



# Brexit and bots: characterizing the behaviour of automated accounts on Twitter during the UK election

Matteo Bruno<sup>1\*</sup> , Renaud Lambiotte<sup>2</sup> and Fabio Saracco<sup>3,1</sup>

\*Correspondence:

[matteo.bruno@imtlucca.it](mailto:matteo.bruno@imtlucca.it)

<sup>1</sup>IMT School for Advanced Studies,  
Piazza S. Francesco 19, 55100 Lucca,  
Italy

Full list of author information is  
available at the end of the article

## Abstract

Online Social Networks (OSNs) offer new means for political communications that have quickly begun to play crucial roles in political campaigns, due to their pervasiveness and communication speed. However, the OSN environment is quite slippery and hides potential risks: many studies presented evidence about the presence of disinformation campaigns and malicious activities by genuine or automated users, putting at severe risk the efficiency of online and offline political campaigns. This phenomenon is particularly evident during crucial political events, as political elections. In the present paper, we provide a comprehensive description of the networks of interactions among users and bots during the UK elections of 2019. In particular, we focus on the polarised discussion about Brexit on Twitter, analysing a data set made of more than 10 millions tweets posted for over a month. We found that the presence of automated accounts infected the debate particularly in the days before the UK national elections, in which we find a steep increase of bots in the discussion; in the days after the election day, their incidence returned to values similar to the ones observed few weeks before the elections. On the other hand, we found that the number of suspended users (i.e. accounts that were removed by the platform for some violation of the Twitter policy) remained constant until the election day, after which it reached significantly higher values. Remarkably, after the TV debate between Boris Johnson and Jeremy Corbyn, we observed the injection of a large number of novel bots whose behaviour is markedly different from that of pre-existing ones. Finally, we explored the bots' political orientation, finding that their activity is spread across the whole political spectrum, although in different proportions, and we studied the different usage of hashtags and URLs by automated accounts and suspended users, targeting the formation of common narratives in different sides of the debate.

**Keywords:** Social networks; Bots; Misinformation

## 1 Introduction

Even if the extent of the impact on election results and on public opinion about pivotal political topics is debated [1–3], new media undeniably represent an extensively used novel channel that has contributed to reshape the political communication strategies. Therefore, due to their social relevance and the technical interest, researchers from different

© The Author(s) 2022, corrected publication 2022. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

disciplines have accordingly focused on the description and analysis of social media structure and dynamics, shedding light on non trivial behaviours occurring in debates in online social networks [4–10]. For instance, the topic of echo chambers, i.e. the creation of polarized groups of users based on the access to similar news sources [11], has received great attention. Even if similar phenomena were already observed in offline contexts before the advent of OSNs [12], echo-chambers' effects seem to be amplified in social networks [7, 11]. In particular, many contributions showed how the formation of echo chambers depends extremely on the kind of platform analysed, showing that some of them -for instance, Twitter or Facebook- are more prone to segregate accounts in groups with similar opinions [13–15]. Nevertheless, there is no complete agreement on the existence or even on the role of echo-chambers in the online debate [1, 16–18].

Beside the different mechanism of opinion dynamics due to the different structures of social networks platforms and the various possibilities for users to interact with each other, it is of crucial importance to investigate artificial manipulations of the behaviour of users. For instance, in recent years a great part of the research literature on social networks has been dedicated to the analysis of the behaviour of social bots (or simply *bots*), i.e. automated social accounts controlled by pieces of software. In general, they can be used to automatically share the latest news or provide information in case of emergencies. Nonetheless, in many occasions malicious activities by automated accounts were observed, with bots often hiding and pretending to be human [5, 8]. Among other usages, it was showed that their employment may foster, precisely, the formation of echo chambers [19] or increase the visibility and credibility of users [6, 8, 20, 21]. The injection of coordinated accounts is particularly common during elections and key events in several countries [22–24]. Detecting the bots and analysing the behaviour of malicious accounts is not, however, an easy task. For now more than a decade, scientists have worked on the detection of automated accounts, using a variety of approaches involving machine learning to analyse linguistic and behavioural features [8, 25, 26] and the analysis of social interactions [27, 28].

In this paper, we focus on the discussion about Brexit on Twitter before and during the UK elections of 2019. The presence and influence of automated accounts in the Twitter Brexit discussion was already noted during the 2016 referendum: in their work, Bastos and Mercea [29] found a botnet of more than 10k bots, propagating fake and hyperpartisan news. They found that the bots are specialising in different activities, with some of them retweeting active users and some other generating cascades of information from other bots' tweets. Similarly, the work of Howard and Kollanyi [30] showed that the bots were mainly focused on retweeting and that their activity was very high compared to genuine users. Furthermore, they found that the bots' activity was more in support of Brexit than against it.

Here, we focus on the UK elections in 2019, as they played a pivotal role for the confirmation of Brexit<sup>1</sup> and they polarised the public opinion about the future relations of the UK with the European Union. To investigate this matter, we have analysed a dataset that includes more than 10 million Brexit-related tweets and 1 million users, over a period of more than one month centered on the day of the general election, held on the 12th of December 2019.

---

<sup>1</sup><https://www.bbc.com/news/uk-politics-50303512>

We summarize our Research Objectives (ROs) as follows.

- *RO1*: quantify the presence of bots talking about Brexit during the UK election process.
- *RO2*: understand the bots' strategies in terms of their positions in the networks of retweets.
- *RO3*: uncover the main Brexit subtopics on which the bots focus most.
- *RO4*: unveil the similarities and differences between suspended users and accounts labeled as bots.

For RO1, we check for the presence and the activity patterns of the bots for our time window centered on the UK election day. Dealing with RO2, by building networks of retweets we are able to see if the bots spread equally through the political groups and if they are positioned in the core or periphery of the network. For RO3, we analyse the hashtags and URLs used mostly by the bots. We also build networks of bots by linking those that show similar patterns of hashtags/URLs usage to find coordinated activity by automated accounts. To tackle RO4, we analyse the categories of bots and suspended users separately for the previous ROs.

## 2 Methods

### 2.1 Data collection

Twitter is commonly adopted in the UK, where it is the second online social network platform after Facebook, according to Statista:<sup>2</sup> Twitter accounts for nearly 23.8% of the online social network market, against 56% of Facebook (May 2021). Even if it does not cover the majority of the population, Twitter has been shown in several contexts to be among the preferred channels for political communications [31].

We downloaded our data using Tweepy,<sup>3</sup> a Python wrapper for the official Twitter API. The period that we considered goes from the 20th of November 2019 to the 23rd of December 2019 (the UK elections were held on the 12th of December). During this period, we collected all the tweets that contained the keyword *Brexit* using the `GET /tweets/search/stream`<sup>4</sup> Twitter API, via Tweepy. We chose to limit the search to a single keyword since it is the specific term to indicate the departure of UK from EU and was ubiquitous in the political debate on this subject, such that as a neologism born in 2012, the word “Brexit” was added to Oxford dictionary in 2016.<sup>5</sup>

In order to overcome possible issues due to the usage of a neologism that appears unchanged in different languages, we filter the stream of tweets to English messages. We hit the 1% rate limit on data streaming only on 12th December, the election day. It could be argued that our filtering procedure is not adequate to analyse the whole discussion, as the UK general election also covered other topics of discussion than Brexit. However, our aim is to investigate the underlying sub-discussion on Brexit, as it played a central and polarising role, but also gained global attention from all over the world.

