p)^{(N-1)-k} \quad (1.5)$$

This formula comes from the selection of the k neighbours of a node from the $N-1$ available other nodes, where each connection occurs with probability p and is discarded with probability $1-p$. Thus the average degree is $\langle k \rangle = p(N-1)$, and the standard deviation is $\sigma^2 = p(1-p)(N-1)$. It is worth noticing that their ratio $\sigma / \langle k \rangle \approx 1/(N-1)^{1/2}$ indicates that as the network size increases, the degree distribution becomes increasingly narrower around the average degree. For large N and a fixed $\langle k \rangle$, the degree distribution is well-approximated by a Poisson distribution $p_k = e^{-\langle k \rangle} \langle k \rangle^k / k!$, which is characterised by its narrow shape. In any case, the distribution does not follow a power law, so misses to capture this characteristic of real networks.

Clustering coefficient

The clustering coefficient C of an Erdős-Rényi model network is determined by the probability p that two neighbours of a node are also connected to each other. Thus, this quantity equal to $C = p = \langle k \rangle / (N-1)$ is quite small if N is large. This low clustering coefficient contrasts with real-world networks, which typically exhibit much higher clustering coefficients. Therefore, the Erdős-Rényi model fails to replicate the high clustering observed in real-world systems.

Average distance

We use here the diameter of a network as a proxy of its average shortest path length. In random graphs, the convention is that the diameter of a graph is defined as the maximum diameter of its connected components [23]. For most values of p , nearly all graphs with the same N and p have an identical diameter, indicating that all graphs with N nodes and a connection probability p have a narrow range in which their diameters D vary, usually concentrated around $D = \ln(N) / \ln(pN)$. Since this diameter value is quite small, the Erdős-Rényi model successfully captures a key characteristic of real networks: a small diameter.

Component sizes

With this model, a giant component can emerge, and Erdős and Rényi found that the transition of apparition of this giant component is determined by the parameters of the model (N, p) . Indeed, when p is sufficiently small relative to N , such that $p = c/N$ with $c < 1$ as N increases, the network consists of small connected components of size $O(\ln(N))$ ¹. However, when $c > 1$, a single giant component of size $O(N)$ emerges coexisting with several smaller isolated components, including a second-largest connected component of size $O(\ln(N))$.

¹The notation O or 'big o' describes the upper bound of the growth rate of a function $f(n)$ in terms of another function $g(n)$. $f(n) = O(g(n))$ means that $f(n)$ grows no faster than a constant multiple of $g(n)$ as $n \rightarrow \infty$

At the critical point where $c = 1$, a second-order transition [24] occurs leading to a giant component of size $O(N^{2/3})$.

Watts-Strogatz Model

While the Erdős–Rényi model is known for its simplicity, it has the disadvantage of having a low clustering coefficient [25]. To solve this issue, Watts and Strogatz introduced a new model (WS) specifically designed to demonstrate the possible presence of high values of clustering coefficient and short distances in the same network.

Description of the model

Inspired by the "six degrees of separation" concept, stating that each person is in the world connected to anyone else with a chain of six acquaintances on average, the objective of Watts and Strogatz was to develop a network model that exhibits small-world properties, i.e. a high clustering coefficient and a short average path length. Their idea [26] was to use a network which already presents a high clustering coefficient, in their case a regular ring lattice which naturally contains a lot of triangles. The model initially consists of N vertices, each connected by k_0 edges. In order to reduce the average path length of the network, the authors introduce a rewiring process where each edge is randomly reconnected with a probability p , meaning that one end does not change while the other end is rewired to a random vertex. The final network is an object between a regular ring lattice where $p = 0$ and a completely random network where $p = 1$.

Degree distribution The degree distribution is typically measured from a version of the Watts-Strogatz model slightly different than the one described above, in which links are not rewired but added to the lattice with a probability p . This model is then similar to an Erdős–Rényi network for which a link exists with a probability p . As the original lattice has an initial number of k_0 edges, the degree distribution is given by a shifted Poisson distribution

$$p_k = e^{-\langle k-k_0 \rangle} \frac{\langle k-k_0 \rangle^{k-k_0}}{(k-k_0)!} \quad (1.6)$$

Thus this model unfortunately does not present a broad degree distribution as real networks, since the Poisson distribution is sharply peaked.

Clustering coefficient

The authors of [27] introduce a new definition of the clustering coefficient at a global level. This quantity C can then be measured as the fraction between the mean number of links between the neighbours of a node and the mean number of possible links between these neighbours. By applying this definition, the authors obtain for the Watts-Strogatz model:

$$C(p) = \frac{3(k-1)}{2(2k-1)}(1-p)^3 \quad (1.7)$$

as $N \rightarrow \infty$. This value remains large for a large interval of p values, ensuring a large clustering coefficient.

Average path length

The average path length $\langle d \rangle$ decreases quickly when p increases, as it introduces shortcuts in the lattice network between the nodes. Indeed, each random rewiring is likely to connect two distant parts of the network, and thus significantly reducing the average path length. Only a low proportion of rewiring is enough to observe an important decrease of the average path length, typically when $p \approx O(1/N)$ [28], while the network retains ordered at the local level. Consequently, this model successfully achieves its objective: for a range of values of p , the network exhibits both the two small-world properties which are a high clustering coefficient and a small path length.

Barabási-Albert Model

The Watts-Strogatz model successfully captures the small-world phenomenon observed in real-world networks, but fails to explain the broad degree distribution characterising these networks. To overcome this challenge, Barabási and Albert introduced a model [29] that generates networks with such properties.

Definition

Earlier synthetic models typically assume that networks have a fixed number of nodes, but in reality they evolve by growing. Building on this concept, Barabási and Albert develop a model, known as the Barabási-Albert (BA) model, belonging to the family of random graphs which evolve through growing processes. Beginning with a small initial number of nodes m_0 , nodes are then added one by one to the network, each connecting to m nodes. To ensure to have a power law degree distribution, the m neighbours are chosen with preferential attachment, meaning that this method biases the selection of the m nodes, promoting the ones which have more connections. The probability that a new node gets attached to the already existing node i is given by:

$$\Pi(k_i) = k_i / \sum_{j \in \Gamma_i} k_j \quad (1.8)$$

The total number of nodes at any given time step t is expressed by $N = m_0 + t$, and the total number of links by $L = mt$. In some studies [30], the initial m_0 nodes form a clique, therefore the number of links is given by $L = mt + m_0(m_0 - 1)/2$.

Properties

Since the probability of selecting a node is proportional to its degree, the degree distribution of the network follows a power law with an exponent of 3 regardless of the value of m , with an average degree of $\langle k \rangle = 2L/N$. Additionally, the average distance $\langle d \rangle$ is given by $\ln(N)/\ln(\ln(N))$, which is consistent with the short average path lengths observed in real networks [31]. However, the clustering coefficient, expressed as $C = \ln(N)^2 m / (8N)$, goes to zero when $N \rightarrow \infty$, unlike the expected high clustering typically observed in real-world networks.

Stochastic Block Model

Another characteristic of real-world networks is the frequent presence of communities, meaning that certain groups of nodes are more densely connected between themselves than with the rest of the network. This observation motivated network scientists in the early 1980s to develop a model to build random modular graphs, which has led to the creation of the Stochastic Block Model (SBM) [32, 33]. In the SBM, nodes are assigned to distinct blocks, and edges are distributed following the Erdős–Rényi model, but the connection probabilities vary depending on whether edges are within or between blocks. This design ensures that nodes within the same block share similar connectivity patterns. The flexibility of the SBM to capture diverse connection schemes among large network blocks makes it a suitable tool in network science, with multiple applications such as graph clustering [33], social network analysis [34], and community detection [35, 36].

1.3 The temporal network approach

Even though static networks are accurate to describe structure characteristics of real-world complex systems, many complex systems are evolving dynamically as their elements and the interactions between them are subject to changes in time. The recent availability of high-temporal-resolution network data sets has stimulated the emergence of the new field of temporal networks [37–39], which has been useful to describe human behaviour [40–42]. In particular, this representation goes beyond the conventional static description of networks [43], as it keeps track of the temporal order of successive interactions between elements and allows for instance to identify notions of potential causality.

1.3.1 Representations

In practical terms, a temporal network $G_T = (V, E_T, T)$ consists of a set of nodes V and a set E_T of events over a time interval T , each event $e(i, j, t) \in E_T$ describes a temporal interac-

tion between two nodes $(i, j) \in V \times V$ at a certain time $t \in T$. A snapshot at a given time t is the static graph containing the nodes and edges present at time t in the temporal network. Therefore, similarly to static networks, single nodes can be characterised by their instantaneous degree (number of neighbours at a given time). At the global level, the temporal paths between nodes represent the series of successive interactions along which information can be transmitted [44, 45]. In addition, the activity of a temporal network stands for its number of events as a function of time.

1.3.2 Properties

Temporal networks present different time-dependent properties at different structural scales. First, links often follow a bursty dynamics, meaning that events occur in rapid, concentrated bursts followed by long periods of inactivity [46, 47]. This characteristic occurs in many real-world temporal networks such as communication networks, social interactions, or even transportation systems. The bursty nature of human behaviour has important implications in the dynamics of spreading processes, like information diffusion or epidemic outbreaks, as burstiness can accelerate or slow down these processes depending on how the events are distributed over time. In addition to this, connections between nodes tend to follow correlated dynamics, such as circadian patterns, which are all typical of human dynamics [48–50]. While different measures such as the degree or the communities change through time, they may evolve at different time scales [51], including overall changes between global states at the macroscopic level [52–54]. In particular, periodic variations can emerge, e.g. driven by the circadian fluctuations of human behaviour [50, 55], regular scheduling in different contexts like in transportation or schools, or the repetition of metabolic reactions in biological systems [56].

1.3.3 Representations and models of temporal networks

Higher order representations of temporal networks

Temporal networks are complex objects, but their study and the investigation of the dynamical processes happening on top of them can be simplified by embedding them into simpler structures. For this purpose, we are using in Chapter 2 two methods, the *Supra-Adjacency* and the *Event-Graph* representations. Both methods embed the information from a temporal network into a static network, enabling them to be applicable to any temporal network. They also provide a lossless mapping that preserves all the essential temporal dynamics while transforming them into a static framework.

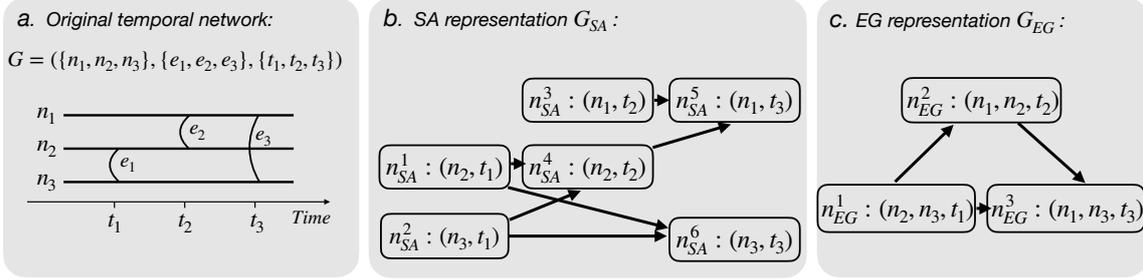


Figure 1.1: Example of the representations of a temporal network (panel a) realised with the Supra-Adjacency (SA) (panel b) and the Event-Graph (EG) (panel c) methods. In the original temporal network G_T , the timelines of the nodes are represented by horizontal lines depicting the activity of each node through time (the direction of the time goes from left to right). Each event occurring between two nodes is symbolized by a line linking those two nodes at the corresponding time. The SA and EG representations are static directed networks. In the SA case, the nodes are pairs (node, timestep) of the temporal network. In the EG, the nodes are events, i.e., triplets (node, node, timestep) of the temporal network.

The *Supra-Adjacency* (SA) representation [57,58] was initially used to investigate dynamical processes on temporal networks, particularly for predicting the evolution and outcomes of such processes with limited information on the dynamics itself. This method transforms a temporal network into a static directed network, efficiently compressing the information. It also encapsulates the paths along which information can propagate in the original temporal network.

In this representation of a temporal network G_T , denoted $G_{SA} = (V_{SA}, E_{SA})$, each node $v_{SA} \in V_{SA}$ represents a pair (node, time) of the original temporal network: the node $(i, t) \in V_{SA}$ denotes that the node $i \in V$ was active at time $t \in T$, i.e., had at least one interaction at t . A directed edge $e_{SA} \in E_{SA}$ between two nodes of V_{SA} , (i, t_a) and (j, t_b) (with $t_a < t_b$), encodes the fact that information can propagate on G_T from node i at t_a to node j at t_b , without intermediary events. If $i = j$, this is possible if t_a and t_b are successive interaction times for i (there is no event involving i at times $t_a < t < t_b$). Edges of type $(i, t_a) \rightarrow (i, t_b)$ in E_{SA} thus simply correspond to following the successive interaction times of i in G_T . For instance in Figure 1.1, node n_2 is involved in the events e_1 and e_2 , happening respectively at times t_1 and t_2 . As it is the same node in two consecutive events, there is an edge in the SA representation between (n_2, t_1) and (n_2, t_2) . For $i \neq j$ instead, the event $(i, j, t_a) \in E_T$ results in two directed edges in E_{SA} : $(i, t_a) \rightarrow (j, t_b)$ and $(j, t_a) \rightarrow (i, t_c)$, where t_b (resp. t_c) is the first time after t_a in which j (resp. i) is active again. In Figure 1.1, there is an edge in the SA representation between (n_2, t_1) and (n_3, t_3) and another edge between (n_3, t_1) and (n_2, t_2) as the nodes n_2 and n_3 are linked at time t_1 in the original temporal networks. The direction of edges in G_{SA} respects the arrow of time, and the set of edges E_{SA} allows to preserve the

information about all possible temporal paths of the original temporal network.

The *Event-Graph (EG)* representation [59, 60] is a static weighted directed acyclic network representation of a temporal network. It encodes all the possible time-respecting paths that information or an infection can follow within the network. The original motivation for using the event-graph is to study percolation in temporal networks with a computationally efficient tool, but it can be applied to other contexts.

The construction of an event-graph network G_{EG} from a temporal network G_T goes as the following: each event in G_T is represented by a node in G_{EG} , and two nodes of G_{EG} are connected if the two corresponding events in G_T were adjacent [59], i.e., share at least one node (in V) and are consecutive. For instance, there is an edge in Figure 1.1 between the nodes of the EG representation corresponding to the events e_1 (n_{EG}^1) and e_2 (n_{EG}^2) as they share the node n_2 . Each edge between two nodes in G_{EG} is directed along the direction of time (from the earlier event to the later one) and is weighted by the time difference between the two corresponding events. Consequently, G_{EG} encodes also all information of time respecting paths emerging in the original temporal network.

The Activity-Driven temporal network Model

The Activity-Driven temporal network model [61] (ADN) is one of the most widely-used frameworks to capture the dynamics of real-world temporal networks. This model is defined by a set of N nodes, each having an intrinsic activity a_i taken from a given distribution $P(a)$ ². At each time step, node i becomes active with probability ηa_i , where η is a rescaling factor and, if active, establishes connections with m other nodes chosen randomly. Connections are erased after each time step thus the model does not present any memory nor correlations between time steps. As a major strength this model explains structural features like the presence of hubs, which arise from the heterogeneous activity levels of the nodes, while remaining simple and flexible. Additional mechanisms driving the link dynamics of nodes can explain further emergent network properties like a high clustering coefficient [63].

Random reference models of temporal networks

On a different note, empirical temporal network data entail structural and temporal correlations of different nature. To explore their characteristics, a common method consists of shuffling the data to create randomized reference models [64] in which specific correlations are destroyed while others are preserved. In other words, through shuffling we create a sample from a uniformly sampled microcanonical ensemble of randomized networks, where

²Here we define the activity as an activation probability. Other works [62] consider a_i as an activation rate, or probability per unit time

certain network properties are kept constrained, while the networks are maximally random otherwise. Specifically, in this work we consider the following random reference models, following the canonical notations introduced in [64]:

- $P_p(\Gamma)$ shuffling: To remove any correlations, the order of the temporal network snapshots is randomly shuffled, keeping fixed the structure of each snapshot. This procedure destroys any structural correlations between consecutive snapshots, removing the effects of structural reorganizations and randomizing also the sequence of the timing of the events between the nodes.
- P_τ shuffling: The method rewires randomly all links in each snapshot of the temporal network. This process keeps the same number of nodes and edges as in the original snapshot. The number of events by time step is thus preserved while the group structure and its changes are removed.

1.4 Dynamical processes

The study of spreading processes on top of networks is a particularly prominent area of network science that gets significant attention from the community. These explorations provide an advanced understanding of various types of propagation, including diseases but also social behaviours, on which I focus in this manuscript. Initially, the models were built assuming that the population was fully mixed, meaning that each individual can interact with any other person in the network [65]. This hypothesis, known as the homogeneous mixing approximation, assumes that the network, for which the nodes represent the individuals, is fully connected. However, most real-world networks are sparse [66], necessitating to use another underlying structure namely complex networks [67].

Models of social contagion on networks commonly describe the propagation as a binary state process [68], attributing to each node a state representing its status regarding the spreading. In this framework, the individuals of the network can be either *susceptible* (S) meaning the contagion has not reached them yet or *infected* (I) indicating they have already been contaminated and may potentially spread the infection. These two states, S and I, are mutually exclusive. The condition to transit from the susceptible to the infected state depends on the chosen model, for which we distinguish two main families: the *simple* and the *complex* contagions. Additionally, infections not caused by the network are modelled as *spontaneous adoption*, accounting for external factors that may influence the infection of the node. Those different contagion types, explained in Figure 1.2, are discussed in the following paragraphs.

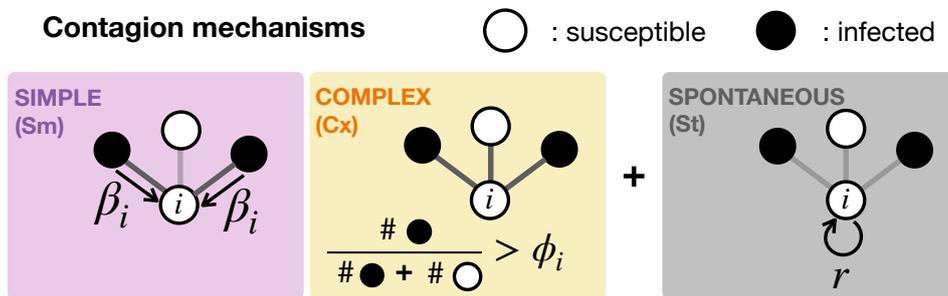


Figure 1.2: Illustration of the three contagion mechanisms: the *simple contagion* parameterised by the adoption probability β_i through a single stimulus; the *complex contagion* parameterised by the threshold ϕ_i of the necessary fraction of adopter neighbours to induce the adoption event; and the *spontaneous adoption* that occurs with probability r .

1.4.1 Simple contagion process (Sm)

One family of mechanisms, commonly termed simple contagion in the social science literature, has first been introduced to model biological epidemic processes [6, 67]. These models are originally used to represent the transmission of a disease from a susceptible to an infected individual after contact. However, their effectiveness has been proved on the spread of other fields such as the propagation of social behaviour like the adoption of applications [69, 70], technologies [71, 72] or the spread of rumors [10, 11]. This contagion type assumes transition probabilities from susceptible to infected state, where a single contact or exposure is taken independently and sufficient to trigger transmission. The term *simple* comes from the fact that an infection may occur from only one exposure. The simplest model that applies this idea, and the one used in this work, is the Susceptible-Infected (SI) model [6].

Susceptible-Infected (SI) model

The concept is the following: each interaction between a susceptible node and an infectious one may independently result in the infection of the susceptible node, following a set probability. At the start, every node is susceptible, except for a small set of infected individuals, called seeds. Any infected individual can spread the infection to one of its susceptible neighbours i with a probability β_i per iteration, as explained in Figure 1.2. If a susceptible node is connected to multiple infected neighbours, it has multiple independent chances of getting infected. The outcome of this process is always the infection of every node after forecasting a logistic curve with an early exponential increase [73].

Other models

In the SI model, the only possible state transition is from susceptible to infected, with no possibility of going back to the susceptible state once infected. Nevertheless, other well-known models allow different transitions, such as the Susceptible-Infected-Susceptible (SIS)

model [67], where infected nodes can revert to being susceptible. Additionally, there are models with additional states, like the Susceptible-Infected-Recovered (SIR) model, where nodes enter a recovered state after infection and cannot be reinfected.

1.4.2 Complex contagion process (Cx)

From simple contagion to complex contagion

Even though the simple contagion effectively models social phenomenon in some cases, the reasons for infection are sometimes more complex. Empirical evidence suggests that peer influence and reinforcement mechanisms play a crucial role in social contagion processes [74,75]. Those indicate the need for an alternative model, called complex contagion [76,77]. In this case, exposures are not independent, but peer pressure can impact in a non-linear way the individual infection probability [13, 17, 78] meaning that contagion requires multiple exposures from different sources to happen. The complex contagion has inspired different models and the main ones are described in the following paragraphs.

Threshold models

Some of the most used models to depict the complex contagion are the threshold models, where individuals change their behaviour once a predefined quantity exceeds a limit. Those thresholds are associated with social contagion phenomena, where the influence of peers can change an individual behaviour after reaching a limit. Interestingly, thresholds are also significant in certain diseases like tuberculosis and dysentery [79] where the infection occurs only when the pathogen concentration in the body surpasses a critical level.

The early fundamental work in this area has been done by Schelling and Granovetter on well-mixed populations. First Schelling [13] defines a model of residential segregation illustrating how an individual decides to leave its neighbourhood based on a threshold on the number of agents of its own colour living close by on a grid. Inspired by the work of Schelling, Granovetter [12] discusses the number of people engaging in a particular behaviour depending on the distribution of individual thresholds, highlighting the critical role of social networks and peer influence in shaping collective outcomes. He claims that this model could explain a range of social behaviours, including the diffusion of innovations, the propagation of rumours or the voting behaviour.

Watts models

The model proposed by Granovetter propagates on top of a fully mixed population, meaning that every individual can potentially interact with any person in the system. The Watts model [17] adopts Granovetter's idea on structured populations represented as networks, thereby removing the fully mixed population assumption. In this model, the determinant quantity is not the number of exposures but the number or proportion of infected peers surrounding a susceptible individual. When this quantity exceeds a certain threshold, the

node gets infected due to the accumulated social influence. We use this model in the rest of the manuscript given its common application in several studies about spreadings of social behaviour [17, 70, 80–82]. We consider the fraction of the infected neighbours as the considered quantity for the infection, and we denoted by ϕ_i the threshold of the node i (see Figure 1.2).

Cascade effect

Depending on the model parameters, the complex contagion mechanism may lead to a cascading phenomenon, where mass infection emerges over a short period of time. This was first shown on Erdős-Rényi networks by Watts [17], while several follow-up studies explored a rich family of similar phenomena in multi-layer [83–85], weighted [86, 87] or temporal networks [47, 88], demonstrating their relevance in real-world settings [70, 74, 89–91].

1.4.3 Spontaneous adoption process (St)

Simple and complex contagions capture network-based adoptions, however, social influence may not always spread on an observable network (e.g., advertisements, news or policy recommendations, etc.). Some studies take such external influences into account by also considering a third mechanism, called spontaneous adoption [70, 92, 93]. This type of infection is agnostic to the underlying network structure and is modelled as an alternative way for a node to become infected in addition to the simple or the complex contagions. We denote by r the rate of infection of the spontaneous adoption (Figure 1.2), which can infect equally at every time step all susceptible nodes.

1.4.4 Comparison and distinguishability of the spreading processes

Most studies focus on understanding the properties and the distinguishability between the simple and the complex contagions, as it helps to better understand and model any contagion process. This comprehension is crucial for example to develop effective strategies in marketing, misinformation control, and public health interventions. Since those works in general do not consider the spontaneous adoption due to its independence from the network, we explore in the next paragraphs the main differences between the simple and the complex contagions.

Speed of the propagations

The speed of the contagion processes differs according to the parameters employed in the simulations, which could help to differentiate between the two processes. In Figure 1.3, I implement separately a simple contagion process in which every node has the same parameter β and a complex contagion process in which all nodes are characterised by the threshold ϕ on Erdős-Rényi networks. High values of β characterise simple contagion processes with

high speed since nodes in this scenario have a higher probability of being infected, commonly after a single stimulus. The opposite effect characterises complex contagion: if ϕ is high, the propagation is slow-down as the proportion of infected neighbours needed for adoption is large.

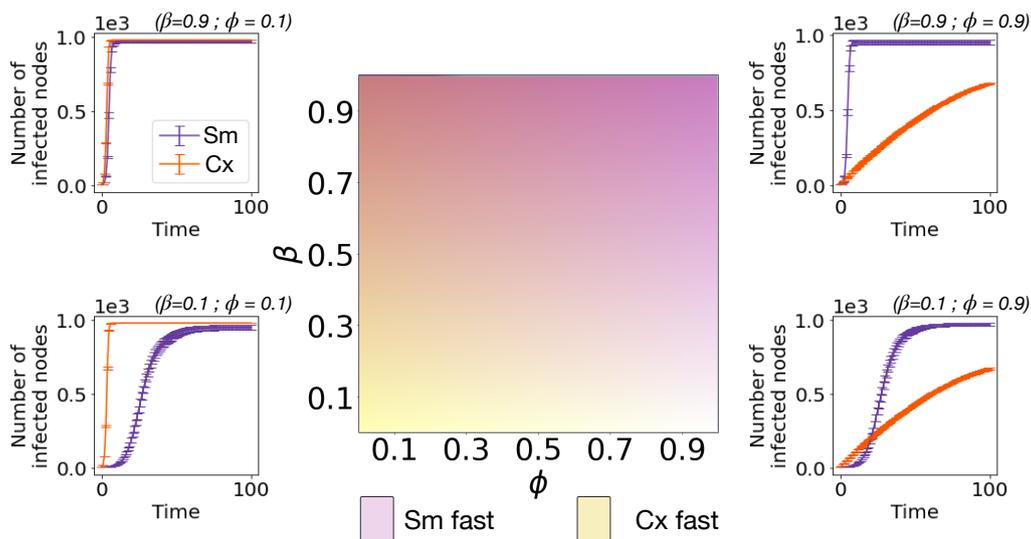


Figure 1.3: (central panel) Speed dependency of the simple and complex contagions in function of the parameters β and ϕ (respectively). The purple and orange colours, which are schematic, display respectively the areas where the simple and the complex contagions are faster. The number of infected nodes through time is shown on the four extremes of the parameter space, in purple for the simple contagion and in orange for the complex contagion. The dynamical processes are implemented separately on an Erdős-Rényi network of 1000 nodes, with an average degree of 4.

Moreover, the logistical curves, which show the number of infected nodes over time, differ between the simple and the complex contagions. In the simple contagion, the curves typically exhibit a fast exponential growth at the beginning as each exposure can independently trigger the spread, followed by a steadier increase. The complex contagion, however, starts more slowly as it requires multiple exposures from multiple sources. Then, once enough reinforcement is achieved, the number of infected nodes suddenly increases, leading to a sharp rise caused by the cascade effect.

Properties of the contagion types

The simple and the complex contagions depend non-trivially on several networks and dynamical characters of an ego and its peers [94]. It has been shown that while simple contagion spreads easier on dense and degree-heterogeneous structures, with high-degree nodes early infected [95], these properties mitigate complex contagion as the threshold of high-degree nodes can be hardly reached [6, 96]. Moreover, while weak ties connecting densely connected communities act as facilitating bridges for simple contagion [12], they slow down complex contagion cascades [76], as they likely deliver non-reinforced social influence to susceptible individuals. In addition, the timing and the order of infection stimuli, their concurrency, and the bursty dynamics of interactions [88, 97, 98] between individuals and their peers influence the adoption dynamics and the macroscopic dynamical outcome of the spreading process as a whole.

Moreover, both contagion types exhibit a property known as the waiting time [99], which refers to the time between when a node becomes infected and when the adoption is actually observed. Since individuals have limited time and energy, there can be a delay between the moment someone wants to adopt a behaviour and the moment it is fully adopted. In the simple and the complex contagion models, this waiting time is always 1 as the ego node adopts immediately in the next time step once its condition of infection is reached. However, in real-world scenarios, this delay can vary, which will limit some of our methods in Chapter 3.

Distinguishability of the mechanisms

Distinguishing between the simple and the complex contagions is an emerging question that several studies have already tackled, classifying a whole spreading process between the different types of contagion.

The first significant study in this area [90] was conducted on the social platform Twitter (now X), using 'social bots' to disseminate information. They describe the simple and the complex dynamics by proposing two Bayesian statistical models which consider the timing of peer stimuli. This method demonstrates that the complex contagion model more accurately represents the observed behaviour of the bots.

Some years later, Hébert-Dufresne et al. [100] compares two interacting simple contagion processes -where being infected by one process increases the probability of being infected by another- with a complex contagion process. They found that these processes are not distinguishable based on macroscopic quantities such as the final fraction of infected nodes or the number of infected neighbours at the moment of the infection.

Another study [96] further distinguishes the simple contagion, the complex contagion and two other types of contagions involving higher-order contagion mechanisms where group dynamics enhances the infection rate. By analyzing how high-degree nodes get infected first with the simple contagion and last with the complex contagion, they measure the cor-

relation between the ranking of the nodes in the contagion with its degree and its clustering coefficient. By training a machine learning algorithm, they successfully classify contagions into four types, achieving accuracies above 0.84.

In Chapter 3, we will also address this issue by classifying infection instances at the node level. Moreover, in chapter 4, we will consider mixed propagation dynamics and determine if these processes are governed by the simple or the complex contagions.

1.5 Outline of the thesis

Revealing mechanisms behind spreading processes is important to understand social behaviours propagation through societies. These dynamical processes are more accurately represented on temporal networks, as they reflect the ever-changing nature of social interactions. However, the evolving characteristic of temporal networks significantly influences propagation processes, hence the need to analyse those objects. In particular, many societal phenomena are cyclic, such as the daily patterns, thus temporal networks present periodic changes that are usually not obvious to measure. In chapter 2, we contribute to this endeavour by defining a new method to measure the periodic time scales of changes in temporal networks. Given a temporal network as input, we first divide it into temporal sub-networks. We then use lossless mappings (see Section 1.3.3) of these temporal sub-networks to get a sequence of static networks and quantify the dissimilarity between them successively to obtain a dissimilarity function describing the changes between the successive temporal sub-networks. We extract the timescales of this function by computing its power spectrum to identify its main frequencies and harmonics. We focus on applying this method to the detection of periodic changes in the link density and group structure of temporal networks.

With a better understanding of the properties of temporal networks, we investigate closer spreading processes on top of networks. The challenge of distinguishing between the simple and the complex contagions is gaining importance in the scientific community, with several studies addressing this issue (see Section 1.4.4). However, these studies commonly make two assumptions limiting their applicability in real-world scenarios. First, they expect full knowledge about both the underlying network structure and the spreading dynamics. Indeed, this is a strong assumption in common real-world scenarios, where information about infection events is typically incomplete or limited to local knowledge, possibly obtained only about an adopting ego and its peers. Second, these studies assume that all individuals follow the same single adoption mechanism; either simple or complex contagion. In contrast, it has been argued that the mechanism driving one's decision to adopt a behaviour during an unfolding social contagion may depend on the intrinsic susceptibility of an individual to the actual behavioural form and the properties of the propagation process itself [101, 102]. Therefore each single adoption event may be driven by different mechanisms that jointly

depend on personal factors [103, 104] (heterogeneous susceptibility and predisposition), the properties of the item being adopted (Gladwell's stickiness [101]), and the particular context (environment, time of adoption, other external factors).

In Chapter 3, we study the distinguishability between simple, complex and spontaneous contagion mechanisms by addressing the challenge that a single social contagion process may involve multiple adoption mechanisms [105]. We frame this question as a classification problem at the ego-network level and explore solutions based on likelihood and random forest approaches. These methods are developed and tested on extensive synthetic simulations, encompassing different spreading scenarios and underlying social structures, ranging from fully controlled experiments to empirical spreading cases observed on Twitter (currently called X).

While this last study is realised on static networks, we consider similar distinguishability questions in a more realistic setting using temporal networks. In Chapter 4, we address this challenge by proposing a new methodological approach to distinguish which contagion type dominates the spreading dynamics by modelling mixed synthetic propagation scenarios on temporal networks. By analysing the simulation curves, we identify three categories of spreading processes: those dominated by the simple contagion during the whole dynamics, those entirely governed by the complex contagion, and mixed dynamics, where simple contagion rules the process at the beginning and then transitions to be dominated by the complex contagion. Through an analytical study, we introduce several methods to identify the transitions between those three regimes.

