Distinguishing mechanisms of social contagion from local network view

Check for updates

Elsa Andres[1], Gergely Ódor[1,2], Iacopo Iacopini[3,4] & Márton Karsai[1,2] ✉

The adoption of individual behavioural patterns is largely determined by stimuli arriving from peers via social interactions or from external sources. Based on these influences, individuals are commonly assumed to follow simple or complex adoption rules, inducing social contagion processes. In reality, multiple adoption rules may coexist even within the same social contagion process, introducing additional complexity to the spreading phenomena. Our goal is to understand whether coexisting adoption mechanisms can be distinguished from a microscopic view at the egocentric network level without requiring global information about the underlying network, or the unfolding spreading process. We formulate this question as a classification problem, and study it through a likelihood approach and with random forest classifiers in various synthetic and data-driven experiments. This study offers a novel perspective on the observations of propagation processes at the egocentric level and a better understanding of landmark contagion mechanisms from a local view.

We influence our peers through our conduct and interactions, thereby impacting their decisions to follow behavioural patterns similar to ours. Such patterns, mediated by social influence, may propagate as a spreading process and lead to macroscopic phenomena of mass adoption of products, ideas, beliefs, or information cascades[1–5]. The relevance of social spreading phenomena has been previously identified[6,7] and arguably explained by simple decision mechanisms on well-mixed populations[8–11]. Meanwhile, the importance of social networks has also been recognised[12–14], as they effectively encode the underlying structure along which social influence travels. Their structure could critically influence the global outcome of social spreading phenomena unfolding on top of them[2,15]. This finding is especially true for temporal networks[16], which capture both the structure and the time of interactions between connected peers, whose time-varying links represent possible events of direct social influence[17,18].

Models of social contagion commonly describe the spreading dynamics as a binary state process[19], in which individuals are identified as nodes of a social network that can be in different states; susceptible nodes (also called ignorants) may adopt a given behaviour and become "infected"—borrowing the term from the literature of infectious disease modelling—or in other words spreaders, or adopters (Note that in this manuscript, we would use these terms interchangeably.) through a cognitive process driven by a variety of contagion mechanisms. One family of mechanisms[1,8,9], commonly termed *simple contagion* in the social science literature[20–24], resembles biological epidemic processes; each interaction between a susceptible node and an infectious one may independently result in an infection event with a predetermined probability, leading to gradually evolving global adoption curves[25].

There is, however, plenty of empirical evidence suggesting that the simple contagion is not sufficient to explain the observed spreading phenomena in certain contexts[13], leading to the concept of an alternative mechanism, called *complex contagion*[24,26,27]. In this case, exposures are not independent, but peer pressure can impact in a non-linear way the individual infection probability, for example by accumulating influence towards an individual adoption threshold[11,15,28]. Depending on the model parameters, the complex contagion mechanism may lead to a cascading phenomenon[29], where mass infection emerges over a short period of time. This was first shown on networks by Watts[15], while several follow-up studies explored a rich family of similar phenomena in multi-layer[30–32], weighted[33,34] or temporal networks[18,35], demonstrating their relevance in real-world settings[36–40]. In this manuscript, we will use the threshold model[15] as a paradigmatic mechanism of complex contagion.

Simple and complex contagion capture network-based adoptions, however, social influence may not always spread on an observable network (e.g., advertisements, news or policy recommendations, etc.). We take such external influences into account by also considering a third mechanism, called *spontaneous adoption*[39,41,42]. Although spontaneous adoption is agnostic to the underlying network structure, infection patterns via the other two mechanisms depend non-trivially on several networks and dynamical

[1]Department of Network and Data Science, Central European University, Vienna, 1100, Austria. [2]National Laboratory for Health Security, HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, 1053, Hungary. [3]Network Science Institute, Northeastern University London, London, E1W 1LP, UK. [4]Department of Physics, Northeastern University, Boston, MA, 02115, USA. ✉e-mail: karsaim@ceu.edu

characters of an ego and its peers[43]. It has been shown that while simple contagion spreads easier on dense and degree-heterogeneous structures, with high-degree nodes early infected[44], these properties mitigate complex contagion as the threshold of high-degree nodes can be hardly reached[2,45]. Moreover, while weak ties connecting densely connected communities act as facilitating bridges for simple contagion[10], they slow down complex contagion cascades[24,46,47], 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[35,48,49] between individuals and their peers influence the adoption dynamics and the macroscopic dynamical outcome of the spreading process as a whole.

Acting alone, all social contagion mechanisms may lead to differentiable infection dynamics at the global scale. In this direction, while distinguishing mechanisms solely from the overall infection dynamics remains a challenge[50], recent methods combining spreading dynamics and network information[45], or considering the timing of peer stimuli[38], or the differences of the spreading dynamics by contagion type in mesoscale structures[51] led to promising results. 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[52–54]. Thus, each single adoption event may be driven by different mechanisms that jointly depend on personal factors[55,56] (heterogeneous susceptibility and predisposition), the properties of the item being adopted (Gladwell's stickiness[52]), and the particular context (environment, time of adoption, other external factors).

In this study, 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[57]. 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 the distinguishability of these mechanisms and to propose solutions that can be readily used in real-world scenarios aimed at understanding social contagion phenomena.

Results
Different mechanisms of social contagion

We study adoption processes on networks, where individuals are represented as a set of nodes V, and their interactions as a set of links E. The number of nodes connecting to a node i (i.e., the number of neighbours) is called the degree of i, denoted by k_i. The most common way to model propagation dynamics on a networked population is to assign a state to every node, which characterises its status with respect to the propagation[19]; a node is either susceptible (S), meaning that it has not yet been reached by the contagion process, or infected (I), if it has already been reached, and thus it can infect others.

We consider three infection mechanisms that can change the state of a node from susceptible to infected (cf. Fig. 1a).

