Temporal dynamics of coordinated online behavior: Stability, archetypes, and influence

Serena Tardelli[a] (ID), Leonardo Nizzoli[a], Maurizio Tesconi[a], Mauro Conti[b], Preslav Nakov[c], Giovanni Da San Martino[b], and Stefano Cresci[a,1] (ID)

Large-scale online campaigns, malicious or otherwise, require a significant degree of coordination among participants, which sparked interest in the study of *coordinated online behavior*. State-of-the-art methods for detecting coordinated behavior perform *static* analyses, disregarding the temporal dynamics of coordination. Here, we carry out a *dynamic* analysis of coordinated behavior. To reach our goal, we build a multiplex temporal network and we perform dynamic community detection to identify groups of users that exhibited coordinated behaviors in time. We find that i) coordinated communities (CCs) feature variable degrees of temporal instability; ii) dynamic analyses are needed to account for such instability, and results of static analyses can be unreliable and scarcely representative of unstable communities; iii) some users exhibit distinct archetypal behaviors that have important practical implications; iv) content and network characteristics contribute to explaining why users leave and join CCs. Our results demonstrate the advantages of dynamic analyses and open up new directions of research on the unfolding of online debates, on the strategies of CCs, and on the patterns of online influence.

coordinated behavior | dynamic networks | social media | online influence | stability

Online platforms offer unprecedented opportunities to organize and carry out large-scale activities, defying the constraints that limit physical interactions. A wide range of activities benefit from such opportunities, including legitimate information campaigns and political protests (1–3), as well as potentially nefarious ones such as disinformation campaigns and targeted harassment (4–8). A common characteristic of large-scale online campaigns is the significant degree of *coordination* among the users involved, which is needed to spread content widely and to let the campaigns obtain significant outreach, ultimately ensuring their success (9).

Due to the pervasiveness of coordinated online behavior and its relevance for the effectiveness of both benign and malicious campaigns, the topic gained much scholarly attention. For example, a few methods were recently proposed for detecting coordinated communities (CCs) and for measuring the extent of coordination between users (9–11). These studies define coordination as an unexpected or exceptional similarity between the actions of two or more users. The analysis is carried out by building a user similarity network based on common user activities (e.g., co-retweets) and by studying it with network science techniques. A limitation of the existing methods is that they are based on *static* analyses. For example, refs. 9–11 each build a single aggregated network that encodes user behaviors occurred throughout many weeks. However, online behaviors are *dynamic* (i.e., time-varying) (12, 13). As such, aggregating behaviors over many weeks (or months) of time represents an oversimplification that risks overshadowing important temporal dynamics. On the contrary, a minority of detection methods solely model time (14, 15), disregarding other important facets of online coordination such as the interactions between users. Due to the above methodological limitations, the temporal dynamics of coordinated behavior are, so far, essentially unexplored. In addition, the majority of the existing literature on coordinated behavior, especially from the area of computer science, focus on *inorganic* and *malicious* coordination. This is because of the nefarious consequences that such forms of coordination have on the online environments and even on the society at large, which make inorganic and malicious coordination worthy of particular attention. However, both foundational works on *offline* coordination (16) as well as more recent works on online coordination (9, 17, 18) remark that coordination can also be *implicit* rather than explicit, and *spontaneous* and *emergent* rather than intentional, inorganic, and organized. In light of this literature, here we refer to coordinated behavior in an intentionally broad and unbiased way. Our choice to encompass all instances of online coordination—malicious and inorganic, or otherwise—allows us to investigate both harmful coordinated groups and neutral ones.

Significance

Social media campaigns shape public opinion in various domains of society. Coordinated behavior is at the heart of successful campaigns, enabling outreach through effective content spreading. This study focuses on the temporal dynamics of coordinated campaigns and characterizes their influence on Twitter during two recent major elections. Analyzing the temporal changes of online behavior uncovers complex dynamic patterns that evolve over time. We reveal different user behaviors and archetypes, measuring how they affect and influence the online environment. Investigating the temporal dimension of online coordination reveals the dynamics of online debates, the strategies of coordinated communities, and the patterns of online influence, with major practical implications for research and policy of online platforms and for the society at large.

Therefore, in addition to providing contributions to the study of online information manipulation, our work also contributes to the understanding of human dynamics and online interactions, providing interesting findings from both computational and social standpoints.

Contributions. We carry out a dynamic analysis of coordinated behavior. Instead of working with static networks, we build and analyze a *multiplex temporal network* (19) and we leverage state-of-the-art dynamic community detection algorithms (20) to find groups of users that exhibited coordinated behaviors in time. We apply our method to two Twitter datasets (21, 22) of politically polarized discussions covering the run up to the 2019 UK general elections (9, 23) and the 2020 USA presidential elections. In addition, we validate our method on a third Twitter dataset (24) featuring known instances of malicious accounts involved in a large information operation (see *Materials and Methods* for details on the datasets and our methodology). We compare the results of our dynamic analysis to the static ones, demonstrating the advantages of the former. Most importantly however, our innovative approach opens up additional directions of research and allows answering to the following research questions:

RQ1. To what extent are CCs stable over time? So far, coordinated behavior has been studied with static methods. However, some CCs might exhibit temporal instability, meaning that, through time, a significant portion of members leave the community while others join it. In this case, time-aggregated static analyses would yield unreliable results by overshadowing the temporal variations. Here, we compare static and dynamic results, and we shed light on the stability of CCs.

RQ2. What are the temporal dynamics of user behavior with respect to the evolving CCs? In case a certain degree of instability existed, some users would belong to different CCs at different points in time. Here, we investigate the dynamics with which user membership to CCs changes through time and their implications on the effectiveness of online campaigns.

RQ3. Why do users belong to different CCs through time? Different patterns of user membership to CCs are indicative of markedly different situations. For instance, users who remain in the same community for a long time might be strong supporters of that community. Conversely, users who abandon a community in favor of another might have been disappointed by the former or persuaded by the latter. In any case, investigating the possible reasons for user shifts between CCs (or lack thereof) is a novel direction of research with important practical implications (e.g., the study of online influence). Here, we analyze users who exhibit different behaviors and we compare them to their respective CCs, gaining insights into why users change community.

Based on the results to the aforementioned RQs, our main contributions are summarized as follows:

- We carry out a dynamic analysis of coordinated online behavior.
- We show that the communities involved in both the UK 2019 and the USA 2020 electoral debates featured variable degrees of *instability*, which motivates dynamic analyses.
- We find that the majority of user shifts from a community to another occurred between *similar* (like-minded) communities, while only a minority involved very different communities.
- We define and characterize three archetypes of users with markedly different behaviors, namely, i) *stationary*, ii) *influenced*, and iii) *volatile* users.
- We find that *content* and *network* characteristics are useful for understanding why users move between communities.

Related Work

The majority of existing approaches for detecting coordinated behavior are based on network science. These works model common activities between users (e.g., co-retweets, temporal and linguistic similarities, etc.) to build user similarity networks and to subsequently analyze them, for example by means of community detection algorithms (10, 11, 18). The typical output of these methods is a network where the CCs are identified. Some network-based methods do not only detect coordination as in a binary classification task but also quantify the extent of coordination between users and communities, thus providing more nuanced results (9). Notably, all aforementioned methods build static networks and employ static community detection algorithms. A few works focused on describing and even predicting temporal changes in online community structure. However, these leveraged explicit communities, such as the scholars belonging to a specific scientific community (25), or the relationship between gamers and teams in certain team-based online games (26). In our work instead, the online communities are not known in advance. Other significant differences with respect to the existing works are related to the temporal granularity of the analysis. For example, some studies measured user migrations between platforms after major events, such as the bans of toxic communities from a certain platform (27, 28). However, such analyses only considered before/after scenarios, without providing a fine-grained temporal modeling of the user migrations nor a nuanced network representation of the communities involved in the event. So far, the dynamic analysis of coordinated behavior is essentially unexplored. To this end, however, several advances were proposed to model time-varying behaviors with multiplex temporal networks (19) and to employ dynamic community detection algorithms (20). Our work applies these techniques to the study of coordinated behavior, thus moving beyond the current state-of-the-art. In addition to network-based approaches, others proposed to detect coordination with temporal point processes where user activities are modeled as the realization of a stochastic process (14, 15). These methods are capable of modeling the latent influence between the coordinated accounts, their strongly organized nature, and possible prior available knowledge. Finally, others adopted traditional feature engineering approaches to find similarities between users (3), or focused on specific user behaviors such as URL sharing (17).

Once coordinated behaviors are detected, subsequent efforts are devoted to characterizing CCs. Characterization tasks are typically aimed at distinguishing between malicious (e.g., disinformation networks) and genuine (e.g., fandoms, activists) forms of coordination (23, 29–31). This can be achieved by analyzing the content shared by the CCs, as done by Hristakieva et al. (23) that estimated the amount of shared propaganda. Others analyzed the structural properties of the coordination networks, finding differences between malicious and benign CCs (9, 29). Finally, another interesting direction of research revolves around estimating the influence that CCs have on other users. However, existing results in this area are still scant and contradictory. For example, Cinelli et al. (32) studied network properties of information cascades on Twitter, finding that coordinated users have a strong influence on the noncoordinated ones that participate in the same cascade. Contrarily, Sharma et al. (14) found that coordinated users have a strong influence on other coordinated users and only a small influence on noncoordinated ones. Here, we show that temporal coordination networks are a valuable tool toward assessing the influence that CCs exert on the users in a network.