This PhD work offers a novel perspective on the observations of temporal networks and spreading processes without prior knowledge about the contagion mechanism of individuals, thereby enhancing our understanding of the dynamics of social contagion processes.

Chapter 2

Detecting periodic time scales of changes in temporal networks

2.1 Introduction

Temporal networks present an accurate framework to study dynamics phenomena, especially human behaviours [40–42], but also biological and ecological systems [106, 107] or public transportation [108, 109]. This representation provides an effective tool to investigate the structure and dynamics of these systems, as well as the potential dynamical processes occurring on top of them [6, 37]. Specifically, this representation preserves the temporal sequence of interactions between the nodes, enabling for example to determine the causality with the temporal paths [44, 45].

Temporal networks have distinct time-dependent characteristics, from node-properties to whole network properties, which may change over different scales, involving shifts at the macroscopic level [52–54]. These changes can be periodic, particularly due to the cyclic pattern of human behaviour, paced for instance by the succession of days and years [50, 55], or the recurrence of metabolic processes in biological systems [56]. Interesting relevant examples of such variations are given by changes in the connection density in the network, or in the way nodes form and dissolve groups or communities. For example, the number and structure of social interactions vary due to daily rhythms and schedules in contexts such as workplaces, scheduled social gatherings or in schools, where students interact within a class during lectures, but also with other classes during breaks [55, 110]. The identification of the temporal scales of periodic variations in a temporal network is an important step for the characterisation and understanding of the system under investigation. However, their measure represents a challenge as they co-appear with other arbitrary non-periodic temporal scales, which appear as noise and hinder the possibility to detect the periodic behaviour by

simply following the temporal evolution of simple network summary measures.

Some recent works have addressed the detection of relevant temporal scales in temporal networks, e.g. by optimizing the overlap between the sets of events on consecutive time intervals [111] or by searching for the precise recurrence of connections between nodes in different time windows [112]. Another approach consists of defining the correlation between instantaneous adjacency matrices of the temporal graph [113]. Finally, computing a whole similarity matrix between all pairs of timestamps can make it possible to detect states in which the network structure remains stable [52–54], but this method requires rather heavy computations.

In this chapter, we introduce a novel approach for measuring the periodic time scales of changes in temporal networks. Our method begins by dividing the temporal network into smaller temporal sub-networks using successive sliding windows. Each sub-network is then transformed into a sequence of static networks through the Supra-Adjacency or the Event-Graph methods. By calculating the dissimilarity between consecutive static networks, we generate a dissimilarity function that captures the changes over time. We analyze the power spectrum of this function to determine the dominant frequencies and harmonics, which indicate the timescales of periodic changes. We are here interested to use our approach to detect periodic changes in terms of link density and group structure of temporal networks. To this aim, first we consider synthetic networks in which we impose periodic variations of density and structure with tunable frequencies. We show that the method is able to retrieve the actual time scales of the networks. We then apply our methods to several empirical temporal networks presenting periodic dynamics. In each case, the method captures correctly the system's main characteristic times, which could most often not be extracted by simple measures of the network's overall activity. Our work opens the door to a better characterisation of the time scales of changes of temporal networks, essential in the understanding of the dynamics of the underlying complex systems. There are also numerous potential applications of understanding the periodic structural and density changes. In communication networks, it can optimize future infrastructure development and bandwidth allocation by identifying peak usage times. The method discussed in this chapter can also help to detect and analyse echo chambers on social networks. Additionally, it can participate to optimize healthcare networks by predicting patient flow patterns and identifying high-interaction periods. Furthermore, it helps to plan strategic interventions during pandemics by uncovering the structural and density patterns of contact networks. Finally, the approach in Chapter 2 contributes to understand how animals in herds live and organize within ecological networks.

2.2 Methods

Let us consider a temporal network $G_T = (V, E_T, T)$, for which we want to measure the periodic time scales of changes. The whole generic pipeline is summarized in Figure 2.1.

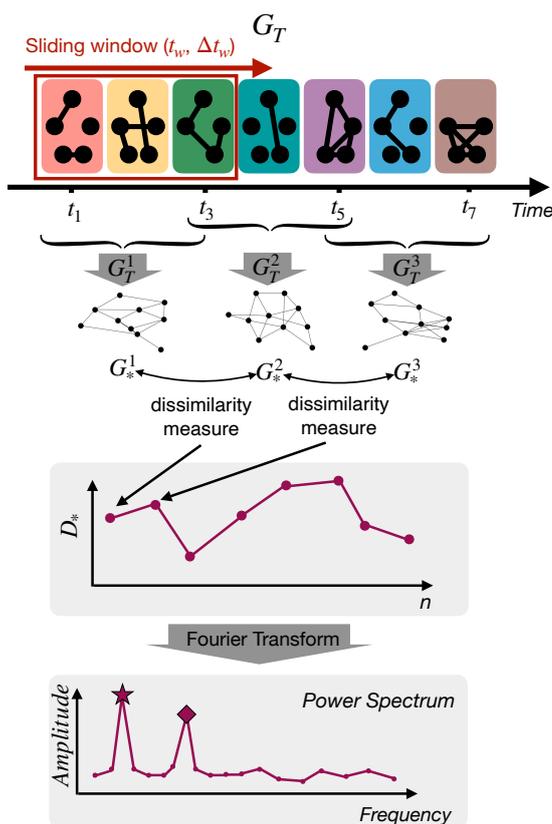


Figure 2.1: Methodology pipeline to measure the time scales of changes of a temporal network G_T . From top to bottom: the initial temporal network is divided into sub-temporal networks through a sliding window of parameters t_w (shift between 2 consecutive windows) and Δt_w (length of the sliding window). The m^{th} sub-network is denoted G_T^m . A static representation of each sub-network (G_*^m) is generated through the method $*$ (either Supra-Adjacency or Event-Graph). Each G_*^m is then described by a 3-dimensional tensor $B_*^m(j, k, \tau)$ that encodes information about the paths and distances in the sub-network. We compare consecutive tensors with a dissimilarity measure, obtaining the dissimilarity function D_* . Finally, we compute the power spectrum of D_* and measure the frequencies of the main harmonics.

Temporal sub-networks. With such a temporal network as an input, we first extract a sequence of temporal sub-networks of G_T by using sliding windows of length Δt_w and stride t_w (shift between the start of successive windows). Successive windows can thus overlap. Specifically, the m^{th} sub-network G_T^m is composed of the nodes of V and of the subset E_{T^m} of events of E_T taking place in the time interval starting at time $m * t_w$ and ending at $m * t_w + \Delta t_w$:

$$G_T^m = (V, E_{T^m}, T^m = [m * t_w : m * t_w + \Delta t_w]) \quad \text{for } m \in \mathbb{N} \quad \text{and} \quad T^m \subseteq T. \quad (2.1)$$

Based on this definition we can obtain a sequence of temporal sub-networks G_T^m to compute a dissimilarity function characterising the dynamical changes in the structure and the overall activities present in the original temporal network G_T . More precisely, we want to compute the dissimilarity between consecutive temporal sub-networks, G_T^m and G_T^{m+1} . However, while there exist several methods to compare static networks, few exist to quantify the dissimilarity between two temporal networks. We thus choose to map the temporal sub-networks into static network representations in order to use an already known and validated comparison method for static networks [114], that we adapt to our purposes.

Static network representations. We thus first map each temporal sub-network onto a static network representation using the two different methods described in Section 1.3.3: the Supra-Adjacency and the Event-Graph representations. We apply them to each temporal sub-network G_T^m defined above to map them into a sequence of static representations G_{SA}^m and G_{EG}^m . In the following, we use the symbol $*$ to refer to the static representation method: it replaces the abbreviation *SA* or *EG*, as every object from now on can be calculated using one method or the other.

Network dissimilarity function. As a next step we compute a dissimilarity function $D_*(m)$ between successive static networks, G_*^m and G_*^{m+1} , for each sequence of static representations $\{G_*^m, m = 1, \dots\}$. To this aim, we adapt a method proposed by Bagrow and Bollt [114], which allows to compare static networks at multiple scales. The first step of this method is to compute, for each static network, its ‘‘portrait’’ B defined as

$$B_{l,k} = \text{number of nodes which have } k \text{ nodes at distance } l .$$

The dissimilarity between two networks is then given by the Kullback-Leibler divergence between their respective portraits. In our case, the static networks that we need to compare are representations of temporal networks, with either the Supra-Adjacency or the Event-Graph method, noted G_*^m . Nodes and edges in these networks contain information about nodes, interactions and times of the original temporal network. To take this into account, we adapt and modify the definition of network portrait, and define the tensor portrait of G_*^m by relying on $BD_*^m(j, k, \tau)$ which is the number of nodes of G_*^m which can reach, in

two hops, j nodes, k events and τ timestamps of the original temporal network. In other words, we consider for each node of the static representation G_*^m its ego-network at distance 2, and count the number of distinct nodes, timestamps and events of the original temporal network G_T^m involved. We then collect this information for all nodes of G_*^m and summarize the resulting histogram as the portrait $B_*^m(j, k, \tau)$. We illustrate this method to compute the tensor portrait $B_*^m(j, k, \tau)$ in Figure 2.2. We also note that the static representations of the

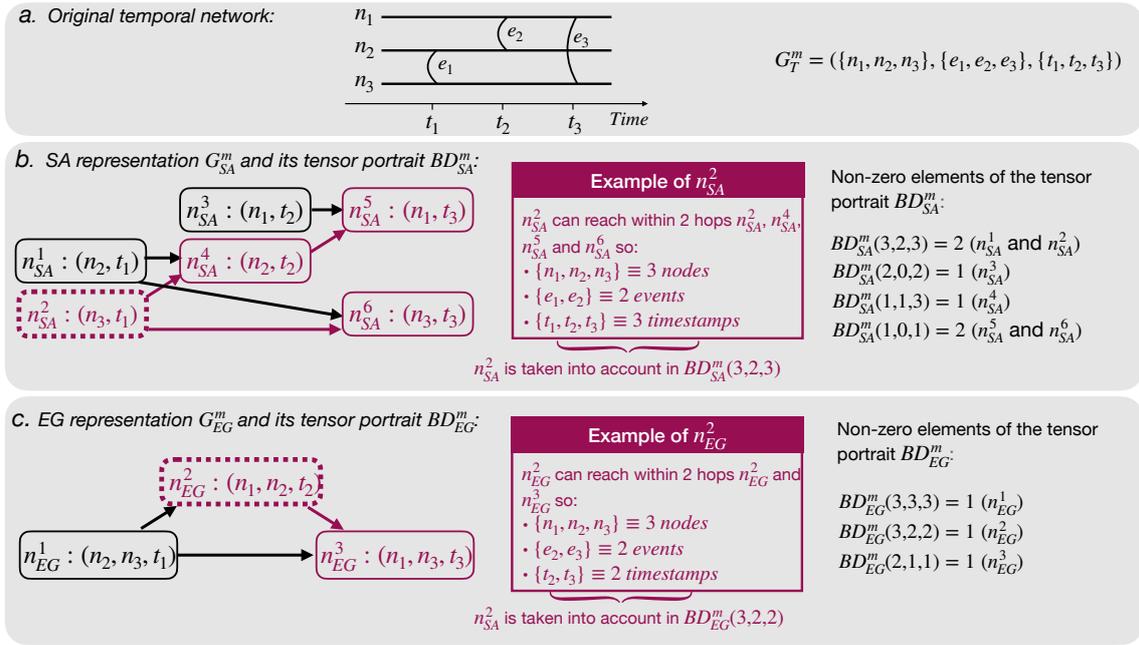


Figure 2.2: Sketch of the method to compute the tensor portraits BD_*^m of the temporal network G_T^m displayed in panel a. The static SA and EG representations G_{SA}^m and G_{EG}^m are shown respectively in panels b and c. We first evaluate the number of nodes, events and timestamps from the temporal network accessible within two hops from each node of the static networks. We illustrate the computation for n_*^2 (purple dashed highlight). We then count the number of nodes of the static representation that can reach j nodes, k events and τ timestamps of the original temporal network to compute the element $BD_*^m(j, k, \tau)$ of the tensor portrait.

temporal networks are directed, with edge directions following the arrow of time. The ego-network of a node of the static representation involves only future timestamps and events. To take also into account how each node can receive information from events in the past, we create for each G_*^m its reversed version $G_{*,R}^m$ by inverting the direction of each edge of the representation and compute its portrait BR_*^m . We then obtain the final tensor B_*^m by summing BD_*^m and BR_*^m .

Finally, we compute the dissimilarity between each pair of consecutive tensors B_*^m and B_*^{m+1} , as their Jensen-Shannon (JS) divergence (if the two tensors B_*^m and B_*^{m+1} differ in size, we adjust the size of the smaller one to the size of the biggest one by filling the missing entries with zeros). The Jensen-Shannon divergence is based on the Kullback-Leider (KL) divergence, also known as the relative entropy, which is a measure of statistical distance that quantifies how two probability distributions differ from each other. This divergence can be interpreted as how ineffective an approximated distribution is to replace the true one. It is defined as the following

$$D_{KL}(P \parallel Q) = \sum_{x \in X} P(x) \ln \left(\frac{P(x)}{Q(x)} \right) \quad (2.2)$$

where X represents the sample space and P and Q are two probability distributions. The KL divergence cannot be negative, and a value of zero means that the two distributions are identical. It is important to note that the KL divergence is not symmetrical between the two distributions. Also, the second distribution Q cannot present a probability of zero over X , otherwise the KL divergence becomes infinite. In that sense, both compared distributions must share the same sample space. To overcome this limitation, the Jensen-Shannon divergence can be used. This consists in using an intermediate distribution M , the average of the two distributions P and Q , and computes:

$$D_{JS}(P, Q) = \frac{1}{2} D_{KL}(P \parallel M) + \frac{1}{2} D_{KL}(Q \parallel M) \quad (2.3)$$

where M is defined as $M = (P + Q)/2$. Note that, in the case of an empty network, the Jensen-Shannon divergence is not defined. We then assign one single event to the corresponding empty temporal sub-network G_T^m . At that step, we have to ensure that the dissimilarity function respects the Nyquist–Shannon sampling theorem [115], i.e. that its sample rate is at least twice bigger than the observation frequency of the temporal network. It is also worth noticing that the method which compares the temporal sub-networks should be asymmetric because it has to be able to differentiate between the transition of the original network from one state to another, from the opposite transition.

Power Spectrum of dissimilarity function. Each dissimilarity function D_* provides an overall signal that reflects the structural and activity changes in the original temporal network. It presents higher values when the network goes through larger and abrupt transformations and takes smaller values when the network is more stable or changing only gradually with time. It can thus provide insights into the time scales of dynamical changes in the original temporal network. In particular, periodic patterns of network changes can be revealed by taking the power spectrum of the dissimilarity function, which should present harmonics at the characteristic frequencies of the temporal network.

Mathematically, according to the Fourier analysis, any signal can be decomposed in a sum of sinusoidal functions, each with its own period, amplitude and phase. These sinusoidal components, called harmonics, have frequencies that constitute either a discrete set if the initial signal is periodic, or a continuous range if it is not cyclical. This decomposition can be studied in the frequency domain where the signal is represented by its frequency components with the Fourier transform. When the initial signals are discrete, we need to use the Discrete Fourier Transform (DFT), given by the following formula

$$X_k = \sum_{n=0}^{N_{sample}-1} x_n e^{-j2\pi kn/N_{sample}}, k \in \llbracket 0, N_{sample} - 1 \rrbracket \quad (2.4)$$

where x_n is the discrete signal sample and N_{sample} is the number of samples. The Fourier transform is invertible, meaning that the original function can be recovered from its frequency spectrum. One of the main algorithms used to compute the DFT is the Fast Fourier Transform due to its computational efficiency. One of the characteristics of the DFT is its magnitude-squared function, which represents the power spectral density of the signal. It illustrates how the power of the signal is distributed over the different frequencies, indicating the amount of power present in each frequency. The magnitude-squared function is given by

$$PS_k = |X_k|^2 \quad (2.5)$$

The power spectral density is invariant to time shifts. One of the goals of the Fourier transform is to measure the correct period of a signal, however this task becomes more challenging when only a sampling from the signal is available, but not the original signal itself. In order to retrieve the correct period, and assure a proper reconstruction of the signal from its discrete sample, we need to apply the Nyquist–Shannon sampling theorem [115], fundamental principle in signal processing. The theorem states that the sampling frequency must be at least twice the highest frequency present in the signal, a rate known as the Nyquist frequency. If the criterion is not respected, the sampled data fails to accurately represent the original signal and the correct period of the signal cannot be measured. Thus, the Nyquist–Shannon criteria is crucial as it ensures that the spectral content of the original signal is accurately preserved when going to a discrete representation.

More precisely, we compute the magnitude-square of the discrete-time Fourier transform of D_* defined as:

$$PS_k = \left| \sum_{j=0}^{N_{sample}} D_*(j) e^{i2\pi kj/N_{sample}} \right|^2 \quad (2.6)$$

where N_{sample} is the length of D_* and $k \in \llbracket 0, N_{sample} - 1 \rrbracket$. The frequency corresponding to the k^{th} harmonic PS_k is $f_k = k/(t_w N_{sample})$ where t_w is the time shift between two successive sub-networks G_T^m . The main harmonics of the PS function (appearing as the largest modes in

the transformed function) correspond to the principal frequencies of the temporal network. Their inverse yield the characteristic time scales of the main periods present in the network dynamics.

In the following, we refer to the full methodology pipelines using respectively the *SA* and *EG* representations as the *SA – method* and *EG – method*.

2.3 Validation on synthetic data sets

To better understand the temporal properties that the above defined dissimilarity functions and their power spectra can capture, we focus on synthetic temporal networks with controlled structural and temporal properties. In particular, we consider networks with tunable changes in activity (number of events per timestamp) and group structure. We utilise the Activity-Driven temporal network (ADN) model [61] for these purposes (Section 1.3.3), considering networks of size $N = 100$ with a power-law node activity distribution with minimum value $\varepsilon = 0.001$ and parameters $m = 4$, $\eta = 4$, $|T| = 9200$ and $\gamma = 1.8$, γ being the exponent of the power-law distribution.

Using these parameters as a baseline, we build three types of periodically varying temporal networks, to model the following settings:

- *Change of activity*: we simulate an ADN in which the density of edges varies periodically in time. We assign to each node i two activity values a_i^1 and a_i^2 , respectively extracted from two power-law distributions with exponents $\gamma_1 = 1.8$ and $\gamma_2 = 2.8$. We then alternate periodically (and synchronously for all nodes) between the two activity values, with a period T_a . This results in periodic changes in the overall activity of the network, as illustrated in Figure 2.3a.
- *Change of grouping*: we consider an ADN model of $N = 100$ nodes forming groups of 5 nodes each, and we periodically alternate, with a period of T_g , between time intervals in which connections are made at random with no restriction as in the baseline and intervals in which only connections within groups are allowed. The average activity is kept constant over time (Figure 2.3b).
- *Change of activity and grouping*: finally, we consider an ADN in which both activity and group structure change periodically over time, by combining the previous two mechanisms, each with its own period, respectively T_a and T_g (see Figure 2.3c).

For each case, we apply the SA and EG methods to compute the power spectra of the resulting temporal networks. As a baseline method, we compute directly the power spectrum

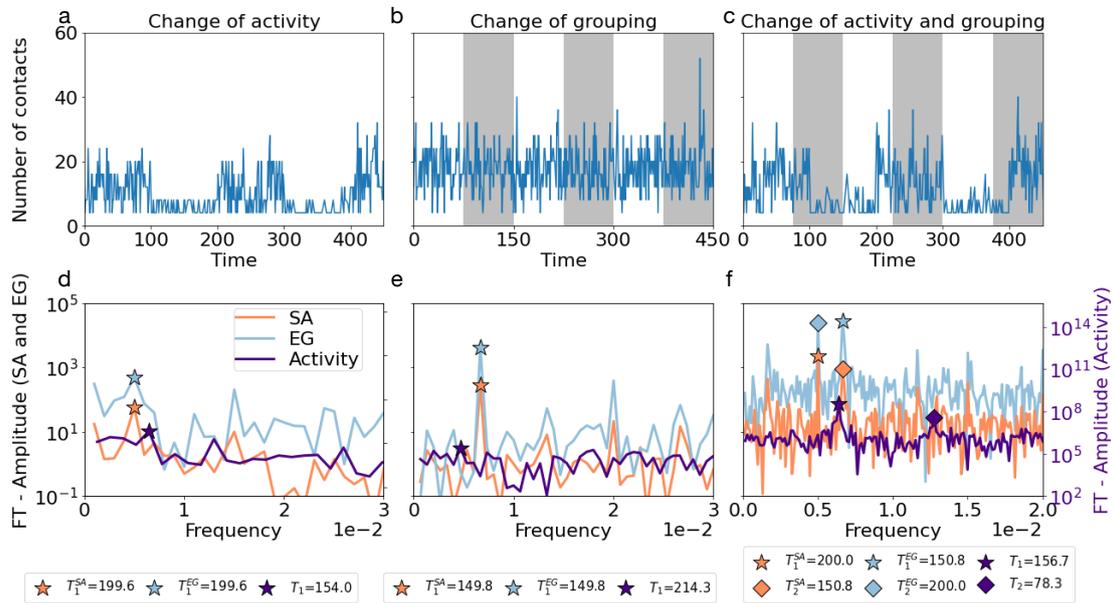


Figure 2.3: Schematic representation of three settings simulated with the Activity-Driven temporal network model with periodic changes of parameters ($N = 100$, $\varepsilon = 0.001$, $\eta = 4$). (a) The *Change of activity* case presents networks with activity periods of $T_a = 200$; (b) the *Change of grouping* case presents recurrent structural changes with period $T_g = 150$; while (c) the *Change of activity and grouping* setting is defined as a mix of both dynamics. Panels (a-c) display the number of events as a function of time for a realization of each experiment; Gray areas in panels b and c indicate the intervals in which interactions can only occur within groups. Panels (d-f) depict the power spectra of these networks obtained respectively through the SA-method and EG-method, as well as the power spectrum of the activity timeline. The first and second harmonics of each power spectrum are shown respectively with a star and a diamond symbol. In each case, the SA-method and EG-method are able to retrieve the correct period of the networks, while the power spectrum of the activity signal fails in measuring temporal structural changes. In the *Change of activity and grouping* case, the SA-method identifies the frequency of activity changes as the main harmonic, while the EG-method detects the structural changes frequency as the dominant one.

of the activity function, that is measured as the link density at each time step of observation (see Figure 2.3). This is a simple summary metric that describes the overall changes in the temporal network and can be computed for any system.

2.3.1 Results

The settings we consider involve either one or two types of periodic changes in the synthetic temporal networks: a periodic fluctuation in the amount of activity and/or in the network structure in terms of inter and intra group interactions. Our first goal is to investigate whether the SA- and EG-methods can uncover the corresponding periods T_a and T_g through the measure of the dominant frequencies in the associated power spectrum. As shown in Figure 2.3d and e, when only one type of periodic change is present, both methods are able to detect the corresponding period. It is evident from the depicted star symbols that indicate the largest mode in the frequency scale, correctly positioned at the right frequency corresponding to the period of the actual periodic changes. At the same time, the baseline method, computed as the *PS* of the activity timeline, strongly underperforms as compared to the other two methods. While in case of activity changes (see panel Figure 2.3d) it at least identifies approximately the value of the period, in case of periodical group changes it does not succeed to capture the rightful period at all. This was expected as in this case the overall activity does not reflect any periodicity but simply fluctuates randomly around a constant value.

When both types of periodic changes are present, an interesting distinction emerges between the results of the SA- and EG-methods. Indeed, both methods correctly detect the T_a and T_g periods as the first two dominant frequencies in the power spectrum. However, in the SA-method the frequency describing the periodic activity changes is identified as the dominant frequency and the periodic group frequency to the second largest value ($T_1^{SA} = T_a = 200$, $T_2^{SA} = T_g = 150$ in Figure 2.3f), while this is reversed for the EG-method ($T_1^{EG} = T_g = 150$, $T_2^{EG} = T_a = 200$ in Figure 2.3f). These results suggest that the SA-method is more sensitive to periodic changes in activity, while the EG-method is more suited to detect periodic structural fluctuations. We also note that the PS of the baseline method yields as dominant timescales $T_1 = 156.7$ and $T_2 = 82.5$, the first one describing approximately the activity periods of the network, while the second one does not correspond to the period of either of the underlying processes.

To check the robustness of the proposed methods against the relative values of the periods, we further investigate this point by exploring systems with different values of T_a and T_g in the *Change of activity and grouping* setting. We generate 100 synthetic temporal networks for each pair of values (T_a, T_g) , compute the dissimilarity function and power spectra of these realizations, and extract the corresponding first two harmonics for each method (SA and EG).

Figure 2.4 summarizes the results by showing in each case the fraction of realizations which detected the periods of T_a, T_g correctly, or failed to detect any of them. These results demonstrate again that the SA-method identifies predominantly T_a (the activity change period) through the first harmonic and T_g (change of group structure) through the second, while the

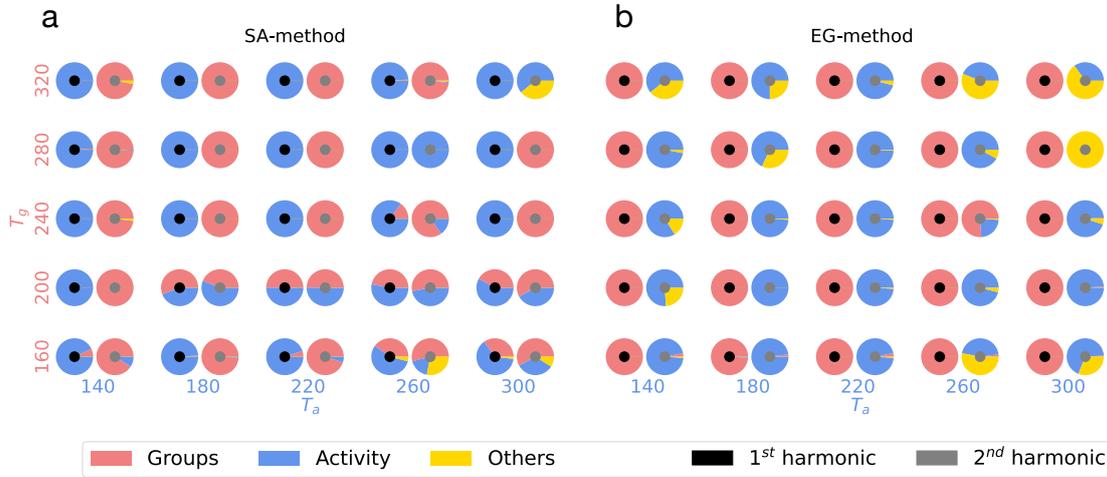


Figure 2.4: Pie charts indicating the proportion of simulations in which the measured periods correspond to either T_a (x-axis) or T_g (y-axis) for periodic synthetic temporal networks generated through the *Change of activity and grouping* setting ($N = 100$, $\varepsilon = 0.001$, $\eta = 3$, $|T| = 9200$). The values correspond to the two first harmonics measured through the SA-method (panel a) and the EG-method (panel b). For each pair of values (T_a, T_g) , we generate 100 realizations of the temporal network and apply the SA- and EG-method to extract the two main harmonics. We show in blue around a small black disk (resp. grey disk) the fraction of realizations in which the main frequency (resp. the second main) corresponds to T_a , in pink the fraction of cases in which it yields T_g , and in yellow the cases in which it corresponds to neither (we consider a tolerance of 10% for both periods). In most cases, both periods are correctly inferred, with the main frequency corresponding to T_a in the SA-method and to T_g in the EG-method.

reverse is observed for the EG-method. Some deviations from this behaviour are observed at large values of the periods and/or when T_a and T_g are close to each other.

2.3.2 Parameter dependencies and limitations

Both the synthetic temporal networks and the analysis method involve some parameters. In particular, we explore their dependencies on the network size, temporal length and ratio between total temporal length and periods of changes T_a or T_g , with results presented in Appendix A.1. We observe that at large network size, both methods identify as main frequency a value corresponding to the half value of the original period. Moreover, evidently, for correct time scale detection the observation period of the temporal network needs to cover at least two full periods of any kind of changes. In addition, regarding the experiment *Change of activity*, both methods are very sensitive to this kind of changes. Even small variations

of the density of links through time, obtained by changing the exponent γ of the ADN, are detected.

The first step of our pipeline moreover involves the definition of sliding windows with stride t_w and length Δt_w . Naturally, these parameters affect the amount of information contained in each sub-temporal network and consequently influence the resulting dissimilarity function [116–118]. We explore the effect of these parameters in Appendix A.1, while keeping $t_w \leq \Delta t_w$ to have a non-zero overlap between successive time windows. We also ensure that the two parameters under study have values below the time span of the network’s period (their maximum value is 20 while the period is 100).

As shown in Appendix A.1 in Figure A.5, the two methods show the best performance if the t_w stride is not too large and if Δt_w length is neither too high nor too small. If the time interval between two temporal sub-networks t_w is too high, we collect less information about the similarity between successive sliding windows. The dissimilarity function is then less precise and our methods perform less well to identify the characteristic temporal scales. Moreover, if Δt_w is too small, each sliding window contains too little information to obtain an accurate measure of the time-scale of the original network. On the opposite, if Δt_w is too large, each temporal sub-network may summarize too much information and lose the specific characteristic of the activity or the structure of the network at a certain time or interval of time. As an observation bias this could smooth dissimilarities between consecutive temporal network slices as they average too much information, and not because the network does not present significant changes through time. It is also worth noticing that both methods measure systematically half of the period as dominant modes for very large values of t_w and Δt_w (Appendix A.1, Figure A.5). In that case, every half-period of the network is covered by a small number of temporal sub-networks, leading to a lack of resolution in the dissimilarity function, in which only the peaks of dissimilarity at half-periods are well marked, leading to the detection of the half-period as typical timescale. Another limitation of these methods is that they cannot measure time scales of networks having periods equal to their time resolutions. It is because it is not feasible to create a sliding window capable of detecting the changes, as it should be smaller than the period, which is the time resolution in this case.

2.4 Applications on real networks

After validating our methods on synthetic networks with controlled properties, to explore further the capabilities of our methods, we consider empirical temporal networks representing different systems. We note that in such systems, in contrast to the cases studied above, several time scales, that correspond both to periodic or non-periodic fluctuations, may co-exist, as well as structural changes of different nature.

2.4.1 Data sets

We consider four temporal networks describing interactions of different nature, with various sizes and over different observation lengths. Their temporal changes of activity are displayed in Figure 2.5.

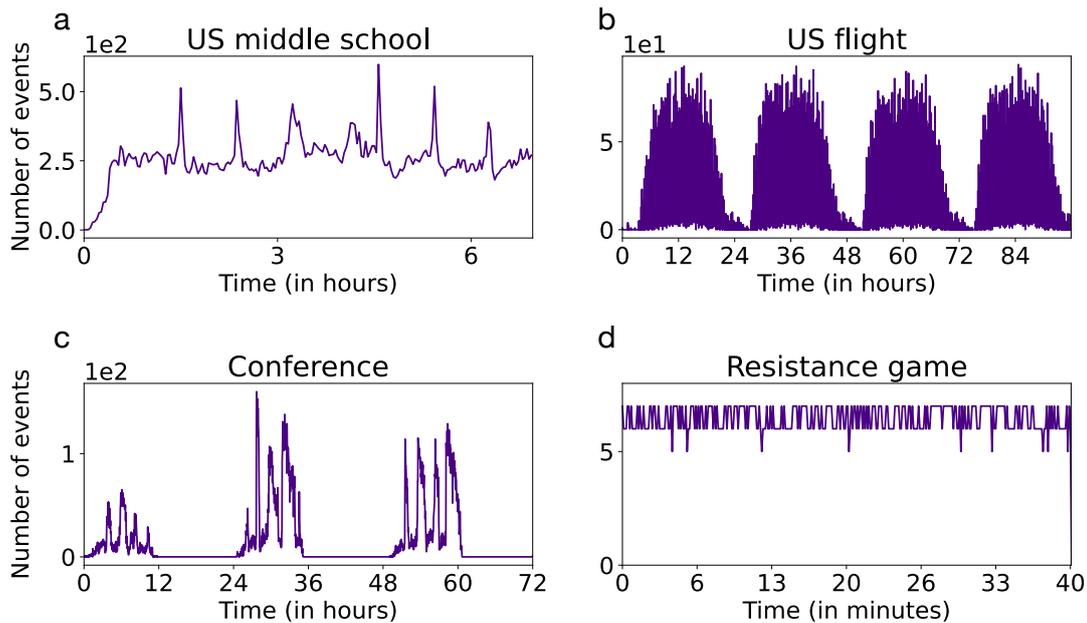


Figure 2.5: Number of events as a function of time for the four data sets: the *US school* (panel a), the *US flight* (panel b), the *Conference* (panel c) and the *Resistance game* (panel d). The *US school* network contains high activity periods during recreational moments of the students' day, while the *US flight* and the *Conference* networks present circadian patterns. The *Resistance game* network does not have particular periodic activity changes.