---

<sup>2</sup><https://www.statista.com/statistics/280295/market-share-held-by-the-leading-social-networks-in-the-united-kingdom-uk/>

<sup>3</sup><https://www.tweepy.org/>

<sup>4</sup><https://developer.twitter.com/en/docs/twitter-api/tweets/filtered-stream/api-reference/get-tweets-search-stream>

<sup>5</sup><https://www.bbc.com/news/uk-england-oxfordshire-38326516>

## 2.2 Automated and suspended accounts

To assign a bot score to each account, we used the Botometer API [32] in its most recent v4 version [33]. Our choice is based on the recent comparison of results of various bot-detection algorithms, showing that Botometer has the best performances over recent annotated data sets [34]. Still, let us remark that Botometer can sometimes detect false positives/negatives and can have some general issues, see [35]. However, in our case we do not have many of the problems that are highlighted in the aforementioned paper, such as the language, which in our case is always English, and we are also using a more recent version of Botometer, which is performing better as also noted in [35]. For the sake of completeness, we provide the list of tweets (see Section Availability of data and materials).

Where not otherwise specified, all unverified users<sup>6</sup> with a Botometer CAP (Complete Automation Probability) score greater or equal than 0.43 are considered bots, as was already chosen in [36]. There is no strict consensus on the threshold value that should be chosen. In our case, we chose the threshold so that the top 7 percentile of users will be labeled as bots. A similar approach was adopted by Ferrara et al. [37]. In Appendix A we show that the choice of the threshold does not modify the observed behaviour of the group, even if it changes the total number of automated accounts.

Furthermore, we decided to treat suspended users differently. In our analysis, we encountered a high number of accounts that were suspended (the Twitter API does not give more information). Dealing with these users, it is not possible to examine their past activity and give them a bot score, so we only had the posts that we downloaded. We decided to treat them as a separate class, because it is not granted that a user was suspended for violating the policies of Twitter on automation. Let us remark that the analysis of suspended users is also independent of the choice of bot detector. However, the analysis presented in [38] showed that the activity of suspended users presents many traits of malicious behaviours.

## 2.3 Network models

*Network projection.* In order to analyse the bipartite networks we built with our data, we apply a recently-proposed algorithm to project the information contained in bipartite networks into monopartite ones [39]. The procedure consists in calculating, for each couple of nodes in one layer, the number of common neighbors on the opposite layer in the real systems, as a measure of similarity between nodes. Next, these observed similarities are compared to the expectations of an entropy-based null-model that accounts for the nodes' degrees: if the observed similarity between the two nodes is statistically significant, a link connecting the two nodes is present in the one-mode projection.

We use the implementation proposed in [40] and the relative code of the package NEMtropy<sup>7</sup> for the computation of the randomization and of the monopartite projection. In the following, the nodes corresponding to one layer will be called with the letter  $r = 1, \dots, R$  as they correspond to the rows layer in a biadjacency matrix representation of the network, and the opposite layer will be indexed as  $c = 1, \dots, C$  as they correspond to

---

<sup>6</sup>Twitter has a procedure to check the identity of some users that are of public interests, as the official account of political parties, newspaper, online newscasts, politicians, journalists and VIPs in general. Due to the verification procedure of Twitter, all verified accounts are considered as genuine by Botometer.

<sup>7</sup><https://github.com/nicoloval/NEMtropy>

columns. We will call the biadjacency matrix representing the network with the letter  $\mathbf{B}$  and its entries with  $b_{rc}$ .

More in detail, in a simple one-mode projection, we can produce a monopartite network of the nodes in one layer of a bipartite network by calculating the number of neighbors they share and setting that measurement as the weight of the link between them. The number of common neighbors is:

$$V_{rr'} = \sum_c b_{rc} b_{r'c} = \sum_c V_{rr'}^c. \quad (1)$$

where we use  $V_{rr'}^c \equiv b_{rc} b_{r'c}$  meaning that  $V_{rr'}^c = 1$  if and only if  $c$  shares a link to both  $r$  and  $r'$ . We then compare the measure of the similarity of the nodes to the similarity that is expected from a null-model that accounts for the nodes' degrees. As a benchmark, we use the Bipartite Configuration Model (BiCM [41]), that belongs to the family of entropy-based null-models, [42–46]. These models arise from the maximisation of the Shannon entropy, which corresponds to maximising the uncertainty about the system, given some constraints, and resulting in a maximally unbiased model. Practically, the constrained maximisation of the Shannon entropy  $S = -\sum_{\mathbf{B}} P(\mathbf{B}) \ln P(\mathbf{B})$  leads to an exponential form for the probability distribution

$$P(\mathbf{B}) = \frac{e^{-H(\vec{\theta}, \vec{C}(\mathbf{B}))}}{Z(\vec{\theta})}, \quad (2)$$

where  $\mathbf{C}(\mathbf{B})$  is the vector of constraints evaluated on the biadjacency matrix  $\mathbf{B}$ ,  $\vec{\theta}$  is the vector of the Lagrangian multipliers associated to the maximisation procedure,  $Z(\vec{\theta})$  is the partition function and  $H(\vec{\theta}, \vec{C}(\mathbf{B}))$  is the Hamiltonian associated to the maximisation problem [42]. For each different case, the actual numerical probability  $P(\mathbf{B})$  is determined by the targeted topological constraints that need to be accounted for [42]. In our case, we need to verify that the similarity between two nodes will be statistically significant when considering the respective degrees, so we employ the BiCM, that is the bipartite entropy-based null-model that takes as constraints the degree sequences of the two layers. With these constraints, the probability of a graph factorises in terms of probabilities per link:

$$P(\mathbf{B}) = \prod_{r=1}^R \prod_{c=1}^C p_{rc}^{m_{rc}} (1 - p_{rc})^{1-m_{rc}}. \quad (3)$$

Each link probability will be a function of the  $(R + C)$ -vector of Lagrange multipliers  $\vec{\theta}$  associated to the degrees as

$$p_{rc} = \frac{x_r y_c}{1 + x_r y_c}, \quad (4)$$

where  $x_r = e^{-\theta_r}$  and  $y_c = e^{-\theta_c}$  are reparameterizations of the Lagrange multipliers.

In order to find the numerical value of the probability, we set the average degrees of the model to the observed ones and solve the set of nonlinear equations

$$\begin{cases} \langle k_r \rangle = \sum_c p_{rc} = k_r^*, & r = 1 \dots R \\ \langle h_c \rangle = \sum_r p_{rc} = h_c^*, & c = 1 \dots C, \end{cases} \quad (5)$$

where  $k_r$  and  $h_c$  are the degree of the node  $r$  and  $c$  respectively, and  $*$  indicates the observed empirical value. More details can be found in [40].