Results

The application of our method (see *Materials and Methods* for a detailed discussion of our methodological approach) to the UK 2019 electoral debate produced a multiplex temporal network G comprising 277 K nodes and 4.1M edges. On average, a layer $G_i \in G$ contains 11 K nodes ($\sigma = 169$) and 164 K edges ($\sigma = 19$K). In total, $\simeq 600$ different CCs were part of G for at least one time window. The analysis of the USA 2020 debate resulted in a multiplex temporal network G comprising 526 K nodes and 7.5M edges. On average, a layer $G_i \in G$ contains 21 K nodes ($\sigma = 2$ K) and 303 K edges ($\sigma = 79$ K). In total, $\simeq 2$ K different CCs were part of G for at least one time window. Out of all the CCs identified in both the UK 2019 and the USA 2020 datasets, only a limited number persisted for all of the time. These also include the vast majority of all users in the networks. For this reason, in the following, we present detailed results only for the top-10 largest CCs from both datasets (*SI Appendix*, Table S3).

Dynamic vs. Static Communities. We first present the results of our dynamic analyses, and we subsequently compare them with the static ones.

Dynamic UK 2019 communities. The dynamic communities that dominated the 2019 UK electoral debate are as follows:

- LAB1: Laborists that supported the party and its leader Jeremy Corbyn, in addition to traditional Labor themes such as healthcare and climate change.
- LAB2: Another community of laborists. Their themes were similar to those of LAB1, but their members exhibited different temporal behaviors (*SI Appendix*, Fig. S1).
- RCH: Users who spread the laborist manifesto and urged others to register for the vote. RCH is mostly similar to LAB2 but, again, featured different temporal behaviors (*SI Appendix*, Fig. S1).
- B60: Activists that used the electoral debate to protest against a pension age equalization law to the detriment of women born in the 1950s (33).
- TVT: A community comprising multiple political parties that pushed for a tactical vote in favor of the laborists, to oppose the conservative party and to stop the Brexit.
- SNP: Supporters of the Scottish National Party (SNP), in favor of the Scottish independence from the UK.
- ASE: Conservative supporters that were mainly engaged in attacking the labor party and Jeremy Corbyn for anti-semitism (34).
- SNPO: Users that debated and opposed Scotland's intention of pushing for a second independence referendum.
- BRX: A community in support of the former Brexit Party.
- CON: Conservatives that supported the party and its leader Boris Johnson, as well as Brexit.

Adding to the content analysis of the themes discussed by each community, we also characterize CCs by their political orientation (*Materials and Methods*). Fig. 1*A* shows the position of each CC within the continuous political spectrum. Overall, the dynamic CCs identified with our method are in line with the 2019 UK political landscape and with the development of the electoral debate (35). CCs represent both large and strongly polarized parties (e.g., laborists and conservatives), as well as smaller ones that teamed up against conservatives with tactical voting. As shown in Fig. 1*A*, each CC also appears to be correctly positioned within the 2019 UK political spectrum, with LAB1, LAB2, and CON holding the extremes of the spectrum, and parties such as liberal democrats and Scottish nationalists (TVT and SNP, respectively) laying toward the middle.
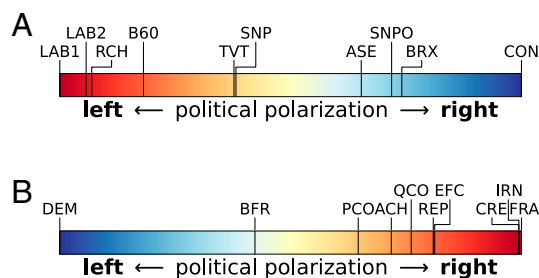

Fig. 1. Position of the CCs in the political spectrum. (*A*) UK 2019. (*B*) USA 2020. The color scheme mirrors the political or ideological colors of the corresponding country.

Dynamic USA 2020 communities. We again characterize the top dynamic communities by analyzing their use of hashtags through time (*SI Appendix*, Fig. S1). The dynamic communities that dominated the 2020 USA electoral debate are as follows:

- DEM: Democrats that supported the party and its leader Joe Biden.
- BFR: A community engaged in promoting Biafra's independence from Nigeria while also supporting Trump (36).
- PCO: Users engaged in Pandemic conspiracy theories and other controversial narratives about the China government responsibility in the COVID-19 pandemic.
- QCO: Users engaged in the QAnon conspiracy theory (37).
- ACH: Hong Kong protesters against China's regime and in support of Trump (38).
- EFC: Users discussing the allegation of election fraud due to postal ballots.
- REP: Republicans that supported the party and its leader Donald Trump.
- FRA: Far-right French users in support of Trump and the voting fraud accusations.
- CRE: A community in support of the Republican party, engaging in conspiracy theories.
- IRN: Users in support of the *Restart* movement, an Iranian political opposition group, which supports Trump for his policy against the Iranian regime (39).

We compute the political polarization of each CC with the same technique adopted for the UK 2019 CCs (*Materials and Methods*). Fig. 1*B* shows the position of each CC within the continuous political spectrum. The dynamic CCs identified with our method are consistent with the political environment in the run up to the 2020 USA presidential elections (40). The main differences between the results obtained for the USA 2020 versus the UK 2019 dataset are due to the inherently different online political environment at the time of the two elections. Specifically, while the online debate about the UK 2019 political election involved communities scattered throughout the whole political spectrum, as shown in Fig. 1*A*, the one about the USA 2020 election involved an overwhelming majority of Republican-aligned communities and only an isolated community of Democrat users, as shown in Fig. 1*B*. This result is also confirmed when analyzing the topic-based pairwise similarities between CCs (*SI Appendix*, Fig. S2).

Comparison with static CCs. Here, we compare the results of our dynamic analyses with those obtained with static analyses on each dataset. Previous studies (9) found 7 static CCs in the UK 2019 dataset, while we found 10 static CCs in the USA 2020 dataset. Fig. 2 presents a heatmap of the overlap between the static CCs
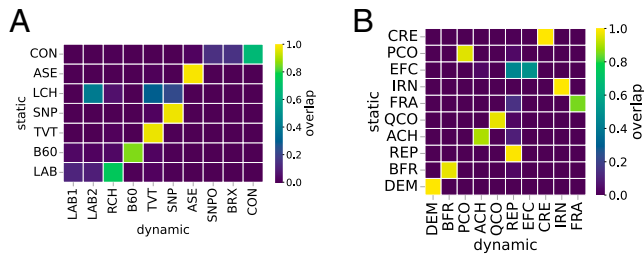
Fig. 2. Mapping between the static CCs found in previous studies and the dynamic CCs found with our method for UK 2019 (*A*) and USA 2020 (*B*). The mapping is based on the overlap between community members. Overall there is a good match between static and dynamic communities.

(*y* axis) and the dynamic ones (*x* axis) found with our method for each dataset. In the figure, the overlap expresses the fraction of users from a static CC that are members of a dynamic one. As shown, there is overall a very good match between static and dynamic CCs. Specifically, 5 out of 7 and 9 out of 10 static CCs have overlap > 60% with a dynamic CC, for UK 2019 and USA 2020, respectively. This result supports the consistency between static and dynamic analyses. Nonetheless, some static communities from each dataset were split into multiple dynamic communities. For UK 2019, the CON, LCH (a small community of activists protesting against an unfair taxation), and LAB static CCs were split, whereas for USA 2020 this occurred to EFC. The results of the dynamic analyses are confirmed by the analysis of the temporal behaviors of these CCs (*SI Appendix*, Fig. S1). As an example, users of the LAB static CC belong to the LAB1, LAB2, and RCH dynamic CCs, which all exhibit overall similar but temporally different behaviors. To this regard, the larger number of CCs found with the dynamic analysis are due to the better modeling of the temporal dimension, which allows identifying behavioral differences that were not observable with static analyses.

RQ1: Temporal Stability of CCs. So far we highlighted that our dynamic analyses provided overall similar—yet more nuanced and fine-grained—results with respect to previous static analyses. Now we leverage our innovative dynamic model of coordinated behavior (*Materials and Methods*) to evaluate the temporal stability of the CCs involved in the online electoral debates, by investigating the extent to which they underwent changes through time. This analysis is relevant for multiple reasons: i) it sheds light on how CCs evolve in time, and ii) it allows assessing how representative are static analyses of coordinated behavior, with respect to the temporal variations of the CCs.