- *US middle school network*: this data set describes close proximity interactions between students of a middle school in the United States, during one day with a temporal resolution of 20 seconds [119], recorded by Radio Frequency Identification (RFID) wearable devices. It involves several low contact periods of class-times and inter-class breaks including two lunch periods with a high number of contacts as students freely mix while changing classrooms or eating together. The network consists of 591 nodes (each node corresponding to a student) and contains 473,755 records of pairwise temporal interactions between them.
- *US flight network*: this air-transportation network describes the direct flight connections between 278 airports in the US [120]. In our observation period we concentrate

on 4 days of data that records 71,315 flights between the airports that we consider as undirected temporal interactions. This network is expected to show strong periodicities in activity, reflecting the daily recurrent flight schedules, while structural changes may not be strong as almost always the same airports are connected every day.

- *Conference network*: these data also describe face-to-face contacts between individuals, with a temporal resolution of 20 seconds, obtained from the SocioPatterns collaboration [121] by RFID devices built on a different architecture. The contacts were measured during a scientific conference, namely the IC2S2 conference that took place in Cologne (Germany) in 2017 [122]. Our observation period spans over the three first days of the conference, and records 229,536 temporal contacts between 274 participants. This data set is expected to show periodic behaviour both in terms of activity and structural changes, by reflecting the circadian pattern characterising the daily life of the participants.
- *Resistance game network*: it is an eye-contact network between participants of the Resistance game [123, 124], which is a role game where some of the players are hidden 'defeaters', and the goal of the other players is to uncover them. The game involves multiple rounds of around 4 minutes each, starting with a discussion involving every participant, and ending with a vote. The recorded network is built from directed events between participants who looked at each other at a given time t . The network is recorded between 8 individuals and contains 52,731 temporal interactions that we deem undirected for simplicity. This network provides an example where the interaction level should not reflect strong periodicity as every player of the game is looking at someone else at each time step but the grouping of participants changes between each session.

2.4.2 Results

After applying our pipeline on each data set using both the SA- and EG- and the baseline methods, in Figure 2.6 we depict the power spectra of the obtained dissimilarity functions, with stars indicating the dominant frequencies. Interestingly, both the SA- and EG-methods identify the relevant timescales of changes in most networks, while the baseline method failed to detect them besides for the *Conference network*. For the *US middle school network*, both methods yield a timescale of about 46 minutes, coherent with the length of a class. Meanwhile, the baseline activity timeline PS would estimate the dominant frequency as corresponding to a period of 139 minutes. In the case of the *US flight network*, where the main changes are expected to be ruled by circadian fluctuations, both SA- and EG-methods also correctly identified periods of around 24 hours. This time-scale is also captured by the baseline method, but recognised only as its third largest harmonic. The two first harmonics

are identified as periods of 5 and 10 minutes, which may correspond to the characteristic times between consecutive departures of planes from the same airport. The *Conference network* also presents strong signs of circadian changes of activity. This is reflected by all computed power spectra, which show a harmonic corresponding to a period of about 24 hours for both the SA-method and the EG-method, captured as well by the baseline method. Finally, regarding the *Resistance game*, which presents only structural changes, the EG-method measures accurately the time-scales of periods characterising a single round in the game, around 4 minutes. Since no periodic change of activities characterise this network, both the SA-method (more sensitive to activity changes) and the baseline method fail to identify any meaningful time-scale. The power spectrum of the SA-method suggests the dominant mode to correspond to 0.53 minute, while the baseline method detects 13 minutes.

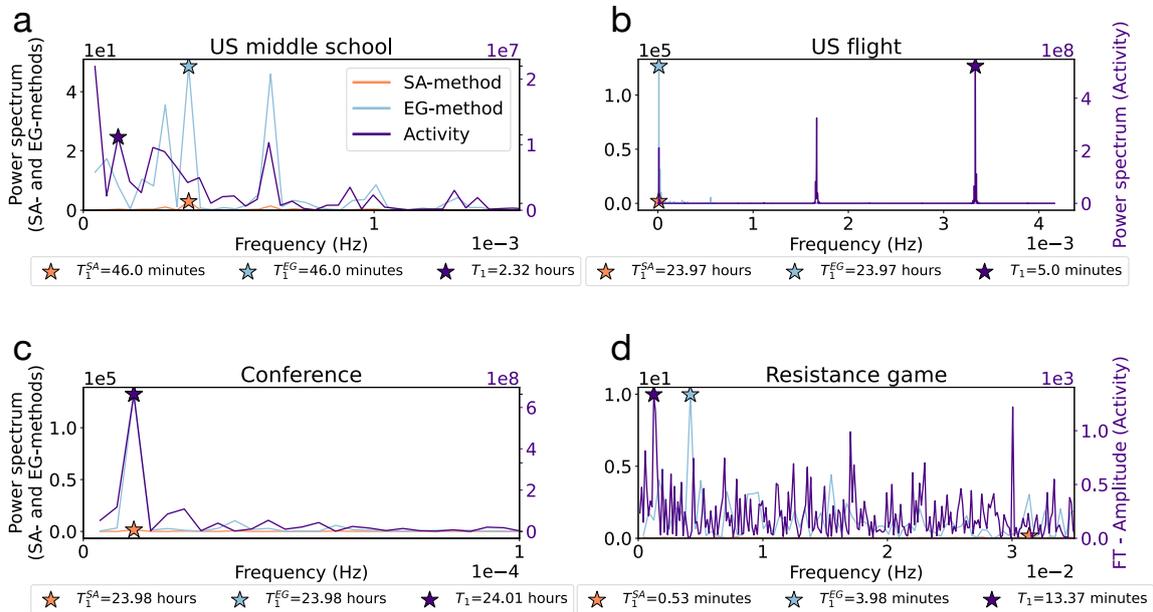


Figure 2.6: Power spectra of dissimilarity and activity functions of four real-world data sets (a) a US middle-school, (b) a US flight network, (c) a conference, and (d) a resistance game network. Dissimilarity functions were calculated by the SA-method (in orange) and the EG-method (in blue), while results computed for the baseline model using activity signals are shown in purple. The highest harmonics are highlighted with a star symbol for each PS, and the corresponding values of the period are indicated below each panel. The parameters of the sliding windows ($t_w, \Delta t_w$) are (2 minutes, 5 minutes) for the *US middle school*, (2 minutes, 10 minutes) for the *US flight*, (2 minutes, 5 minutes) for the *Conference* and (1/3 minute, 1 minute) for the *Resistance game*.

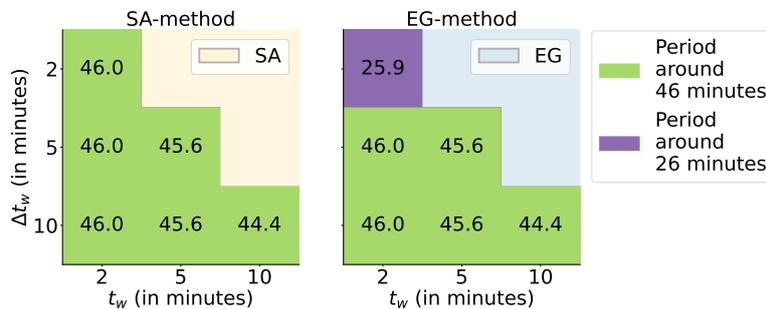


Figure 2.7: Measures of the time-scales of changes of the *US middle school* temporal network with the SA-method (left panel) and the EG-method (right panel) for different parameters $(t_w, \Delta t_w)$. The measure of 46 minutes is robust over the parameter pairs $(t_w, \Delta t_w)$.

It is worth noticing that we have used different sliding window parameters for the different data sets: $(t_w, \Delta t_w)$ are (2 minutes, 5 minutes) for the *US middle school*, (2 minutes, 10 minutes) for the *US flight*, (2 minutes, 5 minutes) for the *Conference* and (1/3 minute, 1 minute) for the *Resistance game*. It is due to the fact that the temporal characteristics of the networks under study vary a lot: from the order of magnitude of the minute for the *Resistance game* to the daily pattern of the *Conference*. The parameters t_w and Δt_w should be chosen in an appropriate manner with respect to the data set and its characteristic time scales, if an a priori knowledge of their order of magnitude is available (cf paragraph 2.3.2). Indeed, Δt_w must be chosen large enough to avoid noise in the measure and small enough to avoid a flattening of the dissimilarity function. Also, the parameter t_w must respect the Nyquist-Shannon theorem. If it is not possible to know a priori which range of parameters to use according to the data set, one should investigate a spectrum of potential parameters $(t_w, \Delta t_w)$ and check that the extracted timescale is robust on a range of parameter values. This has been proceeded for example on the data set *US middle school* (cf Figure 2.7) and the same period of 46 minutes is obtained for most of the parameter pairs $(t_w, \Delta t_w)$. The only different value appears when the parameter $\Delta t_w = 2$ minutes with the *EG-method*. This is because when the sub-networks are too short, the noise perturbs the measure.

2.4.3 Shuffling of the data

To identify which factors influence the measure of the time-scales, we shuffle the data, creating randomized reference models that preserve some correlations while disrupting others. In that purpose, we use the $P_p(\Gamma)$ and the P_τ described in Section 1.3.3. The first approach removes the structure and the activity of the temporal networks, while the second keeps the number of events by time step while erasing the structural organisation. In turn, we apply the SA- and EG-methods to compute the power spectra of the shuffled data and check how

these methods capture the modification of the time scales due to shuffling.

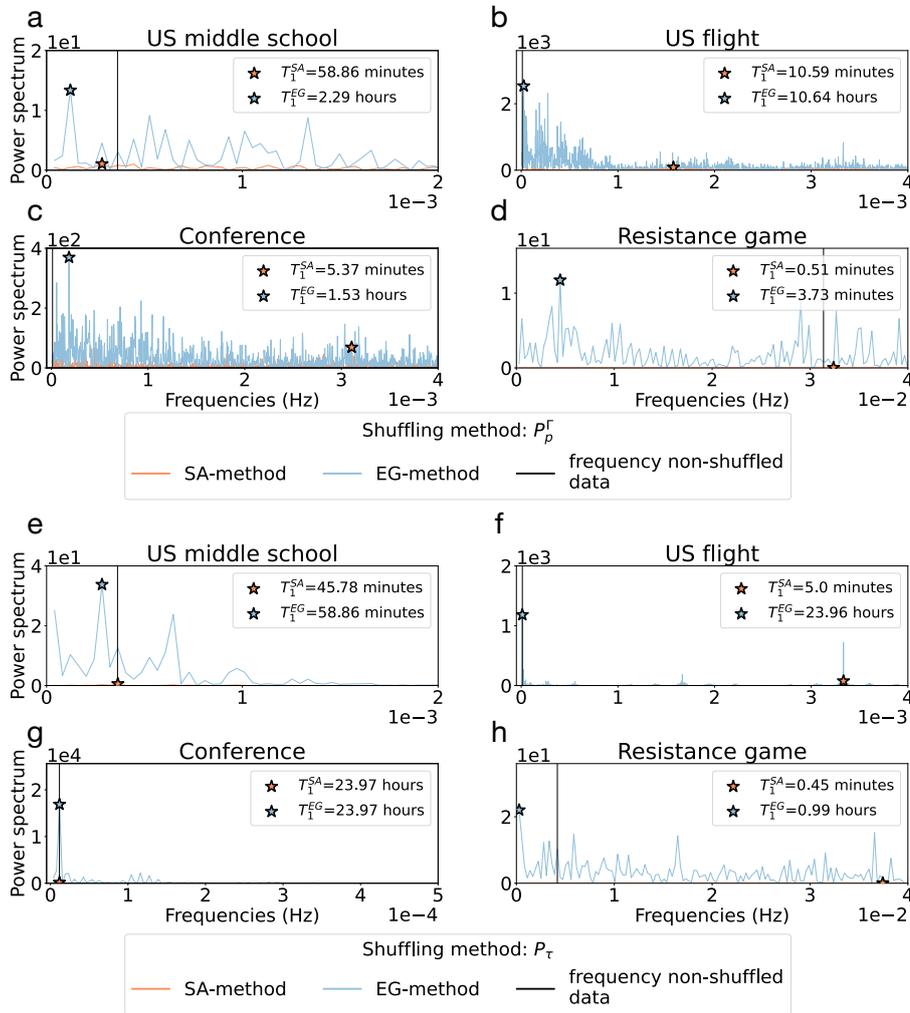


Figure 2.8: Power spectra for the data sets *US school*, *US flight*, *Conference* and *Resistance game* networks shuffled using the two shuffling methods $P_p(\Gamma)$ (panels a-d) and P_t (panels e-h), obtained with the SA-method (orange curve) and the EG-method (blue curve). The period of each original data set is indicated with a black vertical line. For data shuffled using the $P_p(\Gamma)$ method, the original period is never recovered. In the case of the P_t shuffling instead, the SA- and EG methods still measure original periods if the network presents large activity changes (*US flight* and *Conference* data sets). In the case of the *US middle school* network, only the SA-method is able to assess the original time scale as this method performs better to detect activity changes. Finally, none of the methods can measure the original period of the *Resistance game* network shuffled with the P_t method as it does not present any periodic variations.

Results are shown in Figure 2.8 for the four data sets and the two shuffling procedures. When the networks are shuffled with the $P_p(\Gamma)$ procedure (panels a-d), the original periods of changes are not recovered, which is expected since the shuffling destroys any periodicity in the data.

However, when we shuffle the networks using the P_τ method, which removes the structural effects but keeps the fluctuations in the overall activity, our methods present some capacity to identify the residual time scales of changes in some of the data sets. In particular, two of the data sets present large periodic activity variations, i.e. the *US flight* and the *Conference* networks. After shuffling, these regular changes are still present, as the P_τ method preserves their activity timelines, while any other pattern has been destroyed by the shuffling. Consequently, we may still measure their original circadian period from their P_τ -shuffled versions. Indeed, both the SA- and the EG-methods applied to the *Conference* network recover the dominant time scales, while in the case of the *US flight* data set, the EG-method captures the expected time scale of around one day. We also find a time scale of 5 minutes with the SA-method applied to the *US flight* data set, which corresponds to another characteristic time of activity of this network (see Figure 2.6).

In contrast, both the SA- and the EG-methods miss the identification of the original time scales when applied on the P_τ -shuffled *Resistance game*. Since the original network has no periodic fluctuations in terms of activity, neither its shuffled counterpart present any regular changes in terms of activity. Thus the detected time scales are only induced by some noise in the data.

Finally, the *US middle school* network presents activity variations that are not easily assessed even in the original network. Once shuffling with the P_τ method, only the SA-method, which is overall more sensitive to activity changes, retrieves the original period (≈ 46 minutes) in the shuffled network. The EG-method overestimates this time by detecting a period of $5 \approx 9$ minutes.

2.5 Conclusion

In this chapter, we have put forward a new methodology to uncover periodic time-scales of changes in temporal networks. In our proposed pipeline, first we locally aggregate the original temporal network by using a sliding window to build a sequence of temporal sub-networks. Subsequently, we map these temporal sub-networks into a sequence of static networks, using known lossless higher-order temporal network representations, namely supra-adjacency matrices or event-graphs. We further extend a method for the comparison between the consecutive static network samples to define a dissimilarity function that reflects activity and structural changes in the original temporal network. Finally, we take the power spectrum of the dissimilarity function to detect the relevant periodic time-scales from the dominant

frequencies characterising the original network. This work has several applications in network analysis, optimisation and understanding of the diverse dynamic phenomena arising on top of these networks.

We have explored this pipeline, focusing on changes in the activity and group structure of temporal networks. Using synthetic data sets with prescribed changes, we have shown that while both methods are able to recover the time scales of the modelled periodic dynamics, they perform differently in the identification of changes in activity and structure. Specifically, the SA-method is more sensitive to overall activity changes while the EG-method captures better periodic structural fluctuations, which cannot instead be obtained through the power spectrum of the activity timeline. We have also shown that these methods are able to highlight relevant periods in more complex empirical data sets.

The methodology presented here has certain limitations. First, its performance depends on some parameters of the aggregation method and the temporal network observed. The observation needs in particular to span a long enough interval: at least two periods of changes need to be observed. The sliding window parameters also have some impact on the performance: each temporal sub-network should encode enough information but should not be too long to average out relevant changes. The stride should be small enough to keep a reasonable temporal resolution and a substantial overlap between successive windows. Also, the methods presented in this article cannot measure temporal changes if they are smaller or equal to the time of resolution of the network. Overall, if no prior external knowledge is available on the expected order of magnitude of the network timescales, the method should be tested on a range of parameter values, and the robustness of the extracted timescale should be checked.

The proposed methodology pipeline opens the door to the investigation of several interesting extensions and research questions. Possible extensions of the present method could include the consideration of other static representations as well as other similarity measures between successive temporal sub-networks¹, which could potentially be more sensitive to various types of structural changes of the temporal network. For instance, it would be interesting to explore whether changes in the instantaneous core-periphery structure [54] could be uncovered. Future work could also explore extensions to time-varying hypergraphs [125, 126] or the interaction between the detected time scales of the underlying temporal network and ongoing dynamical processes. Our work presents a proof of concept for a new methodological direction that will contribute to the better characterisation of time varying complex structures.

Understanding periodic structural and density changes in communication networks can optimize infrastructure and bandwidth allocation, detect social media echo chambers, and enhance healthcare networks by predicting patient flow patterns. This method also aids in

¹including the tensor portrait defined in Section 2.2 but considering different numbers of hops from each node.

pandemic planning by revealing contact network patterns and helps understand animal organization within ecological networks.

In the work presented in this chapter, I contributed by participating to the design of the methods, conducting all the experiments and exploring with my collaborators their results.

Chapter 3

Distinguishing mechanisms of social contagion from local network view

3.1 Introduction

Building on the understanding of temporal network characteristics, I further explore the dynamics of contagion processes. As previously discussed in Section 1.4, the propagation of social behaviours can be modelled using simple contagion, complex contagion, or spontaneous adoption approaches, and several works provide methods to distinguish between these mechanisms. However, these studies commonly make two assumptions limiting their applicability in real-world scenarios. First, they expect full knowledge about both the underlying network structure and the spreading dynamics. Indeed, this is a strong assumption in common real-world scenarios, where information about infection events is typically incomplete or limited to local knowledge, possibly obtained only about an adopting ego and its peers. Second, these studies assume that all individuals follow the same single adoption mechanism; either simple or complex contagion. In contrast, it has been argued that the mechanism driving one's decision to adopt a behaviour during an unfolding social contagion may depend on the intrinsic susceptibility of an individual to the actual behavioural form and the properties of the propagation process itself [101, 102]. Thus, each single adoption event may be driven by different mechanisms that jointly depend on personal factors [103, 104] (heterogeneous susceptibility and predisposition), the properties of the item being adopted (Gladwell's stickiness [101]), and the particular context (environment, time of adoption, other external factors).

In this chapter, we distinguish between simple, complex and spontaneous contagion mechanisms by addressing both the issue of limited data availability and the challenge that a single

social contagion process may involve multiple adoption mechanisms [105]. We frame this question as a classification problem and explore solutions based on likelihood and random forest approaches. These methods are developed and tested on extensive synthetic simulations, encompassing different spreading scenarios and underlying social structures, ranging from fully controlled experiments to empirical spreading cases observed on Twitter (currently called X). Our ultimate goal is to uncover the fundamental limits of distinguishability of these mechanisms, and to propose solutions that can be readily used in real-world scenarios aimed at understanding social contagion phenomena.

3.2 Results

3.2.1 Methods of classification

The following section presents diverse methods of classification used in this chapter.

Likelihood analysis

In order to perform a classification task, we use in this manuscript a method based on the likelihood analysis [127].

Likelihood definition

The likelihood is used when we want to verify a hypothesis based on observed data. More precisely, it quantifies how well the given hypothesis explains the data, and is expressed as the probability of observing the data D given the hypothesis H :

$$\text{Likelihood} = P(D | H) \tag{3.1}$$

Even if the likelihood $P(D | H)$ is similar to a probability, it is not a probability distribution over the hypothesis space, but instead a function of H for a fixed data set D . Thus, the likelihood should not be seen as the probability of the hypothesis but as a measure of how well the hypothesis explains the data.

Maximum Likelihood Estimation

An important application of the likelihood function is in Maximum Likelihood Estimation (MLE) [128], where the goal is to find the hypothesis H that maximizes the likelihood function. In our case in this chapter, we will have three possible hypotheses and we will choose the one giving the higher likelihood regarding the data.

Markovian hypothesis

Another important concept when calculating the likelihood is the Markovian hypothesis, which states that the future state of a system only depends on its present state, and not on the states preceding it. Thus, the transition from one state to another one is independent of any other transition. In probability terms, this means that we can express the probability of being in one particular state at time t as the product of every transition from one state to another over time until reaching t .

Machine learning methods

One of the first machine learning algorithms was introduced by Samuel in 1959, able to play a game of checker [129]. Since then, research in machine learning has flourished to be today's one of the largest fields of computer science, being able to solve tasks such as classification -which is what we will do in this work-, prediction, clustering, dimensionality reduction and many more [130]. Supervised machine learning algorithms are trained on input and ground truth data pairs to learn patterns and rules to perform intelligent tasks. These algorithms are widely used in various domains, including web search, spam filter, credit scoring, fraud detection or computer design. Although there exist many algorithms, this work mostly employs the random forest algorithm, which is based on the random tree approach. We have however tried several other algorithms described in paragraph ??.

Random tree and random forest

A random tree [131] is a machine learning tool used for classifying data by systematically dividing it into categories based on its features. The tree is trained on labelled data, for which the class of each instance is known, and applied to unlabelled data to estimate their categories. Structurally, a random tree is a binary decision tree, meaning that it is organised with nodes, each node having two child nodes. Initially, the tree begins with a root node containing all instances, then the data is divided into two subsets according to a specific feature, giving the child nodes, which are themselves divided in two and so on. At every division, a feature is selected and the data is split into the subset which has its feature above the threshold, and the one for which the feature falls below it. This process is repeated with different features and thresholds on the resulting subsets until the data is classified as accurately as possible. Thus, in this structure, internal nodes represent tests on specific attributes, branches represent the outcomes of these tests, and terminal nodes, or leaf nodes, represent the final class predictions. Once the tree is built, it can classify new data by following the splits from the root to a leaf node, with each path being interpreted as a logical rule for the classification.

The choice of the feature and the threshold is determined to optimise the purity of the child nodes, meaning that each node must be as homogeneous as possible. There exist several ways to define purity such as the Gini index [132] or the entropy function [133].

There are typically two phases to build a tree: initially, the algorithm recursively generates a partition based on selected attributes, and then it identifies and removes the branches that reflect noise or outliers. The algorithms coding trees are greedy, notably the Hunt's algorithm, which is a recursive algorithm on each child node of the tree.

While the majority of random trees aim to classify data as accurately as possible, some assign an estimation of the class to each leaf node rather than a class label, depending on the frequency of the instances in the leaves. This probabilistic approach can lead to overfitting, especially if a leaf node contains only a few instances, meaning that the frequency may not reflect the reality. To solve this issue, there exist some corrections, like the Laplace correction which moderate the importance of those leaves.

Following the definition of a random tree, a random forest [133] is an ensemble of random trees all built on a different sample of the initial data set, thus they are all slightly different. To classify new data, the random forest collects the inferred classes of every random tree of the forest, and returns the category which appears the most in the results. This approach makes the classification more accurate and reliable than any simple random tree, as combining the results of many trees tends to balance the errors.

Other machine learning algorithms

Even if we use the random forest algorithm to perform our classification task, we initially investigate to find the classification machine learning algorithm that gives the best accuracy. In addition to the *random tree* and the *random forest*, we test the following 7 algorithms:

- *Naives Bayes* [134]: algorithm which classifies instances using the Bayes' theorem under the hypothesis that every pair of features are independent.
- *K-nearest neighbours (Knn)* [135]: the training instances are displayed in a space of the dimension of the number of features. Every instance is assigned to the same category as the majority of its closest neighbours in this feature space.
- *Perceptron* [136]: classifier that learns by iteratively adjusting weights. It uses a threshold function to determine the class of the instances based on the dot product of input features and learned weights.
- *Support Vector Classification (SVM)* [137]: algorithm that identifies an optimal hyperplane to separate data into different classes by maximizing the margin between the classes.
- *Linear Support Vector Classification (Linear SVM)* [137]: variant of Support Vector Classification that specifically employs a linear decision boundary to classify data points into distinct categories.
- *Ada boost* [138]: boosting algorithm that sequentially trains weak learners by emphasizing misclassified instances in subsequent iterations, and thus builds a strong

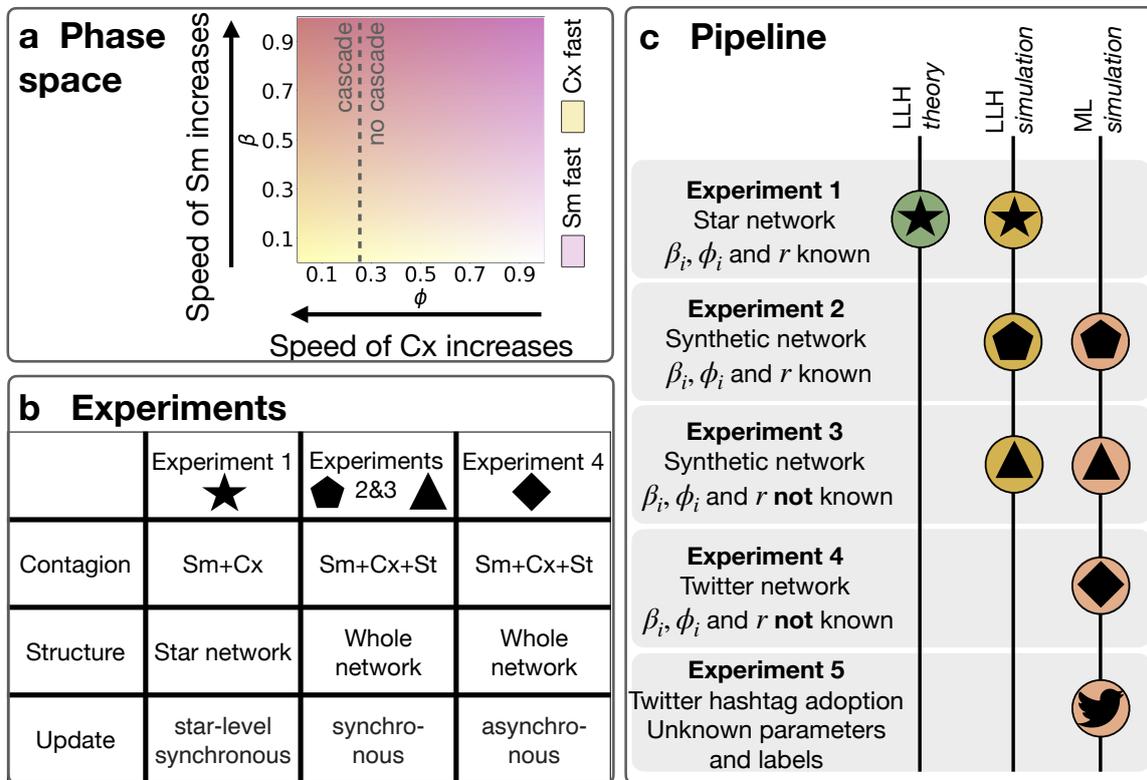


Figure 3.1: Overview of experimental setups. (a) The parameter space (β, ϕ) and the speed dependence of the simple and complex contagion processes shown as a schematic representation for illustration purposes. (b) The different experimental setups that include the considered contagion mechanisms, the complexity of the underlying network, and model update rules. (c) Schematic pipeline for the application of the log-likelihood (LLH) and random forest machine learning (ML) classification approaches to the different experiments.

classifier by combining the predictions of these weak learners.

- *Gradient boosting* [139]: boosting algorithm that sequentially trains weak learners giving more weight to the misclassified instances based on gradients of a loss function.

3.2.2 Different mechanisms of social contagion

We study adoption processes on networks, with the three infection mechanisms (simple, complex and spontaneous contagions) that can change the state of a node from susceptible to infected (cf Figure 1.2). Among the parameters of these spreading approaches, β_i and ϕ_i are crucial in shaping the propagation dynamics. High values of β_i lead to faster adoption

via S_m , while low values of ϕ_i accelerate the adoption rate via C_x , as individual thresholds become easier to reach (see Figure 3.1a and also Section 1.4.4).

The backbone of the chapter is a series of four experiments (Figure 3.1b), where we tackle the problem of distinguishing simulated S_m , C_x and S_t processes based on the infection times of an ego node and its neighbours. The experiments cover a wide range of scenarios, from the simplest configuration on disjoint star networks with β_i and ϕ_i known to the estimator (Experiment 1), to the most involved setup, simulated with co-existing, asynchronous update mechanisms with unknown parameters (Experiment 4). In each experiment, we distinguish the adoption processes using a maximum likelihood approach and a random forest classifier, whenever the method is applicable (Figure 3.1c). The likelihood approach features theoretical guarantees, and the possibility to include prior knowledge about the underlying processes [140]. However, likelihood-based approaches may not be robust if they cannot capture precisely the data from the assumed generative process [141]. In contrast, random forest classifiers tend to be more robust even if the dataset does not fit perfectly to the model, while falling short on the interpretability of the results. Finally, after highlighting the strengths and weaknesses of the two classification approaches, we apply the random forest classifier to real ego-level datasets collected from the Twitter (now called X) micro-blogging and social networking platform.

3.2.3 Process classification with known parameters

We start approaching the proposed classification task in the most elementary case, that is when the parameters $\{\beta_i\}_{i \in N}$, $\{\phi_i\}_{i \in N}$ and r governing the spreading processes are known to the classifier. Even though such information is not available in practical real-world scenarios, this setup represents an ideal starting point to understand the performance of the classifiers in a simple and controlled synthetic context.

Contagion on egocentric networks

Experiment 1. As we aim at classifying contagion mechanisms relying solely on the information available at the level of an ego node and its neighbours, the simplest setting to consider is the case of contagion processes that spread on disjoint star structures that are not part of a larger network structure. We assume knowledge only about egos and their neighbours, that together define a star structure around the central ego. The degrees of the ego (i.e. number of its neighbours) are drawn from a binomial distribution of parameters $(N, p) = (1000, 0.004)$ (which yields a mean of $\langle k \rangle = 4$), excluding the value 0. This was necessary to obtain the same parametrization as the Erdős-Rényi networks that we used in Experiment 2. We assign to each ego-node a predetermined adoption class, simple or complex, with each contagion mechanism having an equal probability of 1/2, with the cor-

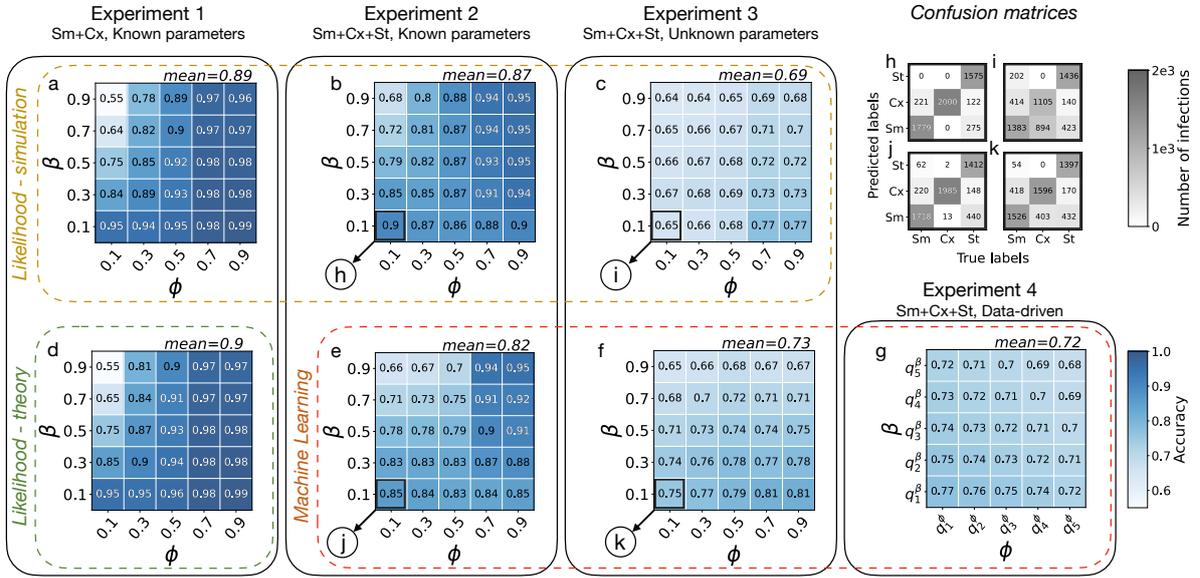


Figure 3.2: Classification accuracy values of the likelihood method (green rectangle (d)) when it is obtained theoretically and yellow rectangle (a-c) when it is obtained by simulation) and of the random forest method (red rectangle (e-g)). Results in the same column are obtained on the same Experiment produced by synthetic models, with model complexity increasing from left to right. In panel (g), the notation $q_n^{parameter}$ represents the n^{th} quintile of the parameter distribution. Panels (h-k) show the confusion matrices, namely the table which displays the counts of true versus predicted labels, with actual labels on the x-axis and predicted labels on the y-axis, associated with the highlighted pairs (β, ϕ) from Experiments 2 and 3. In general, classification accuracy decreases with increasing model complexity, but the accuracy remains well-above the random baseline (0.5 for Experiment 1 and 0.33 for Experiments 2-4). Within one experiment-method pair, accuracy increases with ϕ and decreases with β , which agrees with our intuition that the Sm and the Cx are most difficult to distinguish when both contagions propagate fast in the network.

responding parameter, respectively β or ϕ . Further, we defined the same adoption probability r_T for any neighbour of an ego, mimicking their adoption dynamics as a Bernoulli process. Assuming each node in the ego-network to be susceptible at the outset, neighbours became infected following their Bernoulli dynamics, while egos changed state only when their condition to infect has been satisfied. We simulate this contagion dynamics on 100,000 ego-networks, having 10,000 realisations for each parameter values of β and ϕ taking values from $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ and with parameter $r_T = 0.05$. In this setting the classifier was informed by the β_i , ϕ_i and r parameter values for each instance i .