As for the *simple contagion* (Sm) mechanism, we build on the Susceptible–Infected (SI) model, introduced first in epidemiology[58] and later to characterise the adoption of social behaviours[20–23]. In this model, a susceptible node can become independently infected with a fixed probability during each interaction with an infectious neighbour. Here, we assume that at every time step a susceptible node i could acquire an infection from each infectious node in its neighbourhood with its node-dependent probability $\beta_i \in [0, 1]$ (that could thus be considered alike a heterogeneous susceptibility). After a gradual contamination of the network, the macroscopic
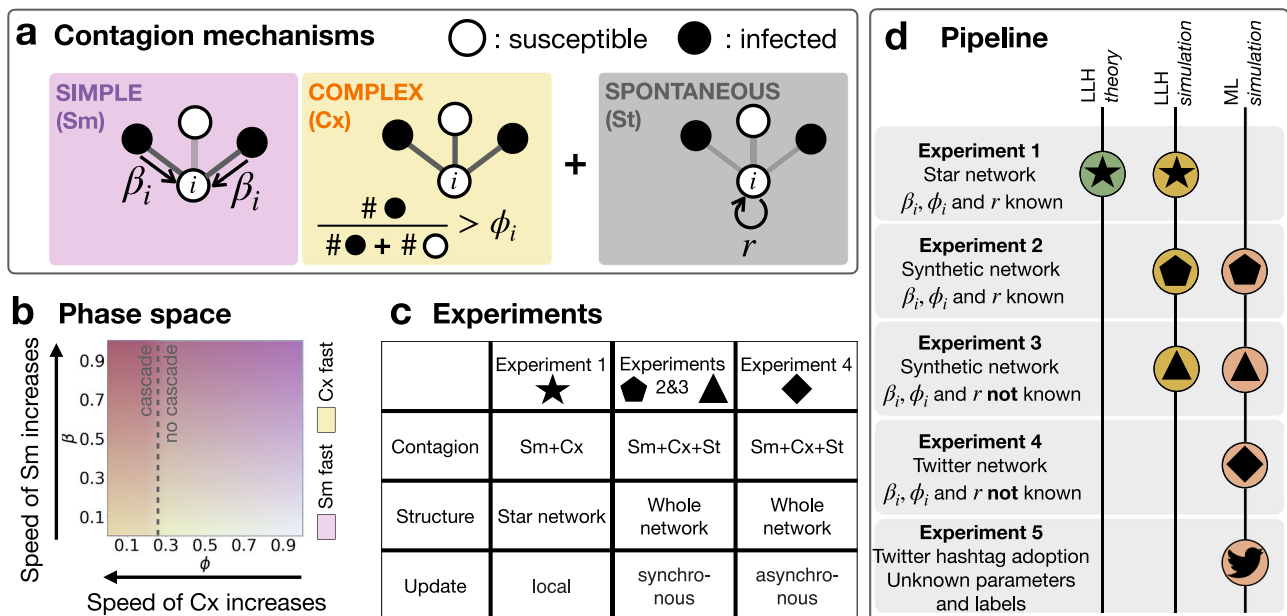

Fig. 1 | Overview of experimental setups. a Illustration of the three contagion mechanisms that are subject to inference: *simple contagion* parameterised by the adoption probability β_i through a single stimuli; *complex contagion* parameterised by the threshold ϕ_i of necessary fraction of adopter neighbours to induce the adoption event; and *spontaneous adoption* that occurs with probability r. **b** The parameter space (β, ϕ) and the speed dependence of the simple and complex contagion processes are shown as a schematic representation for illustration purposes. **c** The different experimental setups that include the considered contagion mechanisms, the complexity of the underlying network, and model update rules. **d** Schematic pipeline for the application of the log-likelihood (LLH) and random forest machine learning (ML) classification approaches to the different experiments.

steady state of an SI contagion process is reached when all nodes become infected. The *Complex contagion* (Cx) mechanism breaks the linearity of the contagion by introducing social reinforcement effects, often found in behavioural patterns: it is the *combined* influence arriving from the neighbours of an ego node, which triggers the adoption. Here, we consider this mechanism by employing a conventional deterministic threshold model introduced by Watts[15], where each susceptible node i becomes infected as soon as its fraction of infectious neighbours exceeds a preassigned intrinsic threshold $\phi_i \in [0, 1]$. This threshold model is known to exhibit rapid cascading behaviour if the necessary conditions on the average degree and the infection threshold are met[15].

Parameters β_i and ϕ_i are crucial in shaping the propagation dynamics. High values of β_i lead to faster adoption via Sm, while low values of ϕ_i accelerate the adoption rate via Cx, as individual thresholds become easier to reach (see Fig. 1b and also Supplementary Material 1).

Finally, we implement a third adoption mechanism called *spontaneous adoption* (St), which models external effects; every susceptible node becomes infected with probability r during any time steps of the process, independently of the state of its neighbours.

The backbone of the paper is a series of four experiments (Fig. 1c), where we tackle the problem of distinguishing simulated Sm, Cx and St 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 (Fig. 1d). The likelihood approach features theoretical guarantees and the possibility to include prior knowledge about the underlying processes[59]. However, likelihood-based approaches may not be robust if they cannot capture precisely the data from the assumed generative process[60]. 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.

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. To isolate the mechanism of the ego node only, we assume that all the neighbours undergo a spontaneous adoption (St mechanism), while the ego can adopt via simple or complex mechanisms, which are randomly assigned at the beginning of each simulation, as well as the β_i and ϕ_i parameters controlling the contagion of each ego node.

After simulating the contagion process for T timesteps, 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 timestep 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 (i.e. the state of the system at a given time only depends on the previous timestep), 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:

$$\mathcal{L}_i(\mathcal{X}) = \prod_{t=0}^{T} \mathbf{P}(\sigma_i(t+1)|\sigma_{i,\mathrm{nb}}(t), \mathcal{X}, \{\beta_i, \phi_i\}), \tag{1}$$

where $\sigma_{i,\mathrm{nb}}(t)$ denotes the trajectories of the ego node and of its neighbours. An observed adoption could then be attributed to the mechanism having the highest likelihood (more details are given in "Likelihood calculations" of the "Methods" section).

Assuming that the star networks have degrees k drawn from a binomial distribution, we display in the heatmap of Fig. 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 $\beta \to 1$ and $\phi \to 0$. 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 (Supplementary Material Fig. S1). 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 as

$$ACC(k, \beta, \phi, r) \approx 1 - \frac{1}{2} \left(\prod_{n=1}^{\lfloor k\phi \rfloor} \frac{p_n - p_n b_n}{b_n + p_n - p_n b_n} \right) b_{\lceil k\phi \rceil}, \tag{2}$$