Temporal evolution of CCs. To assess how CCs changed through time we measured temporal fluctuations in their size, membership, as well as their cumulative influx and outflux of users (*Materials and Methods*). Figs. 3 and 4 present the results of these analyses. Regarding community size, Figs. 3*A* and 4*A* highlight major differences in the temporal evolution of the CCs. Some communities such as SNP, TVT, ASE, IRN, and CRE appear as relatively stable in time, with only limited size fluctuations represented by mostly flat lines. Contrarily, other communities such as LAB1, LAB2, CON, EFC, and ACH exhibit marked variations. In particular for UK 2019, LAB2 almost doubled its initial size between November 21 and 26. CON exhibits a similar trend, albeit with reduced magnitude. For USA 2020 instead, EFC initially expanded and subsequently shrank significantly, while ACH lost more than 50% of its original size in the last couple of weeks before the election. This result demonstrates that some
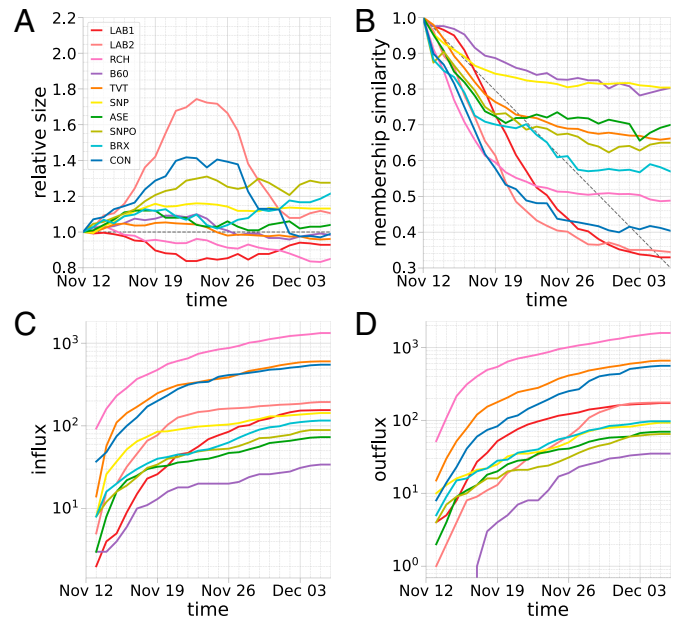

Fig. 3. UK 2019: Temporal stability of the CCs measured in terms of their evolving size (*A*), membership (*B*), and influx (*C*) and outflux (*D*) of users to/from the community. Each tick on the *x* axis corresponds to a 1-wk-long time window. Time windows are offset by one day. Dates on the *x* axis represent the start date of the corresponding time window.

CCs were rather unstable, as they experienced major temporal variations in size. Moreover, our results also demonstrate that some CCs evolved in time in opposite ways. For example, while LAB2 and CON increased their size until around November 23, LAB1 presents an opposite trend. Similar considerations can be made for SNPO, which increased its size until election day, with respect to RCH that progressively shrank, and for other communities involved in the USA 2020 debate.

Figs. 3*B* and 4*B* provide additional results by tracking how the membership—that is, the set of users that belong to a community at a given time—and not just the size, of each CC changes through time. As expected, CCs that experienced major size variations (e.g., LAB2, CON, EFC, ACH) present the largest differences in membership. Interestingly however, also CCs that exhibited moderate size variations (LAB1), or that appeared as overall stable when only considering their size (TVT and ASE), nonetheless feature marked membership differences in time. This result tells us that a stationary size does not necessarily imply stability in terms of members of the community. In fact, some CCs maintained a relatively stable number of members not because of lack of user shifts between CCs, but as a result of a balanced inflow and outflow of members. Fig. 3 *C* and *D* provide detailed results on this aspect for UK 2019, while Fig. 4 *C* and *D* do for USA 2020. As shown, some CCs (B60) had a limited influx and outflux for all the time, which is reflected in Fig. 3*B* by a relatively stable membership. On the flip side, other CCs (RCH, CON) were much more unstable, with a massive influx and outflux of users that translates into unstable membership. Interestingly, there also exist less straightforward situations, such as those of TVT and CRE that featured strong influx and outflux, but a relatively stable size and membership. This result implies that many users joined and left TVT and CRE every time window, but that those who left were likely to come back at a later time and vice versa. In other words, both TVT and CRE were characterized by a restricted set of users who repeatedly joined and left the communities. Overall, our results provide evidence
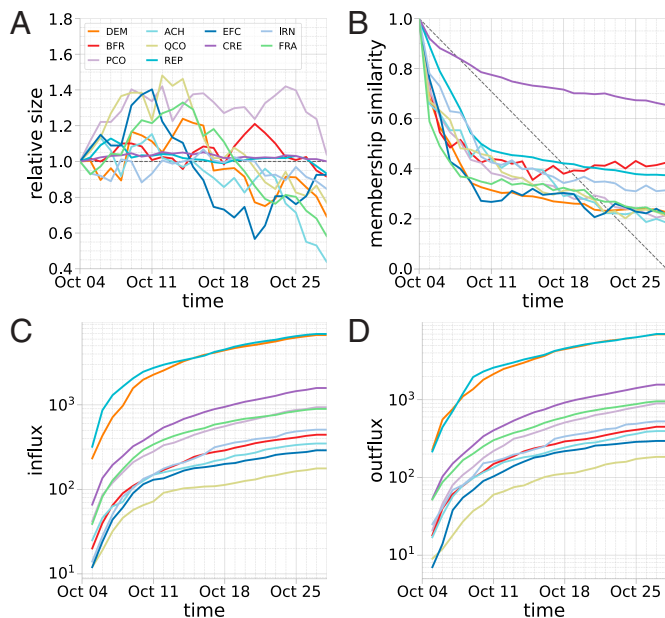
Fig. 4. USA 2020: Temporal stability of the CCs measured in terms of their evolving size (*A*), membership (*B*), and influx (*C*) and outflux (*D*) of users to/from the community. Each tick on the *x* axis corresponds to a 1-wk-long time window. Time windows are offset by 1 d. Dates on the *x* axis represent the start date of the corresponding time window.

of temporal instability and highlight marked differences in the temporal dynamics of some CCs.

Representativeness of static CCs. We conclude our analysis of RQ1 by discussing the implications of our findings about the instability of the CCs, with respect to the static analyses of coordinated behavior. In particular, results in both Figs. 3 and 4 highlighted that some CCs underwent marked changes in time that cannot be captured with time-aggregated static analyses. Figs. 5 and 6 dig deeper into this aspect by comparing the time-evolving size and membership of each dynamic CC to those of its corresponding static CC. In particular, to be strongly representative of a dynamic community, a static CC must have relative size S and membership similarity J both close to 1. The mapping between static and dynamic CCs is in Fig. 2. The results in Figs. 5 and 6 surface some of the drawbacks of static analyses. The only static CCs that are strongly representative of their dynamic counterparts are ASE and SNP for UK 2019, and CRE for USA 2020, as reflected by relative size $S \simeq 1$ and membership similarity $J > 0.6$. The same can be concluded also for TVT, REP, and QCO, although to a much lower extent. Instead,
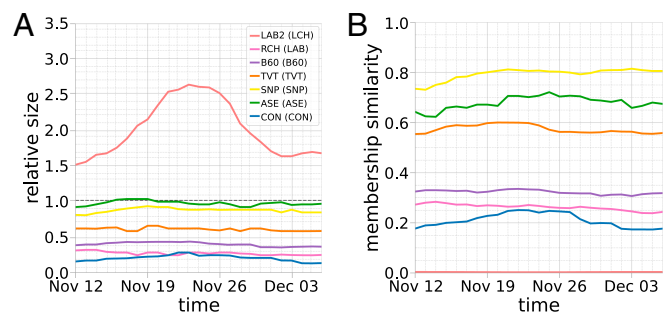

Fig. 5. UK 2019: Comparison between the time-evolving size (*A*) and membership (*B*) of each dynamic CC, and the static size and membership of the corresponding static CC.

Fig. 6. USA 2020: Comparison between the time-evolving size (*A*) and membership (*B*) of each dynamic CC, and the static size and membership of the corresponding static CC.
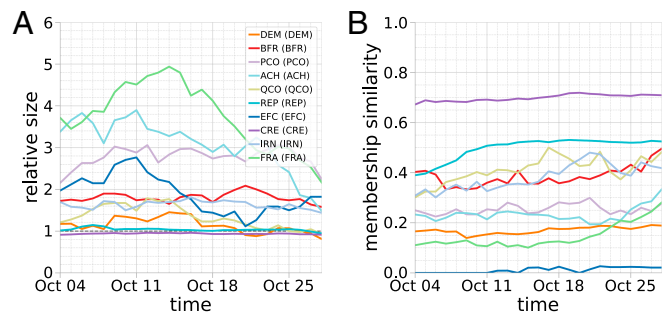
all other static CCs are weakly representative of the corresponding dynamic CCs. In fact, all have relative size far from 1 (e.g., $S \simeq 5$ on October 14 for FRA), and membership similarity $J < 0.4$. Overall, we found several static CCs that are poorly representative of the temporal evolution of the corresponding communities, which could possibly lead to inaccurate or unreliable results.

RQ2: Temporal dynamics of user behavior. The different temporal evolution of the CCs that we observed in RQ1 are due to different user behaviors, which we investigate in this section. Fig. 7 shows the joint and marginal distributions of the number of distinct CCs to which users belonged and the number of user shifts between CCs. As shown, 72% of all UK 2019 users and 50% of all USA 2020 users belonged to two or less CCs for all the time, which is also represented by all marginal distributions being skewed toward small numbers of distinct memberships and shifts. In fact, all four marginal distributions in Fig. 7 are heavy-tailed, as highlighted by the accurate power-law fitting shown with red lines in the marginal histograms. At the same time, however, a minority of users belonged to many CCs, which explains our previous results on the instability of some communities. Fig. 7 also allows evaluating the behavior that we observed in Fig. 3 for TVT and in Fig. 4 for CRE: users repeatedly leaving and rejoining the same community. Such behavior is reflected in Fig. 7 with users having many shifts but few memberships. Fig. 7*A* shows few users with this behavior and, in fact, the distributions of memberships and shifts for UK 2019 are strongly correlated (Pearson's $\rho = 0.73$) meaning that the behavior observed for members of TVT is overall marginal in the whole dataset. On the contrary, Fig. 7*B* depicts a slightly different situation for USA 2020, as demonstrated by the relatively dense (hot) region of the heatmap below the main diagonal of the plot. This means that the behavior observed for CRE is more prominent, albeit still related to a minority of users as reflected by the moderate correlation between the distributions of memberships and shifts for USA 2020 (Pearson's $\rho = 0.38$).