After simulating the contagion process for T time steps, we feed the classification algorithm with the trajectory $\{\sigma_i(t)\}_{t=0}^T$ that takes values 0 (S) or 1 (I) and tracks the status of each

ego node i at each time step t . In order to assess whether the trajectory of an ego has been generated by the Sm or Cx mechanism, we formulate the classification problem under a likelihood framework. Since both contagion processes are Markovian, we can write, for each node i , the likelihood for an observed process to be generated by each mechanism $\mathcal{X} \in \{Sm, Cx\}$ with parameters $\{\beta_i, \phi_i\}$ as the product of the probabilities:

$$\mathcal{L}_i(\mathcal{X}) = \prod_{t=0}^T P(\sigma_i(t+1) | \sigma_{i, \Gamma_i}(t), \mathcal{X}, \{\beta_i, \phi_i\}), \quad (3.2)$$

where $\sigma_{i, \Gamma_i}(t)$ denotes the trajectories of the ego node and of its neighbours. The parameter T , corresponding to the time span of the spreading process, will never be specified as it changes for every simulation and does not influence the results. To compute the probability of observing the ego's state $\sigma_i(t+1)$ conditioned on its state and the states of the neighbours $\sigma_{i, \Gamma_i}(t)$ in the previous time step, we distinguish three cases:

1. ego stays susceptible, formally $\sigma_i(t+1) = \sigma_i(t) = 0$, which we abbreviate as $0 \rightarrow 0$
2. ego becomes infected, formally $\sigma_i(t+1) = 1, \sigma_i(t) = 0$, which we abbreviate as $0 \rightarrow 1$
3. ego stays infected, formally $\sigma_i(t+1) = \sigma_i(t) = 1$, which we abbreviate as $1 \rightarrow 1$.

In case of a simple contagion, the independence of infection probabilities on each edge makes it possible to combine the three cases into a single equation as

$$P(\sigma_i(t+1) | \sigma_{i, \Gamma_i}(t), Sm, \beta) = \begin{cases} \prod_{j \in \Gamma_i} (1 - \beta)^{\sigma_j(t)} & 0 \rightarrow 0 \\ 1 - \prod_{j \in \Gamma_i} (1 - \beta)^{\sigma_j(t)} & 0 \rightarrow 1 \\ 1 & 1 \rightarrow 1 \end{cases} \quad (3.3)$$

In case of a complex contagion, the same likelihood function takes the binary values

$$P(\sigma_i(t+1) | \sigma_{i, \Gamma_i}(t), Cx, \phi) = \begin{cases} \mathbb{K}(\sigma_{i, \Gamma_i}(t)) & 0 \rightarrow 0 \\ 1 - \mathbb{K}(\sigma_{i, \Gamma_i}(t)) & 0 \rightarrow 1 \\ 1 & 1 \rightarrow 1 \end{cases} \quad (3.4)$$

depending on whether the condition

$$\mathbb{K}(\sigma_{i, \Gamma_i}(t)) = \Theta \left(\phi \sum_j A_{ij} - \sum_j \sigma_j(t) A_{ij} \right), \quad (3.5)$$

on the proportion of infected nodes is satisfied or not. In this case A denotes the adjacency matrix of the network, with elements A_{ij} , and Θ denotes the Heaviside step function, which is equal to 1 if the input is positive, 0 otherwise.

An observed adoption could then be classified to the mechanism having the highest likelihood.

Assuming that the star networks have degrees k drawn from a binomial distribution, we display in the heatmap of Figure 3.2a the obtained accuracies (proportion of well-classified nodes) as a function of the respective pair of parameters (β, ϕ) that generated the simulations. We obtain relatively high accuracy values —with a mean of 0.9— over the whole parameter space, with the exception of the portion of the space where Sm and Cx both evolve fast, which corresponds to the parameter extreme when $\beta \rightarrow 1$ and $\phi \rightarrow 0$. In this case, Sm and Cx are very difficult to distinguish; in both cases, the ego node becomes infected most likely one timestep after its first neighbour adopts. This parameter range also corresponds to the least distinguishable scenario at the level of the global epidemic curves, as they both evolve rapidly even in populations with homogeneous adoption mechanisms (Figure 1.3). In this range, the lowest classification accuracy is around 0.55, which is still slightly above the expected accuracy of a random classifier 0.5. Notably, the two processes are highly distinguishable in the opposite case, when $\beta = 0.1$ and $\phi = 0.9$. In this other extreme, ϕ is so high that Cx adoptions are possible only once most of the neighbours of the adopting ego have been spontaneously infected. At the same time, Sm adoptions are still possible via repeated stimuli from a few neighbours, making the two processes easier to distinguish.

A major advantage of this stylised setup on disjoint degree- k star networks is that the likelihood classification accuracy can be approximated analytically. Let us define \hat{X} to be the contagion label that the algorithm assigns, and X to be the true contagion label. Assuming a uniform prior on the contagion labels, the accuracy of the algorithm can be expressed as:

$$\frac{P(\hat{X} = Cx | X = Cx) + P(\hat{X} = Sm | X = Sm)}{2} \quad (3.6)$$

Since for a node infected by the complex contagion, the probability to observe the transition $0 \rightarrow 0$ ($0 \rightarrow 1$) assuming that the complex contagion is at stake, is always equal to 1 as the proportion of infected neighbours is below (above) ϕ (Eq. 3.4). Therefore, we always have $P(\sigma_i(t+1) | \sigma_{i,nb}(t), Cx, \phi) = 1$ and the maximum likelihood approach always classifies complex nodes correctly. Consequently, $P(\hat{X} = Cx | X = Cx) = 1$ always holds.

For the second term, to compute

$$P(\hat{X} = Sm | X = Sm) = 1 - P(\hat{X} = Cx | X = Sm), \quad (3.7)$$

we need to estimate the probability that a node i with degree k becomes infected by the simple contagion immediately after $\lceil k\phi \rceil$ of its neighbours get spontaneously infected, and

therefore it incorrectly becomes classified as complex. Conditioning on the event that the ego has n infected neighbours at time t , we define the following two random variables:

- N_n denotes the number of time steps until a new neighbour gets infected
- E_n denotes the number of time steps until the ego gets infected, assuming that no new neighbour gets infected.

Since at each time step, the probability of a new neighbour spontaneously becoming infected is $p_n = 1 - (1 - r)^{k-n}$, the random variable N_n follows a geometric distribution with success probability p_n . Similarly, since the probability that any of the n neighbours infect the ego node in each time step is $b_n = 1 - (1 - \beta)^n$, the random variable E_n follows a geometric distribution with success probability b_n . Our goal is to compute the probability of the event that the ego becomes infected immediately after $\lceil k\phi \rceil$ of its neighbours get infected, i.e. that $N_n < E_n$ holds for $n < \lceil k\phi \rceil$, but $E_{\lceil k\phi \rceil} = 1$. For each $n < \lceil k\phi \rceil$, the corresponding event probability can be computed based on the well-known formula of two competing geometric random variables. For $n = \lceil k\phi \rceil$, the event probability is simply b_n . Finally, due to the Markov property of the contagion process, assuming that no two neighbours get infected at the same time, we arrive at the final result by computing the product of the event probabilities for each n :

$$P(\hat{X} = Cx \mid X = Sm) \approx \left(\prod_{n=1}^{\lceil k\phi \rceil} \frac{p_n - p_n b_n}{b_n + p_n - p_n b_n} \right) b_{\lceil k\phi \rceil}. \quad (3.8)$$

Our result is an approximation, because we did not account for the low-probability event that two neighbours might be infected at the same time. Despite this limitation, comparing the theoretically estimated accuracies from Eq. (3.8) (visualised in Figure 3.2d) with the simulation outcomes (Figure 3.2a), we observe a very close match, with a maximum difference of 0.01.

Overall, Experiment 1 features a high classification accuracy and precise analytical results, while making strong assumptions on the network structure and the adoption mechanisms. Since the likelihood approach matches the underlying model exactly, it is an optimal estimator, and we omit the application of the random forest approach in this setup. However, since this setting also neglects some of the most important features of realistic social contagions and social structures, it can only be considered as the simplest solvable reference model to be compared with more complex scenarios.

Contagion on random networks

Experiment 2. To generate a more realistic setting, we consider contagion mechanisms that spread over larger network structures. Most of the results in this section were ob-

tained on the giant component of Erdős-Rényi random networks [142] with 1000 nodes and an average degree of 4, but we also present results on random networks with degree heterogeneity, triadic closure and community structure with the same parameters using Watts-Strogatz [143] and Barabási-Albert [29] model networks, Stochastic Block model networks [33], and a real Twitter mention network [144] defined by linked customers if they mutually mentioned each other during the observation period. For computational purposes we filter the Twitter mention network to keep only its largest connected component, i.e. the largest interconnected subset of nodes within a network (370,544 nodes and 1,013,096 links) and we assume it to be undirected by ignoring the directions of its links. Similarly to Experiment 1, we randomly predetermine the contagion mechanism (simple or complex) for each node with a probability $1/2$ and a parameter (β or ϕ) accordingly from the set $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ in order to have all pairs (*process, parameter*) equally distributed in the data set. This time, however, we allow each node to spontaneously adopt during the contagion process, regardless of their predetermined mechanism. This way the contagion does not vanish even on large networks with extreme S_m and C_x contagion parameters, but continues spreading following a linear dynamics. Having all nodes as susceptible at the outset, the propagation initialised by infecting one random node. The spreading process among the rest of the nodes is gradually spreading either by their assigned process of contagion, or through the spontaneous adoption with a rate of r . We stop the contagion process when all of the nodes become infected, except for the Twitter mention network. For this network, we terminate the process when 90% of the nodes are contaminated, as simulating until the entire network is infected is computationally heavy due to its large size.

For each synthetic network model, the propagation is run on 20 independent network realisations, with $r = 0.005$. For each node i , the parameters β_i , ϕ_i and r are assumed to be known by the classifiers. The modification also implies that, since nodes can adopt via the simple, the complex or the spontaneous mechanisms, our classification algorithms need to distinguish between the three hypotheses.

In line with the approach of Experiment 1, we compute the likelihood that each adopter follows a specific contagion mechanism (see Eq. (3.2)) based on the trajectories of the ego nodes and their neighbours. Since the assumption of the independent adoption of the neighbours of an ego does not hold anymore, the likelihood framework becomes an approximation. The calculations are similar to Experiment 1, but instead of two, now they involve three processes: simple, complex and spontaneous adoptions. For clarity, we divide those three processes in four scenarios:

1. The ego, initially assigned with the simple contagion, eventually becomes infected by the simple contagion:

$$P(\sigma_i(t+1)|\sigma_{i,nb}(t), \text{Sm}, \beta) = \quad (3.9)$$

$$\begin{cases} (1-r) \prod_{j \in nb} (1-\beta)^{\sigma_j(t)} & 0 \rightarrow 0 \\ (1-r) \left(1 - \prod_{j \in nb} (1-\beta)^{\sigma_j(t)} \right) & 0 \rightarrow 1 \\ 1 & 1 \rightarrow 1 \end{cases} \quad (3.10)$$

2. The ego, initially assigned with the simple contagion, eventually becomes infected by the spontaneous contagion:

$$P(\sigma_i(t+1)|\sigma_{i,nb}(t), \text{Sm} \rightarrow \text{St}, \beta) = \quad (3.11)$$

$$\begin{cases} (1-r) \prod_{j \in nb} (1-\beta)^{\sigma_j(t)} & 0 \rightarrow 0 \\ r & 0 \rightarrow 1 \\ 1 & 1 \rightarrow 1 \end{cases} \quad (3.12)$$

Note that the $0 \rightarrow 1$ transition can occur at any step due to the Sm or the Sp process, but once it occurs it is irreversible. Therefore we need that

$$\mathbf{P}(0, \text{Sm}|0, \text{Sm}, \beta) + \mathbf{P}(1, \text{Sm}|0, \text{Sm}, \beta) + \mathbf{P}(1, \text{Sp}|0, \text{Sm}, \beta) = 1,$$

which is indeed holds for these equations.

3. The ego, initially assigned with the complex contagion, eventually becomes infected by the complex contagion:

$$P(\sigma_i(t+1)|\sigma_{i,nb}(t), \text{Cx}, \beta) = \quad (3.13)$$

$$\begin{cases} (1-r) (1 - \mathbb{K}(\sigma_{i,nb}(t))) & 0 \rightarrow 0 \\ (1-r) \mathbb{K}(\sigma_{i,nb}(t)) & 0 \rightarrow 1 \\ 1 & 1 \rightarrow 1 \end{cases} \quad (3.14)$$

4. The ego, initially assigned with the complex contagion, eventually becomes infected by the spontaneous contagion:

$$P(\sigma_i(t+1)|\sigma_{i,nb}(t), Cx \rightarrow St, \beta) = \tag{3.15}$$

$$\begin{cases} (1-r)(1-\mathbb{1}(\sigma_{i,nb}(t))) & 0 \rightarrow 0 \\ r & 0 \rightarrow 1 \\ 1 & 1 \rightarrow 1 \end{cases} \tag{3.16}$$

Note that the $0 \rightarrow 1$ transition can occur at any step due to the Cx or the Sp process, but once it occurs it is irreversible. Therefore we need that

$$\mathbf{P}(0, Cx|0, Cx, \beta) + \mathbf{P}(1, Cx|0, Cx, \beta) + \mathbf{P}(1, Sp|0, Cx, \beta) = 1,$$

which is indeed holds for these equations.

Nevertheless, accuracy values for the whole parameter space summarised in Figure 3.2b confirm that this approach can still perform well achieving a mean accuracy of 0.87 —well above the expected accuracy of a random classifier (0.33).

Since the likelihood framework provides an approximate solution for Experiment 2, it calls for alternative approaches. We conduct an extensive classification model selection to find the best machine learning algorithm to distinguish between the simple, complex and spontaneous contagions in Experiment 2. We test 9 algorithms and present the mean accuracies over the whole parameter-space in the classification of the instances from Experiment 2 for each of them in Table 3.1. Among the algorithms displaying the highest accuracies (above 0.82), we opt for the *Random forest* method first due to its significantly faster computation times compared to *SVM*. Additionally, the *Random forest* algorithm, consisting of an ensemble of decision trees whose outcomes are combined, generally outperforms individual *Decision Tree* methods. Finally, we exclude the *Gradient Boosting* algorithm due to its limited explainability.

Naives Bayes	Knn	Perceptron	Linear SVM	SVM	Decision tree	Random forest	Ada boost	Gradient boosting
0.66	0.81	0.68	0.81	0.82	0.82	0.82	0.75	0.83

Table 3.1: Average over the whole parameter-space of the accuracies on the classification of the contagion cases from Experiment 2

In order to strike a balance between performance and interpretability, we train random forest classifiers on the same synthetic dataset as above. After testing on several structural and dynamical features of the ego and its neighbours, we identify eight relevant features for the classification that appear with distinct distributions for different infection mechanisms (cf Appendix B.1). These are (i) the degree, (ii) the proportion of infected neighbours, (iii) the number of infected neighbours, (iv) the sum of received stimuli, (v) the average number of received stimuli by neighbour, (vi) the standard deviation of per neighbour stimuli, (vii) the time since the first infected neighbour and (viii) the time since the last infected neighbour.

We train 25 random forest algorithms, one for each pair of (β, ϕ) by sampling 18,000 instances from Experiment 2, with 6,000 contagion cases from each category. Then we test the models on a set containing 6,000 instances (2,000 instances from each category). The results are averaged over 10 realisations. Each random forest algorithm has 100 trees without any limit on the maximum depth. As mentioned in paragraph 3.2.1, various functions, such as the Gini index or the entropy function, can measure the purity of the leaves. To select the optimal function, we employ the grid search method, which consists in finding the best hyper-parameters (here the best function) for a model by testing all possible combinations.

The random forest approach provides very similar results (see Figure 3.2e) to the likelihood-based calculations (Figure 3.2b), only with slightly worse average accuracy 0.82. The confusion matrix, namely the table evaluating the performance of a classification model by showing the distribution of correct and incorrect predictions across different classes, is displayed in Figure 3.2h and j. While the two methods perform similarly in classifying simple contagion cases, the random forest misclassifies complex and spontaneous instances at a higher rate.

Notably, given the interpretability of the trained random forest classifiers via feature importance, we can further restrict our original eight features to only three, and retain similar accuracies as before (see Appendix B.2). Interestingly, some feature subsets are consistently optimal across the full parameter space. This is reported in Figure 3.3, where we present the number of times a feature appears within the subset of the top-3 optimal features, normalised by the number of possible instances (parameter pairs β, ϕ in the phase space). Overall, the two most recurring features are the times since the first and the last infected neighbours. These can be also easily interpreted within the modelling framework: the time since the first infected neighbour cannot be too high for S_m , as that would mean too many repeated stimuli without an infection event, while for the threshold-based C_x the time since the last infected neighbour has to be necessarily one.

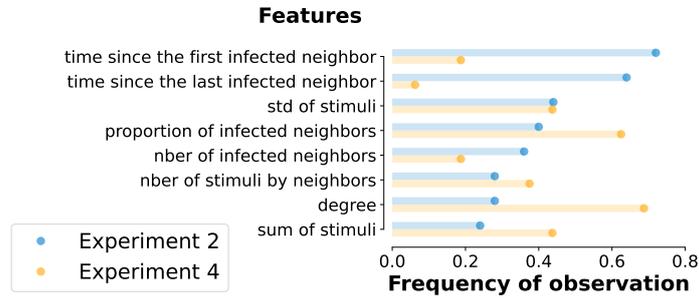


Figure 3.3: Frequency of observation of the features used to train the random forest classifier among the top-3 most important ones across the full parameter space for Experiment 2 (blue) and Experiment 4 (orange). Frequencies are computed as the number of appearances normalised by the number of possible occurrences. The resulting most important features are the *time since the first and the last infected neighbour*.

3.2.4 Process classification with unknown parameters

Up to this point, all the investigated tasks assumed precise knowledge of the parameters β_i , ϕ_i and r governing the different processes. However, in realistic scenarios, these need to be also inferred together with the contagion mechanisms, thus motivating the following experimental setup.

Experiment 3. In this setting we classify the contagion instances from Experiment 2 assuming unknown contagion parameters, which means distinguishing mechanisms without knowledge of the parameters that governed them. In the likelihood approach, we use the same equations to compute the likelihood that the contagion instance i is simple, complex or spontaneous as before, except we also estimate the values of β_i , ϕ_i and r . We set the value of $\hat{\beta}_i$ as the inverse of the number of received stimuli by the node i , and the value of $\hat{\phi}_i$ as the proportion of infected neighbours at the time of the infection of node i . The value of \hat{r} is calculated as the fraction of time spent by a node in the S state with at least one infected neighbour, averaged on every node in that case.

Figure 3.2c shows that we still classify the adoption mechanisms with high accuracy, especially considering the increased difficulty of the classification problem compared to the earlier settings. The mean accuracy was found to be 0.69, well above the reference value of a random classifier (0.33). We observe the worse performance for low values of ϕ , due to the high rate of confusion between complex and simple contagion cases (Figure 3.2i). Those nodes are generally infected just after the appearance of an infectious neighbour, making it difficult for the model to distinguish between the two peer-driven mechanisms S_m and C_x . The accuracy is the highest for large values of ϕ and low values of β . As before, we gain the

most information about the processes when both of them are progressing slowly.

We also test the random forest approach in this experiment by using the same features used in Experiment 2, but training instead one unique model over the whole phase space as the parameters are not known anymore. The training data set contains 18,000 instances in total (6,000 instances in each category), regardless of the parameters. The results are averaged over 10 realisations. Each random forest algorithm has 100 trees without any limit on the maximum depth. The use of the Gini function or the entropy function is determined by grid search.

Interestingly, this solution provides slightly more accurate results (see Figure 3.2f) than the likelihood method (see Figure 3.2c), especially for low values of β . Reading the confusion matrices (in Figure 3.2k and i resp.), this improvement mostly comes from the better classification of complex contagion instances, that were commonly classified as simple by the likelihood approach. Nevertheless, the overall accuracy of the random forest classifier is lower for Experiment 3 as compared to Experiment 2, which is expected, as the estimators receive less information.

To understand how network structure influences process distinguishability, we apply the classification methods on Experiments 2 and 3 on various networks (Figure 3.4). The values of accuracies remain consistent across the Barabási-Albert, Watts-Strogatz and Stochastic Block Model networks. This suggests that the global network structure has a limited impact on the local differentiation of contagion processes in each performed experiment. However, we observe a decrease of 0.02 on the accuracy average considering a true Twitter network, but with the machine learning method with unknown parameters. Indeed, one of the most important features of this method is the degree (Figure 3.3), which presents a larger variation with the Twitter network.

3.2.5 Case study: adoption mechanisms on Twitter

After demonstrating the validity of our methods in controlled synthetic settings, we now turn our focus toward real contagion processes to showcase the applicability of the devised approach to empirical scenarios. To this end, we rely on an ego-level dataset of adoptions from Twitter [145] (now called X), a micro-blogging and social networking platform, where users can follow each other, and share short messages, or tweets. The dataset contains all tweets posted by 8527 selected users (egos who are interested in French politics) and the people they follow (whom we call followees, or the members of the ego network) between May 1 2018 to May 31 2019 (for more details about the data collection see [146]). This amounts up to a total of 1,844,978 timelines, i.e., the timely ordered personal stream of tweets posted by all these users. This dataset allows us to identify the time of adoption of a given hashtag by an ego together with the time of all incoming stimuli from its neighbours that previously

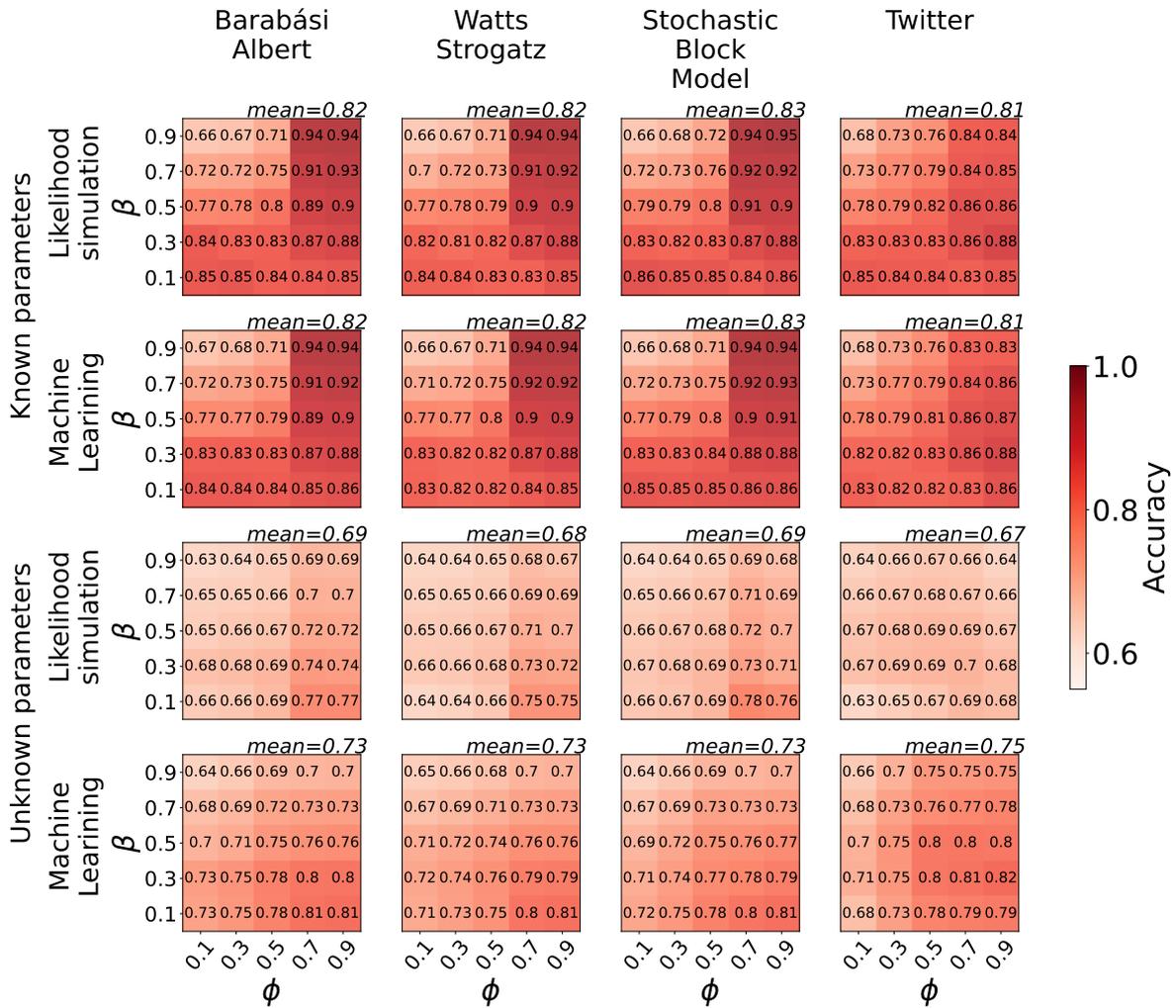


Figure 3.4: Accuracies obtained by classifying the infection instances from Experiments 2 and 3 on different networks (x-axis) and with different methods (y-axis). The networks are created with 1000 nodes and an average degree of 4. The model parameters are chosen to maintain this average degree, with $m = 2$ for the Barabasi-Albert model and $k_0 = 4$ for the Watts-Strogatz model. In addition, in the Watts-Strogatz model, 5% of the edges are rewired. In the Stochastic Block Model, nodes are gathered into 20 groups, with a probability to have an edge between groups of 0.001. To maintain an average degree of 4, the probability of an edge within a group is adjusted, resulting in a value of 0.066. The results are averaged over 20 realisations. The values of accuracies do not depend on the structure of the network.

posted the same hashtag. These tweets cover multiple topics, which may correspond to the spreading of various co-occurring social contagion processes. Since we are interested in analysing each contagion process separately, we filter messages that contain a given set

	Sm	Cx	St
Random forest	970	349	4955
Likelihood	4440	1447	387

Table 3.2: Number of instances of contagion mechanisms inferred by the likelihood and random forest methods on the #GiletsJaunes Twitter dataset.

of hashtags within the same topic. We choose to focus on the hashtag #GiletsJaunes and its variants¹, characterising a political uprising in France that induced a significant social contagion unfolding on Twitter. We first identify egos who adopted a related hashtag, and observe the posts of their followees over the preceding week, limiting in this way the effect of influence to the recent past only. As per the synthetic cases, we can define the degree of an ego as the number of its followees who have posted at least one tweet during the week preceding the adoption. In addition, user activity on Twitter is not linear in time—as in our previous simulations—but it is driven by circadian fluctuations, bursty patterns, and individual preferences. We thus move from real-time to event-time simulations. In this setting, a time step for an ego (leading to potential adoption cases) is counted as the number of tweets by the followees, regardless of whether they contain the hashtag of interest; every time an alter posts content containing the selected hashtag, the ego will receive a stimulus.

Empirical traces of social contagion set a particularly difficult problem for classification because neither the parameters of the different contagion mechanisms are known, nor any ground truth is available for validation of the classification results. In the following, we propose pathways that yet allow us to learn about the distinguishability of contagion mechanisms in the Twitter dataset.

As a starting point, we applied our classifiers designed for Experiment 3, where we have no information about the adoption parameters. Table 3.2 shows that the two methods give rather unbalanced results, with the random forest detecting a large number of spontaneous adoptions and the likelihood approach being biased towards simple contagion. This discrepancy in the results suggests that one or both of the models might not be capturing the interaction patterns within the Twitter data sufficiently well.

When it comes to empirical adoption data collected via social media, one of the largest biases is induced by the waiting time t_w [85], that is the time gap between the moment someone becomes convinced by an idea (upon exposure) and the moment we can actually observe it through an active adoption event (posting). We report the waiting time distribution for the Twitter dataset in Figure 3.5a, where one time step corresponds to the time span between two consecutive tweets. This $t_w = t_a - t_e$ lag between the exposure t_e and the adoption t_a time can depend on individual user characteristics. It biases our observations as

¹We target every user who has posted one of those hashtags: #GiletsJaunes, #giletsjaunes, #Giletsjaunes, #GiletJaune, #Giletjaune, #giletjaune, #giletsjaune, #Giletsjaune, #GJ

during this t_w time further exposures can appear, that in principle could not be even necessary for the subsequent adoption (“incubation”). Nevertheless, the only observation we can make is about the sequence of influencing tweets, as we can not know the exact tweet that triggered the adoption. The effects of such biases have been studied earlier in other scenarios of online adoption [85, 99]. In light of these observations, it is clear from the likelihood computations and from the feature importance ranks shown in Figure 3.3 that both the approaches used so far are ill-suited in this case since they heavily rely on precise adoption times —assuming no waiting time. To steer our classification algorithms away from making estimates based on this hard assumption, we now introduce a synthetic contagion process evolving on an activity driven temporal network model parameterised from data, and where waiting times can be measured. The goal of this following model is to obtain a representation as close to reality as possible, enabling the training of a random forest algorithm to classify real contagion cases.

Activity driven networks with asynchronous dynamics

Experiment 4. This experiment is inspired by the Activity Driven network model (ADN) and has been created to represent the propagation of a hashtag on the Twitter platform. Here we use the largest connected component of an undirected mutual follower network from Twitter [144] and concentrate on the propagation of the hashtags related to the political movement #GiletsJaunes. For computational purposes, we iteratively filter this network to reduce its size. At the outset, the filtered network only contains one randomly selected node from the initial network. Subsequently, a neighbour of the initial node is selected with a probability inversely proportional to the node’s degree. Once a neighbour is selected, it is incorporated into the filtered network along with its edge. This approach allows us to sample a connected subset without bias toward hubs. Subsequently we reproduce this process, each time selecting a neighbour from the newly integrated node and its edge, until we achieve a network size of 100,000 nodes.

We assume that nodes can be in three distinct states: susceptible (not yet infected), aware (they are already infected, but that has not been observed yet through an active post), and detected (they are infected and this has been observed). Every node i is assigned with an activity $\hat{a}_i \in [0, 1]$ mimicking its level of participation on the Twitter platform. As the distribution of the number of tweets posted by each user during a week depends on its degree and because those distributions along a certain degree range are not part of the typical known distributions (Figure 3.5c), we sample the assigned activity of each node with a normal distribution centered on the average number of tweets posted by each user corresponding to its degree and constrained between 0 and 1. They are also attributed homogeneously to a contagion process, either simple or complex, which determines their adoption mechanism.