with $p_n = 1 - (1-r)^{k-n}$ and $b_n = 1 - (1-\beta)^n$ (see the "Methods" section for the details of the calculation). Comparing the theoretically estimated accuracies from Eq. (2) (visualised in Fig. 2d) with the simulation outcomes (Fig. 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 obtained on the giant component of Erdös–Rényi random networks[61] 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. Similarly to Experiment 1, we randomly predetermine the contagion mechanism (simple or complex) for each node. 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
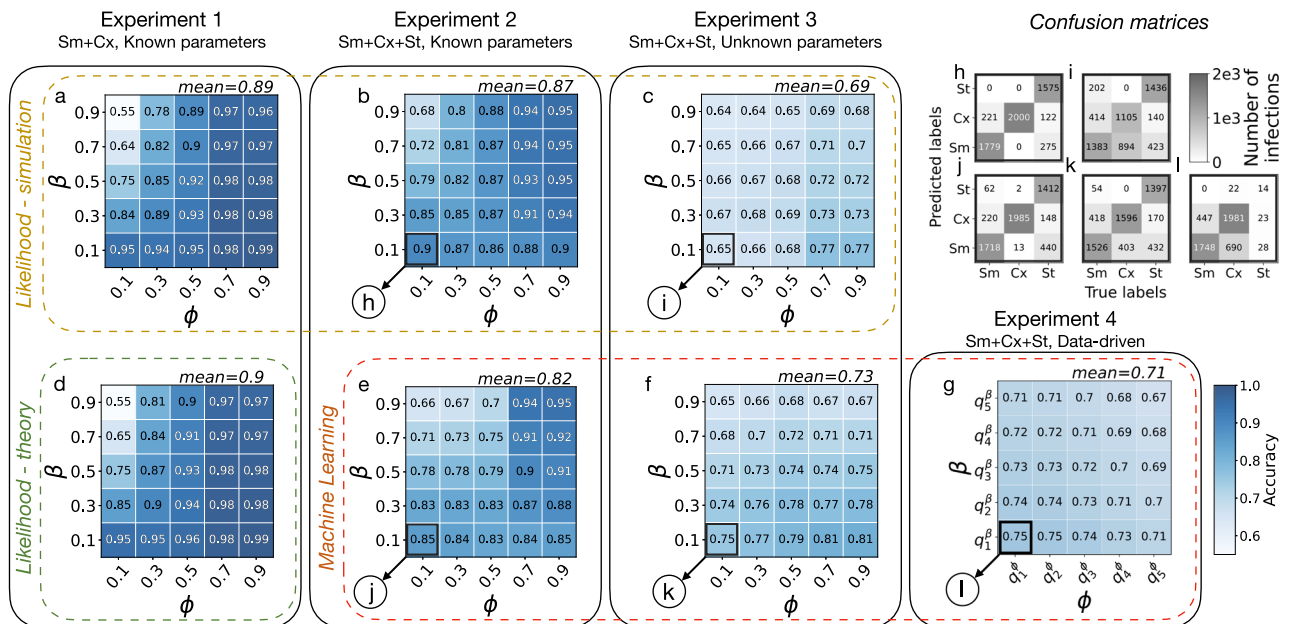
Fig. 2 | Comparison of the classification accuracies across the different methods and experiments. Classification accuracy values of the likelihood method (green rectangle (**d**) when it is obtained theoretically and yellow rectangle (**a–c**) when it is obtained by simulation) and of the random forest 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 $q_n^{\text{parameter}}$ represents the nth quintile of the parameter distribution. Panels **h–l** show the confusion matrices associated to the highlighted pairs (β, ϕ) from Experiments 2–4. 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.
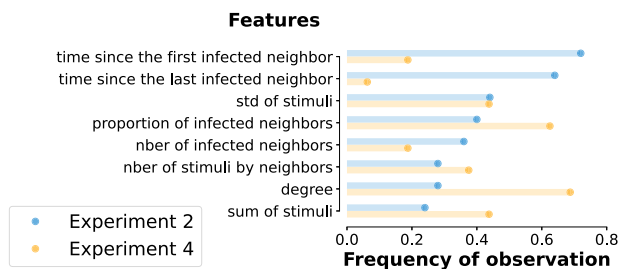

Fig. 3 | Feature importance for Experiments 2 and 4. Frequency of observation of the features used to train the random forest classifier among the top-3 most important ones across the full parameter space for Experiment 2 (blue) and Experiment 4 (orange). Frequencies are computed as the number of appearances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

on large networks with extreme Sm and Cx contagion parameters but continues spreading following linear dynamics. The modification also implies that, since nodes can be adopted via simple, complex or spontaneous mechanisms, our classification algorithms need to distinguish between the three hypotheses (see the "Methods" section).

In line with the approach of Experiment 1, we compute the likelihood that each adopter follows a specific contagion mechanism (see Eq. (1)) based on the trajectories of the ego nodes and their neighbours. Since the assumption on the independent adoption of the neighbours of ego does not hold anymore, the likelihood framework becomes an approximation (see the subsection "Likelihood calculations" of the "Methods" section for the detailed derivation). Nevertheless, accuracy values for the whole parameter space summarised in Fig. 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. After an extensive

classification model selection (cf. Supplementary Material 2), we selected a random forest approach as the consistently best-performing classifier. In order to strike a balance between performance and interpretability, we train random forest classifiers on the same synthetic dataset as above. After testing 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. Supplementary Material 3). 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 a random forest model using these input features for each adopted node that appeared during a simulated contagion with Sm and Cx with parameters β and ϕ. The random forest approach provides very similar results (see Fig. 2e) to the likelihood-based calculations (Fig. 2b), only with a slightly worse average accuracy of 0.82. According to the confusion matrices shown in Fig. 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 Supplementary Material 3). Interestingly, some feature subsets are consistently optimal across the full parameter space. This is reported in Fig. 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 also be easily interpreted within the modelling framework: the time since the first infected neighbour cannot be too high for Sm, as that would mean too many repeated stimuli without an infection event, while for the threshold-based Cx the time since the last infected neighbour has to be necessarily one.

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 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 (see the "Methods" section for more details).

Figure 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 (Fig. 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 Sm and Cx. 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. Interestingly, this solution provides slightly more accurate results (see Fig. 2f) than the likelihood method (see Fig. 2c), especially for low values of β. Reading the confusion matrices (in Fig. 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.

Note that we conducted Experiments 2 and 3 on various types of random networks including Erdős–Rényi[62] (presented above), Barabási–Albert[63], Watts–Strogatz[64] and stochastic block model[65] networks (see Supplementary Material 4) with very similar results. This suggests that the global network structure has limited impact on the local differentiation of contagion processes in each performed experiment.

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[66] (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 ref. 67). This mounts 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 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 of hashtags within the same topic. We choose to focus on the hashtag

Table 1 | Number of instances of contagion mechanisms inferred by the likelihood and random forest methods on the #GiletsJaunes Twitter dataset

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

#GiletsJaunes and its variants (We target every user who has posted one of those hashtags: #GiletsJaunes, #giletsjaunes, #Giletsjaunes, #GiletJaune, #Giletjaune, #giletjaune, #giletsjaune, #Giletsjaune, #GJ.), 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 weather 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 1 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. Indeed, both methods are made for the characteristics of synthetic data, assuming that a complex contagion adoption always occurs a time step after the infection of a neighbour (time since the last infected neighbour = 1). However, in real data, a delay could occur between the infection of a node and its actual observation, making the classification of the complex instances inaccurate and misclassifying it as simple or spontaneous contagion. This observation suggests that we need to relax our assumptions, especially on complex contagion, and introduce the possibility of delays in the adoption times into our models.