The analysis of Fig. 7 only considers the number of memberships and shifts between CCs. However, not all shifts are the same, as moving between two opposite communities (e.g., at the extremes of the political spectrum) entails a much bigger change—a farther leap—than moving between two similar ones. To account for this facet we assign a weight $w_{k,j}$ to all shifts $s_{k \to j}$ between any origin community C_k and any destination community C_j, based on the (dis)similarity between C_k and C_j (*Materials and Methods*). When considering also the distance between the origin and destination CCs involved in user shifts, results show that the majority of shifts occur between politically similar communities. This finding is in line with the theories
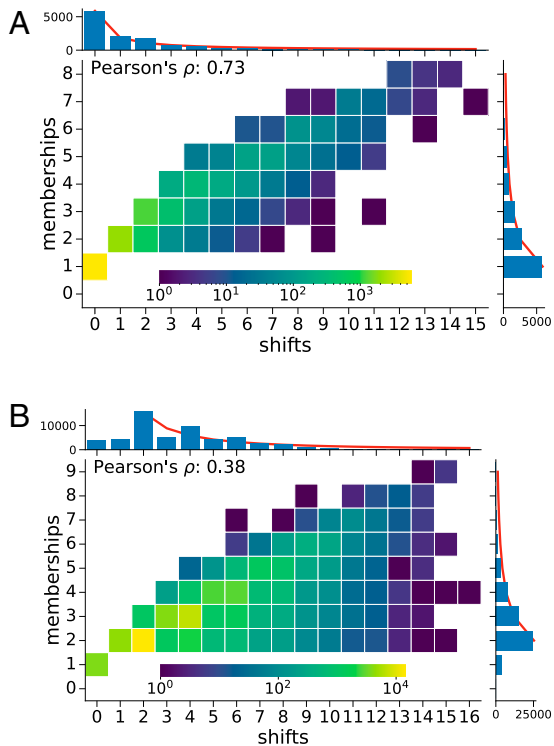
Fig. 7. Joint discrete distribution of the number of user shifts between CCs and user memberships to CCs, with marginal univariate distributions. (*A*) UK 2019. (*B*) USA 2020. Notably: i) 51% and 93% of all users belonged to more than one CC for UK 2019 and USA 2020, respectively, ii) memberships and shifts are strongly correlated ($\rho = 0.73$) in UK 2019 and moderately so ($\rho = 0.38$) in USA 2020, iii) the marginal distributions of memberships and shifts are heavy-tailed in both datasets.

about political polarization and echo chambers (41). Nonetheless, a minority of users experienced major ideological shifts by moving across the political spectrum, for both the UK 2019 and the USA 2020 election (*SI Appendix*, Fig. S3). Another interesting observation derived from this analysis is that the vast majority of shifts for UK 2019 occurred toward the left of the political spectrum. This means that, overall, the users involved in the UK 2019 online electoral debate ideologically moved toward the left as the debate unfolded (see *SI Appendix*, Fig. S3 and corresponding discussion).

RQ3: Archetypes and Drivers of User Behavior. When investigating user behaviors in RQ2, Fig. 7 surfaced heavy-tailed distributions for both user memberships to CCs and shifts between CCs, which imply heterogeneous user behaviors. On one hand, these distributions represent a bulk of stationary users with few shifts and memberships. On the other hand, however, their long tails also admit the existence of some volatile users characterized by a multitude of shifts, and of all other behaviors in between the "stationary" and "volatile" extremes. Following this observation, we introduce three archetypes of users, each corresponding to different temporal behaviors with important practical implications. For each archetype, we i) propose an operative definition, ii) apply the definition to measure the presence of such users in our CCs, and iii) explore possible motivations for their behavior.

Archetype 1: Stationary. *Stationary users are those who belong to the same community for all the time.* This archetype straightforwardly emerges from Fig. 7 and the above discussion. The analysis of stationary users is relevant because their behavior could imply that

they are strong supporters and core members of their CC (42). Fig. 8 shows the proportion of stationary users, as well as of the users of the remaining archetypes, in the CCs. Users that do not match the definition of any archetype introduced in this section are grouped as "others." As shown, the proportion of stationary users is in the region of 50% for UK 2019, while it varies significantly for the CCs involved in USA 2020. In detail, center-leaning communities in UK 2019—such as B60, TVT, and SNP—featured a large share of stationary users: between 68% and 79% of all community members. Contrarily, stationary users constituted only $\simeq$42% of the strongly polarized CCs such as LAB1, LAB2, and CON. This finding suggests that politically polarized communities were more unstable than moderate ones during the UK 2019 online electoral debate. Interestingly, also our results in Fig. 3 support this conclusion, with politically extreme CCs appearing as overall more unstable. This result is particularly relevant also in light of the many studies that specifically focus on strongly polarized communities, such as those on political polarization, fringe and extreme behaviors, and far-right online groups (43–45). The main difference between UK 2019 and USA 2020 is represented by the overall lower fraction of stationary users in the former online debate. Within this context, the two Republican CCs were the most stable ones with REP (Republicans) containing 39% stationary users and CRE (Conspiracist Republicans) containing as much as 62%. This result is again consistent with those reported in Fig. 4 *A* and *B* about the temporal stability of CCs. Stationary users for all other CCs were in the region of 18%. This marked difference between UK 2019 and USA 2020 is motivated by the different political polarization of the communities that took part in the two online debates, shown in Fig. 1. Indeed, while the UK 2019 CCs almost uniformly spanned the whole political spectrum, eight out of ten of those involved in USA 2020 laid on the right-hand side of the spectrum. In turn, this resulted in more shifts between communities and consequently in less stationary users.

To gain insights into why stationary users never leave their CC, we analyze their use of hashtags as a proxy for their interests and viewpoints, in relation to those of their CC and of all other communities (see *SI Appendix*, Fig. S4 and related discussion). The rationale for this analysis stems from the literature on echo chambers and political polarization as these users might be disincentivized to change community because they already belong to the CC that mostly represents their interests and viewpoints (41). Results show that 94% of all UK 2019 stationary users and 97% of all USA 2020 stationary users are mostly similar to the CC to which they belong, as highlighted in the figure by the great prevalence of users along the main diagonal.
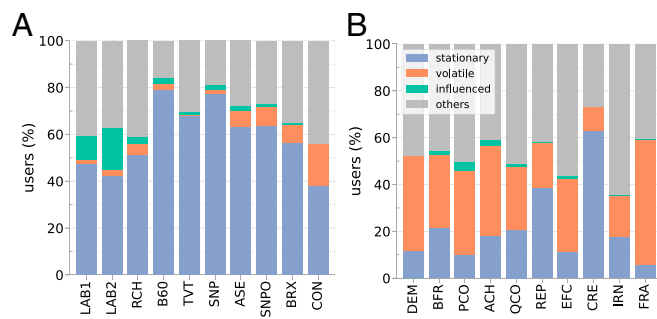

Fig. 8. Membership composition of the UK 2019 (*A*) and USA 2020 (*B*) CCs in terms of the different archetypes of users. CCs are ordered on the *x* axis according to their political leaning.

Overall, our results confirm the similarity between the interests and viewpoints of stationary users and those of the CC to which they belong, consistently with the established literature on echo chambers (46).

Archetype 2: Influenced. *Influenced users are those who change community and remain in the destination community for a relatively long time.* One of the advantages of our dynamic analysis lies in the possibility to investigate user shifts between CCs. In the case of polarized or controversial online debates, shifts—disregarded in previous studies—provide valuable information on the evolving stance of the users with respect to the sides involved in the debate. To this end, users who abandon a community to join another one for a long time might have been persuaded or influenced by the latter. Hence, detecting and characterizing influenced users could have important practical implications for the study of online campaigns. For our subsequent analyses, we consider as influenced all those users who never left the destination community C_j after a shift $s_{k \to j}$. Moreover, we also impose that each influenced user remained in C_j for at least one third of the time covered in the dataset, so as to avoid considering as influenced those users who stayed in the destination community for just a few days before the election. Fig. 8*A* shows marked differences in the proportion of influenced users within the UK 2019 CCs. Out of all the communities, LAB2 was the one with the highest share (18%), followed by LAB1, B60, and RCH. These are all left and center-left leaning communities. Opposite results emerge for right-leaning communities, where CON had the lowest share of influenced users, followed by BRX, TVT, and SNPO. These results extend and reinforce those related to the net flows of users between CCs (*SI Appendix,* Fig. S3). Together, they surface a strong political imbalance, in favor of the left, in the capacity to attract and hold users by the CCs involved in the electoral debate. The proportion of influenced users reveals interesting patterns also for the USA 2020 scenario, as shown in Fig. 8*B*. Here, the CCs with the highest share of influenced users were PCO and ACH, followed by BFR, QCO, and EFC. In the context of the USA 2020 online debate, these were the most center-leaning communities. Similarly to the UK 2019 case, the analysis of influenced users in USA 2020 unveiled an interesting pattern in the capacity of some CCs to pull and hold users.