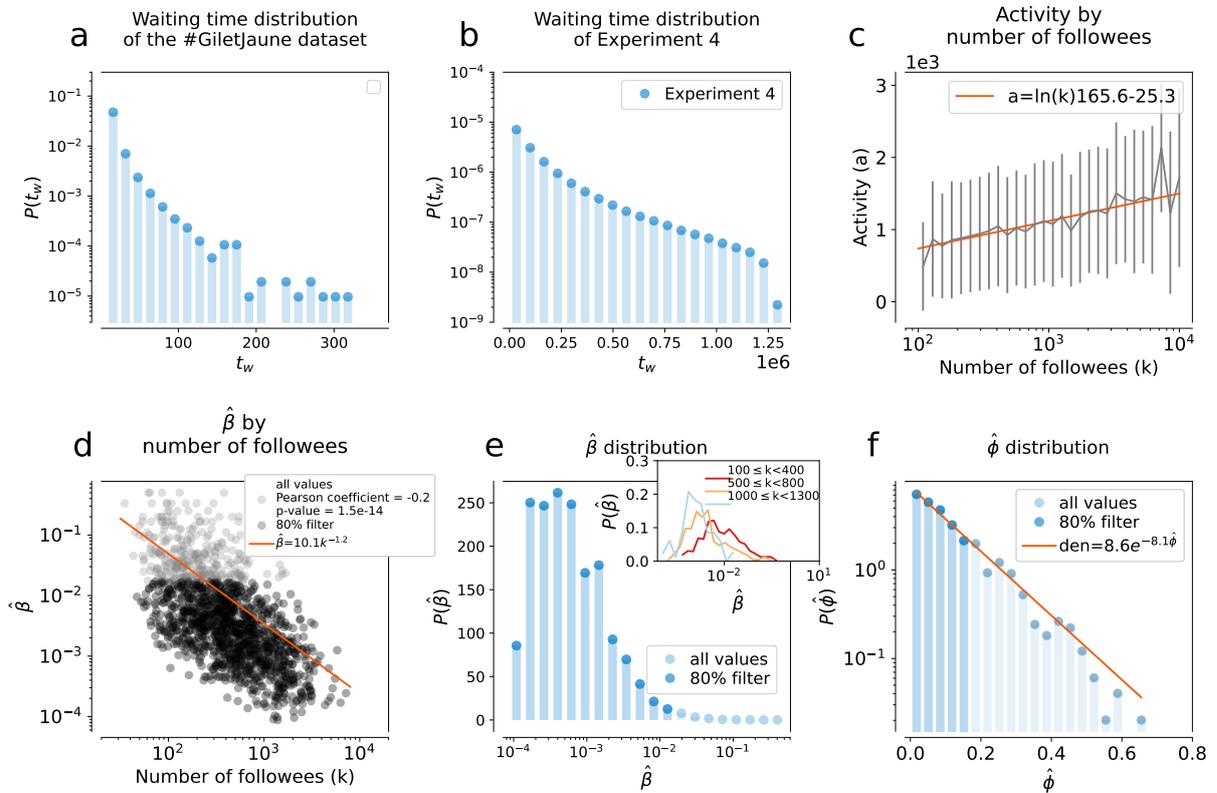


Figure 3.5: Parameter distributions and dependencies of Experiment 4 inferred from the #GiletsJaunes Twitter dataset. The waiting time distributions observed in the (a) #GiletsJaunes dataset and (b) in Experiment 4. (c) Correlation between the activities and degrees of nodes in the synthetic propagation inferred from the distribution of the number of tweets posted during the week before adoption as a function of the number of active followees in the Twitter dataset. (d) Correlation between the inferred simple contagion parameter $\hat{\beta} = 1/(\text{number stimuli})$ and node degrees observed for egos in the #GiletsJaunes dataset. (e) Distribution of the inferred simple contagion parameter $\hat{\beta}$. The inset depicts the same distribution stratified by degree. (f) Distribution of the $\hat{\phi}$ complex contagion parameter inferred as the proportion of infected neighbours at the time of adoption of an ego in the #GiletsJaunes dataset ($\hat{\phi}$). Since the $P(\hat{\beta})$ and $P(\hat{\phi})$ are broad, we apply a filter to retain the 80% of their smallest values.

Further, nodes are endowed with parameters $\hat{\beta}_i$ or $\hat{\phi}_i$. $\hat{\beta}$ are defined as the inverse of the number of times a hashtag appeared in the timeline of an observed ego's neighbours, one week before the ego's adoption. Note that we consider cases of infected egos who have at least one infected neighbour at the time of adoption. Since the $\hat{\beta}$ parameter shows correlation with the node degree (see Figure 3.5d), we decided to account for this dependency when sampling

$\hat{\beta}$ values for egos. We group nodes by their degrees and assume that each $P(\hat{\beta})_k$ distribution for a degree class can be approximated by a log-normal distribution with an average characterising the actual degree class (see Figure 3.5e and its inset). Thus for each node i with degree k to obtain a $\hat{\beta}_i$ we simply sample the corresponding log-normal distribution. At the same time, the parameter $\hat{\phi}_i$ for the complex contagion mechanism is measured as the fraction of infected neighbours of an ego that adopted a hashtag. The distribution of $P(\hat{\phi})$ (in Figure 3.5f) is measured from adoption cases where the last infected neighbour of the ego before its adoption was a newly infected neighbour. We assign a parameter $\hat{\phi}_i$ to a node i by sampling this distribution $P(\hat{\phi})$ shown in Figure 3.5f. Finally, to avoid the sampling of extreme values, since the distributions $P(\hat{\beta})$ and $P(\hat{\phi})$ appeared as broad distributions, we filter them by keeping 80% of their lowest values for parameter sampling. For a robustness analysis on the effect of filtered fraction of inferred parameters see the Appendix B.3.

After infecting a uniformly randomly selected seed node to launch the spreading process, we iteratively execute the following protocol at each time step: first a node is selected randomly with a probability proportional to its activity, indicating that this node posts a tweet. If the node is susceptible, it can become adopted with probability r , mimicking the possibility to post the hashtags spontaneously. Otherwise the susceptible node can get infected through its assigned adoption mechanism. If a node is active but susceptible, its post will not count towards the influence of its neighbour. However, if the node is aware, at the time of its next post it becomes detected. Once aware or detected we assume that at each future activity of a node it will post the spreading hashtag. If a post of a node includes the hashtag, it counts as a stimulus to all of its neighbours, which can become aware if they are susceptible and their condition of infection is reached. In our simulations we modelled the contagion processes in the network until they reached 90% of the nodes and used the observed adoption instances for the training of a random forest classifier that was not aware of the contagion parameters. The resulting waiting time, measured for each infected node as the time between the aware and detected state, follows a broad distribution (Figure 3.5b), similar to the empirical observations.

The complexity of Experiment 4 makes the application of the likelihood method unfeasible, so we continue our investigation only through the random forest approach, using the same feature set as in the previous experiments, and assuming unknown contagion parameters. As before, we pre-assign an adoption mechanism to each node in the modelled Activity-Driven network and compute the classification accuracy. Results, shown in Figure 3.2g, demonstrate that despite the increased complexity of this data-driven experiment, the random forest can achieve good classification accuracy all across the parameter space, with average accuracy 0.72. In this experiment, the spontaneous adoptions are the hardest to classify since they appear with a very low rate (see Appendix B.3, Table B.1). It is worth noticing that the importance of the features is different from the one previously shown for Experiment 2 (Figure 3.3). While the feature *time since the last infected neighbours* dimin-

ishes in importance due to the presence of a waiting time, the *proportion of the infected neighbours*, and particularly the *degree of the central ego gain significance* (Figure 3.3).

Classification of Twitter hashtags

Experiment 5. To conclude our case study on the Twitter dataset, we apply the trained models from Experiment 4 on the adoption cases of #GiletsJaunes and related hashtags. The inset of Figure 3.6 shows that most adoption cases are classified as simple as opposed to complex. This suggests that more people adopt #GiletsJaunes through a repeated influence from their contacts than through combined influence mechanisms. The less detected class is the one of spontaneous adoptions, suggesting the limited influence of external sources with respect to peer-induced contagion within the platform.

Since no ground truth exists for this dataset, instead of visualizing the accuracy values on the (β, ϕ) phase space, I show in Figure 3.6 the full distribution of inferred adoption mechanisms stratified by their inferred contagion parameters $\hat{\beta}$ and $\hat{\phi}$ (aggregated in deciles). We can see that ego nodes with high $\hat{\beta}$ and low $\hat{\phi}$ values are more likely to be classified as Cx, whereas egos with low $\hat{\beta}$ and high $\hat{\phi}$ tend to be classified as Sm. However, Figure 3.6 also suggests that the two inferred parameters, $\hat{\beta}$ and $\hat{\phi}$, cannot capture the complexity of the classification problem on their own. Indeed, both Sm and Cx adoptions appear throughout the parameter space, highlighting the added value of the random forest classifiers trained in our modelling framework. Finally, we observe that the certainty of the classification algorithm improves with lower $\hat{\beta}$ and higher $\hat{\phi}$ values, which can be explained by the increased number of stimuli, and therefore a richer dataset, in this parameter range.

3.3 Discussion

Our goal in this chapter was to infer social contagion mechanisms leading to the adoption of products, ideas, information, or behaviours. We restricted the focus to three complementary contagion mechanisms potentially determining the behaviour of an ego node, whether adopting spontaneously (exogenous influence) or due to transmission on a social network (endogenous influence) via simple or complex contagion mechanisms. The general problem of distinguishing social contagion mechanisms in networked populations has recently been addressed by analysing macroscopic spreading curves at the population level [90, 96, 100], typically assuming that only one a single mechanism is exclusively present during the contagion process. In this work, we overcome these assumptions by (i) considering only microscopic information at the level of the adopter and their peers and (ii) allowing different contagion mechanisms to be simultaneously present —with different parameters— during the same spreading phenomenon. Under these assumptions, we tackled the inference ques-

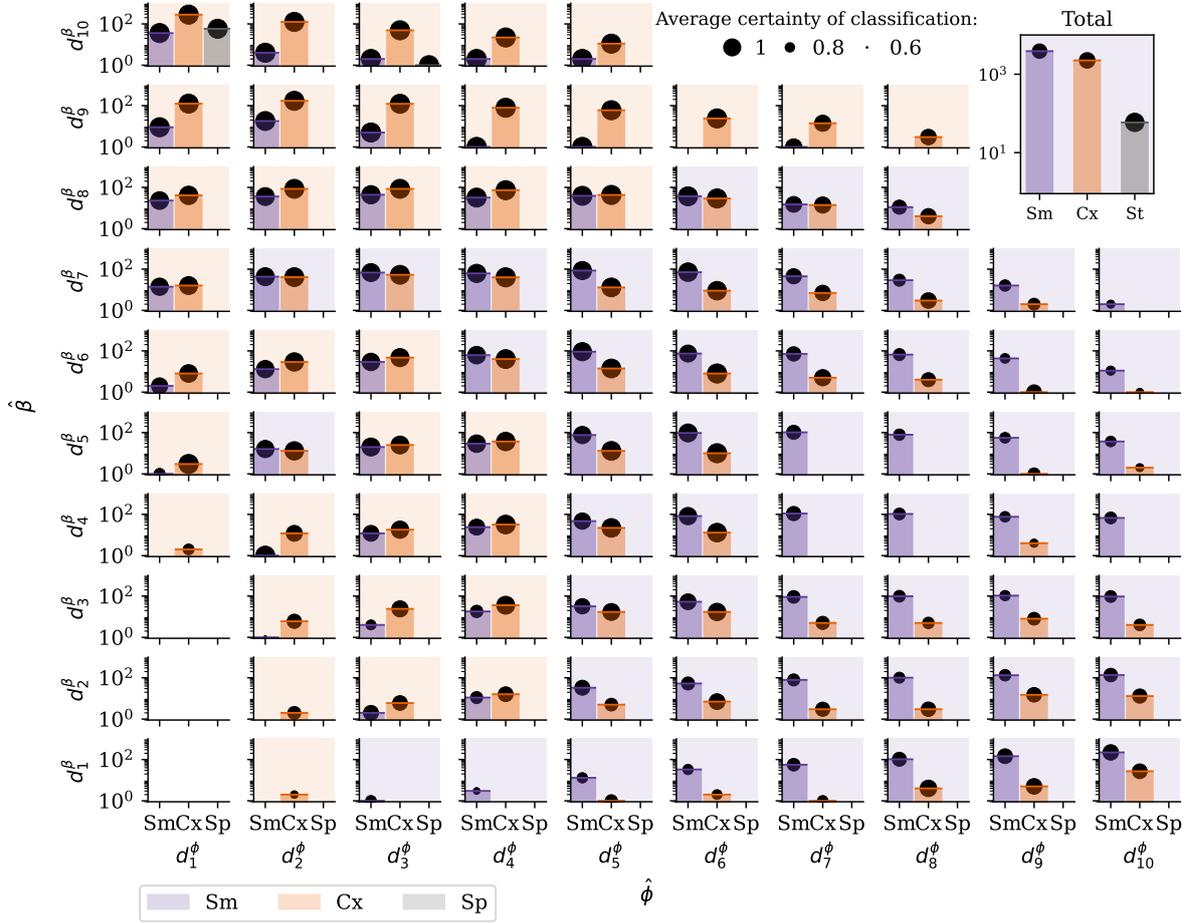


Figure 3.6: Classification of contagion mechanisms of the #GiletsJaunes Twitter dataset as the function of $\hat{\phi}$ (x-axis) and $\hat{\beta}$ (y-axis) parameters. The notation $d_n^{parameter}$ represents the n^{th} deciles of the parameter distribution from the #GiletsJaunes dataset from Figure 3.5. The classification results of each instance i are shown at the corresponding location of the decile of its inferred $\hat{\phi}_i$ and $\hat{\beta}_i$ parameters sampled from the $P(\hat{\phi})$ and $P(\hat{\beta})$ distributions. The background colour of each panel indicates the dominating classified mechanism that characterise the given parameters (purple for Sm, orange for Cx and blue for Sp). The certainty of classification, displayed with black circles, is defined as the proportion of trees in the random forest that have classified an instance into the assigned contagion type, averaged over the set of instances classified in that contagion type. Most of the infection cases are classified as simple if their $\hat{\beta}$ are in the 8th decile or below and their proportion of infected neighbours is greater than d_5^ϕ , and as complex otherwise.

tion as a classification problem under a likelihood and a random forest approach over a sequence of experiments with increasing levels of complexity. We showed, in controlled synthetic settings, that the limited information available from an ego and its peers is generally enough to distinguish the specific adoption dynamics with varying levels of accuracy depending on the contagion parameters. The lines between the mechanisms become more blurred in cases when one infectious neighbour is enough to induce the adoption of an ego. This can happen for strongly infectious items spreading via simple contagion (akin to high individual susceptibility) or low individual thresholds in adoptions triggered by complex contagion, both cases leading to an immediate local transmission and rapid global spreading. Interestingly, in the simplest experiments performed via simulations on synthetic static networks, we found little impact of the network structure on the accuracy of the classification task. Recent results have shown that simple contagion leads to similar infection patterns across different network models, while the patterns associated with complex contagion mechanisms are less robust [94, 147, 148]. This could explain the fact in Experiments 2 and 3 we do not observe major differences in the distinguishability of the mechanisms over different network structures, from Erdős-Rényi graphs to those generated via Barabási-Albert, Watts-Strogatz, and Stochastic Block Model approaches. Increasing the level of realism, we demonstrated that simplistic models fail to capture the full complexity proper of real-world transmissions, such as waiting times, or the non-static structure of empirical social networks. The challenges arising in these scenarios confirm the inherent difficulty that comes with these tasks when several internal and external factors are at play at both the dynamical and structural level, as also highlighted in other recent studies that tackled the inference problem in different contexts [149]. Nevertheless, even in these realistic settings when mechanistic approaches seem to be out of reach, a random forest classifier trained on a carefully parametrized synthetic model can give interpretable results.

Despite the comprehensive approach to the inference problem in this paper, our results presented here have certain limitations. First, for simplicity reasons we only consider static network structures, while in reality social influence is mediated via temporal interactions. Further, we assumed that the effects of external influence (like advertisements and news) do not vary in time, that is clearly an approximation. During our likelihood formulation we assume each contagion instance to be independent from each other, which is only an approximation, that is accounted for in the random forest approach. Finally, since no real dataset is available with ground truth information regarding the adoption mechanisms of a social contagion, it prevents us to validate our findings in our final experimental setting. Such datasets are challenging to collect and require careful experimental design. However, we envision that our learned labels could be validated even without extensive data collection, if they were treated as hidden variables used to predict adoption times. More explicitly, under the assumption that certain spreading processes or certain individuals predominantly follow simple or complex contagion, our inferred labels on past datasets could predict future labels, which could improve native estimates of future adoption times. Given the great number of

difficulties and unknowns, we leave this task for future research.

Beyond accounting for these limitations, possible extensions of the present method could include the analysis of the spreading of different items on the same population; or to classify different infection mechanisms [74, 81] even beyond pairwise exposures [150–152], as considered in a recent work [96]. Another potential direction for future research is to explore the competition between simple and complex contagions, where the adoption mechanism of a node is not predetermined but depends on the circumstances [153]. One could also integrate homophily, and conceive a model in which nodes of a same group are more likely to adopt through the same contagion process or nodes within the same mechanism have a higher probability to form connections between themselves [154]. Such a scenario would create correlated inferences, potentially affecting the accuracy of the classification.

We believe that our results open the door to the investigation of microscopic social contagion mechanisms at the local network level. In one way, our study aims to contribute to the understanding of how seemingly similar macroscopic processes can be differentiated at the microscopic level. In another way, we hope to lay down a path to study social contagion processes at the level of individuals, that is more feasible from a real data perspective and can lead us to a more fine-grained understanding of how local decision mechanisms lead to system level global phenomena in social contagion processes.

In the work presented in this chapter, I contributed by designing the likelihood method and the machine learning approach with my collaborators, making the synthetic propagations, running the five experiments and exploring the results and finding ways to improve them with my collaborators.

Chapter 4

Competition between simple and complex contagion on temporal networks

4.1 Introduction

Following the previous study on the distinguishability between the simple and the complex contagions in static networks, this chapter extends the analysis to temporal networks, which offer a more realistic framework for modelling human interactions. The addition of a temporal component changes the approach. Indeed, the time-varying nature of social interactions influences heavily the contagion dynamics, as it determines the time-respecting paths along which information or influence can be disseminated. The lifetime of links [155], the frequency of interactions [85, 156], the limited waiting times of processes at nodes [157], the causally related adjacent events [158], the memory length of influence [47, 88, 159], or the heterogeneous bursty interaction dynamics [47, 97, 156, 159] all have been identified important to influence spreading processes on temporal networks.

As it has been recognised in Chapter 3, both the simple and the complex contagions are present in one single spreading process. Several studies follow this approach: in some of them, contagion mechanisms vary across network layers or community structures [154, 160], while other researches assign individual thresholds to nodes, specifying the number of exposures needed for adoption [85, 104, 105, 161, 162]. Across all these works, the contagion mechanism of a node is always assigned beforehand. The infectiousness of the spreading process and the underlying network structure often determine which mechanism dominates [163]. In a case of a mixture of simple or complex adoption cases, the easier contracted simple contagion cases govern the spreading initially and they trigger in turn the complex contagion adoptions that are conditional to a fraction of peers to be infected a priori [105].

However, in real-world scenarios, individuals are not limited to a single contagion mechanism: they might adopt behaviours through both simple and complex contagions depending on their social context. For example, adoption might occur from a single intimate contact (simple contagion) or through repeated interactions with distant acquaintances (complex contagion). Therefore, considering the interplay between these mechanisms is pivotal to approximating more realistic social phenomena. In this chapter, we propose a new approach to identify whether simple, complex, or mixed contagion processes dominate at different stages of spreading dynamics in temporal networks. We classify processes into three categories and introduce methods to analyse transitions between them based on various parameters. In our simulations an individual is not endowed with a pre-assigned contagion mechanism but, depending on the peers involved, they could adopt via both simple or complex contagions.

4.2 Model definition

We simulate the dynamics of time-varying interactions by building on the Activity-Driven network (ADN) model, since it provides a flexible modelling framework of temporal networks. The network evolves through asynchronous iterations of N microscopic time steps of duration $\Delta t = 1/N$, in which each node is updated once on average.

To model the binary state contagion process [68] on the top of an ADN model, we assume as previously that each node at any time step can be in one of the two mutually exclusive states: susceptible (S), or infected (I). We set all nodes initially as susceptible and start the spreading from a small set of randomly selected infected seed nodes. Once a susceptible node is contaminated during the spreading, it remains in this state, thus the system is evolving towards a fully infected absorbing state. Susceptible nodes can become infected in two ways: either get contaminated through the simple contagion mechanism, or through the complex contagion mechanism.

We schematically summarize the network dynamics and the two contagion mechanisms in Fig. 4.1. We initially infect a randomly selected 1% of seed nodes. Subsequently, at each time step, if an activated node is susceptible, it will follow the simple contagion mechanism with probability p , and the complex contagion one with probability $1 - p$. In the former case, the node may get infected with probability β from any of its infected neighbours. Otherwise, it gets infected if the fraction of the infected neighbours exceeds the threshold ϕ . In practical terms, if n_i is the number of infected nearest neighbours of the susceptible node i , complex contagion takes place if $n_i \geq z$, where $z = \lceil m\phi \rceil$. These steps are repeated until every node is in the infected state.

In the following, we denote by $I(t)$ the number of infected nodes at time t and by $\rho(t) = I(t)/N$ (ρ for brevity) the fraction of infected nodes (prevalence), while ρ_0 being this quantity at the start of the propagation. Likewise, ρ_s and ρ_c indicate the fractions of nodes

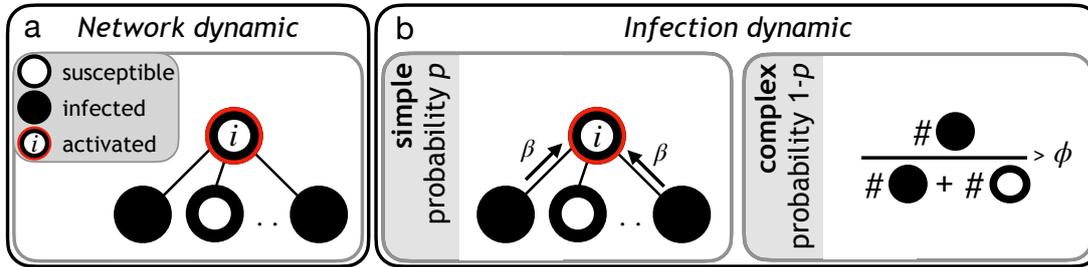


Figure 4.1: Network and contagion dynamics. (a) A node i is activated with probability a_i and connects to m randomly selected nodes. If the activated node i is susceptible, with probability p it follows the simple contagion mechanism (b) and gets infected by each infected neighbor with probability β . With probability $(1 - p)$ it follows the complex contagion rule (c) and gets infected if the fraction of its infected neighbors is above ϕ .

infected by the simple or complex contagion, respectively. To simplify the mathematical description, we formulate some assumptions about the propagation process. We model a homogeneous network dynamics by assuming that all nodes have the same activity a . We consider the heterogeneous case in Section 4.4. We also consider that all nodes share the same parameters β and ϕ . Further, we study the cases of $z = 2$ and $z = 3$. The analytical treatment for more general cases with any value of z is reported in the Appendix C.1.

4.3 Analytical study

To shed analytical light on the model, we consider its description in terms of a mean-field rate equation, inspired by Ref. [61]. We consider the time evolution of the number of infected individuals $I(t)$. In a microscopic time step Δt , we choose a node at random (the ego node), which is susceptible with probability $(N - I)/N$. The ego node becomes active with probability a , in which case with probability p it follows a simple contagion, and with probability $1 - p$ complex contagion to potentially become infected. Therefore, within a mean-field approximation [61], the number of infected nodes at time $t + \Delta t$ can be written as

$$I(t + \Delta t) = I(t) + \frac{N - I(t)}{N} a [p \Delta_{m,\beta}(\rho) + (1 - p) P_z(\rho)], \quad (4.1)$$

where $\Delta_{m,\beta}(\rho)$ and $P_z(\rho)$ are the probabilities that the focal node becomes infected by the simple and complex processes, respectively. For the simple contagion process, the ego node is connected to m other nodes, each one infected with probability ρ . Each infected neighbour transmits the infection to the ego node with probability β . Thus, the probability that any one

of the infected neighbours infects the ego is

$$\Delta_{m,\beta}(\rho) = 1 - (1 - \rho\beta)^m. \quad (4.2)$$

On the other hand, for complex contagion to take place, we need the ego to be neighbour of at least z infected nodes, an event that happens with probability

$$P_z(\rho) = \sum_{n=z}^m \binom{m}{n} \rho^n (1 - \rho)^{m-n} \equiv I_\rho(z, m + 1 - z), \quad (4.3)$$

where $I_x(a, b)$ is the regularized incomplete beta function [164].

Assuming that the time interval $\Delta t = 1/N$, in such a way that a whole update of the network corresponds to one Monte Carlo time step¹, we can take the thermodynamic limit $N \rightarrow \infty$ in Eq. (4.1) to write the differential rate equation

$$\frac{d\rho}{dt} = a(1 - \rho) [p \times \Delta_{m,\beta}(\rho) + (1 - p) \times P_z(\rho)]. \quad (4.4)$$

In the following, we analyse the beginning of the propagation process, in the limit $t \rightarrow 0$, $\rho \ll 1$, considering different cases.

4.3.1 Pure complex contagion

We first consider the case where only complex contagion is present, corresponding to $p = 0$. In this scenario, using Eqs (4.4) and (4.3), we obtain

$$\frac{d\rho}{dt} = a(1 - \rho) I_\rho(z, m + 1 - z). \quad (4.5)$$

Since we are interested in the behaviour for small t and ρ , we can use the power expansion of the regularized incomplete beta function for integer z [164]

$$I_\rho(z, m + 1 - z) \sim \rho^z \binom{m}{z}, \quad \rho \rightarrow 0. \quad (4.6)$$

Thus, keeping only the leading terms, Eq. (4.5) can be written as

$$\frac{d\rho}{dt} = C\rho^z \quad (4.7)$$

¹In particle physics, a Monte Carlo time step is the time it takes for each particle in a system, on average, to have one chance to change state in a simulation.

where the constant C is equal to

$$C = a \binom{m}{z} \quad (4.8)$$

The solution of this equation, in terms of the initial density of infected seeds ρ_0 , is

$$\rho^{z-1}(t) = \frac{1}{\rho_0^{1-z} - (z-1)Ct}. \quad (4.9)$$

This solution shows a linear decreasing behaviour in time of the function $1/\rho^{z-1}$, with a divergence at a time $t = \rho_0^{1-z}/[C(z-1)]$. The time of the divergence is smaller when ρ_0 is higher, as initially there are more infected nodes, and when z is smaller, since the condition to be infected is easily reached.

This divergence is non-physical since it appears from a lower order approximation; adding other terms in the full equation curbs the divergence and leads to a prevalence $\rho(t) \leq 1$. However, we can interpret this apparent divergence in opposition to the behaviour of pure simple contagion, in which prevalence grows exponentially in the linear approximation [165]. Complex contagion operates instead in cascades, in which a large fraction of nodes become infected in a very short period of time [17]. Thus, we can identify the divergence time as the onset of the cascading behaviour of complex contagion.

4.3.2 Mixed simple and complex contagion

We consider now the case of mixed simple and complex contagions. To simplify our calculations, we focus here on the case of $z = 2$. We thus have, for the simple contagion infection probability,

$$\Delta_{m,\beta}(\rho) = 1 - (1 - \rho\beta)^m \simeq m\beta\rho - \rho^2 m(m-1) \frac{\beta^2}{2}, \quad (4.10)$$

where we have kept only the lower order terms in ρ . For the complex contagion probability, from Eq. (4.6),

$$P_2(\rho) \simeq \rho^2 \binom{m}{2} = \rho^2 \frac{m(m-1)}{2}, \quad (4.11)$$

where again we have kept only the leading terms in ρ . Inserting Eqs. (4.10) and (4.11) in Eq. (4.4), we obtain

$$\frac{d\rho}{dt} = a\rho m\beta(\rho + B\rho^2) \quad (4.12)$$

up to order ρ^2 , and where we have defined the constant

$$B = \frac{(1-p)(m-1)}{2p\beta} - \frac{(m-1)\beta}{2} - 1. \quad (4.13)$$

The solution of Eq. (4.12) in terms of the initial density of infected nodes is

$$\rho(t) = \frac{1}{Ae^{-tapm\beta} - B}, \quad (4.14)$$

where we have defined the constant

$$A = B + \frac{1}{\rho_0}. \quad (4.15)$$

The sign of B determines the behaviour of the prevalence. If B is negative (and ρ_0 sufficiently small in such a way that $A > 0$), then $\rho(t)$ grows at short times, until it saturates to the value $1/|B|$. If B is positive, on the other hand, $\rho(t)$ diverges at the time

$$t_{\text{casc}} = \frac{1}{apm\beta} \ln \left(\frac{A}{B} \right), \quad (4.16)$$

which serves as a proxy to indicate the time of the cascade, namely the time of a sudden increase of the density of infected nodes.

These different behaviours, depending on the sign of B can be understood as follows. For $B < 0$, the second term on the r.h.s of Eq. (4.12) is negative, in agreement with the second order expansion of the rate equation for pure simple contagion (see Eq.(4.10)). Otherwise, if $B > 0$, the second term on the r.h.s of Eq. (4.12) is positive, as corresponding to the pure complex contagion rule (see Eq. (4.11)). We can interpret the value $B = 0$, corresponding to the probability

$$p_c(\beta) = \frac{m-1}{2\beta + (1+\beta^2)(m-1)}, \quad (4.17)$$

as the boundary separating a dominating simple contagion (for $p > p_c$) from a dominating complex contagion (for $p < p_c$). Within the complex contagion dominated phase, the divergence time t_{casc} is a proxy of the time at which complex contagion takes over from the simple contagion prevalent at very short times.

The divergence time t_{casc} diminishes as ρ_0 increases, since the increase of the proportion of infected nodes occurs earlier when there are initially more infected nodes. The parameter β also makes the critical time decreasing when it is increasing: as there is a higher probability to be infected for the simple contagion, the number of infected nodes is greater, thus the increase of ρ is earlier.

In the complex contagion dominated region with $B > 0$, another way to estimate when this dynamics takes over simple contagion emerges from the analysis of Eq. (4.12). In this equation, we have a linear term, arising from simple contagion, and a quadratic term, with components from the complex and simple contagions. Assuming that complex contagion dominates over simple contagion when the second order term becomes larger than the linear one, we can define the threshold density of infected nodes (see Appendix C.1 for the expression of ρ_{eq} in the general case $z > 2$)

$$\rho_{eq} = \frac{1}{B}, \quad (4.18)$$

such that at the time t_{eq} , corresponding to $\rho(t_{eq}) = \rho_{eq}$, the second order term overcomes the first order one, and we expect complex contagion to prevail. From the solution Eq. (4.14), we can estimate, within small time approximation,

$$t_{eq} = \frac{1}{apm\beta} \ln\left(\frac{A}{2B}\right) \quad (4.19)$$

such that $t_{eq} < t_{casc}$. In the case that $\rho_0 > \rho_{eq}$, the quadratic term containing elements from the complex and simple contagions dominates from the initial instant of the dynamics. This scenario corresponds to $t_{eq} = 0$ and marks the region below which the complex contagion is fully predominant from the early stage of the dynamics.

To sum up the behaviour of $z = 2$, for $p > p_c(\beta)$, contagion is dominated by the simple mechanism. For $p < p_c(\beta)$, contagion is dominated, at large times by the complex mechanism. At short times, however, simple contagion is prevalent, since it is a larger average transmission. At the time t_{eq} , complex contagion takes over from the initially predominant simple mechanism.

4.4 Numerical simulations

We contrast our analytical results with numerical simulations, with parameters $N = 1000$, $a = 1$, $m = 5$ and $\phi = 0.25$ for $z = 2$, and $\phi = 0.5$ for $z = 3$. We average simulations over 100 runs. We fix $z = 2$ (if not noted otherwise) while varying β and p .