When it comes to empirical adoption data collected via social media, one of the largest biases is induced by the waiting time t_w[32], 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 Fig. 4a, 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 during this t_w time further exposures can appear that, in principle, could not even be 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[32,68]. In light of these observations, it is clear from the likelihood computations and from the feature, importance ranks shown in Fig. 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
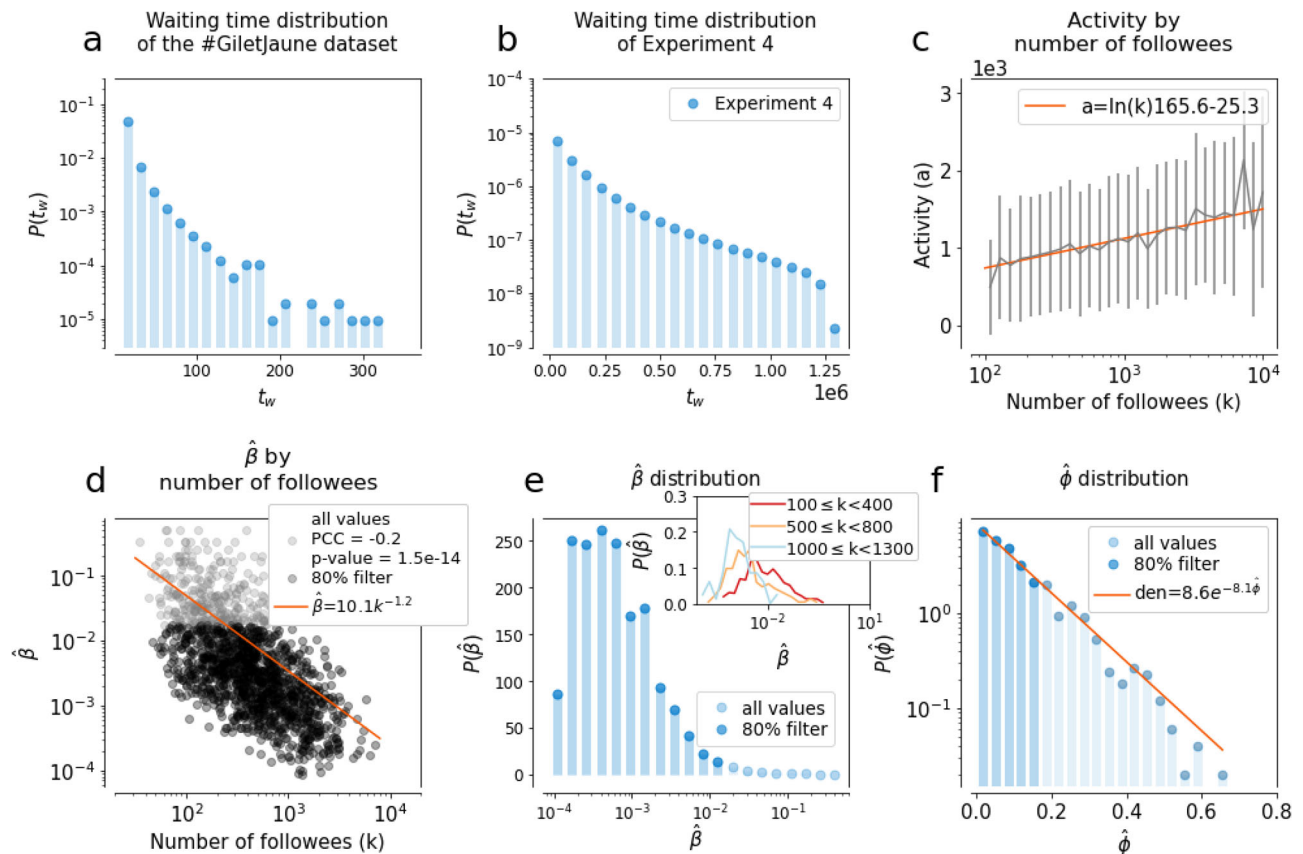
Fig. 4 | Parameter distributions and dependencies of Experiment 4 inferred from the #GiletsJaunes Twitter dataset. The waiting time distributions observed in the **a** #GiletsJaunes dataset and **b** in Experiment 4. **c** Correlation between the activities and degrees of nodes in the synthetic propagation 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 $\hat{\beta} = 1/(\text{numberstimuli})$ 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 80% of their smallest values.

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. We employ a connected and undirected sample of the follower Twitter network as the underlying structure for the contagion process (for more details about the network creation see subsection "Experiment" in the section "Methods", s). 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 $\hat{a}_i \in [0, 1]$ sampled from a truncated normal distribution with values constrained between 0 and 1, and an average activity that characterises nodes coming from the same degree group as node i (Fig. 4c). They are also attributed to a contagion process, either simple or complex, which determines their adoption mechanism. Further, nodes are endowed with parameters $\hat{\beta}_i$ or $\hat{\phi}_i$, respectively, sampled from the empirical distributions $P(\hat{\beta})$ and $P(\hat{\phi})$ shown in Fig. 4d, f. Since these distributions are broad, we filtered them and kept only samples from their lowest 80% (more details about sampling and filtering in the "Methods" section and Supplementary Material 5).

At every step, a node is selected with a probability proportional to its activity, modelling its action of posting. If the selected node is susceptible, we assume its post induces no influence on its neighbours. Once a node is infected via one of the considered mechanisms, it enters the aware state and no further stimuli are necessary for adoption —yet to be observed. The next time the node is selected for an interaction, it becomes detected. If a node is aware or detected, its posts are considered as influencing events to its neighbours. The resulting waiting time, measured for each infected node as the time between the aware and detected state, follows a broad distribution (Fig. 4b), similar to the empirical observations. More details about the model definition and evaluation are explained in "Methods" section.