Next, we explore possible drivers for the behavior of influenced users. For each influenced user u, we investigate whether there exist signals capable of explaining its shift $s_{k \to j}$ from the origin community C_k to the destination community C_j. For this analysis, we consider the time-evolving relationship between u, C_k, and C_j from a twofold perspective: i) topic-based similarity and ii) topological position in the network. With the former, we assess whether influenced users exhibited a time-increasing topic-based similarity with their destination communities. With the latter, we evaluate whether the position of the influenced users in the temporal network became gradually closer to their destination community. The temporal trends (see *SI Appendix,* Figs. S5 and S6, and related discussions) reveal that influenced users indeed exhibited an increasing similarity—be it topic- or network-based—to their destination community as time went by. Meanwhile, they also became increasingly dissimilar to their origin community. This finding is confirmed in both datasets and holds under all the viewpoints considered in our analyses, revealing that content and network characteristics provide insights into why users shift between CCs (see *SI Appendix* for details).

Archetype 3: Volatile. *Volatile users are those who repeatedly change community, staying in each community only for a limited amount of time.* At the opposite of stationary users are those whose behavior is very erratic. Volatile users often shift between CCs without attaching to any, if not for a very limited time. The identification and characterization of volatile users in online debates are relevant, as they might represent undecided users who have not already taken a definitive position about the discussed topic (47). Here, we operationalize volatile users as those who performed three or more shifts, and that spent less than one third of the total time in each CC to which they belonged. Fig. 8*A* shows that the proportion of volatile users in the UK 2019 CCs is skewed toward the right-hand side of the political spectrum. In detail, CON had by far the largest share of volatile users (18%), followed by SNPO, BRX, and ASE—which are all right-leaning communities. Volatile users in each of the remaining CCs accounted for less than 5% of each community's members. This result is the counterpart of what we discussed about influenced users for the left-leaning CCs. In fact, while we previously found evidence of the effectiveness of the left-leaning communities in attracting (i.e., influencing) users as the electoral debate unfolded, here we find evidence of the weakness of the right-leaning CCs. Interestingly, this result appears to be in contrast to the majority of the existing literature on the use of online platforms by political groups, where right-leaning communities were described as tech-savvy and capable of making the most out of social media campaigning (45, 48). The distribution of volatile users within the USA 2020 CCs is relatively uniform, with the exception of CRE and FRA. About the former, all the results are consistent in highlighting that CC as the overall more stable. This result is particularly relevant in comparison to the overall instability of the majority of CCs involved in the USA 2020 debate. The latter community instead features an opposite behavior, as it is characterized by a marked instability (Fig. 6*A*), emphasized by a majority of volatile users and very few stationary ones (Fig. 8*B*). All other CCs involved in the USA 2020 debate have a considerable fraction of volatile users in the region of 40%. Overall, these results indicate that the USA 2020 online debate was characterized by much more instability than the UK 2019 one, as demonstrated by both the results on the fraction of stationary, influenced, and volatile users (Fig. 8) as well as by the analysis of temporal stability of the CCs involved in the two debates (Fig. 5 vs. Fig. 6).

When analyzing possible drivers for the behavior of volatile users, we found that the shifts by volatile users are typically very short (*SI Appendix,* Fig. S7). This result is again consistent with the echo chamber theory (46) and explains why volatile users change community so often. Interestingly, the same analysis also shows that some influenced users permanently join politically distant communities.

Validation. As a final case study, we apply our method to a Twitter dataset related to a large-scale information operation (see *Materials and Methods* and *SI Appendix* for details on the dataset). The dataset is composed of the inauthentic accounts that took part in the operation and of a comparable set of genuine accounts. As such, it represents a labeled benchmark suitable for validating the efficacy of our method at distinguishing between inauthentic and inorganic forms of coordination versus genuine ones. Specifically, we run our method with different combinations of parameters, evaluating at each run the extent to which the inauthentic accounts are grouped in distinct communities with respect to the genuine ones. To measure the extent of separation between inauthentic and genuine accounts, we adapt the well-known F1 score to our problem (*Materials and Methods*). Fig. 9*A* shows an excerpt of the results that we obtained with a grid search
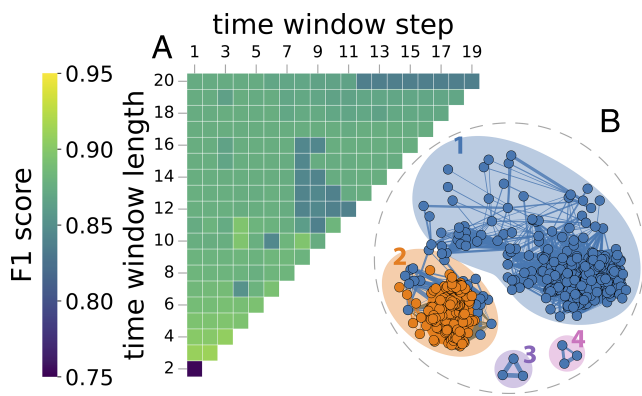
Fig. 9. Honduras 2019. (*A*) Efficacy (F1 score) of our method at detecting CCs involved in an information operation, based on the choice of parameters *time window length* and *time window step*. The resolution parameter γ of the community detection algorithm is set to the optimal value $\gamma = 10^{-5}$. (*B*) Snapshot (layer) of the Honduras 2019 multiplex temporal network showing that all inauthentic accounts (orange-colored) are clustered in a single CC, together with only a few genuine accounts (blue-colored).

over the two main parameters of our method (i.e., the *length* and *step* of the time windows), when fixing the resolution parameter of the community detection algorithm. As shown, the application of our method yields CCs in which the inauthentic accounts are generally separated from the genuine ones, as testified by $0.75 \leq F1 \leq 0.91$. The analysis reveals a trend where larger F1 scores are obtained for small values of the two parameters. In fact, the best F1 score is obtained with time window length = 3 d and step = 2 d. Fig. 9*B* shows one layer of the multiplex temporal network where the nodes are colored so as to represent the inauthentic (orange-colored) and genuine (blue-colored) accounts, and where the different CCs are highlighted. The network in the figure shows a good separation between the inauthentic and genuine accounts, where the former are all grouped in a single community with only a few genuine nodes. For reference, applying the traditional static coordination detection approach to this problem would result in F1 = 0.82, demonstrating the advantage of a properly configured dynamic approach.

Discussion

We investigated the temporal dynamics of coordinated online behavior in the context of two recent major elections and an information operation. Our approach, grounded on a multiplex temporal network and dynamic community detection, identified more CCs than those found in previous works, surfacing temporal nuances of coordination that would not be observable with the traditional static approach. Our analysis produced key findings in multiple areas. Regarding CCs, we found that the majority were rather unstable and experienced significant changes in size and membership through time. Regarding user membership to CCs, we found evidence of heavy-tailed distributions implying the existence of a bulk of relatively stationary users and of a long tail of volatile ones. The majority of user shifts from a community to another occurred between politically aligned communities according to the echo chamber theory, although we found a subset of users who crossed the political spectrum. Finally, we found that content and network characteristics convey useful information for understanding why users move between certain CCs.

Implications. Our results bear important implications about the limitations of static analyses of online coordination, the strategies

of CCs, the patterns of online influence, and for the research and policy of online platforms.

Instability and influence. The first major implication of our study is the increased awareness of the instability of CCs. To this end, our results open up new directions of research on the temporal dynamics of coordinated online behavior, which was so far almost exclusively analyzed from a static rather than a *dynamic* standpoint (10, 32). However, our results show the limitations of the former approach, suggesting that additional efforts should be devoted to devising nuanced and reliable dynamic analyses. Then, our results about the broad array of diverse user behaviors, and the resulting user archetypes that we identified, have significant implications for the study of online interactions and influence (49). For example, future computational and social research should extensively investigate the motivations for such heterogeneous behaviors, for which we provided but some initial explanations. Our results in this area can be particularly useful toward the ongoing research on online influence and persuasion, as the shifts detected via dynamic analyses of coordination could contribute to identifying successful cases of influence over users or communities in a network (50). Similarly, our scientific approach could be carried over to investigate the temporal dynamics of online polarization and the possible temporal evolution of echo chambers (46, 51). Our findings are also particularly impactful toward understanding, and possibly even predicting, the outcome of polarized and controversial online debates, such as those preceding major elections (52). Specifically, if verified in other contexts and platforms, our results could provide useful information for nowcasting and forecasting the dynamics of group influence in online debates.

From observational to predictive studies. In the future, it would be possible to progressively shift from observational to *predictive* analyses, provided the availability of adequate ground-truths, which currently represents a limiting factor for all studies in the area of coordinated online behavior (29). In addition to opening up the possibility to predict the dynamics of online influence, our results show interesting temporal correlations between significant changes in the structure of CCs and real-world events (see *SI Appendix*, Fig. S8 and related discussion). Therefore, while additional research is needed to design the methods and validate the approach, our results open the door to i) leveraging external knowledge about real-world events to estimate their online impact in terms of the changes experienced by the interested communities; ii) leveraging significant changes in community structure to identify events that have been noteworthy in a specific period of time and for a specific audience (53).