Figure 4.2 shows the contagion curves for different β and p (see Figure C.1 in the Appendix C.2 for the full parameter space). When p is large, dynamics is determined by simple contagion, as we can see by the fast growth of the prevalence, which is faster for larger values of β , see Figure 4.2a and 4.2b. When p is small, on the other hand, the complex contagion prevails, see Figure 4.2c-d. In this regime, the prevalence grows initially slowly, followed by a sudden increase indicative of cascading. Interestingly, as shown in

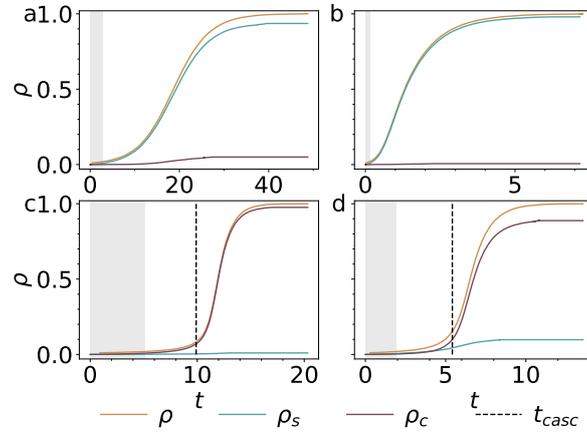


Figure 4.2: Fraction of infected nodes ρ (yellow line), proportion of nodes infected by simple ρ_s (blue) and complex ρ_c (purple) contagion as a function of time, for $z = 2$. Panels show different (β, p) parameters: a) $(0.05, 0.99)$, b) $(0.99, 0.99)$, c) $(0.05, 0.05)$, and d) $(0.99, 0.05)$. The simple (complex) contagion dominates the propagation when p is high (low), with a minor influence from β which can be observed on Figure C.1. The grey area indicates the early period of the contagion up to t_{init} , when $\rho = 0.02$. Dashed vertical lines indicate t_{casc} , corresponding to the expected outbreak time if B is positive (see Eq. (4.16)), in panels c) and d).

Figure 4.2d, the ρ_c cascade emerges earlier when rare ($p = 0.05$) but likely transmitting ($\beta = 0.99$) simple contagion events are present. Since the initial seeds may not be eligible to induce a complex contagion outbreak, simple contagion cases build up the necessary initial conditions to trigger complex contagion.

Figure 4.2 and Figure C.1 (Appendix C.2) show that the analytical divergence time t_{casc} generally predicts quite closely the time when the infection curve $\rho(t)$ starts taking off, signalling the onset of cascades. Nevertheless, t_{casc} slightly underestimates the real outbreak time if both β and p have high values, while t_{casc} is somewhat late if β is low and p is high.

Further, we note here that a non-constant activity distribution can affect the propagation speed. In Figure 4.3, we aim to compare the results of two activity distributions: a dirac distribution (every node has the same activity, used in the main manuscript) and a power law distribution. Both distributions have the same average of 0.1, thus the exponent of the power law is 1.14. We observe that the simulations made with the dirac distribution are at first slower than the ones made with a power law distribution, but reach faster the final state (Figures 4.3, panels a to d). Indeed, as some nodes have a high activity with the power-law distribution, they are infected first and accelerate the beginning of the process. However, as the majority nodes have a low activity, the process is slow to reach the contagion of all the nodes.

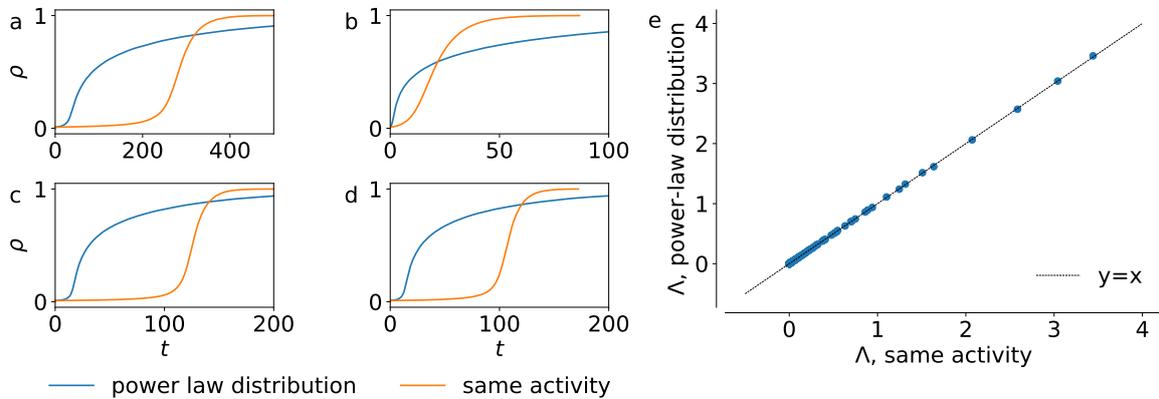


Figure 4.3: Comparison of the contagion curves (panels a to d) and the final ratio Λ (panel e) when every node has the same activity and when the activities are sampled from a power law. Both distributions have the same average, i.e., 0.1. The exponent of the power law distribution is 1.14. The simulations on the left part are parameterised with (β, p) : a - (0.05, 0.99), b - (0.99, 0.99), c - (0.05, 0.05), and d - (0.99, 0.05). The propagation is considerably slower when the nodes have the same activity, but reaches faster the final state. On panel e, every scatter point stands for a couple of parameters (β, p) . The final ratio Λ is identical for the two activity distributions.

Following this qualitative analysis, we focus now on disentangling the effective mechanisms that rule the evolution of the spreading process. As we have mentioned above, the dominant mechanism can change during the process in the mixed simple/complex scenario. We thus consider the early and late stages of the propagation separately. The early stage corresponds up to the time t_{init} , at which the prevalence ρ fulfils the condition $\rho = 0.02$, while the late stage encompasses from t_{init} to the time when all nodes are infected, t_{end} . To identify the early and late contagion mechanisms we consider the ratios $\lambda = \rho_c(t_{init})/\rho_s(t_{init})$ and $\Lambda = \rho_c(t_{end})/\rho_s(t_{end})$. It is worth noticing that the outcome Λ of the two compared activity distributions on Figure 4.3 (the dirac distribution and the power law distribution) are identical: the final ratio between the number of nodes infected by the simple and the complex contagions is the same (Figure 4.3e). If the simple (complex) contagion dominates the entire process, both quantities remain below (above) 1. If the simple contagion dominates the early stage but the complex one takes over in the late stage, then we expect $\Lambda > 1$ and $\lambda < 1$ respectively.

Since the case $\Lambda < 1$ and $\lambda > 1$ cannot be observed (once the complex contagion is triggered, it propagates much faster than the simple one), we classify the spreading dynamics into three categories: pure simple contagion ($\Lambda < 1$ and $\lambda < 1$), pure complex contagion ($\Lambda > 1$ and $\lambda > 1$), and mixed contagion ($\Lambda > 1$ and $\lambda < 1$). We thus expect two transitions in the parameter space (β, p) . The first, pure simple to mixed contagion, and the second from mixed

to pure complex. The first transition takes place at a threshold $p_c(\beta)$, given by Eq. (4.17) for $z = 2$, separating the phase in which simple contagion is always dominating from the phase in which complex contagion dominates at large times. In the mixed phase, while complex contagion is dominant at large times, at short times simple contagion is prevalent. The second transition separates this mixed phase from the pure complex contagion phase, in which even at short times complex contagion is predominant.

In the following, we propose two different methods to identify the transition point from mixed to pure complex phases.

Method 1: In the complex dominated phase, we measure the time t_{eq} and prevalence $\rho_{eq} = \rho(t_{eq})$ when the two terms in the rhs of Eq. (4.12) become equal, indicating when the simple contagion term takes over the complex contagion one. These quantities can be computed analytically from Eqs. (4.18) and (4.19), and numerically by evaluating in simulations when the first and second terms in the rhs of Eq. (4.12) become equal. The time t_{eq} signals thus the transition between a dynamics initially dominated by simple contagion and the dynamics asymptotically dominated by complex contagion. In the parameter space (β, p) , the transition from the mixed phase to the pure complex contagion phase should thus correspond to $t_{eq}(\beta, p) = 0$ (noted t_{eq} null in the following figures), that is, when right at the initial time step the contagion is dominated by the complex mechanism.

Method 2: The second approach relies on the results in Eq. (4.9), suggesting that the function $1/\rho^{z-1}$ should decrease linearly with time t , if the spreading is governed by pure complex contagion. We demonstrate this behaviour in Figure 4.4a and b (for $z = 2$ and $z = 3$ respectively) by showing the curve $1/\rho^{z-1}$ for simulations corresponding to the (β, p) pairs indicated by the black points in Figure 4.4c and d respectively. In the case where both simple and complex contagions are present, looking at the short time behaviour of $1/\rho^{z-1}$ is also linear with t , see Eq. (4.14), as it can be seen expanding the exponential in the denominator regardless the sign of B . However, this linear trend breaks down earlier, since the expansion of the exponential fails sooner. As a result, the complex contagion dominates longer, as marked by a longer linear decrease. To quantify this effect, we first consider an initial regime up to small initial time t_{init} , in which we assume a linear behavior for $1/\rho^{z-1}$. To take equally different values of z , we define t_{init} as the time when ρ satisfied the condition $\frac{1}{\rho^{z-1}} / \frac{1}{\rho_0^{z-1}} = 0.5$ (if $z = 2$ and $\rho_0 = 0.01$, this condition corresponds to $\rho = 0.02$). We then fit a linear function to $1/\rho(t)^{z-1}$ in the interval $[0, t_{init}]$. Finally, we find the value t_{lim} at which the linear behavior breaks down. This is defined by the time at which the function $1/\rho(t)^{z-1}$ differs from the initial linear fit by a value larger or equal than $\varepsilon = 0.05/\rho_0^{z-1}$. The prevalence at this time is denoted as $\rho_{lim} = \rho(t_{lim})$.

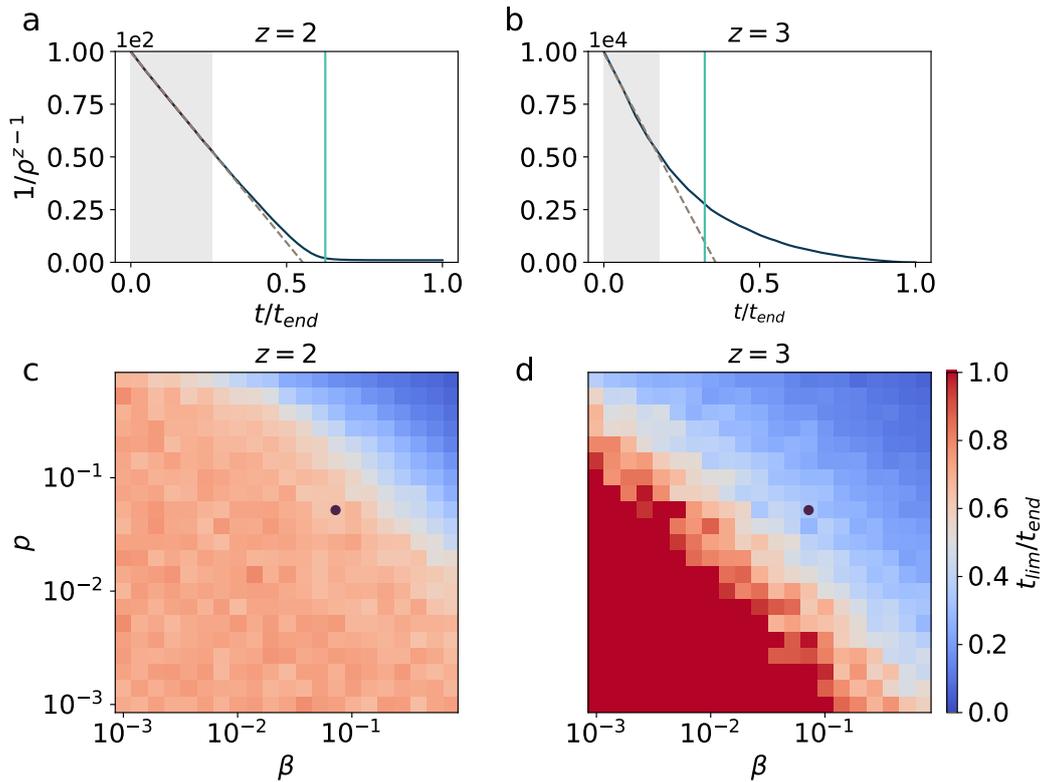


Figure 4.4: The $1/\rho^{z-1}$ curve (dark blue line) obtained from simulation averages with $(\beta, p) = (0.037, 0.037)$, fitted by a linear function (black dotted line) on the period before t_{init} (grey area) for $z = 2$ (panel a) and $z = 3$ (panel b). The t_{lim} time point when the two curves diverge is corresponding to an $\epsilon > 0.15/\rho_0^{z-1}$ difference, indicated by a vertical light blue line. The different t_{lim} obtained across the whole parameter-space (β, p) are displayed for $z = 2$ (panel c) and $z = 3$ (panel d). $1/\rho^{z-1}$ stops being linear at early times when the propagation is fast i.e., when both β and p are high, while this linearity persists longer when the contagion process is slow (low β , and high p for $z = 2$). Simulations results were calculated as averages over 100 realisations.

In Figure 4.4 we show an example of the application of Method 2. Indeed, as Figure 4.4a depicts, t_{lim} takes a larger value as the $1/\rho^{z-1}$ curve scales linearly over an extended regime, suggesting dominating complex contagion. On the contrary, in Figure 4.4b, the simulations are dominated by simple contagion, thus t_{lim} is small as a shorter linear scaling is observable. We display the rescaled time t_{lim}/t_{end} for an extended parameter space in Figures 4.4c and Figures 4.4d for z equal to 2 and 3, respectively. According to these results our method can well separate the regime where the complex contagion dominates, characterised by large t_{lim}/t_{end} values and corresponding to lower values of β and p , from the simple contagion dominated regime, with small values of t_{lim}/t_{end} and for higher values of β and p .

Fig. 4.5 shows how the analytical and numerical methods capture the transitions from pure simple contagion to mixed, and from mixed to pure complex, in the (β, p) parameter space. Figs. 4.5a and 4.5c show a phase diagram depicting the different regions in the (β, p) space for $z = 2$ and $z = 3$, respectively. The regions are determined by comparison of the values λ and Λ defined above: pure simple contagion for $\Lambda, \lambda < 1$, pure complex contagion for $\Lambda, \lambda > 1$, and mixed contagion for $\Lambda > 1, \lambda < 1$. For $z = 2$, the phase boundary between pure complex (brown area) and mixed (pink area) is well captured by Method 1, corresponding to a null t_{eq} line in red. The boundary between mixed and pure simple (green area) is slightly shifted from the prediction $p_c(\beta)$ (green line). In the case of $z = 3$, again Method 1 provides an excellent approximation for the boundary between pure complex/mixed phases, which seems to take place at smaller values of (β, p) than for $z = 2$. In this case, we do not have an analytical prediction for the boundary mixed/pure simple, leaving us alone with the results of numerical simulations. However, it seems to appear again for smaller (β, p) values.

In Figs. 4.5b and 4.5d, we present the predictions by Method 2 for the location of the boundaries between pure complex/mixed phases for the cases of $z = 2$ and $z = 3$, respectively. In these plots we depict as color maps the density ρ_{lim} for the whole range of the (β, p) parameter space. As we can see, the collapse of the boundary extracted from Method 2 with the classification made in terms of Λ and λ ratios is excellent, while also being in very good agreement with the prediction of Method 1 represented by t_{eq} null.

4.5 Conclusion

In this chapter, we introduced a mixed model of social contagion on temporal networks, in which nodes can be infected by either simple or complex contagion dynamics simultaneously. We focused on the simplest modelling scenario of a homogeneous Activity-Driven network with all nodes having the same activity. We tackled the solution of the model analytically using a mean-field rate equation for the total density of infected nodes as a function of time, and numerically through extensive simulations. This way we proposed two ways to differentiate between processes of different dominant contagion processes.

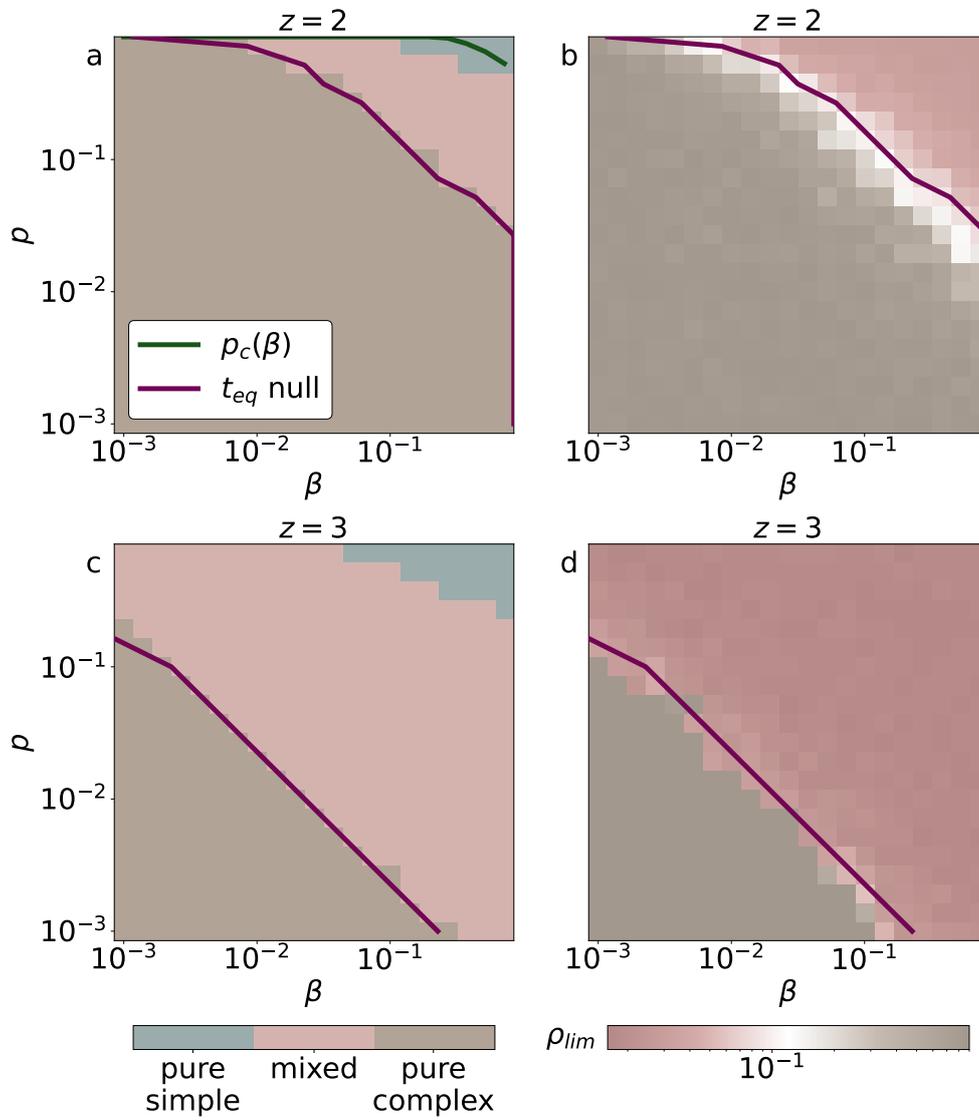


Figure 4.5: Areas corresponding to the three categories in the parameter-space (β, ρ) on panel a for $z = 2$ and on panel b for $z = 3$. The quantities $p_c(\beta)$ and the limit of t_{eq} null, indicated by a green and a purple line respectively, match the transitions between those areas. The values of ρ_{lim} , displayed in the second column, are also marked by the transition between pure complex and mixed.

We identified three phases of contagion according to the two main parameters of the model, the probability of adopting by simple contagion p and the infection probability of simple contagion β . For small (large) values of (β, p) , the dynamics is ruled purely by simple (complex) contagion. For intermediate values of (β, p) , instead, the dynamics is mixed: initially dominated by simple contagion but following complex contagion at large times. Finally, for large (β, p) parameter values the spreading is dominated by the simple contagion mechanism. We proposed a criterion to determine the phases of the system, confirmed by analytical expressions for one of the boundaries and by numerical methods for all of them.

Future work should be dedicated to extending the analytical study to larger z values, to find a critical adoption probability p_c separating simple and mixed contagion in these cases. Furthermore, we explored the effect of heterogeneity of nodes —with respect to their activity rate— only qualitatively, by comparing contagion curve profiles between constant and power-law activity distributions. Extending the analytical and numerical frameworks presented here to the case of heterogeneous networks represents a significant avenue for future research.

We believe that our results open the door to the investigation of mixed-propagation in which a same node can adopt by both the simple and the complex contagion mechanisms. Our aim is to contribute to the understanding of the spreading of social behaviour taking into account the differences between individuals.

In the work presented in this chapter, I contributed by designing the research question and the model, conducting the mathematical analysis with my collaborators, making the simulations and exploring the results with my collaborators.

Chapter 5

Conclusion

Understanding the underlying mechanisms of spreading processes is crucial for explaining how social behaviours propagate in our societies. Temporal networks, which reflect the dynamic nature of social interactions, provide an accurate framework to study processes arising on top of them. Since these evolving structures influence propagation, the coupled analysis of the interactions and the process dynamics is necessary to understand the spreading phenomena. Especially, many temporal characters, such as daily routines, exhibit cyclic patterns that are often difficult to detect when they are represented by temporal networks. In Chapter 2, we proposed a method to measure the periodic time scales of these temporal objects that we applied to assess periodic changes in link density and group structures. We proved the efficiency of the methods by applying them to synthetic and real networks. We then shifted focus to social behaviour propagation and we aimed to distinguish between the spontaneous, simple and complex contagion mechanisms. Many studies assume that only one single contagion mechanism occurs in a whole spreading process, but in reality different ones can coexist depending on the individuals. In Chapter 3, we addressed this classification problem on the microscopic level, using likelihood and random forest approaches that we tested on synthetic simulations and real-world data. Finally, Chapter 4 extended this analysis to temporal networks, categorizing the simulations from a mixed propagation model into pure simple, pure complex, or mixed dynamics and identifying the dominant contagion type at the macroscopic scale through different methods.

The limitations of this work are very similar to those encountered in complex systems. Studying spreading processes on top of networks especially involves making several assumptions. While it is certain that interactions between individuals play a role in the spreading of behaviours, it is a significant assumption to claim that every exchange is about the adoption behaviour. In reality, there is a competition on every different topic two individuals can discuss, which creates uncertainty about the type of content shared during their interactions. Moreover, it is difficult to distinguish if the spreading is caused by social influence

or other factors like homophily. The Activity-Driven Model also assumes that any node can interact with any other node if they are randomly selected, which is unrealistic as people tend to interact in limited social circles.

Another challenge faced in this work is about the availability of data needed to study those dynamical processes. In Chapter 3, we assumed that all the local information of the spreading of an ego-network is known. In reality, it is difficult to be exhaustive in enumerating every social interaction and understanding the exact time of infection of the neighbours. Even on online platforms like Twitter, the underlying algorithms remain unknown meaning that we cannot be certain about every interaction since users are shown selected posts. For those reasons, getting a complete data set of a whole propagation process is even harder, but would be useful to apply the methods of Chapter 4. Additionally, the size of the contagion data sets can be significantly large, which can be problematic for running the algorithms efficiently. In Chapters 2, we assumed a complete knowledge of the temporal network. However, it is complicated to gather data over a long period of time, unless dealing with specific scheduled events like airplane flights.

Finally, the analytical methods used in the diverse chapters of this work come with their own limitations. Several simplifying assumptions were needed to be able to calculate the equations, which made the results less accurate.

This work opens the door to studying propagation processes at the individual level, allowing nodes to adopt through different mechanisms. This approach provides a more realistic view by reflecting the differences between individuals and leads to more accuracy in the modellings of propagation as long as a better understanding of how humans influence each other. I hope this thesis will encourage more research in this direction, potentially uncovering new phenomena.

Similarly, in the same way we combine spreading processes with temporal networks, future work could extend these approaches to cumulate in addition to the temporal component other types of networks, such as directed networks or hyper-graphs. This would not only increase the complexity of the propagation models but also make them more realistic.

Moreover, this work could be applied to more real-world datasets, particularly for the studies presented in Chapter 3 and Chapter 4, as both lack ground-truth data for which the contagion process at stake is known. It would be worth considering the creation of datasets specifically designed to distinguish between the simple and the complex contagion processes. The COVID-19 pandemic has been extensively studied, resulting in the creation of large datasets in epidemiology, and it is promising that we may also see a rise of datasets on social behaviours, especially nowadays when information spreads more widely and quickly than before.

To conclude, this study offers a perspective on both temporal networks and social contagion, enhancing our understanding of spreading processes without prior assumptions about individual behaviours.

Appendix A

Detecting periodic time scales of changes in temporal networks

A.1 Sensitivity analysis: size and length of the temporal network, sliding window parameters

To evaluate the reliability of our results, we perform a sensitivity analysis on the parameters of the experiments *Change of activity* and *Change of grouping*. We change one parameter in both experiments while keeping the other constant ($N = 100$, $\varepsilon = 0.001$, $\eta = 4$, γ oscillating between 1.8 and 2.8 for the *Change of activity* and $\gamma = 2.8$ for *Change of grouping*, $T_a = T_g = 100$, $|T|$ is adjusted to have 12 periods, $\Delta t_w = 5$ and $t_w = 2$).

When we vary the number of nodes of the networks (Figure A.1), the original period is almost always measured properly except when the number of nodes is high. In that case, the measured period is 50 (corresponding to a frequency of 0.02), which is half of the original period. In fact, when the networks change from a high-activity state to a low-activity state, we observe a peak in the dissimilarity function. This situation happens twice in a period: when changing from low to high and from high to low activity. The measured period is then the half period (Note that this happens at all sizes, and the half-period is indeed always recovered as one of the harmonics, but it seems here to become dominant at large sizes).

The same analysis has been realized by changing the length of the period (Figure A.2) and in every case, the original period is correctly measured: the length of the period does not influence the observation. Instead, when we change the number of periods observed (Figure A.3), we observe that we need a minimum of 2 periods to measure the original time scale. Finally, changing the exponents in the *Change of activity* experiment does not have any effect on the measure of the correct time scale (Figure A.4) but they must obviously be

different, which is necessary to have a period network.

We finally study the influence of the parameters of the sliding window t_w and Δt_w on the results. In Figure A.5, we compute the power spectra of the AD network with the SA-method and the EG-method having different parameters t_w and Δt_w . Indeed, in the case of empirical data sets, if no prior information is available on the orders of magnitude, one should explore a range of possible parameters and check the robustness of the extracted timescales. The correct period (100) is properly measured if t_w is not too large and if Δt_w has an intermediate value.

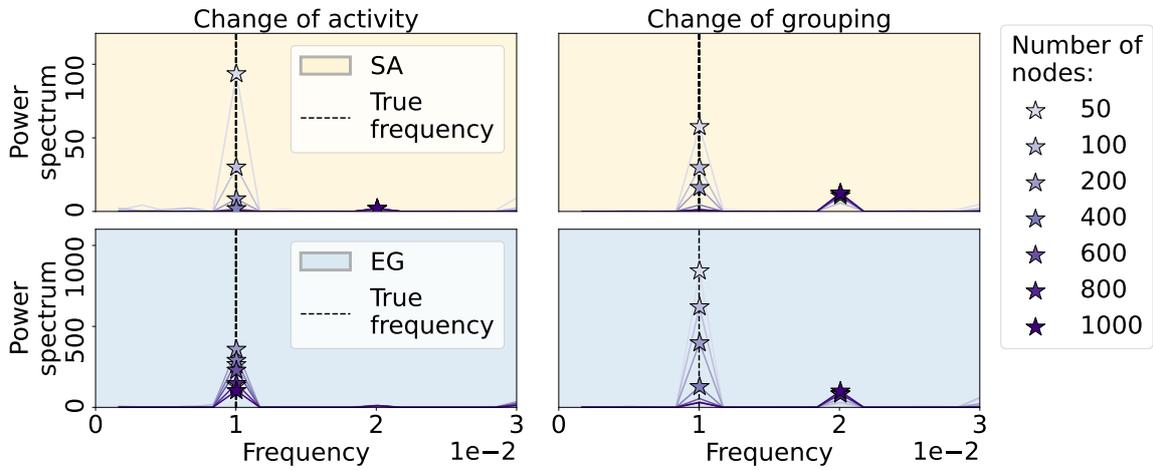


Figure A.1: Power spectra of the temporal network from the *Change of activity* experiment (left column) and the *Change of grouping* experiment (right column) measured from the SA-method (yellow background) and the EG-method (blue background). The number of nodes of the AD networks varies from 50 to 1000. The correct frequencies are indicated with vertical dashed lines. Those original periods are well-measured in the majority of the cases. However, when the number of nodes is too important, the method measures the semi-period.

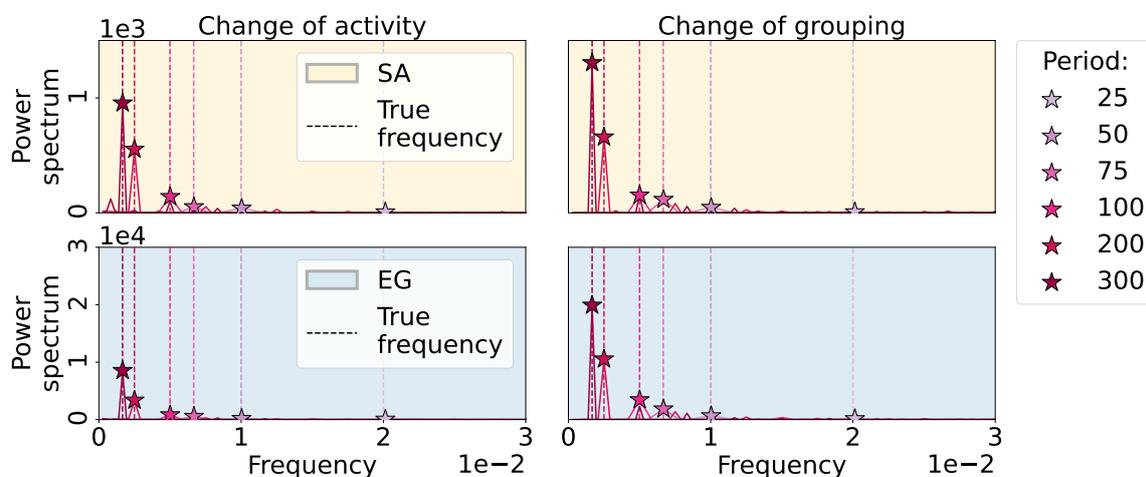


Figure A.2: Power spectra of the temporal network from the *Change of activity* experiment (left column) and the *Change of grouping* experiment (right column) measured from the SA-method (yellow background) and the EG-method (blue background). The period varies from 25 to 300. The correct frequencies are indicated with vertical dashed lines and are here always well recovered.

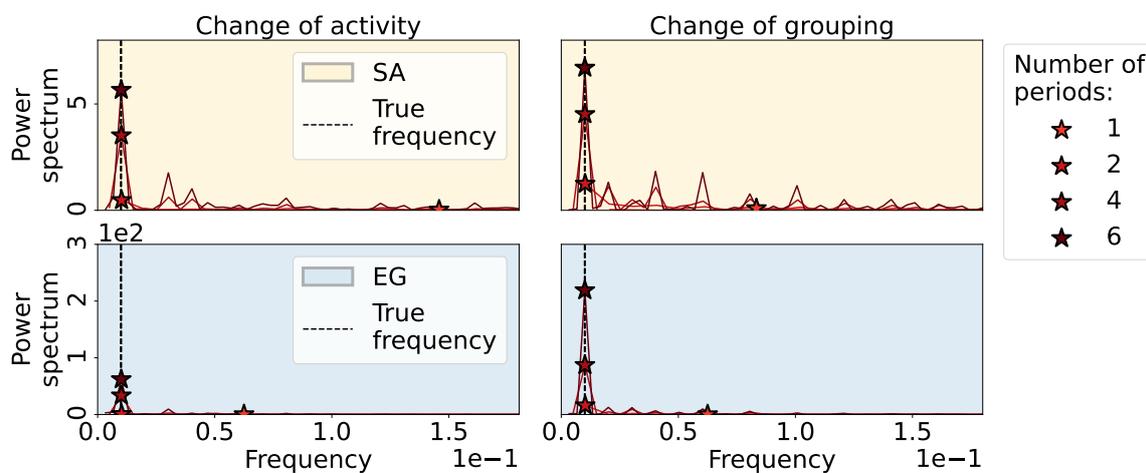


Figure A.3: Power spectra of the temporal network from the *Change of activity* experiment (left column) and the *Change of grouping* experiment (right column) measured from the SA-method (yellow background) and the EG-method (blue background). The number of periods of the AD networks observed during T varies from 1 to 6. The correct frequencies are indicated with vertical dashed lines. Those proper periods are well-measured as long as the data set contains at least two periods.