The entropy-based null-model allows to treat links as independent random variables, so the probability of a common neighbor reads  $P(V_{rr'}^c = 1) = p_{rc}p_{r'c}$ . The distribution of a  $V_{rr'}$  is then a sum of independent Bernoulli random variables with different coefficients, resulting in a Poisson-Binomial distribution. We then measure the statistical significance of the similarity of two nodes  $r$  and  $r'$  by calculating the p-value of the observed  $V_{rr'}^*$  with respect to the so obtained distribution. To perform this computation, given the large size of our datasets, we approximate the Poisson-Binomial distribution with a Poisson one with the same mean. The error of this approximation is controlled by Le Cam's theorem [47–49]:

$$\sum_{k=0}^C \left| f_{PB}(V_{rr'} = k) - \frac{\mu^k \exp(-\mu)}{k!} \right| < 2 \sum_{c=1}^C (p_{rc}p_{r'c})^2. \quad (6)$$

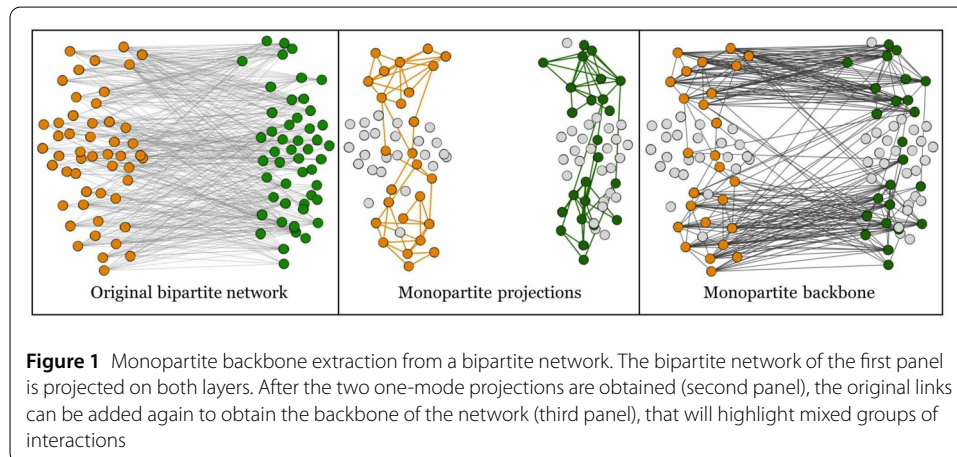
To validate links, we set a threshold that will distinguish the p-values that are significant from the ones to be disregarded. We apply the False Discovery Rate (FDR) procedure [50]. With  $H_1, \dots, H_N$  different hypotheses associated to different p-values, the recipe of FDR prescribes to sort the p-values in increasing order and then look for the largest integer  $i$  such that

$$\text{p-value}_i \leq \frac{\hat{it}}{N} \quad (7)$$

where  $t$  is a single-test significance level (we use  $t = 0.05$  unless otherwise stated). Then we reject all hypotheses  $H_i$  such that  $i < \hat{i}$ . In our case, the procedure validates all links with a p-value lower than the threshold found via the FDR.

*Monopartite backbone.* After obtaining the two one-mode projections of a bipartite network, we add again the links of the original bipartite network to connect the nodes that remained in the projections. This way, we obtain a monopartite network that will show the groups of similar nodes of the same type, and link the groups if they were originally related. A representation of this procedure is shown in Fig. 1.

*Discursive communities.* A method for extracting memberships of users to different discursive communities in social networks was presented in [51] and later refined in [21, 52].





The method is based on the behaviour of verified users, i.e. the accounts that are certified by a Twitter official procedure: these are property of newspapers, newscasts, journalists, politicians, political parties and VIPs in general that can be of public interest. Verified users have a stronger tweeting (i.e. generation of original contents) activity than retweeting (i.e. sharing posts published by others) activity, on average. In this sense, we can leverage on the presence of verified users to extract discursive communities. In fact, we can use the validated projection described earlier and apply it to a bipartite network of verified and unverified users, in which a link is present if one of the users retweeted the other one at least once. In this sense, we are stating that two verified users are perceived as similar if they both interact with a considerable number of unverified accounts. After obtaining the validated projection, we can find different communities of verified users using the Louvain algorithm (see Sect. 2.4). With the so found community labels, we can obtain groups of similar unverified users in the original network by applying a label propagation algorithm. In this work, we simply label a user with the most common label among its neighbours, removing one edge at random in case of ties, using the algorithm proposed in [53] and considering the labels of verified users as fixed. Iterating the procedure several times until consensus is reached and no node will change its label, we obtain a labeling for all users. Since this procedure can yield different results for the random choice we make when breaking ties, we repeat the whole label propagation 1000 times and take the most frequent outcome. More details on this procedure can be found in [21, 51, 52].

## 2.4 Community detection

To investigate the community structure of the networks, we use modularity-optimisation strategies. For monopartite networks, we look for optimal modularity partitions using the Louvain algorithm [54]. The Louvain community detection algorithm is order dependent [55], so we rerun the algorithm after reshuffling the order of the nodes. We then consider the partition in community displaying the largest value of the modularity.

## 2.5 Core-periphery measures

To measure if nodes belong to the core or periphery of their communities, we employ a strategy similar to the one found in Guimerá et al. [56]. We use two measures: the first one, coming directly from this original paper, is called *participation score* and reads

$$P(i) = 1 - \sum_{c=1}^C \left( \frac{k_{ic}}{k_i} \right)^2 \quad (8)$$

where  $i$  is a node of the graph,  $C$  is the number of communities,  $k_i$  is the degree of node  $i$  and  $k_{ic}$  is the degree of node  $i$  towards the community  $c$ , i.e. the number of neighbors of  $i$  belonging to community  $c$ . The participation score will be 0 for a node that has all its neighbors in its own community, and  $\frac{\bar{c}-1}{\bar{c}}$  if its neighbors are all equally distributed inside  $\bar{c}$  communities.

We also evaluate how relevant a node is inside its community, in accordance to the strategy of [56], but using a slightly different measure. In the original reference, in order to compare the degree of the nodes towards other nodes of their community, the authors propose the z-score of the in-degrees inside a community  $c_i$  in the observed distribution.

Actually, due to the non-Gaussian distribution of  $k_{ic_i}$ , the z-score is ill-defined and we prefer to introduce a new measure, that we call *relevance score*:

$$R(i) = -\log \text{pval}(k_{ic_i}) = -\log \mathbf{P}(k_{ic_i} \geq k_{ic_i}). \quad (9)$$

### 3 Results

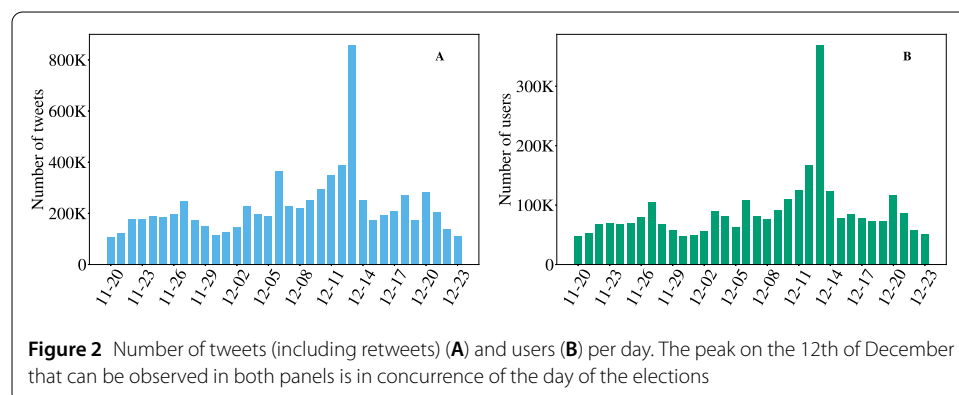
The number of posts per day fluctuates between 100K and 300K for most of the period we consider (Fig. 2), and it jumps to 900K on election day. In total, we analysed more than 10 million posts. The number of users follows a similar trend, being almost never lower than 50K and rising as high as 350K on the election day, for a total number of over one million users.