Validation. The lack of reference datasets and authoritative ground truths on coordinated online behavior currently represents one of the strongest limiting factors to the research in this area. In addition to carrying out manual investigations (see *SI Appendix*, Fig. S8 and the related analysis), here we circumvented this problem by validating our approach on a labeled dataset that contains both inauthentic accounts involved in an information operation and genuine ones. This is a favorable approach to validating coordination detection methods for multiple reasons. First, it provides a way to test the efficacy of a method at capturing the behavioral patterns of different accounts, including the inauthentic ones that should stand out from the rest. To this end, our results revealed that our method allows distinguishing inorganic CCs from organic ones, to a large extent. Second, the same analysis can also inform the choice of parameters of the method, which represents a further outstanding

challenge in the field. Finally, it demonstrates the practical usefulness of the dynamic analysis of coordinated behavior for a relevant computational social science task.

Profiling. While we provided multiple findings toward improving our understanding of online debates and online human dynamics at large (54), we only scratched the surface of a complex and multifaceted phenomenon. Another aspect worthy of discussion is the potential for using our proposed methodology to study and contrast online information manipulation or other nefarious instances of online coordination (4). *Profiling coordination*—that is, inferring the main characteristics, peculiarities, and possibly even the intent behind different groups of coordinated users, turns out to be particularly challenging (29). However, a certain degree of success at profiling coordination is needed for being able to tell the difference between inauthentic or harmful coordination and unintentional coordination among independent users (e.g., fandoms or other grassroots movements). Here, we intentionally kept a neutral stance with respect to the many existing forms of online coordination, enabling the study of diverse instances of coordination without any inherent bias. This was because our main goal for this study was to investigate the temporal dynamics of coordination, leaving the task of distinguishing between inauthentic/harmful and authentic/harmless coordination for subsequent analyses. These would ideally be conducted by human analysts, possibly domain experts, who could draw upon the insights and characteristics of the detected communities, as provided by this and other studies (9, 10). Nevertheless, we acknowledge the importance of investigating inauthentic and harmful communities as part of future dynamic analyses on online coordination. For example, many harmful instances of coordination, such as strategic information operations, exploit multiple tactics and resources to boost their chances of success (5). Studying online debates by means of dynamic analyses could allow spotting early signals of harmful coordination, possibly enhancing timely responses to what threatens the safety and integrity of the online environment (55).

Limitations and Future Work. The main drawback of our study stems from its relatively limited scope, with respect to the breadth and depth of a complex phenomenon such as online coordination. For example, we do not make any assumption on the possible inauthenticity or harmfulness of the identified CCs of UK 2019 and USA 2020. However, inauthentic and harmful coordination could be used in the context of online political debates as a means to influence the electoral outcome (56). Here, our results for the USA 2020 presidential election surfaced the activity of multiple coordinated groups supporting election fraud narratives and other conspiracy theories (57). Likewise, we also investigated the behavior of foreign influence groups that participated in the electoral debate. Nevertheless, our results do not specifically address the impact that inauthentic and harmful groups could have had on the analyzed online debates, for which our study does not provide conclusive results, but rather calls for additional research. Moreover, our validation of the methodology, while showing limited sensitivity to small variations in the time window length and step parameters, revealed our method's sensitivity to the resolution parameter of the underlying community detection algorithm (*SI Appendix*). However, this is a widely recognized phenomenon in the literature (58, 59). In fact, the optimal value for the resolution parameter varies based on network characteristics and the goal of the analysis, leaving its selection to the analyst discretion. Finally, our analysis is

based on data collected from a single platform, while online coordination often involves activities intertwined across multiple platforms (60). As such, we might have missed significant coordinated efforts that have occurred on platforms other than Twitter. This limitation is shared with the vast majority of the existing literature on the subject, mainly due to the challenges of acquiring related and comparable datasets across multiple platforms. In light of this widespread limitation, however, future research on coordinated online behavior should strive to collect and analyze multiplatform datasets, for that could reveal patterns of coordination that would otherwise remain hidden. Unfortunately, however, the limitations related to data availability extend beyond the challenges of studying multiplatform coordination. The recent changes in API availability enforced by Twitter (61), particularly after its transition to X.com, represent a paramount example. The discontinuation of Twitter Academic APIs and the prohibitive costs of all other options pose significant obstacles to data availability, with nefarious consequences in terms of greatly reduced platform transparency and reproducibility of scientific results. While the general outlook on social media data availability remains grim at the time of writing, the European Digital Services Act (DSA) could turn the tide in the struggle for access to platform data, by providing legal and technical facilities for submitting data access requests for research purposes (62). Finally, future research should also deviate from the traditional focus of online coordination studies (e.g., political discussions and inauthentic coordination) to encompass topics, temporal dynamics, and coordination patterns that are typical of nonpolarized or noncontroversial online interactions.

Materials and Methods

Data. The data for our study cover two recent major political events–the 2019 UK general elections and the 2020 USA presidential elections–and a large-scale information operation.

UK 2019 General Election. We leverage a publicly available reference dataset related to the online Twitter debate about the 2019 UK general election.* This dataset is relevant for our present study since it has already been the subject of static analyses of coordinated behavior (9, 23). The dataset was built in ref. 9 via the Twitter Streaming API during the last month before the UK 2019 election day, namely between November 12 and December 12, 2019. It contains 11,264,820 distinct tweets (left- and right-leaning, as well as neutral) about the 2019 UK general election published by 1,179,659 distinct users. The tweets were selected according to a set of hashtags or relationships to party accounts (see *SI Appendix* for details).

USA 2020 Presidential Election. To assess the generalizability of our findings we collected and publicly shared a second dataset of tweets related to the USA 2020 presidential election.† Similarly to the UK 2019 dataset, this data collection spanned the month leading up to the election day, encompassing the period from October 4 to November 3, 2020. Likewise, this collection process relied on a combination of election-related hashtags, party hashtags, and official party and leader accounts (see *SI Appendix* for details). In total, this dataset comprises 263,518,037 distinct tweets that pertain to the online discourse surrounding the election. The tweets were generated by 15,288,527 distinct users.

Honduras 2019 information operation. To validate our method we use a publicly available dataset related to a large-scale information operation promoted by the government of Honduras in 2019.‡ The dataset includes both the inauthentic accounts who took part in the operation, as detected by Twitter Moderation Research Consortium,§ and a comparable set of legitimate accounts (see *SI Appendix* for details). As such, it can be used as a ground-truth

---

and benchmark to test the efficacy of our method at detecting the coordinated inauthentic accounts. Similarly to the previous cases, we used one month of data, spanning from November 11 to December 11, 2019. This portion of the dataset contains 251,191 tweets shared by 75,845 distinct users.

Dynamic Analysis of Coordinated Online Behavior. To detect and study coordinated behavior, we broadly follow the state-of-the-art network analysis frameworks recently proposed by refs. 9–11, which consider repeated similarities in user behaviors as a proxy for coordination. However, we differentiate from the existing methods by building a dynamic multiplex network instead of a static one, and by analyzing it with a dynamic community detection algorithm. Our detailed methodology is presented in the following.

Preliminaries. To enable comparisons with previous works, we compute user similarity based on *co-retweets*–the action of retweeting the same tweet by different users, and we analyze *superspreaders*–the top 1% users with the most retweets.¶ We computed user similarities based on co-retweets also because our study specifically focuses on superspreaders, who are characterized by a marked retweeting behavior (63), and because of the excellent results obtained with co-retweets in recent related literature (10, 11, 64) and in our previous static analyses of online coordination (9).

Dynamic network modeling. We build our dynamic user similarity network as a multiplex temporal network $G = \langle G_0, \dots, G_{N-1} \rangle$ where each layer G_i models user behaviors occurred during a given time window t_i. We adopted this network representation over other options, such as discrete-time dynamic graphs and embeddings methods, since multiplex temporal networks are state-of-the-art for observational and descriptive tasks (20). For example, this representation makes detecting and interpreting perturbations of the network topology (i.e., investigating the stability of CCs) advantageous. We work with a sequence of discrete and overlapping time windows: each t_i has a duration $d = 7$ days and an offset (step) $\delta = 1$ day from t_{i-1}. After a set of preliminary tests, we ended up using overlapping time windows instead of nonoverlapping ones, because the latter neglect all interactions occurring across any two adjacent windows. On the contrary, overlapping time windows allow to consider all relevant interactions, while also guaranteeing smooth transitions between different time steps. Regarding the size of the time window, the larger it is, the smoother are the transitions between the different snapshots of the multiplex temporal network. However, very large windows hide changes that occur within that time frame, up to the point that particularly large ones make the dynamic analysis collapse to a static one, negating the advantages of the former. On the contrary, particularly small windows may fail to collect meaningful statistics at each time step. Here, our choice of time window length and step is supported by the favorable results obtained for that parameter configuration in our grid search validation. As sketched in Fig. 10, for each t_i, we build a weighted undirected user similarity network $G_i = (V, E, W)$. To obtain G_i, we first model each user $v \in V$ with the TF-IDF weighted vector of the tweets it retweeted during t_i. An edge $e \in E$ between two users exists if they retweeted at least one common tweet during t_i. Edge weights $w \in W$ are computed as the cosine similarity between pairs of user vectors. The TF-IDF weighting scheme discounts retweets of viral tweets and emphasizes user similarities due to unpopular tweets, contributing to highlighting interesting behaviors. Finally, for each network G_i, we retain only the statistically significant edges by computing its multiscale backbone (65).