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 Fig. 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 an average accuracy of 0.71. In this experiment, the spontaneous adoptions are the hardest to classify since they appear with a very low rate (see the confusion matrix Fig. 2 panel l and the Supplementary Material, Table S2). It is worth noticing that the importance of the features is different from the one previously shown for Experiment 2 (Fig. 3). While the feature *time since the last infected*
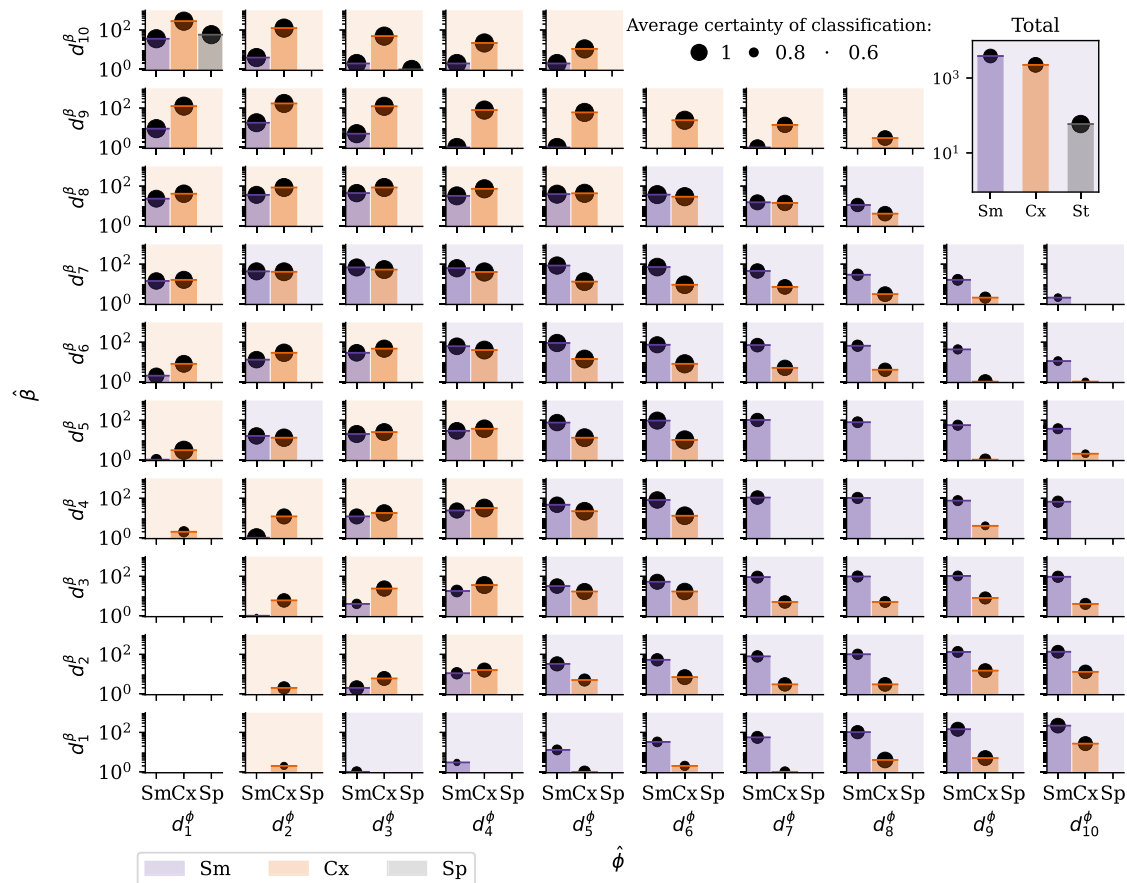
Fig. 5 | Classification of contagion mechanisms of the #GiletsJaunes Twitter dataset as the function of $\hat{\phi}$ (x-axis) and $\hat{\beta}$ (y-axis) parameters. The notation $d_n^{\text{parameter}}$ represents the nth deciles of the parameter distribution from the #Gilets-Jaunes dataset from Fig. 4. 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 $> d_5^{\phi}$, and as complex otherwise.

neighbours diminishes 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 (Fig. 3).

Classification of Twitter hashtags
Experiment 5. To conclude our case study on the Twitter dataset, we apply the trained models from Experiment 4 to the adoption cases of #Gilets-Jaunes and related hashtags. The inset of Fig. 5 shows that most adoption cases are classified as simple as opposed to complex. This suggests that more people adopt #GiletsJaunes through repeated influence from their contacts than through combined influence mechanisms. The less detected class is one of the 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 visualising the accuracy values on the (β, ϕ) phase space, we show in Fig. 5 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, Fig. 5 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 database in this parameter range.

Discussion
Our goal in this work 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[38,45,50], typically assuming that only one 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 question 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 the 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[43,69,70]. 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 levels, as also highlighted in other recent studies that tackled the inference problem in different contexts[71]. Nevertheless, even in these realistic settings, when mechanistic approaches seem to be out of reach, a random forest classifier trained on a carefully parametrised 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 from validating 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[36,72] even beyond pairwise exposures[73–75], as considered in a recent work[45]. 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[20]. One could also integrate homophily and conceive a model in which nodes of the same group are more likely to adopt through the same contagion process or nodes within the same mechanism have a higher probability of forming connections between themselves[76]. Such scenarios 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 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, which 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.

Methods
Experiments
To study the distinguishability of the Sm, Cx and Sp contagion processes we defined three experimental settings with increasing complexity:

Experiment 1—Classification on egocentric networks. In Experiment 1 we assume no underlying network structure to disseminate the spreading process but we operated only with isolated ego networks. We assume knowledge only about egos and their neighbours, that together defined 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 Erdös–Rényi networks that we used in Experiment 2. We assign to each ego-node a predetermined adoption class, simple or complex, with the corresponding parameter, respectively, β or ϕ. Further, we defined the same adoption probability r_{nb} 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 had been satisfied. We simulate this contagion dynamics on 100, 000 ego-networks, having 10,000 realisations for each parameter values β and ϕ taking values from {0.1, 0.3, 0.5, 0.7, 0.9} and with parameter $r_{nb} = 0.05$. In this setting, the classifier was informed by the β_i, ϕ_i and r parameter values for each instance i.

Experiment 2—Classification of random networks with known parameters. Experiment 2 is conducted on an Erdös–Rényi model network[61], with 1000 nodes and average degree 4. For comparison purposes, in Supplementary Material 4, we also demonstrate our results using Watts–Strogatz[64] and Barabási–Albert[63] model networks, stochastic block model networks[65], and a real Twitter mention network[77] 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. As in Experiment 1, we assign all nodes beforehand with a contagion process (Sm or Cx) 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. 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, where the process is terminated when 90% of the nodes become infected. 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.

Experiment 3—Classification of random networks with unknown parameters. Experiment 3 is aiming to solve the classification of the same contagion instances than Experiment 2 but without prior knowledge about the parameters of β_i, ϕ_i and r.

Experiment 4—Classification of real networks with known parameters. Experiment 4 is inspired by the activity-driven network model[78] 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[77] and concentrate on the

propagation of the hashtags related to the political movement called #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. 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.

Parameter sampling. First of all, in this setting each node is assigned with an activity, 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, we sample the assigned activity of each node with a normal distribution centred on the average number of tweets posted by each user corresponding to its degree.

Before inferring distributions of $\hat{\beta}$ and $\hat{\phi}$, we assign to each infected node an adoption process using the following heuristic. Events in which a susceptible node becomes infected without having any infectious neighbours at the time of transition are classified as spontaneous adoptions. Additionally, instances where the last event before the ego becomes infected is a new infected neighbour are classified as complex contagion. This classification is motivated by the fact that the newly infected neighbour increases the proportion of infected neighbours, potentially allowing the threshold ϕ to be overpassed. All other instances are classified as simple contagion. This pre-classification step reduces errors in inferring $\hat{\beta}$ (resp., $\hat{\phi}$) from complex (resp., simple) contagion instances, resulting in more accurate distributions.