#### 3.1 Bots statistics

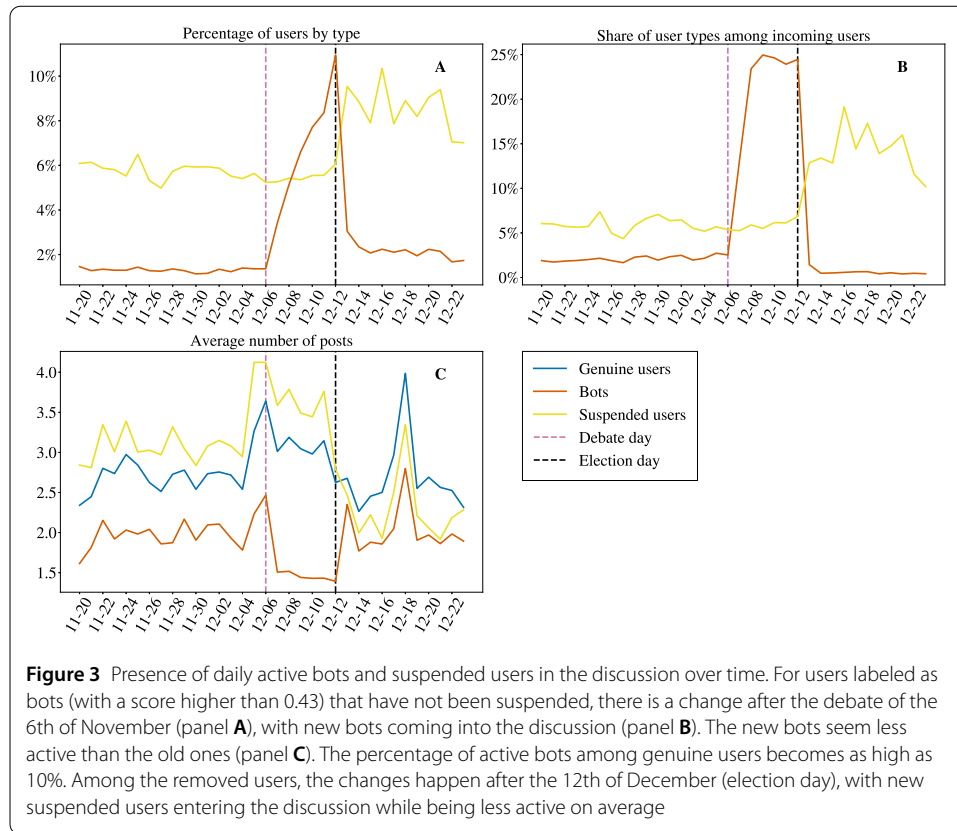
In the following section, we quantify the number of bots in the dataset to attain RO1. The average percentage of bots in the discussion fluctuates around 2% for the first part of our time period (Fig. 3). Interestingly, though, the number of bots in the dataset increases steeply after the 6th of December. It rises as high as 11% on the election day, almost 10% higher than one week before, but rapidly drops after the election. The number of suspended users also shows a change of behaviour around that date, but a different one: their percentage goes up only moderately, from 6% to 9% after the election day, and this increase is not followed by a rapid drop. It can be observed that the increase in the bots' presence comes with a relative decrease in their activity, as per Fig. 3, panel C: bots were, on average, tweeting more before the 6th of December, while the separation between users and bots increases from the 8th and they start to tweet less in the period of their increase before the elections. This behaviour comes mostly from retweet activity, so the bots that enter the Brexit discussion after the 6th of December seem to be more silent than the ones that were already present in the dataset.

For suspended users, the activity trends are almost inverted: they are more active than a normal user and considerably more than a bot, perhaps making it easier for Twitter to detect such accounts. Their activity does not show the same pattern of the bots but instead goes down after the election day.

As shown in Table 1, bots and suspended users tend to retweet malicious accounts more than genuine users, but they are not very supportive of each other and their activity focuses mainly on retweeting genuine accounts, thus generating noise and fostering discussions







**Table 1** Percentage of retweets and quote tweets by users, divided by type. Bots and suspended users are not very supportive of each other and their activity focuses on retweeting genuine accounts, thus generating noise and fostering discussions that are already present

User type	Tot %	Retweeted user type				Quoted user type			
		Suspended	Bots	Non-bots	Verified	Suspended	Bots	Non-bots	Verified
Suspended	8.58%	7.55%	1.18%	48.35%	42.92%	7.31%	0.82%	43.54%	48.33%
Bots	5.6%	3.86%	3.69%	43.85%	48.6%	2.18%	6.4%	34.17%	57.25%
Non-bots	84.51%	3.45%	0.72%	54.11%	41.72%	3.02%	0.56%	47.19%	49.23%
Verified	1.32%	0.82%	0.47%	34.43%	64.28%	1.25%	0.42%	26.06%	72.27%

that are already present. Also, they are not much retweeted or quoted by genuine users, with only around 4% of the retweet/quote activity of genuine unverified users interacting with suspended users or bots, and even less for verified accounts. This suggests that the bots' strategy is to mainly increase the visibility, and implicitly the credibility, of genuine accounts by retweeting their messages. The predilection of social bots for retweeting activities was also observed in US Covid-19 debate on Twitter [37]: the retweeting activity of social bots is more than 10 times their creation of new contents.

The quantity of new bots that appear in the discussion can be seen in panel B of Fig. 3. The figure focuses on the new users and bots that appear for the first time in our dataset. To quantify the users that are appearing for the first time in the first days we have also used the tweets from the days before the 20th of November. The increase in the percentage of bots and suspended users comes with a higher relative percentage of new users of the corresponding type introduced in the discussion. This rapid, and large, increase of a new kind of bots in the dataset appears to be rather unique. With our analysis, we found