Dynamic community detection. The multiplex temporal network $G = \langle G_0, \dots, G_{N-1} \rangle$ is suitable for being analyzed with a dynamic community detection algorithm. Leiden is a state-of-the-art community detection algorithm for multiplex networks that improves the well-known Louvain algorithm by identifying higher quality and well-connected communities (66). It allows community detection on multiplex networks by jointly considering the internal edges in each layer (solid edges in Fig. 10), as well as the edges that connect nodes across layers (dashed edges in Fig. 10). Leiden is therefore a cross-time algorithm that identifies communities based on the full temporal evolution of the network (20). An important implication is the possibility to assign nodes to different communities depending on the time window. As such, it is particularly suitable for studying the temporal evolution of user behaviors and of the CCs.

¶ Despite accounting for only 1% of all users, superspreaders produced ~39% of all tweets and ~44% of all retweets.
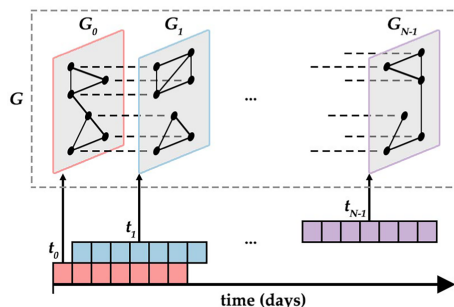

Fig. 10. Overview of our methodological approach for building the multiplex temporal network G. Data from the overlapping time windows t_i yield the weighted undirected user similarity networks G_i that constitute the layers of G. Then, dynamic community detection is performed on the multiplex temporal network G.

Notably, it would have been possible to use a static community detection algorithm such as Louvain on each snapshot (i.e., layer) of our multiplex temporal network. However, such an approach would have later required the application of a community tracking algorithm, which would have made the overall process more convoluted and error-prone.

Political Polarization. We compute a polarization score for each CC based on the political polarization of the hashtags used by its members. We obtain a polarization score for each hashtag in the dataset by applying a label propagation algorithm. In detail, the score for any given hashtag is iteratively inferred from its co-occurrences with other hashtags of known polarity. We initialize the algorithm with the hashtags used for collecting the tweets as the seeds of known polarity (*SI Appendix*, Tables S1 and S2). Once hashtag polarities are computed, the polarity of a CC is obtained as the TF weighted average of the polarities of the hashtags used by members of that community. As a result, each CC is assigned a polarity score $p \in [-1, +1]$. Finally, the scores are normalized so that the rightmost-leaning CC in each dataset has $p = +1$ and the leftmost-leaning one has $p = -1$. The most polarized communities are CON ($p = +1$) and LAB1 ($p = -1$) for UK 2019, and IRN ($p = +1$) and DEM ($p = -1$) for USA 2020.

Temporal Community Monitoring. To assess how CCs changed through time we measure, for each community C_k, and for each t_i with $i = 0, \dots, N - 1$:

- the size of C_k relative to t_0: $S(k, i) = \frac{\text{size}_k(t_i)}{\text{size}_k(t_0)}$;
- the Jaccard similarity of the membership of C_k relative to t_0: $J(k, i) = \frac{|\text{mem}_k(t_i) \cap \text{mem}_k(t_0)|}{|\text{mem}_k(t_i) \cup \text{mem}_k(t_0)|}$, where $\text{mem}_k(t_i)$ is the set of users that belong to C_k at time t_i;
- the influx $F_{\text{in}}(k, i)$ and outflux $F_{\text{out}}(k, i)$ of C_k, respectively expressed as the cumulative number of users who joined and left the community up to t_i.

Similarities between Communities. We assign a weight $w_{k,j}$ to all shifts $s_{k \to j}$ between any origin community C_k and any destination community C_j, based on the (dis)similarity between C_k and C_j. We compute the similarity between two CCs as the cosine similarity of the TF weighted vectors of the hashtags used by the communities. Then, we weight shifts proportionally to the dissimilarity of the involved CCs: $w_{k,j} = 1 - \text{sim}(k, j)$. We also compute *net shifts* between CCs, as in *SI Appendix*, Fig. S3 where an edge $C_k \to C_j$ exists only if there is a positive net user flow $F_{k \to j}$ from C_k to C_j: $F_{k \to j} = \sum s_{k \to j} - \sum s_{j \to k} > 0$. Then, edge thickness in the figure is proportional to $w_{k,j} \times F_{k \to j}$.

Detecting Accounts Involved in an Information Operation. We measure the extent to which our method is capable of detecting the inauthentic accounts who took part in an information operation by adapting the well-known F1 score to our problem. For each layer G_n of the multiplex temporal network G, we first discard all communities that do not contain any inauthentic account. For

each remaining community C_i^n with $i = 1, \ldots, M$, we compute the Precision P_i^n, Recall R_i^n, and F1 score $F1_i^n = 2\frac{P_i^n \times R_i^n}{P_i^n + R_i^n}$. F1 score is the harmonic mean of Precision and Recall and measures the extent to which the community C_i^n contains all and only the inauthentic accounts. The F1 score for G_n is the weighted mean of the F1 scores of its communities: $F1_n = \frac{1}{M}\sum_{i=1}^{M} w_i F1_i^n$, where w_i is proportional to the fraction of users in C_i^n over all users in G_n, so that larger communities contribute more toward $F1_n$. Finally, the overall F1 score is computed as the mean of the $F1_n$ scores of each layer $G_n \in G$.

Data, Materials, and Software Availability. Previously published data were used for this work (https://doi.org/10.5281/zenodo.4647893, https://doi.org/10.5281/zenodo.7358386, and https://doi.org/10.5281/zenodo.10650967) (21, 22, 24).

Author affiliations: [a]Institute of Informatics and Telematics, National Research Council, Pisa 56124, Italy; [b]Department of Mathematics, University of Padua, Padua 35122, Italy; and [c]Department of Natural Language Processing, Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi 54115, United Arab Emirates