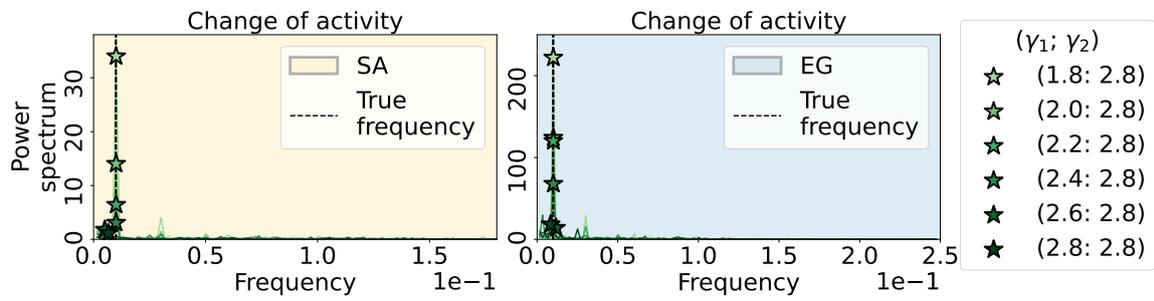


Figure A.4: Power spectra of the temporal network from the *Change of activity* experiment measured from the SA-method (left column, yellow background) and the EG-method (right column, blue background). The parameter γ_1 of the AD networks corresponding to periods of high activity varies from 1.8 to 2.8, while the other parameter γ_2 characterising the low-activity state remains constant equal to 2.8. The correct frequencies are indicated with vertical dashed lines. Those proper periods are well-measured as long as γ_1 is different than γ_2 .

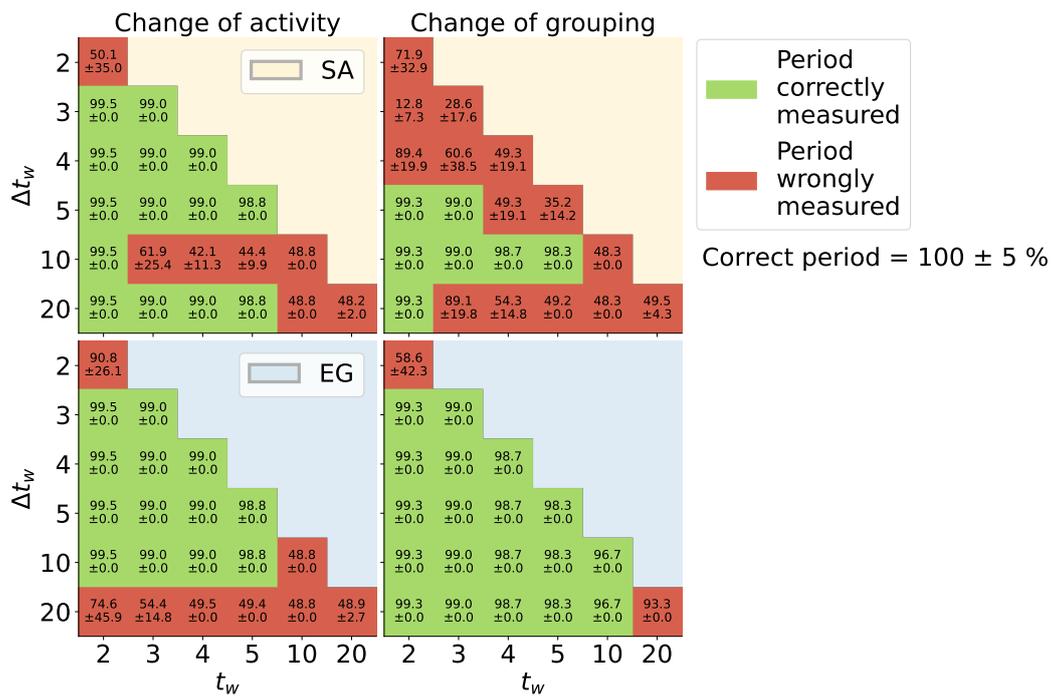


Figure A.5: Period measured through the SA-method (orange background) and the EG-method (blue background) for the *Change of activity* experiment (first column) and the *Change of grouping* experiment (second column). We change the parameters of the sliding window: the x-axis presents different values of the stride t_w and the y-axis different values of time-window lengths (Δt_w). The period of the initial networks is 100 and the results are averaged over 10 realisations. The results are displayed as the average period over the different realisations \pm the standard deviation.

Appendix B

Distinguishing mechanisms of social contagion from local network view

B.1 Distribution of the features of the random forest of Experiment 2

The features of the random forest have been chosen to present different values according to the mechanisms of adoption. As depicted in Figure B.1, the distributions of most features differ for the simple, complex and spontaneous adoptions. The degree is the only feature which is not related to the propagation itself but to the structure of the network. While it does not present significant differences in the distributions within Erdős-Rényi networks, we keep it due to the potential influence of a node's degree in other type of networks.

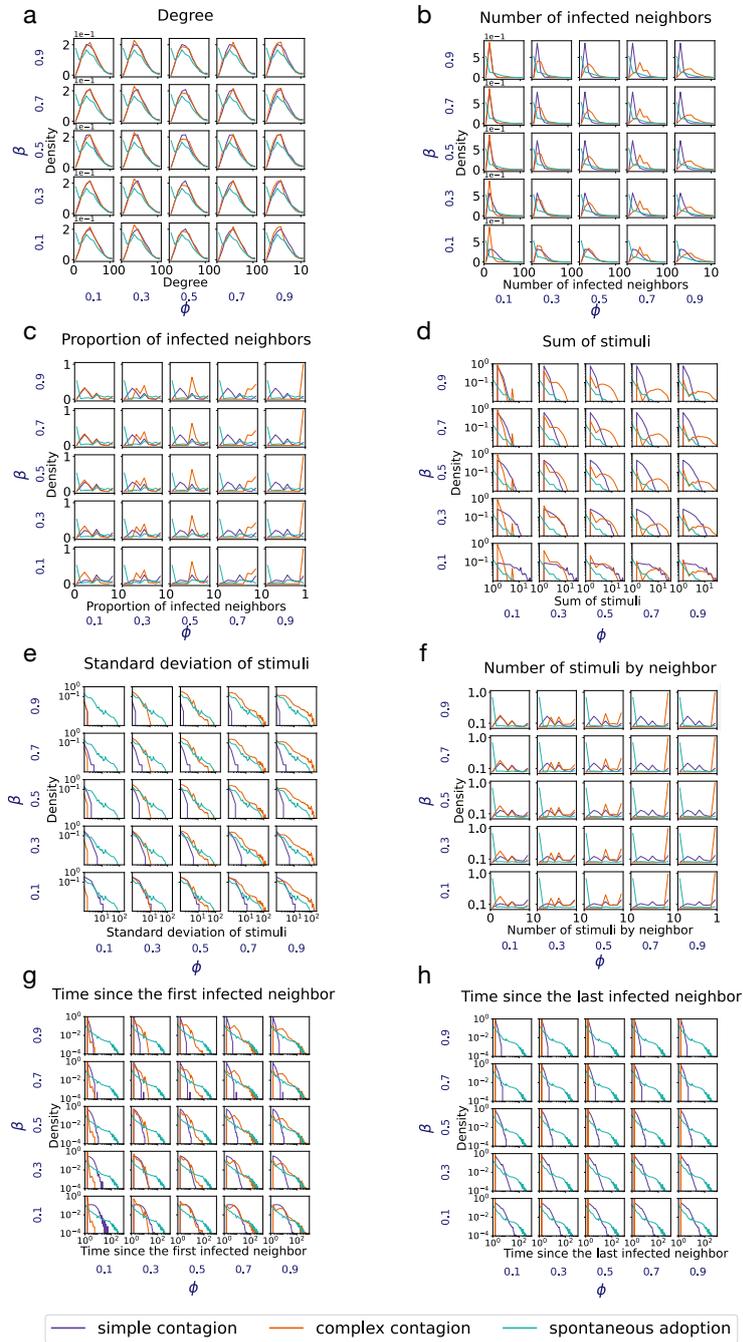


Figure B.1: Distribution of the features of the random forest algorithms across the parameter space: degree (panel a), number of infected neighbours (panel b), proportion of infected neighbours (panel c), sum of stimuli (panel d), standard deviation of stimuli (panel e), number of stimuli by neighbours (panel f), time since the first infected neighbour (panel g) and time since the last infected neighbour (panel h). The displayed values, taken from Experiment 2, are grouped by their dynamical processes (simple, complex or spontaneous).

B.2 Best subset of features for the random forest on Experiment 2

To evaluate the significance of the features of the random forest method on Experiments 2 and 4, we train algorithms with all possible subsets of the eight features. Figures B.2 and B.3 present the subset with the highest accuracy (y-axis) for each subset length (x-axis), respectively for Experiments 2 and 4, across the whole parameter-space, with corresponding accuracy values indicated in blue. Looking at the results from Experiment 2, enlarging the feature set from one to three increases the accuracy, but a plateau is reached for subsets larger than four. In other words, in most regions of the parameter space, only three features are necessary to achieve the same accuracy as with more features. However, this optimal subset varies through the parameter space. Also, adding features increases the accuracy when ϕ is high, but does not have any effect when the value of ϕ is small. Differently, the set of features for the classification of Experiment 4 does not have a great influence on the accuracies, as the obtained values for different lengths of subsets are very similar. In contrast, the selection of features for the classification in Experiment 4 has a limited influence on the accuracies, as the obtained values for different subset lengths are very similar.

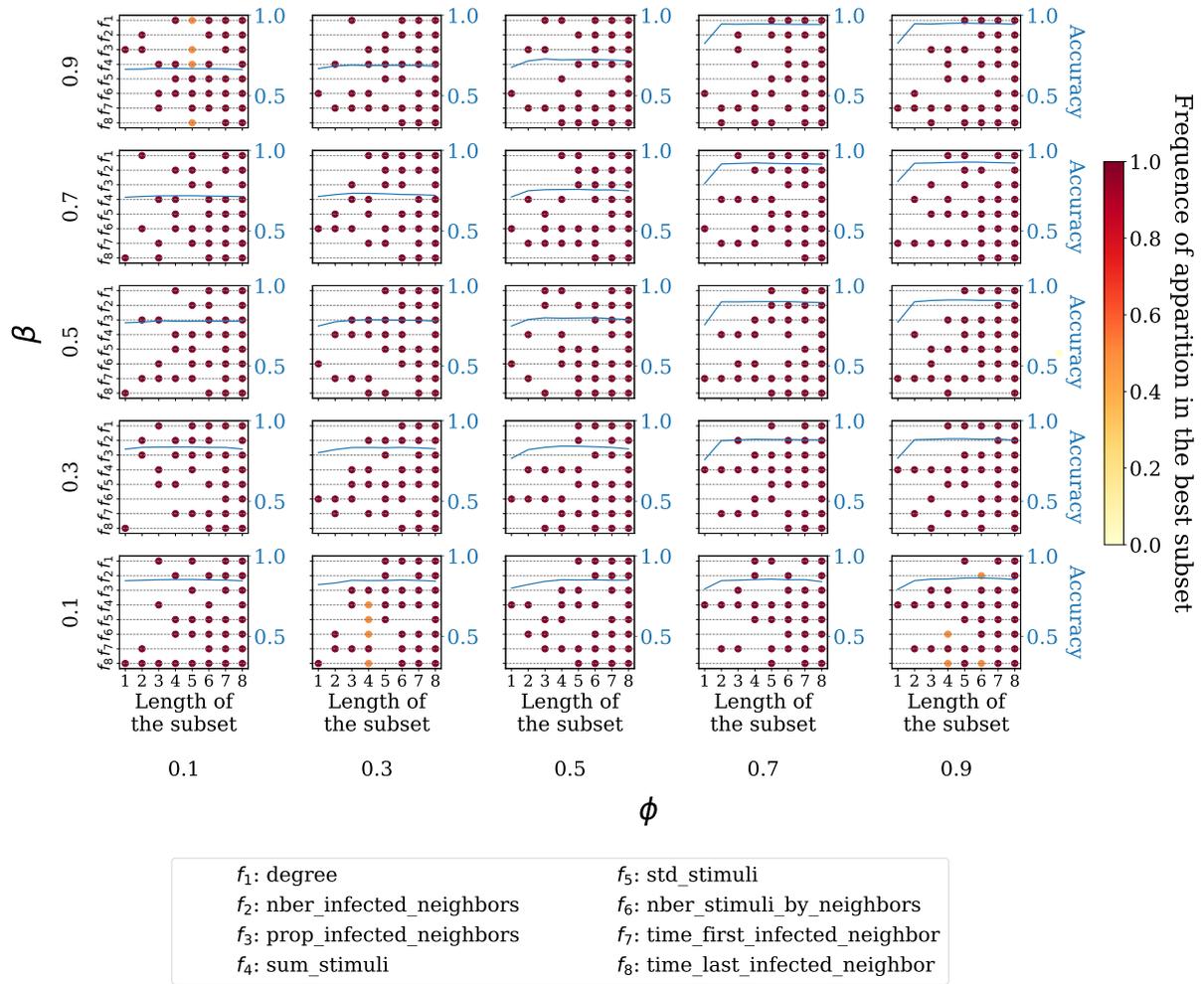


Figure B.2: Subsets of features giving the best accuracies (y-axis of each subplot) in the parameter space (β, ϕ) for a certain length of subset (x-axis of each subplot) in the classification with the random forest of Experiment 2. The corresponding accuracies are displayed in blue. If several subsets give the same best accuracies, we compute the frequency of apparition of each feature in those subsets. In most of the cases, only three features are enough to obtain the same accuracy values as with the total set of features, but those three features are different across the parameter space.

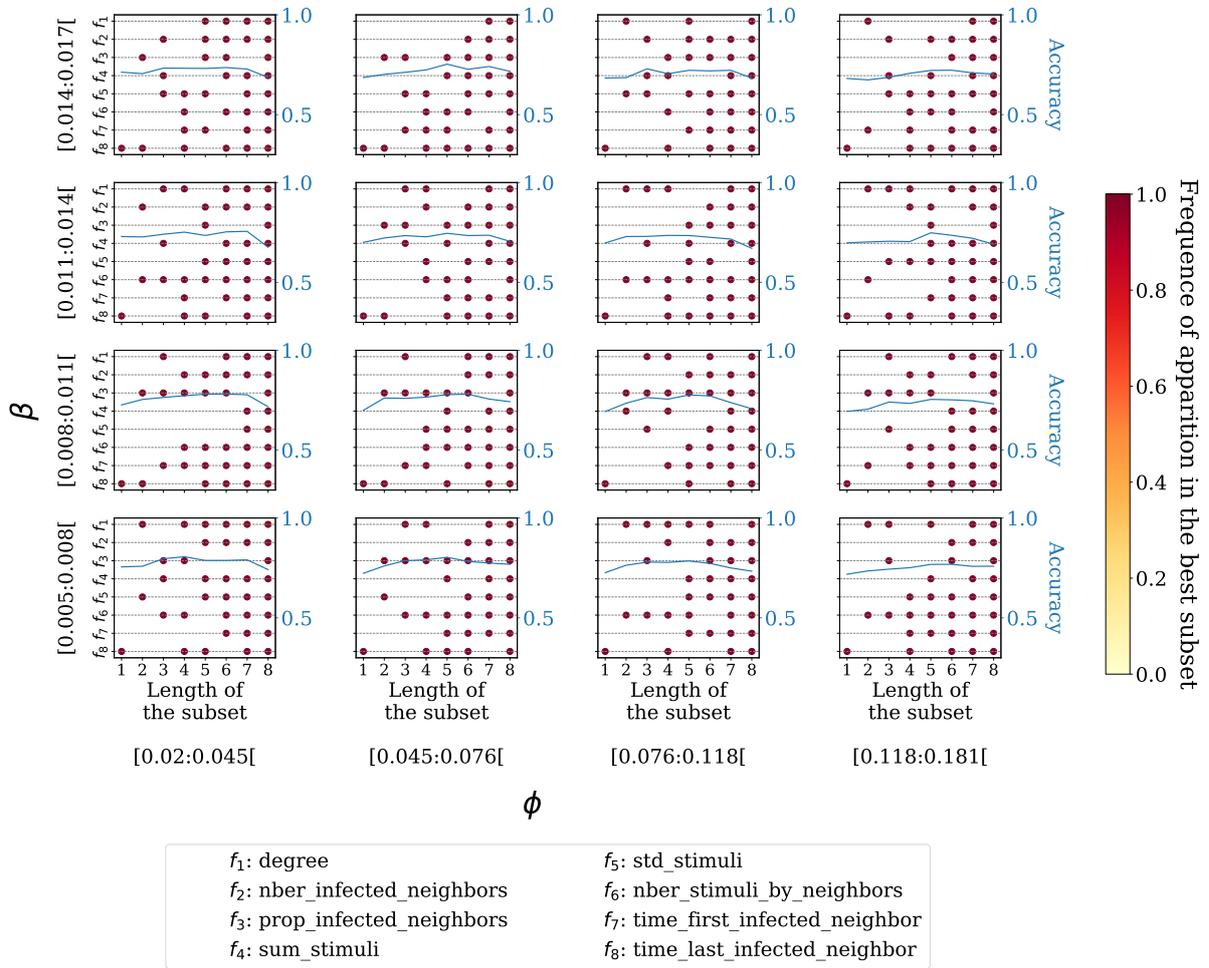


Figure B.3: Subsets of features giving the best accuracies (y-axis of each subplot) in the parameter space (β, ϕ) for a certain length of subset (x-axis of each subplot) in the classification with the random forest of Experiment 4. The corresponding accuracies are displayed in blue. If several subsets give the same best accuracies, we compute the frequency of apparition of each feature in those subsets. In most of the cases, only three features are enough to obtain the same accuracy values than with the total set of features, but those three features are different across the parameter space.

B.3 Accuracies of the classification of spontaneous adoption on Experiment 3

We assess the classification accuracies of the simple, complex and spontaneous cases from Experiment 4 with the random forest algorithm, using different values of filtering on the values of $\hat{\beta}$ and $\hat{\phi}$ (40%, 60%, 80% and 100%). The accuracies of the classification of the simple and complex instances increase while the percentage of the filtering diminishes (Figure B.4), even though all the obtained values remain above the accuracy of the random classification (0.33). We choose to work with a filter of 80% which presents accuracies above 0.65 while keeping most of the values of the distribution. The accuracies of the classification of the spontaneous instances (Table B.1) are consistently low, regardless of the percentage of data filtering employed in Experiment 3. This is attributed to the inability to assess the rate of infection r .

Percentage of filtering	40	60	80	100
Accuracy of the St	0.07	0.12	0.23	0.50

Table B.1: Accuracy of the classification of the spontaneous adoptions on Experiment 4 with the random forest.

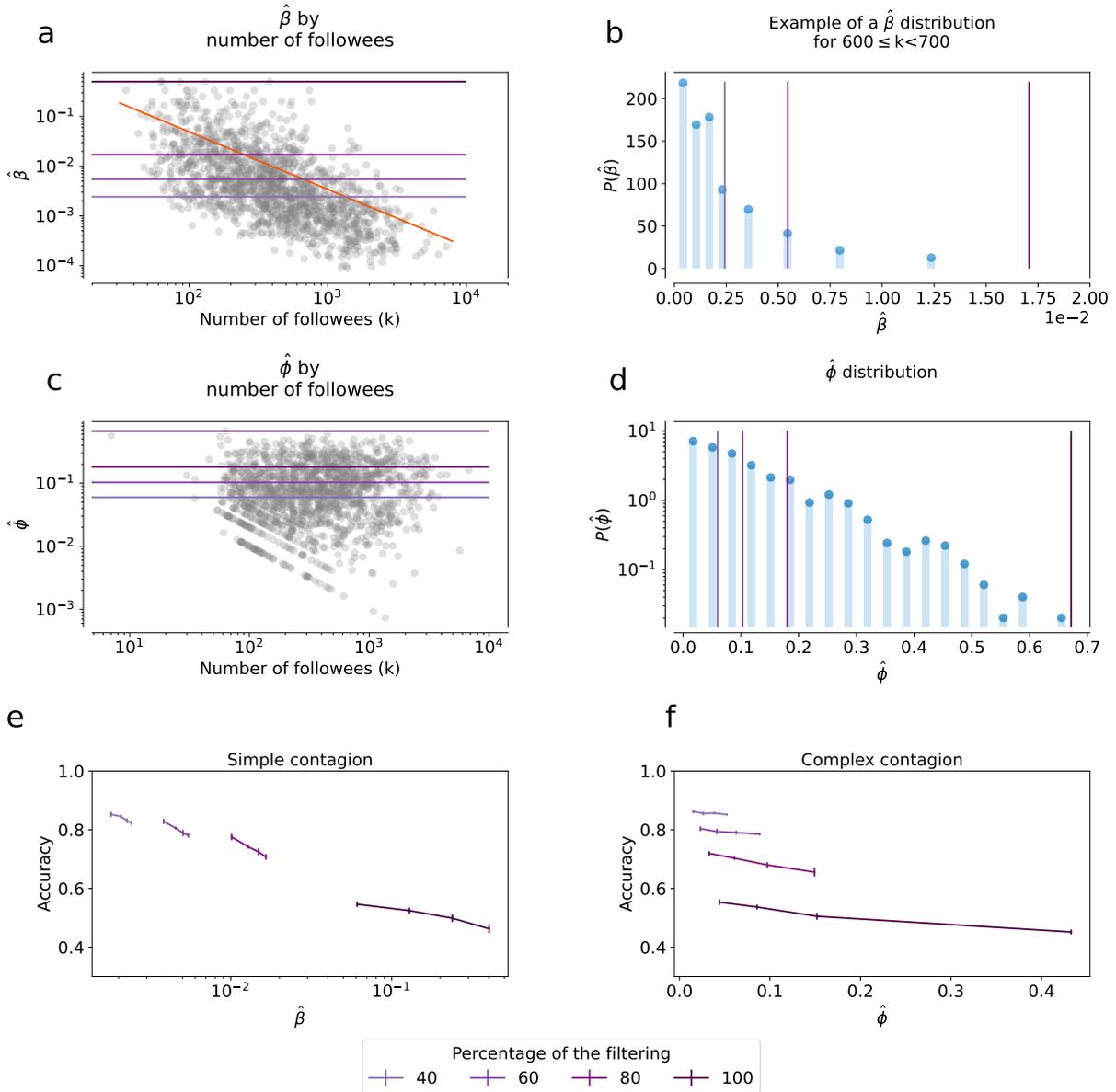


Figure B.4: Parametrization and accuracy of the classification of Experiment 4 constructed based on the #GiletsJaunes Twitter dataset. The distributions of $\hat{\beta}$ and $\hat{\phi}$ respectively panels a-b and c-d are filtered keeping their 40%, 60%, 80% or 100% lower values. The accuracy values of the classification of the simple contagion (panel e) and the complex contagion (panel f) increase while the percentage of filtering increases.

Appendix C

Competition between simple and complex contagion on temporal networks

C.1 General case $z > 2$

We can also calculate ρ_{eq} when $z > 2$. To do so, we first aim to calculate the expression of $P(n \geq z)$ and $\Delta_{m,\beta}(\rho)$ in the general case in which z can take any value. We first prove that $P(n \geq z)$ is a polynomial for which the z^{th} order Taylor expansion has a lower degree of z , meaning that $P(n \geq z)$ is governed by the term ρ^z . Indeed, the Taylor expansion of the term $(1 - \rho)^{m-n}$ in Eq. (4.3) gives the following.

$$P(n \geq z)(\rho) = \sum_{n=z}^m \binom{m}{n} \rho^n \left[1 + \sum_{i=1}^z \frac{(-1)^i \rho^i}{i!} (m-n) \dots (m-n-i+1) \right] \quad (C.1)$$

In order to prove that the lowest degree of $P(n \geq z)$ is z , we consider the term ρ^α , with $\alpha < z$ and demonstrate that its coefficient C_α is null. From Equation C.1, if α is equal to 0, the coefficient is $1 - \binom{m}{0} \rho^0$, which is null, otherwise the coefficient C_α is the following.

$$C_\alpha = -\binom{m}{\alpha} - \sum_{n=0}^{\alpha-1} \binom{m}{n} \frac{(-1)^{\alpha-n}}{(\alpha-n)!} (m-n) \dots (m-\alpha+1) \quad (C.2)$$

By reorganising the terms, Eq. (C.5) is then equal to the following.

$$C_\alpha = -\binom{m}{\alpha} \left[-1 - \sum_{n=0}^{\alpha-1} \binom{\alpha}{n} (-1)^{\alpha-n} \right] \quad (\text{C.3})$$

The right term of Eq. (C.3) can be expressed as $-1 - \left[\sum_{n=0}^{\alpha} \binom{\alpha}{n} (-1)^{\alpha-n} 1^n - 1 \right]$. Using the Newton's binomial, we show that that term is null, thus C^α is also null, which demonstrates that the z^{th} order Taylor expansion of the polynomial $P(n \geq z)$ has a lower degree of z . The coefficient of the term in z is then the following:

$$C_z = - \sum_{n=0}^{z-1} \binom{m}{n} \frac{(-1)^{z-n}}{(z-n)!} (m-n) \dots (m-z+1) \quad (\text{C.4})$$

By reorganising the term, we prove that $C_z = \binom{m}{z}$, and then $P(n \geq z) = \binom{m}{z} \rho^z$.

The general term of $\Delta_{m,\beta}(\rho)$ is given by its Taylor expansion:

$$\Delta_{m,\beta}(\rho) = - \sum_{i=1}^z \binom{m}{i} (-1)^i \beta^i \rho^i \quad (\text{C.5})$$

We then use those expressions in Eq. (4.4). In the right part of the equation, the term in ρ^z is $a \binom{m}{z} [p(-\beta)^{z-1} [\beta + \frac{z}{m-z+1}] + 1 - p] \rho^z$ and the term in ρ is $-apm\beta\rho$. Those two terms are equivalent when:

$$\rho_{eq} \approx \sqrt[z-1]{\frac{mp\beta}{\binom{m}{z} [p(-\beta)^{z-1} [\beta + \frac{z}{m-z+1}] + 1 - p]}} \quad (\text{C.6})$$

C.2 Simulations on the extended parameter-space (β, p)

The contagion curves and the expected times of increase t_{casc} , displayed on Figure 4.2 in the main text for four different parametrisations (β, p) , are shown for a larger range of values on Figure C.1. We can observe that large values of p and small values of β lead to a slow dynamics, as the model attempts but fails to infect nodes through the simple contagion. In contrast, the fastest contagion processes are for high values of both p and β as the nodes in this setting are mainly infected successfully by the simple contagion.

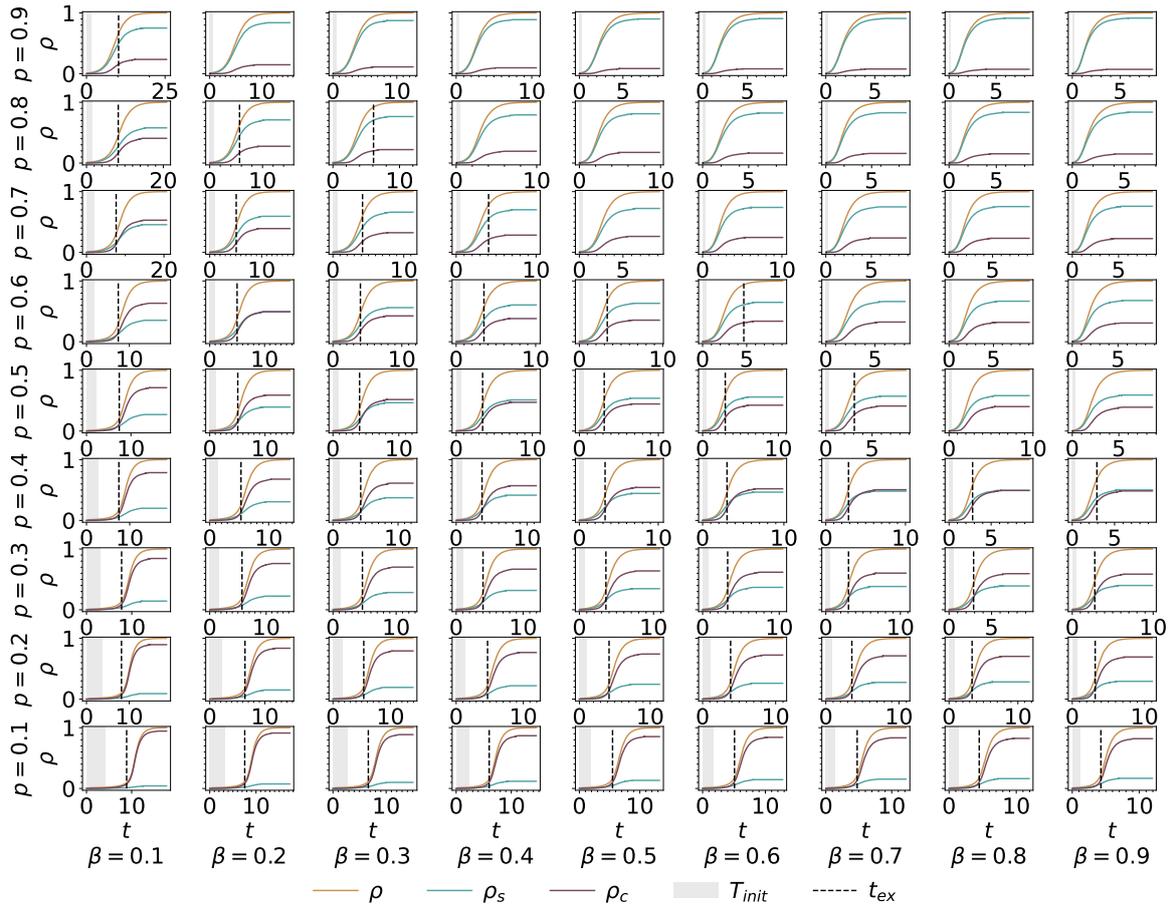


Figure C.1: Fraction of infected nodes, $\rho(t)$, as the function of time for simulated spreading scenarios (yellow line), together with the proportion of nodes infected by the simple, ρ_s , (blue curve) and the complex, ρ_c , (purple curve) contagions for $z = 2$. Panels show simulation results averaged over 100 realisations, for different values of β (x-axis) and p (y-axis).

We explore then the proportion of nodes infected by each process through time, namely ρ_s and ρ_c on Figure 4.2 (extended parameter-space of Figure 4.2). Complex contagion dominates the propagation for small values p , as nodes are more likely infected by the complex contagion in this setting. Also, in line with previous results, the simple contagion governs the spreading when p is high. The influence of β is minor but observable, leading to higher ρ_s when β increases. In particular, in the simulations for $p = 0.5$, $p = 0.6$ and $p = 0.7$, raising β changes the contagion process dominating the simulation, from the complex to the simple contagions.

Figure C.2 exemplifies the second method to evaluate the dominance of the simple or complex contagions at the beginning of the process with an extended parameter-space compared to Figure 4.4 in the main text. The curve ρ^{z-1} is linear decreasing if the complex contagion dominates, which is the case for low values of β and p . We measure the linearity of the early times of the curves by fitting the curves with a line and showing the time when the simulations and the fits have a difference higher than ε . As expected, the linear trend is not present for high values of β and p , for which the simple contagion governs.