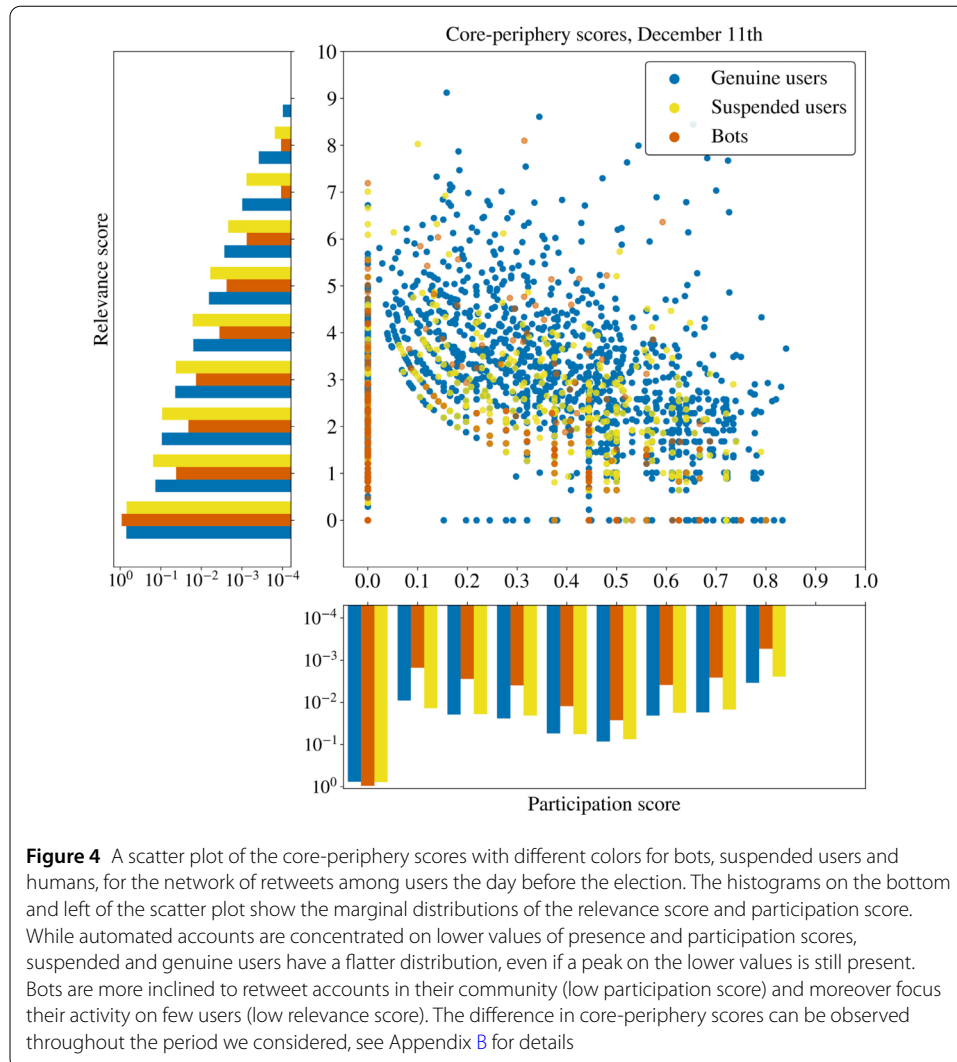
differences between the Twitter activity of the already present bots and of the new bots. Remarkably, while the number of automated and suspended accounts increases (panel A of Fig. 3), their average activity does not present an analogous steep increase (panel C of Fig. 3): even if bots and suspended users enter overwhelmingly in the discussion, their activity remains nevertheless limited. Interestingly, the slopes of curves in panel C are quite similar, even if the levels are different. Otherwise stated, the different kinds of users seem to get a nearly constant share of total messages, while the total volume of posts changes. We also found significant differences in both panel A and B between the slopes of bots and suspended users, contributing towards RO4.

### 3.2 Core-periphery structure and the bots' position

To pursue RO2, we analysed the position of bots and suspended users in the networks of retweets. We obtained each network of retweets by linking two users if one of them has retweeted the other at least once, without considering the direction of the retweet nor their number. In this sense, the resulting network will be undirected and binary. By not considering the number of retweets, we avoid noise from accounts whose activity is much higher with respect to the others.

Our aim is to understand whether the retweet activity of the bots is strategically determined to fuel echo chambers or to target undecided users, by analysing the position of the bots in the network. We partially follow the approach used in [56] to measure the position of the nodes in the networks, and specifically whether the users lie in the core or periphery of their community. We employ a *participation score* (Eq. (8)) to measure how much a user restricts its retweet activity on only one community, and a *relevance score* (Eq. (9)) to measure how relevant a user is inside its community, see the Methods Sect. 2.5 for more details. The results are shown in Fig. 4: the bots, regardless of the community of belonging, seem to have a clear preference for focusing on retweeting users of their same community more than the other users. In the figure, we can see the distributions for the 11th of December. Indeed, the distributions of the *participation score* are very different when distinguishing between genuine users and bots, with a seemingly higher preference of bots for retweeting only in their own community, thus having a lower participation. The *relevance score* reveals instead a comparatively low centrality of the bots, meaning that the bots are usually on the periphery of one community, but focusing only on that single community with a low number of retweets, confirming the findings of [57]. The low scores of bots in both participation and relevance, however, can also be dependent on the degree distribution as well, since they very often tend to limit their activity to a single retweet or a few retweets. In Appendix B we show that the differences in the network position are consistent over time.

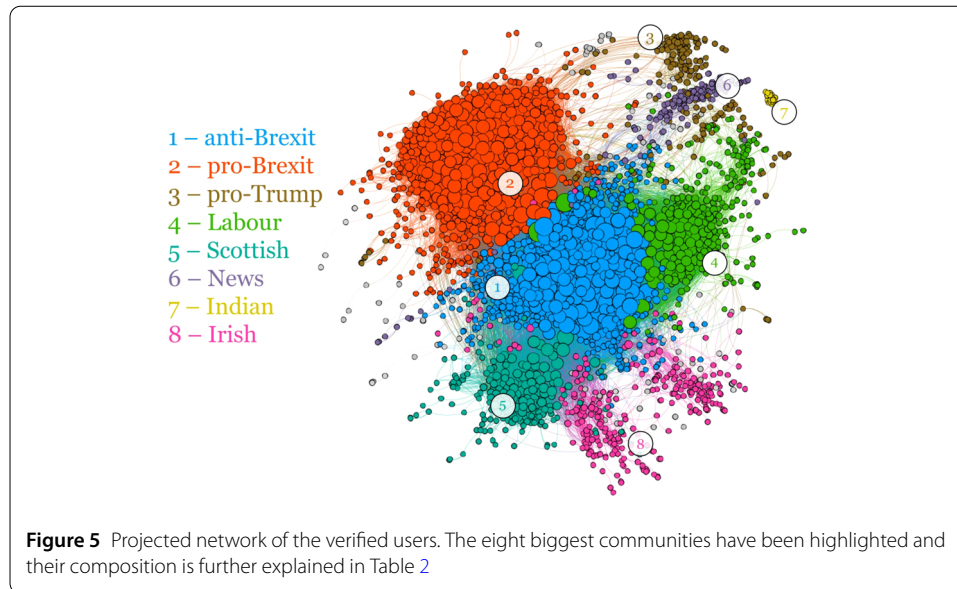
These results on core-periphery positions of the bots can also be explained by considering that the retweet activity of all users is not only focused on users of their same kind, as shown in Table 1. Indeed, bots and non-bots quote suspended users for a total of about 3-3.5% of their activity, while suspended users retweet and quote other suspended users for more than 7% of their total activity. A similar scenario appears for bots, whose retweet and quote activity focuses on bots much more (3.69% and 6.4%) than the activity of suspended and genuine users, who only direct towards bots around 1% of their total retweets. However, the automated accounts do not do enough to push a consistent part of them to the core of the retweet networks, leaving them more to the periphery than the other users.



### 3.3 Discursive communities

To better understand the position of automated accounts and towards RO2, we analysed the bipartite network of verified and unverified users, which we obtained by linking two users if the unverified one has retweeted the verified user at least once. Understanding the structure of such network is useful to characterize the behaviour of accounts with respect to pivotal topics and their interactions with different political sides. With such a network, we analyse the discursive communities that were introduced in the Methods section.

We first project the network on the layer of verified users by using the statistical projection described in Methods. We thus obtain a monopartite network of verified users (Fig. 5), that tells us which couples of verified accounts are retweeted by a significant number of common unverified users. With this network, we obtain labels for each verified user according to the community structure that we find by running the Louvain algorithm [54]. We gave a name to the communities *a posteriori* by looking directly at the users therein: indeed, the information about verified users is certified directly by Twitter platform and thus can be used to assess the nature of these communities. Then we are able to obtain the labels of unverified users by running a simple label propagation algorithm on the bipartite



verified-unverified network. With the so found communities, we quantify the presence of bots and suspended users in each of them.