1. S. González-Bailón, J. Borge-Holthoefer, A. Rivero, Y. Moreno, The dynamics of protest recruitment through an online network. *Sci. Rep.* **1**, 1–7 (2011).
2. Z. C. Steinert-Threlkeld, D. Mocanu, A. Vespignani, J. Fowler, Online social networks and offline protest. *EPJ Data Sci.* **4**, 1–9 (2015).
3. C. Francois, V. Barash, J. Kelly, Measuring coordinated versus spontaneous activity in online social movements. *New Media Soc.* **25**, 3065–3092 (2021).
4. D. M. Lazer *et al.*, The science of fake news. *Science* **359**, 1094–1096 (2018).
5. K. Starbird, Disinformation's spread: Bots, trolls and all of us. *Nature* **571**, 449–450 (2019).
6. A. Bovet, H. A. Makse, Influence of fake news in Twitter during the 2016 US presidential election. *Nat. Commun.* **10**, 7 (2019).
7. E. Mariconti *et al.*, "You know what to do": Proactive detection of YouTube videos targeted by coordinated hate attacks. *Proc. ACM Hum.–Comput. Interact.* **3**, 1–21 (2019).
8. R. Di Pietro, M. Caprolu, S. Raponi, S. Cresci, "New dimensions of information warfare" in *Advances in Information Security* (Springer, 2021), vol. 84.
9. L. Nizzoli, S. Tardelli, M. Avvenuti, S. Cresci, M. Tesconi, "Coordinated behavior on social media in 2019 UK general election" in *Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM'21)* (2021), vol. 15, pp. 443–454.
10. D. Pacheco *et al.*, "Uncovering coordinated networks on social media: Methods and case studies" in *Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM'21)* (2021), pp. 455–466.
11. D. Weber, F. Neumann, Amplifying influence through coordinated behaviour in social networks. *Soc. Network Anal. Mining* **11**, 111 (2021).
12. V. Sekara, A. Stopczynski, S. Lehmann, Fundamental structures of dynamic social networks. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9977–9982 (2016).
13. D. Weber, L. Falzon, Temporal nuances of coordination networks. arXiv [Preprint] (2021). https://doi.org/10.48550/arXiv.2107.02588 (Accessed 17 April 2024).
14. K. Sharma, Y. Zhang, E. Ferrara, Y. Liu, "Identifying coordinated accounts on social media through hidden influence and group behaviours" in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (2021), pp. 1441–1451.
15. Y. Zhang, K. Sharma, Y. Liu, "Vigdet: Knowledge informed neural temporal point process for coordination detection on social media" in *Advances in Neural Information Processing Systems (NeurIPS'21)* (2021), pp. 3218–3231.
16. T. W. Malone, K. Crowston, The interdisciplinary study of coordination. *ACM Comput. Surv.* **26**, 87–119 (1994).
17. F. Giglietto, N. Righetti, L. Rossi, G. Marino, It takes a village to manipulate the media: Coordinated link sharing behavior during 2018 and 2019 Italian elections. *Inf. Commun. Soc.* **23**, 867–891 (2020).
18. T. Magelinski, L. Ng, K. Carley, A synchronized action framework for detection of coordination on social media. *J. Online Trust Safety* **1**, 1–24 (2022).
19. P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, J. P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).
20. G. Rossetti, R. Cazabet, Community discovery in dynamic networks: A survey. *ACM Comput. Surv.* **51**, 1–37 (2018).
21. L. Nizzoli *et al.*, Coordinated Behavior on Social Media in 2019 UK General Election [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4647893. Deposited 30 March 2021.
22. S. Tardelli *et al.*, Multifaceted Online Coordinated Behavior in the 2020 US Presidential Election [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7358386. Deposited 1 June 2023.
23. K. Hristakieva, S. Cresci, G. Da San Martino, M. Conti, P. Nakov, "The spread of propaganda by coordinated communities on social media" in *Proceedings of the 14th ACM Web Science Conference (WebSci'22)* (2022) pp. 191–201.
24. L. Cima *et al.*, Twitter dataset about Information Operations in Honduras and UAE [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10650967. Deposited 12 February 2024.
25. L. Backstrom, D. Huttenlocher, J. Kleinberg, X. Lan, "Group formation in large social networks: Membership, growth, and evolution" in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'06)* (2006) pp. 44–54.
26. A. Patil, J. Liu, J. Gao, "Predicting group stability in online social networks" in *Proceedings of the 22nd International Conference on World Wide Web (WWW'13)* (2013) pp. 1021–1030.
27. E. Newell *et al.*, "User migration in online social networks: A case study on Reddit during a period of community unrest" in *Proceedings of the 10th International AAAI Conference on Web and Social Media (ICWSM'16)* (2016), vol. 10, pp. 279–288.
28. M. Horta Ribeiro *et al.*, Do platform migrations compromise content moderation? Evidence from r/The_Donald and r/Incels. *Proc. ACM Hum.–Comput. Interact.* **5**, 1–24 (2021).
29. L. Vargas, P. Emami, P. Traynor, "On the detection of disinformation campaign activity with network analysis" in *The 2020 ACM SIGSAC Conference on Cloud Computing Security Workshop (CCSW'20)* (2020), pp. 133–146.
30. M. Mendoza, M. Tesconi, S. Cresci, Bots in social and interaction networks: Detection and impact estimation. *ACM Trans. Inf. Syst.* **39**, 1–32 (2020).
31. A. Simchon, W. J. Brady, J. J. Van Bavel, Troll and divide: The language of online polarization. *PNAS Nexus* **1**, pgac019 (2022).
32. M. Cinelli, S. Cresci, W. Quattrociocchi, M. Tesconi, P. Zola, Coordinated inauthentic behavior and information spreading on Twitter. *Decis. Supp. Syst.* **160**, 113819 (2022).
33. S. Smith, Backto60 granted leave to appeal. *Pensionsage* (2020). https://www.pensionsage.com/pa/Backto60-granted-leave-to-appeal.php. Accessed 01 October 2023.
34. G. Pogrund, G. Arbuthnott, Revealed: The depth of labour anti-semitism. *The Times* (2019). http://www.thetimes.co.uk/article/revealed-the-depth-of-labour-anti-semitism-bb57h9pdz. Accessed 19 September 2023.
35. D. Jackson, E. Thorsen, D. Lilleker, N. Weidhase, "UK election analysis 2019: Media, voters and the campaign" in *Bournemouth University Centre for Comparative Politics and Media Research* (Technical report, 2019).
36. E. Akinwotu, 'He just says it as it is': Why many Nigerians support Donald Trump. *The Guardian* (2020). https://www.theguardian.com/world/2020/oct/31/he-just-says-it-as-it-is-why-many-nigerians-support-donald-trump. Accessed 19 September 2023.
37. M. Wendling, QAnon: What is it and where did it come from? *BBC* (2021). https://www.bbc.com/news/53498434. Accessed 19 September 2023.
38. H. Davidson, Why are some Hong Kong democracy activists supporting Trump's bid to cling to power? *The Guardian* (2020). https://www.theguardian.com/us-news/2020/nov/13/trump-presidency-hong-kong-pro-democracy-movement. Accessed 19 September 2023.
39. A. Tabatabai, QAnon: Goes to Iran. *Foreignpolicy* (2020). https://foreignpolicy.com/2020/07/15/qanon-goes-to-iran/. Accessed 19 September 2023.
40. E. Ferrara, H. Chang, E. Chen, G. Muric, J. Patel, Characterizing social media manipulation in the 2020 US presidential election. *First Monday* **25**, 1–32 (2020).
41. K. Garimella, G. De Francisci Morales, A. Gionis, M. Mathioudakis, "Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship" in *Proceedings of the 27th International Conference on World Wide Web (WWW'18)* (2018), pp. 913–922.
42. A. Trujillo, S. Cresci, "Make Reddit great again: Assessing community effects of moderation interventions on r/The_Donald" in *Proceedings of the 25th ACM Conference On Computer-Supported Cooperative Work And Social Computing (CSCW'22)* (ACM, 2022), pp. 1–28.
43. P. Agathangelou, I. Katakis, L. Rori, D. Gunopulos, B. Richards, "Understanding online political networks: The case of the far-right and far-left in Greece" in *Proceedings of the 9th International Conference on Social Informatics (SocInfo'17)* (Springer, 2017), pp. 162–177.
44. S. Zannettou *et al.*, "On the origins of memes by means of fringe Web communities" in *Proceedings of the Internet Measurement Conference (IMC'18)* (2018), pp. 188–202.
45. F. Morstatter, Y. Shao, A. Galstyan, S. Karunasekera, "From alt-right to alt-rechts: Twitter analysis of the 2017 German federal election" in *Companion Proceedings of the The Web Conference 2018* (2018), pp. 621–628.
46. M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, M. Starnini, The echo chamber effect on social media. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023301118 (2021).
47. J. Lenti, G. Ruffo, Ensemble of opinion dynamics models to understand the role of the undecided about vaccines. *J. Complex Networks* **10**, cnac018 (2022).
48. S. González-Bailón, V. d'Andrea, D. Freelon, M. De Domenico, The advantage of the right in social media news sharing. *PNAS Nexus* **1**, pgac137 (2022).
49. S. Rathje, J. J. Van Bavel, S. Van Der Linden, Out-group animosity drives engagement on social media. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2024292118 (2021).
50. S. T. Smith *et al.*, Automatic detection of influential actors in disinformation networks. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011216118 (2021).
51. C. K. Tokita, A. M. Guess, C. E. Tarnita, Polarized information ecosystems can reorganize social networks via information cascades. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102147118 (2021).
52. P. T. Metaxas, E. Mustafaraj, Social media and the elections. *Science* **338**, 472–473 (2012).
53. J. Kalyanam, M. Quezada, B. Poblete, G. Lanckriet, Prediction and characterization of high-activity events in social media triggered by real-world news. *PLoS ONE* **11**, e0166694 (2016).
54. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, M. Tesconi, Emergent properties, models and laws of behavioral similarities within groups of Twitter users. *Comput. Commun.* **150**, 47–61 (2020).
55. J. B. Bak-Coleman *et al.*, Stewardship of global collective behavior. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025764118 (2021).
56. C. A. Bail *et al.*, Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 243–250 (2020).

57. P. S. Vishnuprasad *et al.*, "Tracking fringe and coordinated activity on Twitter leading up to the US Capitol attack" in *The 18th International AAAI Conference on Web and Social Media (ICWSM'24)* (AAAI, 2024).

58. V. L. Dao, C. Bothorel, P. Lenca, Community structure: A comparative evaluation of community detection methods. *Network Sci.* **8**, 1–41 (2020).

59. S. Fortunato, M. Barthelemy, Resolution limit in community detection. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 36–41 (2007).

60. L. Ng, I. J. Cruickshank, K. M. Carley, Coordinating narratives framework for cross-platform analysis in the 2021 US Capitol riots. *Comput. Math. Organ. Theory* **29**, 1–17 (2022).

61. J. Calma, Twitter just closed the book on academic research. *The Verge* (2023). https://www.theverge.com/2023/5/31/23739084/twitter-elon-musk-api-policy-chilling-academic-research. Accessed 19 September 2023.

62. European Commission, Proposal for a regulation on a single market for digital services. *Digital Services Act and Amending Directive* (2020). https://eur-lex.europa.eu/legal-content/en/TXT/?uri= COM:2020:825:FIN. Accessed 17 April 2024.

63. S. Pei, L. Muchnik, J. S. Andrade Jr, Z. Zheng, H. A. Makse, Searching for superspreaders of information in real-world social media. *Sci. Rep.* **4**, 5547 (2014).

64. D. Schoch, F. B. Keller, S. Stier, J. Yang, Coordination patterns reveal online political astroturfing across the world. *Sci. Rep.* **12**, 4572 (2022).

65. M. Á. Serrano, M. Boguná, A. Vespignani, Extracting the multiscale backbone of complex weighted networks. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6483–6488 (2009).

66. V. A. Traag, L. Waltman, N. J. Van Eck, From Louvain to Leiden: Guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).

67. D. Powers, Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *J. Mach. Learn. Technol.* **2**, 37–63 (2011).