The parameters are sampled for each node depending on the pre-assigned mechanism. For simple contagion, parameter values for $\hat{\beta}$ 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 a correlation with the node degree (see Fig. 4d), we 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 Fig. 4e 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 Fig. 4f) 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 Fig. 4f. 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 Supplementary Material.

Contagion model with waiting time. Beyond the realistic data-driven parametrisation of the network and adoption mechanisms, our main goal with this experiment is to simulate spreading scenarios to study the effects of waiting times between node adoption and its observation on the inference of spreading mechanisms. For this reason, we assume that every node of the network can be in one of the three following states: susceptible (not infected), aware (infected, but the infection cannot be observed) and detected (the infection can be observed). 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 of posting 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.

Experiment 5—Classification of hashtag adoption with unknown ground truth. Since we cannot obtain the contagion mechanisms as ground truth labels for real data, we use the trained model in Experiment 4 for the classification of the empirical adoption instances. With this model, which has been trained on data-driven model data closest to reality, we explore the whole $(\hat{\beta}, \hat{\phi})$ parameter space, as shown in Fig. 5. We then use this algorithm to analyse various social contagion processes by focusing on tweets with specific sets of hashtags corresponding to distinct topics. We use a dataset collected by[67] from Twitter, now known as X, a social media platform where users can follow each other and share brief posts, or tweets. This dataset spans from May 1, 2018, to May 31, 2019, and includes all tweets from 8527 selected users interested in the European election of 2019 (denoted as egos) and the accounts they follow (denoted as followees). In total, this comprises 1,844,978 timelines, representing the chronological history of tweets from these users. These tweets cover a range of topics, referenced with keywords called hashtags.

Among all the possible hashtags, we choose to focus on #GiletsJaunes and its variations (#GiletsJaunes, #giletsjaunes, #Giletsjaunes, #GiletJaune, #Giletjaune, #giletjaune, #giletjaune, #Giletsjaune, #GJ), which are linked to a political movement in France that causes social contagion on Twitter. Our first step is to identify users who have adopted one of these hashtags and then examine the tweets from their followees. We consider that individuals remember influences only from their recent past, thus we study tweets of the ego and its followees on the last week before the adoption. We define a user's degree by the number of followees who posted at least one tweet in the preceding week and we use event time instead of real time for our analysis. Event time counts the number of tweets by followees, regardless of whether they contain the hashtag. We define a stimulus as a tweet posted by followee containing the hashtag.

Likelihood calculations

The classification with the likelihood approach follows the same protocol for each experiment: we first compute the likelihood that a given observed adoption case has been caused by each mechanism, being simple, complex or spontaneous, and then we classify the adoption event into the category which maximises the likelihood.

Experiment 1. We determine the likelihood that a node i has been infected either through simple or complex contagion using Eq. (1), which expresses the likelihood of the entire process as a product of the likelihoods of each time step (Markov property). We call $\sigma_i(t)$ the state of a node i at time t, being 0 (S) or 1 (I). To compute the likelihood of observing the ego's state $\sigma_i(t+1)$ conditioned on its state and the states of the neighbours $\sigma_{i,\text{nb}}(t)$ in the previous timestep, we distinguish three cases:

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

In the 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

$$\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Sm, \beta) = \begin{cases} \prod_{j \in nb} (1-\beta)^{\sigma_j(t)} & 0 \to 0 \\ 1 - \prod_{j \in nb} (1-\beta)^{\sigma_j(t)} & 0 \to 1 \\ 1 & 1 \to 1 \end{cases}$$

where nb is the set of the neighbours of the ego.

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

$$\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Cx, \phi) = \begin{cases} \mathbb{1}(\sigma_{i,nb}(t)) & 0 \to 0 \\ 1 - \mathbb{1}(\sigma_{i,nb}(t)) & 0 \to 1 \\ 1 & 1 \to 1 \end{cases}$$

depending on whether the condition

$$\mathbb{1}(\sigma_{i,nb}(t)) = \Theta\left(\sum_j \sigma_j(t) A_{ij} - \phi \sum_j A_{ij}\right),$$

on whether 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 if positive, 0 otherwise.

Accuracy estimation for Experiment 1. In Experiment 1, the accuracies of the maximum-likelihood classification algorithm can be computed analytically across the phase space. 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

$$\frac{P(\hat{X} = Cx \mid X = Cx) + P(\hat{X} = Sm \mid X = Sm)}{2}.$$

Since for a node infected by the complex contagion, we always have $\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Cx, \phi) = 1$, 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

$$P(\hat{X} = Sm \mid X = Sm) = 1 - P(\hat{X} = Cx \mid X = Sm),$$

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:

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

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, the outcomes closely align with the accuracy values observed in the simulations (see Fig. 2, panel d).

Experiment 2—Classification with known parameters. The calculations of the likelihoods of Experiment 2 are similar to Experiment 1, but instead of two, they now involve three processes: simple, complex and spontaneous adoptions. For clarity, we divide those three processes into four scenarios:

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

$$\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Sm, \beta) = \tag{3}$$

$$\begin{cases} (1-r)\prod_{j \in nb}(1-\beta)^{\sigma_j(t)} & 0 \to 0 \\ (1-r)\left(1 - \prod_{j \in nb}(1-\beta)^{\sigma_j(t)}\right) & 0 \to 1 \\ 1 & 1 \to 1 \end{cases} \tag{4}$$

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

$$\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Sm \to St, \beta) = \tag{5}$$

$$\begin{cases} (1-r)\prod_{j \in nb}(1-\beta)^{\sigma_j(t)} & 0 \to 0 \\ r\prod_{j \in nb}(1-\beta)^{\sigma_j(t)} & 0 \to 1 \\ 1 & 1 \to 1 \end{cases} \tag{6}$$

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

$$\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Cx, \beta) = \tag{7}$$

$$\begin{cases} (1-r)\left(1 - \mathbb{1}(\sigma_{i,nb}(t))\right) & 0 \to 0 \\ \mathbb{1}(\sigma_{i,nb}(t)) & 0 \to 1 \\ 1 & 1 \to 1 \end{cases} \tag{8}$$

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

$$\mathcal{L}(\sigma_i(t+1)|\sigma_{i,nb}(t), Cx \to St, \beta) = \tag{9}$$

$$\begin{cases} (1-r)\left(1 - \mathbb{1}(\sigma_{i,nb}(t))\right) & 0 \to 0 \\ r\left(1 - \mathbb{1}(\sigma_{i,nb}(t))\right) & 0 \to 1 \\ 1 & 1 \to 1 \end{cases} \tag{10}$$

Experiment 3—Classification with unknown parameters. In this case, we assume that parameter values are not known for the classifier and we employ the same formulas as in Eqs. (3)–(10) used for classifying contagion instances from Experiment 2 with known parameters. However, here the parameters β and ϕ and r are no longer the true values but are instead inferred from the modelled spreading process: $\hat{\beta}$ as the inverse of the number of stimuli, $\hat{\phi}$ as the proportion of infected neighbours and $\hat{r}$ 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.