The results of this procedure can be read in Table 2. The communities have very straightforward interpretations, each containing verified users that can be associated with a stance on Brexit or with a country such as India or Ireland. Bots and suspended users are spread across the network and can be found in all communities, although their share varies consistently. The biggest community is the one supporting EU and against Brexit, with mixed users involved, including actors and famous people, but also politicians and parties such as *LibDems*. Then a pro-Brexit group includes the Conservatives and the Brexit Party. The pro-Brexit community has some interesting features. It presents a high percentage of suspended users, 11.76%, that also appear to be organized since their validation % is the highest among all groups. The validation percentage columns tell us how many users of a certain type are not discarded in the projection of the bots-hashtags and suspended users-hashtags networks (see Sect. 3.4), which in turn is a proxy for the structural coordination of the category of users. Moreover, the Brexit community has the highest link density (i.e. the ratio of the number of links over the number of users in the community) among all communities, meaning that the participation of users in this community is the highest. In general, the highest link densities are found in the groups that are directly involved with Brexit, i.e. pro- and anti-Brexit and the Scottish community; the Labour community link density is then surprisingly low. Remarkably, there is also a Trump supportive community, separated by the other groups and contains a staggering 23% of suspended users and a high coordination of the automated accounts shown by the 12% of validated bots, highest ratio among all groups, and the 7% of validated suspended users, second highest. The Labour party is remarkably forming a community by itself (i.e. separated from the no-Brexit one), and includes also some news accounts. It makes sense to find the Labour community far from the pro-Brexit side but on its own: it highlights the fact that the stance on Brexit of the Labour party was rather unique, since they were proposing a renegotiation of the Brexit deal and a new referendum. Many news accounts lie in a community of their own, linked to communities of both pro- and anti-Brexit stances, but it is not rare to find some

**Table 2** Composition of the main discursive communities of the verified-unverified network after the label propagation. The projection on the verified users is shown in Fig. 5: from the projected network, the so-found communities are fixed labels in the verified-unverified bipartite network, and are then assigned to the rest of the users via label propagation. The *validated bots %* and *validated suspended %* columns show the percentages of bots and suspended users in the community that share a significant number of commonly retweeted hashtags with other users of the same type and therefore appear in the hashtags-users projections of Fig. 7: the higher these percentages are, the more the automated accounts are coordinated

#	Type	Size	Most retweeted verified users	% of bots	% of suspended users	% of verified users	% of validated bots	% of validated suspended users	Link density
1	anti-Brexit	479,718	davidschneider, DavidLammy, mrjamesob, Femi_Sorry, JimMFelton, HackedOffHugh, Channel4News, Keir_Starmer, LibDems	4.66%	4.71%	1.33%	3.16%	4.68%	7.9
2	pro-Brexit	166,254	BorisJohnson, Nigel_Farage, LeaveEUOfficial, Conservatives, brexitparty_uk, darrengrimes_, GoodwinMJ, KTHopkins, JuliaHB1	6.91%	11.76%	1.03%	2.91%	7.34%	12.72
3	pro-Trump	112,766	realDonaldTrump, ScottPresler, DiamondandSilk, RealCandaceO, ddale8, greggutfeld, DineshDSouza, IngrahamAngle, mtracey	3.7%	23.09%	0.71%	12.14%	7%	2.64
4	Labour	106,227	jeremycorbyn, OwenJones84, BBCPolitics, UKLabour, PeterStefanovi2, PeoplesMomentum, yanisvaroufakis, bbcquestiontime	8.79%	6.78%	0.75%	0.71%	1.92%	1.63
5	Scottish	20,258	theSNP, dannywallace, joannaccherry, NicolaSturgeon, IrvineWelsh, Feorlean, PeteWishart, AngusMacNeilSNP	5.35%	4.24%	1.6%	1.94%	5.47%	5.78
6	News	13,297	Reuters, BBCEBreaking, business Brexit, TheEconomist, AJEnglish, AFP, nytimes, FT, CNN	8.92%	7.97%	4.93%	3.63%	1.13%	1.64
7	Indian	8908	gauravcsawant, swapan55, lyervval, abhijitmajumder, AdityaRajKaul, TarekFatah, TVMohandasPai, WIONews, republic, samirsaran	1.95%	13.07%	0.67%	6.32%	4.3%	1.27
8	Irish	3323	SJAMcBride, naomi_long, sinnfeinireland, SenatorMarkDaly, GerryAdamsSF, ClaireHanna, Mr_JSheffield, cstross, glynmoody, rtenews	5.09%	3.49%	11.95%	3.55%	0%	1.41

**Table 3** Average total number of retweets received by users of discursive communities of Fig. 5. In some communities, bots or suspended users are able to make themselves more credible and be retweeted much more: the bots of the pro-Trump community, for instance, are retweeted much more than in the other communities, while the suspended users of pro- and anti-Brexit communities get more attention than the others

#	Type	Average number of retweets received			
		Verified	Genuine	Bots	Suspended
1	anti-Brexit	1255.91	21.28	3.52	11.9
2	pro-Brexit	2924.08	23.14	5.71	19.53
3	pro-Trump	1777.63	4.34	14.59	3.5
4	Labour	3051.15	8.77	0.29	3.88
5	Scottish	1387.31	12.92	0.52	7.35
6	News	127.83	0.54	0.49	0.17
7	Indian	247.1	4.5	2.66	0.29
8	Irish	108.91	0.56	0.17	0.06

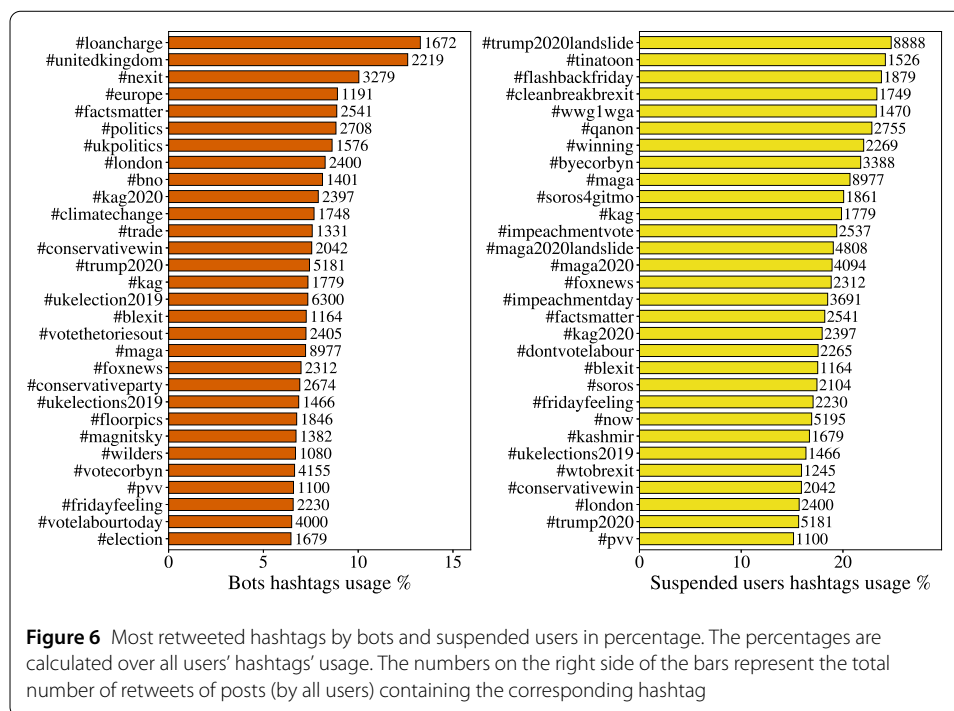
news accounts inside other groups. Other smaller main communities are related to nationalities: Scottish, Indian and Irish in particular.