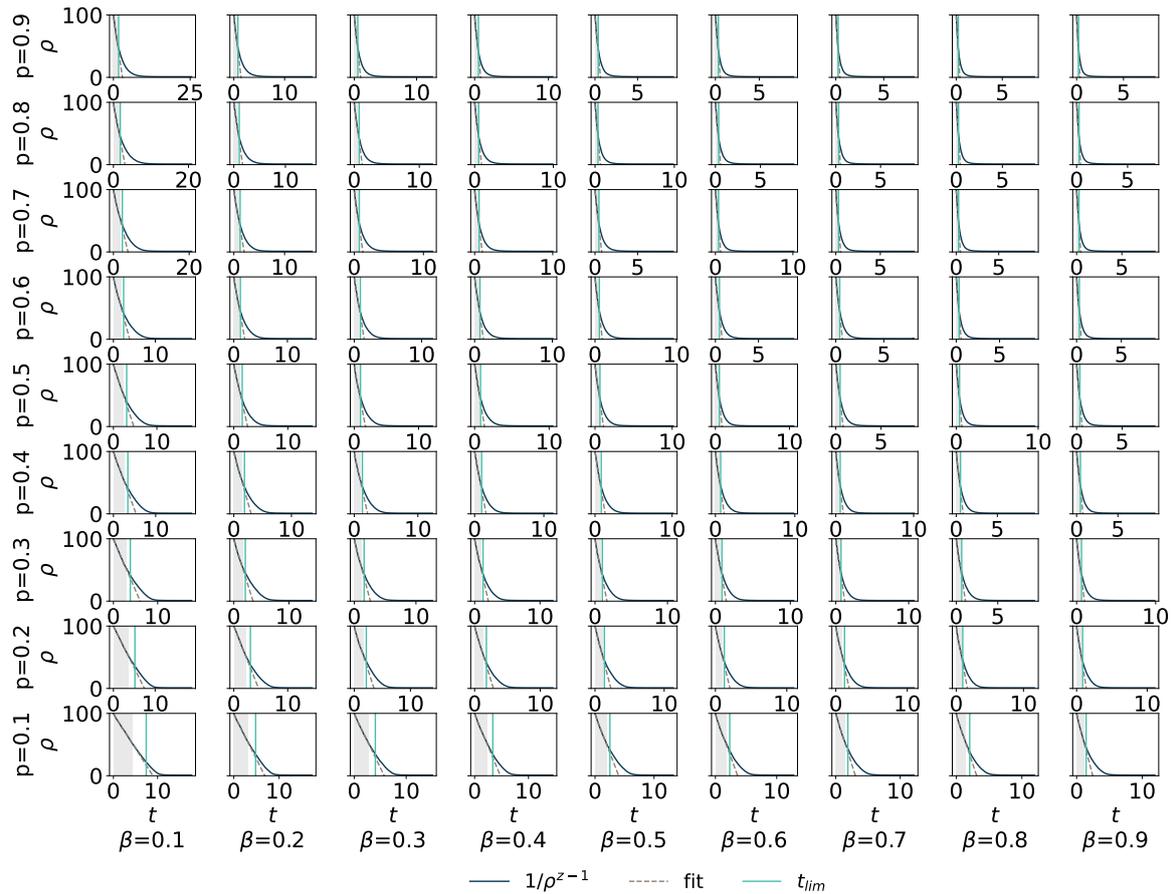


Figure C.2: Inverse of the proportion of infected neighbours to the power $z-1$ (dark blue line), fitted with a linear function on the first part of the propagation (black dotted line). The x-axis stands for different values of β , while the y-axis represents the values of ρ . The time t_{lim} when the difference between $1/\rho^z$ and its fit is higher than $\epsilon > 5$ is indicated with a light blue line.

Bibliography

- [1] R. Axelrod, “The dissemination of culture: A model with local convergence and global polarization,” *Journal of conflict resolution*, vol. 41, no. 2, pp. 203–226, 1997. Publisher: Sage Periodicals Press 2455 Teller Road, Thousand Oaks, CA 91320.
- [2] S. Wichmann, “The emerging field of language dynamics,” *Language and Linguistics Compass*, vol. 2, no. 3, pp. 442–455, 2008. Publisher: Wiley Online Library.
- [3] J. K. Parrish, W. Hamner, and W. M. Hamner, *Animal groups in three dimensions: how species aggregate*. Cambridge University Press, 1997.
- [4] I. D. Chase, “Social process and hierarchy formation in small groups: a comparative perspective,” *American Sociological Review*, pp. 905–924, 1980. Publisher: JSTOR.
- [5] C. Castellano, S. Fortunato, and V. Loreto, “Statistical physics of social dynamics,” *Reviews of modern physics*, vol. 81, no. 2, p. 591, 2009.
- [6] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008.
- [7] S. Jose, “Covid vaccine and generation z—a study of factors influencing adoption,” *Young Consumers*, vol. 23, no. 1, pp. 16–32, 2022.
- [8] N. A. Christakis and J. H. Fowler, “The collective dynamics of smoking in a large social network,” *New England journal of medicine*, vol. 358, no. 21, pp. 2249–2258, 2008.
- [9] B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee, “Econophysics and socio-physics: trends and perspectives,” 2006. Publisher: John Wiley & Sons.
- [10] D. J. Daley and D. G. Kendall, “Stochastic rumours,” *IMA Journal of Applied Mathematics*, vol. 1, no. 1, pp. 42–55, 1965.
- [11] D. P. Maki and M. Thompson, “Mathematical models and applications prentice-hall,” *Englewood Cliffs (NJ)*, 1973.

- [12] M. S. Granovetter, “The strength of weak ties,” *American journal of sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.
- [13] T. C. Schelling, “Dynamic models of segregation,” *Journal of mathematical sociology*, vol. 1, no. 2, pp. 143–186, 1971.
- [14] W. Chen, C. Castillo, and L. V. Lakshmanan, *Information and influence propagation in social networks*. Springer Nature, 2022.
- [15] D. Centola, “The spread of behavior in an online social network experiment,” *science*, vol. 329, no. 5996, pp. 1194–1197, 2010.
- [16] Y. Moreno, M. Nekovee, and A. F. Pacheco, “Dynamics of rumor spreading in complex networks,” *Physical review E*, vol. 69, no. 6, p. 066130, 2004.
- [17] D. J. Watts, “A simple model of global cascades on random networks,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 5766–5771, 2002.
- [18] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics reports*, vol. 424, no. 4-5, pp. 175–308, 2006.
- [19] S. Milgram, “Six degrees of separation,” *Psychology Today*, vol. 2, pp. 60–64, 1967.
- [20] P. S. Dodds, R. Muhamad, and D. J. Watts, “An experimental study of search in global social networks,” *science*, vol. 301, no. 5634, pp. 827–829, 2003.
- [21] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, “Four degrees of separation,” in *Proceedings of the 4th annual ACM Web science conference*, pp. 33–42, 2012.
- [22] P. Erdős and A. Rényi, “On the evolution of random graphs,” *Publ. math. inst. hung. acad. sci.*, vol. 5, no. 1, pp. 17–60, 1960.
- [23] F. Chung and L. Lu, “The diameter of sparse random graphs,” *Advances in Applied Mathematics*, vol. 26, no. 4, pp. 257–279, 2001. Publisher: Elsevier.
- [24] H. E. Stanley and G. Ahlers, “Introduction to phase transitions and critical phenomena,” 1973. Publisher: American Institute of Physics.
- [25] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.
- [26] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *nature*, vol. 393, no. 6684, pp. 440–442, 1998. Publisher: Nature Publishing Group.

- [27] A. Barrat and M. Weigt, “On the properties of small-world network models,” *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 13, pp. 547–560, 2000. Publisher: Springer.
- [28] Y. W. Chen, L. F. Zhang, and J. P. Huang, “The watts–strogatz network model developed by including degree distribution: theory and computer simulation,” *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 29, p. 8237, 2007.
- [29] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [30] G. G. Piva, F. L. Ribeiro, and A. S. Mata, “Networks with growth and preferential attachment: modelling and applications,” *Journal of Complex Networks*, vol. 9, no. 1, p. cnab008, 2021.
- [31] R. Cohen and S. Havlin, “Scale-free networks are ultrasmall,” *Physical review letters*, vol. 90, no. 5, p. 058701, 2003.
- [32] P. W. Holland, K. B. Laskey, and S. Leinhardt, “Stochastic blockmodels: First steps,” *Social networks*, vol. 5, no. 2, pp. 109–137, 1983.
- [33] C. Lee and D. J. Wilkinson, “A review of stochastic block models and extensions for graph clustering,” *Applied Network Science*, vol. 4, no. 1, pp. 1–50, 2019.
- [34] E. A. Hobson, M. J. Silk, N. H. Fefferman, D. B. Larremore, P. Rombach, S. Shai, and N. Pinter-Wollman, “A guide to choosing and implementing reference models for social network analysis,” *Biological Reviews*, vol. 96, no. 6, pp. 2716–2734, 2021.
- [35] E. Abbe, “Community detection and stochastic block models: recent developments,” *Journal of Machine Learning Research*, vol. 18, no. 177, pp. 1–86, 2018.
- [36] L. Peel, D. B. Larremore, and A. Clauset, “The ground truth about metadata and community detection in networks,” *Science advances*, vol. 3, no. 5, p. e1602548, 2017.
- [37] P. Holme and J. Saramäki, “Temporal networks,” *Physics reports*, vol. 519, no. 3, pp. 97–125, 2012.
- [38] P. Holme, “Modern temporal network theory: a colloquium,” *The European Physical Journal B*, vol. 88, pp. 1–30, 2015.
- [39] N. Masuda and R. Lambiotte, *A guide to temporal networks*. World Scientific, 2016.
- [40] A. Barrat and C. Cattuto, “Temporal networks of face-to-face human interactions,” *Temporal networks*, pp. 191–216, 2013.

- [41] S. Lehmann, “Fundamental structures in temporal communication networks,” *Temporal Network Theory*, pp. 25–48, 2019.
- [42] M. Saqr and S. López-Pernas, “The why, the what and the how to model a dynamic relational learning process with temporal networks,” in *Proceedings of the NetSciLA22 workshop*, 2022.
- [43] D. Braha and Y. Bar-Yam, “Time-dependent complex networks: Dynamic centrality, dynamic motifs, and cycles of social interactions,” in *Adaptive networks: Theory, models and applications*, pp. 39–50, Springer, 2009.
- [44] R. K. Pan and J. Saramäki, “Path lengths, correlations, and centrality in temporal networks,” *Physical Review E*, vol. 84, no. 1, p. 016105, 2011.
- [45] A. Li, S. P. Cornelius, Y.-Y. Liu, L. Wang, and A.-L. Barabási, “The fundamental advantages of temporal networks,” *Science*, vol. 358, no. 6366, pp. 1042–1046, 2017.
- [46] A. Moinet, M. Starnini, and R. Pastor-Satorras, “Burstiness and aging in social temporal networks,” *Physical review letters*, vol. 114, no. 10, p. 108701, 2015.
- [47] S. Unicomb, G. Iñiguez, J. P. Gleeson, and M. Karsai, “Dynamics of cascades on burstiness-controlled temporal networks,” *Nature communications*, vol. 12, no. 1, p. 133, 2021.
- [48] A.-L. Barabasi, “The origin of bursts and heavy tails in human dynamics,” *Nature*, vol. 435, no. 7039, pp. 207–211, 2005.
- [49] M. Karsai, K. Kaski, and J. Kertész, “Correlated dynamics in egocentric communication networks,” *Plos one*, vol. 7, no. 7, p. e40612, 2012.
- [50] T. Aledavood, S. Lehmann, and J. Saramäki, “Digital daily cycles of individuals,” *Frontiers in Physics*, vol. 3, p. 73, 2015.
- [51] J. Saramäki and E. Moro, “From seconds to months: an overview of multi-scale dynamics of mobile telephone calls,” *The European Physical Journal B*, vol. 88, pp. 1–10, 2015.
- [52] V. Gelardi, J. Fagot, A. Barrat, and N. Claidière, “Detecting social (in) stability in primates from their temporal co-presence network,” *Animal Behaviour*, vol. 157, pp. 239–254, 2019.
- [53] N. Masuda and P. Holme, “Detecting sequences of system states in temporal networks,” *Scientific reports*, vol. 9, no. 1, pp. 1–11, 2019.

- [54] N. Pedreschi, C. Bernard, W. Clawson, P. Quilichini, A. Barrat, and D. Battaglia, “Dynamic core-periphery structure of information sharing networks in entorhinal cortex and hippocampus,” *Network Neuroscience*, vol. 4, no. 3, pp. 946–975, 2020.
- [55] J. Fournet and A. Barrat, “Contact patterns among high school students,” *PloS one*, vol. 9, no. 9, p. e107878, 2014.
- [56] M. Lucas, A. Morris, A. Townsend-Teague, L. Tichit, B. Habermann, and A. Barrat, “Inferring cell cycle phases from a partially temporal network of protein interactions,” *Cell Reports Methods*, vol. 3, no. 2, 2023.
- [57] E. Valdano, L. Ferreri, C. Poletto, and V. Colizza, “Analytical computation of the epidemic threshold on temporal networks,” *Physical Review X*, vol. 5, no. 2, p. 021005, 2015.
- [58] K. Sato, M. Oka, A. Barrat, and C. Cattuto, “Dyane: dynamics-aware node embedding for temporal networks,” *arXiv preprint arXiv:1909.05976*, 2019.
- [59] M. Kivelä, J. Cambe, J. Saramäki, and M. Karsai, “Mapping temporal-network percolation to weighted, static event graphs,” *Scientific reports*, vol. 8, no. 1, pp. 1–9, 2018.
- [60] A. Mellor, “Event graphs: Advances and applications of second-order time-unfolded temporal network models,” *Advances in Complex Systems*, vol. 22, no. 03, p. 1950006, 2019.
- [61] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, “Activity driven modeling of time varying networks,” *Scientific reports*, vol. 2, no. 1, pp. 1–7, 2012.
- [62] M. Starnini and R. Pastor-Satorras, “Topological properties of a time-integrated activity-driven network,” *Physical Review E*, vol. 87, no. 6, p. 062807, 2013.
- [63] G. Laurent, J. Saramäki, and M. Karsai, “From calls to communities: a model for time-varying social networks,” *The European Physical Journal B*, vol. 88, pp. 1–10, 2015.
- [64] L. Gauvin, M. Génois, M. Karsai, M. Kivelä, T. Takaguchi, E. Valdano, and C. L. Vestergaard, “Randomized reference models for temporal networks,” *SIAM Review*, vol. 64, no. 4, pp. 763–830, 2022.
- [65] R. M. Anderson and R. M. May, *Infectious diseases of humans: dynamics and control*. Oxford university press, 1991.
- [66] M. E. Newman, “The structure and function of complex networks,” *SIAM review*, vol. 45, no. 2, pp. 167–256, 2003.

- [67] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” *Reviews of modern physics*, vol. 87, no. 3, p. 925, 2015.
- [68] J. P. Gleeson, “Binary-state dynamics on complex networks: Pair approximation and beyond,” *Physical Review X*, vol. 3, no. 2, p. 021004, 2013.
- [69] J.-P. Onnela and F. Reed-Tsochas, “Spontaneous emergence of social influence in online systems,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 43, pp. 18375–18380, 2010.
- [70] M. Karsai, G. Iniguez, K. Kaski, and J. Kertész, “Complex contagion process in spreading of online innovation,” *Journal of The Royal Society Interface*, vol. 11, no. 101, p. 20140694, 2014.
- [71] E. Oster and R. Thornton, “Determinants of technology adoption: Peer effects in menstrual cup take-up,” *Journal of the European Economic Association*, vol. 10, no. 6, pp. 1263–1293, 2012.
- [72] O. Bandiera and I. Rasul, “Social networks and technology adoption in northern mozambique,” *The economic journal*, vol. 116, no. 514, pp. 869–902, 2006.
- [73] S. Cure, F. Pflug, and S. Pigolotti, “Rate of epidemic spreading on complex networks,” *arXiv preprint arXiv:2406.15449*, 2024.
- [74] N. O. Hodas and K. Lerman, “The simple rules of social contagion,” *Scientific reports*, vol. 4, no. 1, p. 4343, 2014.
- [75] F. L. Pinheiro, M. D. Santos, F. C. Santos, and J. M. Pacheco, “Origin of peer influence in social networks,” *Physical review letters*, vol. 112, no. 9, p. 098702, 2014.
- [76] D. Centola and M. Macy, “Complex contagions and the weakness of long ties,” *American journal of Sociology*, vol. 113, no. 3, pp. 702–734, 2007.
- [77] D. Guilbeault, J. Becker, and D. Centola, “Complex contagions: A decade in review,” *Complex spreading phenomena in social systems: Influence and contagion in real-world social networks*, pp. 3–25, 2018.
- [78] M. Granovetter, “Threshold models of collective behavior,” *American journal of sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.
- [79] R. I. Joh, H. Wang, H. Weiss, and J. S. Weitz, “Dynamics of indirectly transmitted infectious diseases with immunological threshold,” *Bulletin of mathematical biology*, vol. 71, pp. 845–862, 2009.

- [80] S. Melnik, J. A. Ward, J. P. Gleeson, and M. A. Porter, “Multi-stage complex contagions,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 23, no. 1, 2013.
- [81] Z. Ruan, G. Iniguez, M. Karsai, and J. Kertész, “Kinetics of social contagion,” *Physical review letters*, vol. 115, no. 21, p. 218702, 2015.
- [82] L. Böttcher, J. Nagler, and H. J. Herrmann, “Critical behaviors in contagion dynamics,” *Physical review letters*, vol. 118, no. 8, p. 088301, 2017.
- [83] G. F. de Arruda, F. A. Rodrigues, and Y. Moreno, “Fundamentals of spreading processes in single and multilayer complex networks,” *Physics Reports*, vol. 756, pp. 1–59, 2018.
- [84] B. Min and M. San Miguel, “Threshold cascade dynamics in coevolving networks,” *Entropy*, vol. 25, no. 6, p. 929, 2023.
- [85] M. Karsai, G. Iniguez, R. Kikas, K. Kaski, and J. Kertész, “Local cascades induced global contagion: How heterogeneous thresholds, exogenous effects, and unconcerned behaviour govern online adoption spreading,” *Scientific reports*, vol. 6, no. 1, p. 27178, 2016.
- [86] S. Unicomb, G. Iniguez, and M. Karsai, “Threshold driven contagion on weighted networks,” *Scientific reports*, vol. 8, no. 1, p. 3094, 2018.
- [87] X. Li, P. Wang, X.-J. Xu, and G. Xiao, “Universal behavior of the linear threshold model on weighted networks,” *Journal of Parallel and Distributed Computing*, vol. 123, pp. 223–229, 2019.
- [88] F. Karimi and P. Holme, “Threshold model of cascades in empirical temporal networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 16, pp. 3476–3483, 2013.
- [89] V. V. Vasconcelos, S. A. Levin, and F. L. Pinheiro, “Consensus and polarization in competing complex contagion processes,” *Journal of the Royal Society Interface*, vol. 16, no. 155, p. 20190196, 2019.
- [90] B. Mønsted, P. Sapiezynski, E. Ferrara, and S. Lehmann, “Evidence of complex contagion of information in social media: An experiment using twitter bots,” *PloS one*, vol. 12, no. 9, p. e0184148, 2017.
- [91] J. Borge-Holthoefer, A. Rivero, I. García, E. Cauhé, A. Ferrer, D. Ferrer, D. Francos, D. Iniguez, M. P. Pérez, G. Ruiz, *et al.*, “Structural and dynamical patterns on online social networks: the spanish may 15th movement as a case study,” *PloS one*, vol. 6, no. 8, p. e23883, 2011.

- [92] A. L. Hill, D. G. Rand, M. A. Nowak, and N. A. Christakis, “Infectious disease modeling of social contagion in networks,” *PLOS computational biology*, vol. 6, no. 11, p. e1000968, 2010.
- [93] J. L. Toole, M. Cha, and M. C. González, “Modeling the adoption of innovations in the presence of geographic and media influences,” *PloS one*, vol. 7, no. 1, p. e29528, 2012.
- [94] J. Ugander, L. Backstrom, C. Marlow, and J. Kleinberg, “Structural diversity in social contagion,” *Proceedings of the national academy of sciences*, vol. 109, no. 16, pp. 5962–5966, 2012.
- [95] H. Yang, M. Tang, and T. Gross, “Large epidemic thresholds emerge in heterogeneous networks of heterogeneous nodes,” *Scientific reports*, vol. 5, no. 1, p. 13122, 2015.
- [96] G. Cencetti, D. A. Contreras, M. Mancastropa, and A. Barrat, “Distinguishing simple and complex contagion processes on networks,” *Physical Review Letters*, vol. 130, no. 24, p. 247401, 2023.
- [97] M. Karsai, H.-H. Jo, K. Kaski, *et al.*, *Bursty human dynamics*. Springer, 2018.
- [98] N. Masuda, J. C. Miller, and P. Holme, “Concurrency measures in the era of temporal network epidemiology: A review,” *Journal of The Royal Society Interface*, vol. 18, no. 179, p. 20210019, 2021.
- [99] W. Wang, H. E. Stanley, and L. A. Braunstein, “Effects of time-delays in the dynamics of social contagions,” *New Journal of Physics*, vol. 20, no. 1, p. 013034, 2018.
- [100] L. Hébert-Dufresne, S. V. Scarpino, and J.-G. Young, “Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement,” *Nature Physics*, vol. 16, no. 4, pp. 426–431, 2020.
- [101] M. Gladwell, *The tipping point: How little things can make a big difference*. Little, Brown, 2006.
- [102] S. Aral and C. Nicolaides, “Exercise contagion in a global social network,” *Nature communications*, vol. 8, no. 1, p. 14753, 2017.
- [103] B. State and L. Adamic, “The diffusion of support in an online social movement: Evidence from the adoption of equal-sign profile pictures,” in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 1741–1750, 2015.

- [104] P. D. Karampourniotis, S. Sreenivasan, B. K. Szymanski, and G. Korniss, “The impact of heterogeneous thresholds on social contagion with multiple initiators,” *PloS one*, vol. 10, no. 11, p. e0143020, 2015.
- [105] B. Min and M. San Miguel, “Competing contagion processes: Complex contagion triggered by simple contagion,” *Scientific reports*, vol. 8, no. 1, p. 10422, 2018.
- [106] K. Trøjelsgaard and J. M. Olesen, “Ecological networks in motion: Micro-and macroscopic variability across scales,” *Functional Ecology*, vol. 30, no. 12, pp. 1926–1935, 2016.
- [107] M. M. Hosseinzadeh, M. Cannataro, P. H. Guzzi, and R. Dondi, “Temporal networks in biology and medicine: a survey on models, algorithms, and tools,” *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 12, no. 1, p. 10, 2022.
- [108] N. Huynh and J. Barthelemy, “A comparative study of topological analysis and temporal network analysis of a public transport system,” *International Journal of Transportation Science and Technology*, vol. 11, no. 2, pp. 392–405, 2022.
- [109] M. Salama, M. Ezzeldin, W. El-Dakhakhni, and M. Tait, “Temporal networks: A review and opportunities for infrastructure simulation,” *Sustainable and resilient infrastructure*, vol. 7, no. 1, pp. 40–55, 2022.
- [110] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, *et al.*, “High-resolution measurements of face-to-face contact patterns in a primary school,” *PloS one*, vol. 6, no. 8, p. e23176, 2011.
- [111] R. K. Darst, C. Granell, A. Arenas, S. Gómez, J. Saramäki, and S. Fortunato, “Detection of timescales in evolving complex systems,” *Scientific reports*, vol. 6, no. 1, p. 39713, 2016.
- [112] K. Sugishita and N. Masuda, “Recurrence in the evolution of air transport networks,” *Scientific reports*, vol. 11, no. 1, pp. 1–15, 2021.
- [113] L. Lacasa, J. P. Rodriguez, and V. M. Eguiluz, “Correlations of network trajectories,” *Physical Review Research*, vol. 4, no. 4, p. L042008, 2022.
- [114] J. P. Bagrow and E. M. Bollt, “An information-theoretic, all-scales approach to comparing networks,” *Applied Network Science*, vol. 4, no. 1, pp. 1–15, 2019.
- [115] C. Shannon, “Communication in the presence of noise,” *Proceedings of the IRE*, vol. 37, pp. 10–21, jan 1949.

- [116] R. Sulo, T. Berger-Wolf, and R. Grossman, “Meaningful selection of temporal resolution for dynamic networks,” in *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, pp. 127–136, 2010.
- [117] G. Krings, M. Karsai, S. Bernhardsson, V. D. Blondel, and J. Saramäki, “Effects of time window size and placement on the structure of an aggregated communication network,” *EPJ Data Science*, vol. 1, no. 1, pp. 1–16, 2012.
- [118] M. Kivelä and M. A. Porter, “Estimating interevent time distributions from finite observation periods in communication networks,” *Physical Review E*, vol. 92, no. 5, p. 052813, 2015.
- [119] D. J. Toth, M. Leecaster, W. B. Pettey, A. V. Gundlapalli, H. Gao, J. J. Rainey, A. Uzicanin, and M. H. Samore, “The role of heterogeneity in contact timing and duration in network models of influenza spread in schools,” *Journal of The Royal Society Interface*, vol. 12, no. 108, p. 20150279, 2015.
- [120] “Bureau of transportation statistics, bureau of transportation statistics website (2017).,”
- [121] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani, “Dynamics of person-to-person interactions from distributed rfid sensor networks,” *PloS one*, vol. 5, no. 7, p. e11596, 2010.
- [122] M. Génois, M. Zens, M. Oliveira, C. M. Lechner, J. Schaible, and M. Strohmaier, “Combining sensors and surveys to study social interactions: A case of four science conferences,” *Personality Science*, vol. 4, pp. 1–24, 2023.
- [123] C. Bai, S. Kumar, J. Leskovec, M. Metzger, J. F. Nunamaker Jr, and V. Subrahmanian, “Predicting the visual focus of attention in multi-person discussion videos.,” in *IJCAI*, pp. 4504–4510, 2019.
- [124] S. Kumar, C. Bai, V. Subrahmanian, and J. Leskovec, “Deception detection in group video conversations using dynamic interaction networks.,” in *ICWSM*, pp. 339–350, 2021.
- [125] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, “Networks beyond pairwise interactions: Structure and dynamics,” *Phys. Rep.*, vol. 874, pp. 1–92, 2020.
- [126] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. Murray, T. Peixoto, F. Vaccarino, and G. Petri, “The physics of higher-order interactions in complex systems,” *Nat. Phys.*, vol. 17, no. 10, pp. 1093–1098, 2021.

- [127] R. van de Schoot, S. Depaoli, R. King, B. Kramer, K. Märtens, M. G. Tadesse, M. Vannucci, A. Gelman, D. Veen, J. Willemsen, *et al.*, “Bayesian statistics and modelling,” *Nature Reviews Methods Primers*, vol. 1, no. 1, p. 1, 2021.
- [128] D. G. Kleinbaum, M. Klein, D. G. Kleinbaum, and M. Klein, “Maximum likelihood techniques: An overview,” *Logistic regression: A self-learning text*, pp. 103–127, 2010.
- [129] A. L. Samuel, “Some studies in machine learning using the game of checkers,” *IBM Journal of research and development*, vol. 3, no. 3, pp. 210–229, 1959.
- [130] W. I. D. Mining, *Introduction to data mining*. Springer, 2006.
- [131] Y.-Y. Song and L. Ying, “Decision tree methods: applications for classification and prediction,” *Shanghai archives of psychiatry*, vol. 27, no. 2, p. 130, 2015.
- [132] L. Ceriani and P. Verme, “The origins of the gini index: extracts from variabilità e mutabilità (1912) by corrado gini,” *The Journal of Economic Inequality*, vol. 10, pp. 421–443, 2012.
- [133] L. Breiman, “Random forests,” *Machine learning*, vol. 45, pp. 5–32, 2001.
- [134] H. Zhang, “The optimality of naive bayes,” *Aa*, vol. 1, no. 2, p. 3, 2004.
- [135] A. Mucherino, P. J. Papajorgji, P. M. Pardalos, A. Mucherino, P. J. Papajorgji, and P. M. Pardalos, “K-nearest neighbor classification,” *Data mining in agriculture*, pp. 83–106, 2009.
- [136] S. I. Gallant *et al.*, “Perceptron-based learning algorithms,” *IEEE Transactions on neural networks*, vol. 1, no. 2, pp. 179–191, 1990.
- [137] C.-W. Hsu, C.-C. Chang, C.-J. Lin, *et al.*, “A practical guide to support vector classification,” 2003.
- [138] R. Wang, “Adaboost for feature selection, classification and its relation with svm, a review,” *Physics Procedia*, vol. 25, pp. 800–807, 2012.
- [139] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Frontiers in neuro-robotics*, vol. 7, p. 21, 2013.
- [140] L. F. Price, C. C. Drovandi, A. Lee, and D. J. Nott, “Bayesian synthetic likelihood,” *Journal of Computational and Graphical Statistics*, vol. 27, no. 1, pp. 1–11, 2018.
- [141] M. H. Hansen and B. Yu, “Model selection and the principle of minimum description length,” *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 746–774, 2001.

- [142] P. Erdős, A. Rényi, *et al.*, “On the evolution of random graphs,” *Publ. math. inst. hung. acad. sci.*, vol. 5, no. 1, pp. 17–60, 1960.
- [143] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [144] S. Unicomb, G. Iñiguez, J. Kertész, and M. Karsai, “Reentrant phase transitions in threshold driven contagion on multiplex networks,” *Physical Review E*, vol. 100, no. 4, p. 040301, 2019.
- [145] J. F. de Oliveira, H. T. Marques-Neto, and M. Karsai, “Measuring the effects of repeated and diversified influence mechanism for information adoption on twitter,” *Social Network Analysis and Mining*, vol. 12, no. 1, p. 16, 2022.
- [146] J. F. de Oliveira, H. T. Marques-Neto, and M. Karsai, “Information adoption via repeated or diversified social influence on twitter,” in *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 237–241, IEEE, 2020.
- [147] D. A. Contreras, G. Cencetti, and A. Barrat, “Infection patterns in simple and complex contagion processes on networks,” *PLoS Computational Biology*, vol. 20, no. 6, p. e1012206, 2024.
- [148] T. W. Valente and G. G. Vega Yon, “Diffusion/contagion processes on social networks,” *Health Education & Behavior*, vol. 47, no. 2, pp. 235–248, 2020.
- [149] G. St-Onge, L. Hébert-Dufresne, and A. Allard, “Nonlinear bias toward complex contagion in uncertain transmission settings,” *Proceedings of the National Academy of Sciences*, vol. 121, no. 1, p. e2312202121, 2024.
- [150] I. Iacopini, G. Petri, A. Barrat, and V. Latora, “Simplicial models of social contagion,” *Nature communications*, vol. 10, no. 1, p. 2485, 2019.
- [151] G. St-Onge, I. Iacopini, V. Latora, A. Barrat, G. Petri, A. Allard, and L. Hébert-Dufresne, “Influential groups for seeding and sustaining nonlinear contagion in heterogeneous hypergraphs,” *Communications Physics*, vol. 5, no. 1, p. 25, 2022.
- [152] G. Ferraz de Arruda, G. Petri, P. M. Rodriguez, and Y. Moreno, “Multistability, intermittency, and hybrid transitions in social contagion models on hypergraphs,” *Nature Communications*, vol. 14, no. 1, p. 1375, 2023.
- [153] J. P. Gleeson, K. P. O’Sullivan, R. A. Baños, and Y. Moreno, “Effects of network structure, competition and memory time on social spreading phenomena,” *Physical Review X*, vol. 6, no. 2, p. 021019, 2016.

- [154] F. Diaz-Diaz, M. San Miguel, and S. Meloni, “Echo chambers and information transmission biases in homophilic and heterophilic networks,” *Scientific Reports*, vol. 12, no. 1, p. 9350, 2022.
- [155] P. Holme and F. Liljeros, “Birth and death of links control disease spreading in empirical contact networks,” *Scientific reports*, vol. 4, no. 1, p. 4999, 2014.
- [156] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, “Small but slow world: How network topology and burstiness slow down spreading,” *Physical Review E*, vol. 83, no. 2, p. 025102, 2011.
- [157] A. Badie-Modiri, A. K. Rizi, M. Karsai, and M. Kivelä, “Directed percolation in temporal networks,” *Physical Review Research*, vol. 4, no. 2, p. L022047, 2022.
- [158] M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, J. Saramäki, and M. Karsai, “Multiscale analysis of spreading in a large communication network,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2012, no. 03, p. P03005, 2012.
- [159] T. Takaguchi, N. Masuda, and P. Holme, “Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics,” *PloS one*, vol. 8, no. 7, p. e68629, 2013.
- [160] A. Czaplicka, R. Toral, and M. San Miguel, “Competition of simple and complex adoption on interdependent networks,” *Physical Review E*, vol. 94, no. 6, p. 062301, 2016.
- [161] P. S. Dodds and D. J. Watts, “Universal behavior in a generalized model of contagion,” *Physical review letters*, vol. 92, no. 21, p. 218701, 2004.
- [162] W. Wang, M. Tang, P. Shu, and Z. Wang, “Dynamics of social contagions with heterogeneous adoption thresholds: crossover phenomena in phase transition,” *New Journal of Physics*, vol. 18, no. 1, p. 013029, 2016.
- [163] N. Horsevad, D. Mateo, R. E. Kooij, A. Barrat, and R. Bouffanais, “Transition from simple to complex contagion in collective decision-making,” *Nature communications*, vol. 13, no. 1, p. 1442, 2022.
- [164] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*. New York: Dover, 1972.
- [165] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, “Velocity and hierarchical spread of epidemic outbreaks in scale-free networks,” *Phys. Rev. Lett.*, vol. 92, p. 178701, Apr 2004.