Random forest classification

Experiment 2—Classification with known parameters. We train 25 random forest algorithms, one for each pair of (β, ϕ) by sampling 18,000 instances from Experiment 2, with 6000 contagion cases from each category. Then we test the models on a set containing 6000 instances (2000 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. The use of the Gini function or the entropy function is determined by grid search.

Experiment 3—Classification with unknown parameters. We train a unique random forest model on a sample of Experiment 2, which contains 18,000 instances in total (6000 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.

Data availability

The filtered network structure and distributions of the inferred parameters are shared at an open data repository at https://github.com/ElsaA05/DistinguishSimpleComplex/tree/main/data. and at https://zenodo.org/records/14518761.

Code availability

We made available the code of the different experiments at an open source code repository at https://github.com/ElsaA05/DistinguishSimpleComplex/tree/main/analysis and at https://zenodo.org/records/14518761.

References

1. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591 (2009).
2. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge university Press, 2008).
3. Jose, S. Covid vaccine and generation z—a study of factors influencing adoption. *Young Consum.* **23**, 16–32 (2022).
4. Christakis, N. A. & Fowler, J. H. The collective dynamics of smoking in a large social network. *N. Engl. J. Med.* **358**, 2249–2258 (2008).
5. Centola, D. & Lord, T. *The Truth About Behavioral Change* (MIT Sloan Management Review, 2018).
6. Bass, F. M. A new product growth for model consumer durables. *Manag. Sci.* **15**, 215–227 (1969).
7. Rogers, E. M., Singhal, A. & Quinlan, M. M. Diffusion of innovations. In *An Integrated Approach to Communication Theory and Research* 432–448 (Routledge, 2014).
8. Maki, D. P. & Thompson, M. *Mathematical Models and Applications* (Prentice Hall, Englewood Cliffs, NJ, 1973).
9. Daley, D. J. & Kendall, D. G. Stochastic rumours. *IMA J. Appl. Math.* **1**, 42–55 (1965).
10. Granovetter, M. S. The strength of weak ties. *Am. J. Sociol.* **78**, 1360–1380 (1973).
11. Schelling, T. C. Dynamic models of segregation. *J. Math. Sociol.* **1**, 143–186 (1971).
12. Chen, W., Castillo, C. & Lakshmanan, L. V. *Information and Influence Propagation in Social Networks* (Springer Nature, 2022).
13. Centola, D. The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197 (2010).
14. Moreno, Y., Nekovee, M. & Pacheco, A. F. Dynamics of rumor spreading in complex networks. *Phys. Rev. E* **69**, 066130 (2004).
15. Watts, D. J. A simple model of global cascades on random networks. *Proc. Natl Acad. Sci. USA* **99**, 5766–5771 (2002).
16. Holme, P. & Saramäki, J. Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).
17. Masuda, N. & Holme, P. *Introduction to Temporal Network Epidemiology* (Springer, 2017).
18. Unicomb, S., Iñiguez, G., Gleeson, J. P. & Karsai, M. Dynamics of cascades on burstiness-controlled temporal networks. *Nat. Commun.* **12**, 133 (2021).
19. Gleeson, J. P. Binary-state dynamics on complex networks: pair approximation and beyond. *Phys. Rev. X* **3**, 021004 (2013).
20. Gleeson, J. P., O'Sullivan, K. P., Baños, R. A. & Moreno, Y. Effects of network structure, competition and memory time on social spreading phenomena. *Phys. Rev. X* **6**, 021019 (2016).
21. Crane, R. & Sornette, D. Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl Acad. Sci. USA* **105**, 15649–15653 (2008).
22. Sreenivasan, S., Chan, K. S., Swami, A., Korniss, G. & Szymanski, B. K. Information cascades in feed-based networks of users with limited attention. *IEEE Trans. Netw. Sci. Eng.* **4**, 120–128 (2016).
23. Gleeson, J. P., Ward, J. A., O'sullivan, K. P. & Lee, W. T. Competition-induced criticality in a model of meme popularity. *Phys. Rev. Lett.* **112**, 048701 (2014).
24. Centola, D. & Macy, M. Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**, 702–734 (2007).
25. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925 (2015).
26. Guilbeault, D., Becker, J. & Centola, D. Complex contagions: a decade in review. In: Lehmann, S. & Ahn, Y. Y. (eds) *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-world Social Networks* 3–25 (Springer Nature, 2018).
27. Galesic, M. et al. Beyond collective intelligence: collective adaptation. *J. R. Soc. interface* **20**, 20220736 (2023).
28. Granovetter, M. Threshold models of collective behavior. *Am. J. Sociol.* **83**, 1420–1443 (1978).
29. Gleeson, J. P. Cascades on correlated and modular random networks. *Phys. Rev. E* 046117 (2008).
30. de Arruda, G. F., Rodrigues, F. A. & Moreno, Y. Fundamentals of spreading processes in single and multilayer complex networks. *Phys. Rep.* **756**, 1–59 (2018).
31. Min, B. & San Miguel, M. Threshold cascade dynamics in coevolving networks. *Entropy* **25**, 929 (2023).
32. Karsai, M., Iñiguez, G., Kikas, R., Kaski, K. & Kertész, J. Local cascades induced global contagion: how heterogeneous thresholds, exogenous effects, and unconcerned behaviour govern online adoption spreading. *Sci. Rep.* **6**, 27178 (2016).
33. Unicomb, S., Iñiguez, G. & Karsai, M. Threshold driven contagion on weighted networks. *Sci. Rep.* **8**, 3094 (2018).
34. Li, X., Wang, P., Xu, X.-J. & Xiao, G. Universal behavior of the linear threshold model on weighted networks. *J. Parallel Distrib. Comput.* **123**, 223–229 (2019).
35. Karimi, F. & Holme, P. Threshold model of cascades in empirical temporal networks. *Physica A* **392**, 3476–3483 (2013).
36. Hodas, N. O. & Lerman, K. The simple rules of social contagion. *Sci. Rep.* **4**, 4343 (2014).

37. Vasconcelos, V. V., Levin, S. A. & Pinheiro, F. L. Consensus and polarization in competing complex contagion processes. *J. R. Soc. Interface* **16**, 20190196 (2019).

38. Mønsted, B., Sapieżyński, P., Ferrara, E. & Lehmann, S. Evidence of complex contagion of information in social media: an experiment using Twitter bots. *PLoS ONE* **12**, e0184148 (2017).

39. Karsai, M., Iniguez, G., Kaski, K. & Kertész, J. Complex contagion process in spreading of online innovation. *J. R. Soc. Interface* **11**, 20140694 (2014).

40. Borge-Holthoefer, J. et al. Structural and dynamical patterns on online social networks: the spanish May 15th movement as a case study. *PLoS ONE* **6**, e23883 (2011).