Furthermore, Table 3 shows that the presence of automated accounts has also a different impact in different communities. In detail, the average total number of retweets per user across all our period is presented. When the number of retweets per user of bots and suspended users is high, the automated accounts are pushed more to the core of the discussion and can be more influential. Interestingly, in the pro-Trump community we find the bots with the highest number of retweets to their posts among all communities, with an average of almost 15 retweets to each bot. The second in terms of this measure is the pro-Brexit community that “only” has 5.7 retweets per bot. Instead, again the pro-Brexit community has the highest average number of retweets per suspended user.

### 3.4 What do the bots say?

Understanding the language used in online social networks is a crucial task with plenty of implications, especially in the case of automated accounts: in fact, a detailed characterization of the language used by social bots permits to study how automated accounts influence humans [19, 58]. Even if a detailed analysis of the language used by the different categories and its implications is beyond the scope of the present paper, we present some brief insights about the frequencies of hashtags used by the various user groups.

To understand what the bots are posting about for RO3, we observed the use of hashtags and URLs by users in our dataset. In Fig. 6 we show the most retweeted hashtags (in percentage) by bots and suspended users. On the right side of the bars, we also wrote the total number of posts found with each hashtag, to let the reader understand which hashtags, among those pushed by the bots/suspended users, are more central in the discussion. Interestingly, we can find topics that are external to Brexit, for example Trump 2020 campaign-related tags, suggesting an external connection of the Brexit debate to Trump’s campaign. Those hashtags are quite aggressive and target some US misinformation arguments (#Qanon, #soros4gitmo, #soros) or refer to the extreme right propaganda (#tinatoon is a popular comic strip of US alt-right; #blexit are black and latin-American Trump supporters that renamed themselves after the Brexit campaign). Nevertheless, suspended users’ hashtags referring to UK elections are supporting conservatives (#byecorbyn, #dontvotelabour, #conservativewin). Note that the use of popular hashtags such as #fridayfeeling in combination with political topics was observed also in [38].

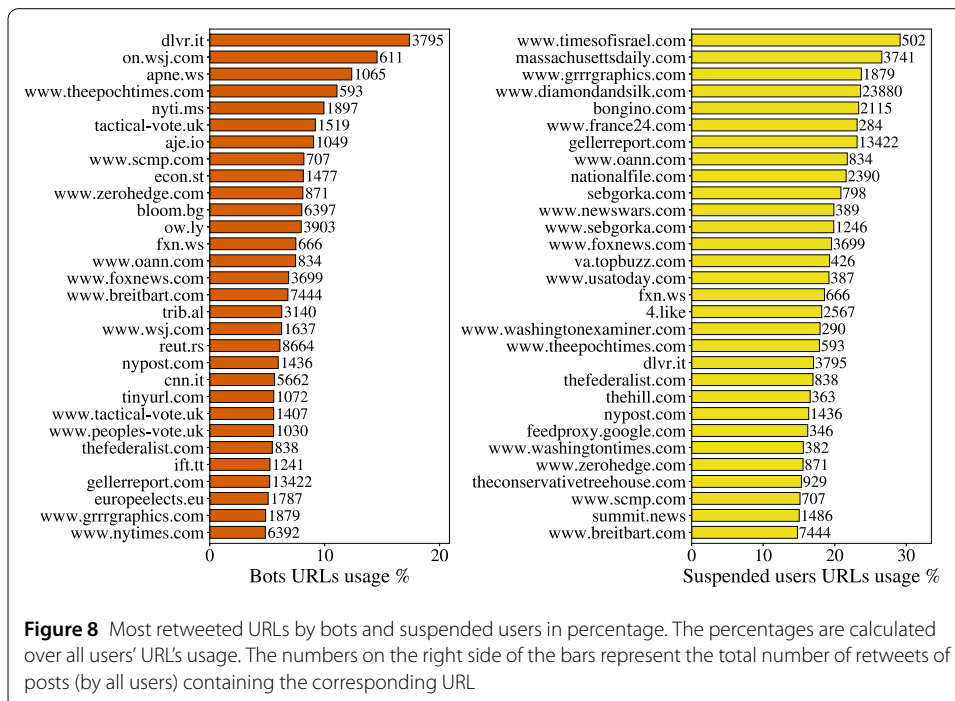
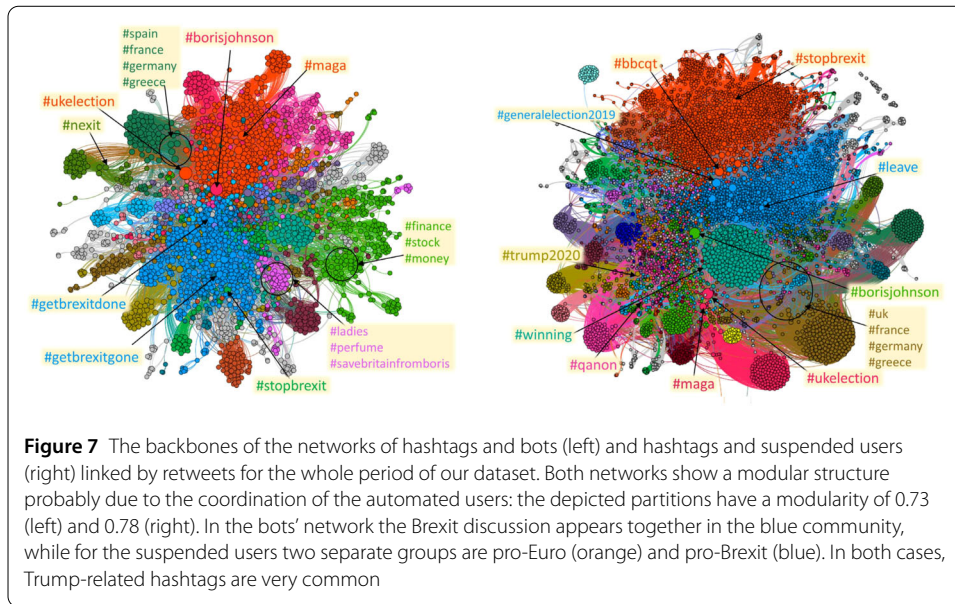


In the activity of bots we still find several Trump-related hashtags (#kag2020, #trump2020, #blexit, #maga, #foxnews), even though we can see more focus on the actual theme of the UK election. In this case, the hashtags used by bots seem to be less polarised: in fact, among the most used ones, we find, #votecorbyn, #conservativewin and #ukpolitics. #votethetoriesout and #conservativewin have nearly the same percentage. It is also worth to mention the presence of #nexit: the word Nexit is the equivalent of Brexit for the Netherlands.

To improve our understanding of the bots' organization, we built two bipartite networks connecting, respectively, bots and hashtags, and suspended users and hashtags. A connection is present between an account and a hashtag if the latter was retweeted at least once by the user. A first inspection suggests a separation of the activity of bots or suspended users in groups. In order to strengthen the signal, we extracted a backbone of the networks by projecting the bipartite network on its layers and re-linking the nodes of the two monopartite projections with the original links (see the Methods section for details). The two backbone networks can be seen in Fig. 7, with bots-hashtags on the left, and suspended-hashtags on the right. Our methodology reveals the dense cliques of hashtags that have common retweeters among bots and suspended users. We then search for communities in the network by running the Louvain algorithm for modularity optimization.