41. Hill, A. L., Rand, D. G., Nowak, M. A. & Christakis, N. A. Infectious disease modeling of social contagion in networks. *PLoS Comput. Biol.* **6**, e1000968 (2010).

42. Toole, J. L., Cha, M. & González, M. C. Modeling the adoption of innovations in the presence of geographic and media influences. *PLoS ONE* **7**, e29528 (2012).

43. Ugander, J., Backstrom, L., Marlow, C. & Kleinberg, J. Structural diversity in social contagion. *Proc. Natl Acad. Sci. USA* **109**, 5962–5966 (2012).

44. Yang, H., Tang, M. & Gross, T. Large epidemic thresholds emerge in heterogeneous networks of heterogeneous nodes. *Sci. Rep.* **5**, 13122 (2015).

45. Cencetti, G., Contreras, D. A., Mancastroppa, M. & Barrat, A. Distinguishing simple and complex contagion processes on networks. *Phys. Rev. Lett.* **130**, 247401 (2023).

46. Weng, L., Menczer, F. & Ahn, Y.-Y. Predicting successful memes using network and community structure. In: Adar, E., Resnick, P., Choudhury De, M., Hogan, B. & Alice, Oh. (eds) *Proceedings of the International AAAI Conference on Web and Social Media* Vol. 8, 535–544 (AAAI, 2014).

47. Nematzadeh, A., Ferrara, E., Flammini, A. & Ahn, Y.-Y. Optimal network modularity for information diffusion. *Phys. Rev. Lett.* **113**, 088701 (2014).

48. Karsai, M. et al. *Bursty Human Dynamics* (Springer, 2018).

49. Masuda, N., Miller, J. C. & Holme, P. Concurrency measures in the era of temporal network epidemiology: a review. *J. R. Soc. Interface* **18**, 20210019 (2021).

50. Hébert-Dufresne, L., Scarpino, S. V. & Young, J.-G. Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement. *Nat. Phys.* **16**, 426–431 (2020).

51. Weng, L., Menczer, F. & Ahn, Y.-Y. Virality prediction and community structure in social networks. *Sci. Rep.* **3**, 1–6 (2013).

52. Gladwell, M. *The Tipping Point: How Little Things Can Make a Big Difference* (Little, Brown, 2006).

53. Aral, S. & Nicolaides, C. Exercise contagion in a global social network. *Nat. Commun.* **8**, 14753 (2017).

54. Aiyappa, R., Flammini, A. & Ahn, Y.-Y. Emergence of simple and complex contagion dynamics from weighted belief networks. *Sci. Adv.* **10**, eadh4439 (2024).

55. State, B. & Adamic, L. The diffusion of support in an online social movement: Evidence from the adoption of equal-sign profile pictures. In: Cosley, D., Forte, A., Ciolfi, L. & McDonald D. (eds) *Proc. 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* 1741–1750 (Association for Computing Machinery, 2015).

56. Karampourniotis, P. D., Sreenivasan, S., Szymanski, B. K. & Korniss, G. The impact of heterogeneous thresholds on social contagion with multiple initiators. *PLoS ONE* **10**, e0143020 (2015).

57. Min, B. & San Miguel, M. Competing contagion processes: Complex contagion triggered by simple contagion. *Sci. Rep.* **8**, 10422 (2018).

58. Stattner, E. & Vidot, N. Social network analysis in epidemiology: current trends and perspectives. In: Abraham, A. (ed) *2011 Fifth International Conference on Research Challenges in Information Science* 1–11 (IEEE, 2011).

59. Price, L. F., Drovandi, C. C., Lee, A. & Nott, D. J. Bayesian synthetic likelihood. *J. Comput. Graph. Stat.* **27**, 1–11 (2018).

60. Hansen, M. H. & Yu, B. Model selection and the principle of minimum description length. *J. Am. Stat. Assoc.* **96**, 746–774 (2001).

61. Rényi, A. et al. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60 (1960).

62. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47 (2002).

63. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).

64. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).

65. Lee, C. & Wilkinson, D. J. A review of stochastic block models and extensions for graph clustering. *Appl. Netw. Sci.* **4**, 1–50 (2019).

66. de Oliveira, J. F., Marques-Neto, H. T. & Karsai, M. Measuring the effects of repeated and diversified influence mechanism for information adoption on Twitter. *Soc. Netw. Anal. Min.* **12**, 16 (2022).

67. de Oliveira, J. F., Marques-Neto, H. T. & Karsai, M. Information adoption via repeated or diversified social influence on Twitter. In: Atzmuller, M., Coscia, M. & Missaoui, R. (eds) *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* 237–241 (Institute of Electrical and Electronics Engineers Inc., 2020).

68. Wang, W., Stanley, H. E. & Braunstein, L. A. Effects of time-delays in the dynamics of social contagions. *N. J. Phys.* **20**, 013034 (2018).

69. Contreras, D. A., Cencetti, G. & Barrat, A. Infection patterns in simple and complex contagion processes on networks. *PLoS Comput. Biol.* **20**, e1012206 (2024).

70. Valente, T. W. & Vega Yon, G. G. Diffusion/contagion processes on social networks. *Health Educ. Behav.* **47**, 235–248 (2020).

71. St-Onge, G., Hébert-Dufresne, L. & Allard, A. Nonlinear bias toward complex contagion in uncertain transmission settings. *Proc. Natl Acad. Sci. USA* **121**, e2312202121 (2024).

72. Ruan, Z., Iniguez, G., Karsai, M. & Kertész, J. Kinetics of social contagion. *Phys. Rev. Lett.* **115**, 218702 (2015).

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

74. St-Onge, G. et al. Influential groups for seeding and sustaining nonlinear contagion in heterogeneous hypergraphs. *Commun. Phys.* **5**, 25 (2022).

75. Ferraz de Arruda, G., Petri, G., Rodriguez, P. M. & Moreno, Y. Multistability, intermittency, and hybrid transitions in social contagion models on hypergraphs. *Nat. Commun.* **14**, 1375 (2023).

76. Diaz-Diaz, F., San Miguel, M. & Meloni, S. Echo chambers and information transmission biases in homophilic and heterophilic networks. *Sci. Rep.* **12**, 9350 (2022).

77. Unicomb, S., Iñiguez, G., Kertész, J. & Karsai, M. Reentrant phase transitions in threshold driven contagion on multiplex networks. *Phys. Rev. E* **100**, 040301 (2019).

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

Acknowledgements

Author contributions

All authors contributed to the development of the research design. E.A. performed the numerical simulations and data analysis. E.A., G.Ó., and I.I. developed the statistical analysis. All authors wrote the first draft of the

manuscript, interpreted the results, and edited and approved the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44260-025-00034-2.

Correspondence and requests for materials should be addressed to Márton Karsai.

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

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