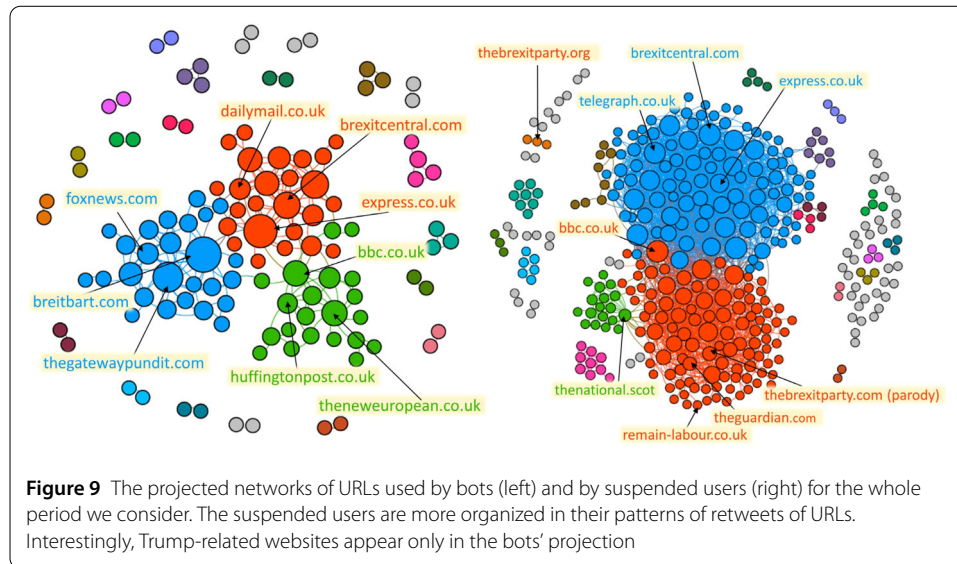
Among the mixed suspended users-hashtags communities highlighted in the right-side network of Fig. 7, there are some that are naturally linked together, and some that are most likely corresponding to an unusual activity of bots. For example, the orange and the blue communities, that are the largest ones, contain hashtags that refer to the elections and each one leans towards one direction (conservative/Brexit for the blue one, labour/pro-EU for the orange one). Meanwhile, some smaller but denser communities hashtags related to Qanon and the Trump campaign (pink, golden and purple, bottom left) or contain links to other EU countries such as France, Germany and Greece. For the network of bots, on





the left of Fig. 7, the situation is similar but this time the discussion on Brexit is mixed in one community (blue), while the other communities contain either groups similar to those found for the suspended users, or other generic topics, maybe injecting noise or advertising products.

Then we analysed the use of URLs by bots and suspended users. In Fig. 8 we present the URLs most used by bots and suspended users compared to genuine users. As in the case of hashtags, on the right side of each bar we also gave the total number of posts containing the corresponding URLs, to better show which of them are more central in the entire discussion. For example, *timesofisrael.com* appears in only a few hundreds retweets in to-



tal, most of them being retweets of a few tweets and is therefore not of particular interest in our analysis; conversely, *diamondandsilk.com* appears in more than 23K retweets. The retweet activity of both categories, but particularly suspended users, spams many websites that are close to Trump's campaign, such as *diamondandsilk.com*, *gellerreport.com*, *breitbart*, *grrrgraphics*.

We also built the bots-URLs retweet network (Fig. 9) similarly to what we did for the hashtags, and in the same spirit we projected the network on the layer of the URLs via a statistical projection. Interestingly, on the projection from the URL-suspended users network we can find a separation of the URLs in the pro-EU and pro-Brexit links, and that is reflected also in newspapers URLs, with the most conservative ones leaning towards the blue community while the most pro-EU lie in the orange community. The green community on the right-side network in Fig. 9 is mostly made of Scottish websites. For the bots-URLs projection (left-side network of the same figure), although the projection is less populated, the story is similar, with a separation in pro-Euro (green) and pro-Brexit (blue) groups, but this time there is a community of Trump supportive websites of comparable size, including URLs and disinformation websites such as *diamondandsilk.com*, *gellerreport.com*, *breitbart.com* and such. This is analogous to what happened in the networks of bots/suspended users and hashtags. In this case, the projection on the users' layer did not validate links between bots or suspended users since the use of URLs by users is way less diversified.

#### 4 Discussion and conclusion

Our analysis confirms that bots played a central role in the Brexit Twitter discussion.

We attained RO1 by quantifying the presence of bots in the discussion, in Sect. 3.1. We have observed that the percentage of bots went from less than 2% to more than 11% in one week, that is crucially the one just before the elections. This radical increase is observed right after the TV debate between the leaders of the main political parties, Boris Johnson and Jeremy Corbyn, which happened six days before the election. Such findings suggest that the injection of bots in the Brexit discussion was not a coincidence, but had the purpose of amplifying the importance of Brexit as a topic during the general election.

Unfortunately, due to the limited time-window of our data, we cannot say if these bots are new to the whole political discussion or if they were simply shifting their attention on Brexit. Furthermore, we have observed differences in the temporal patterns of bots and suspended users. The presence of the suspended users is not increasing at the same time of bots, suggesting that one wave of bots went under the radar in Twitter's bans, and that different types of automated accounts populate the Twitter landscape.

In Sect. 3.2, we contributed to RO2 by highlighting the differences in the activity of bots and genuine accounts and by analysing their different positions in the network structure. We started by considering the core-periphery structure of the network, partially using the approach of [56]: we analysed the position of the bots in the communities found in the networks of retweets, finding that automated accounts tend to be embedded in one single community and to stay in the periphery of it, confirming previous findings [19, 21, 57]. This is in part due to a large share of bots having low degree, but also to the fact that there were not many bots retweeting other bots, thus not being able to push some of the automated accounts to the core of communities. We also showed that the statistical differences between bots and humans increased with the mass insertion of bots. Furthermore, in Sect. 3.3, by considering the retweet patterns of unverified users towards verified users we were able to highlight groups of users tweeting about similar topics, and we analysed the participation of bots and suspended users in such groups. We found that the activity of bots is not limited to specific discussions, but instead they are present everywhere, although in different proportions. For suspended users instead, we found that a large part of their activity is focused on the Trump campaign. It is worth noting that Trump supportive bots have been observed even in French presidential election. Remarkably, in [9], the author detected that some automated accounts that supported 2016's Trump presidential campaign and remained inactive in the following months, were later re-animated to support a disinformation campaign against Macron during French presidential campaign. Such a behaviour suggests, as noticed by the author, the possible existence of a black market for reusable political disinformation bots, also suggested in [59]. In the present manuscript we cannot state if the Trump-supporting bots and suspended users that we found have been used in other contexts, since we do not have any lists of bots active in other debates. Nonetheless, it is interesting to notice the presence this group displaying a so high number of particular efficient (due to their frequency of validated users) suspended accounts, again in a foreign political debate. Anyway, independently of their origin, we observed that bots and suspended accounts supporting Trump campaigns had a greater impact in the UK debate: in fact, Table 3 shows that the number of retweets per bots is much higher in the pro-Trump community than in any other one, even UK based. The different effects of US bots may be, as already suggested by the author of [9], due to the language continuity of UK, since it was absent in the case of French election campaign.

In Sect. 3.4 we dealt with RO3 by analysing the use of hashtags and URLs by bots. Our analysis has revealed that while there was a big activity by automated accounts, the streams of coordinated bots do not look to be only specifically linked to the Brexit discussion, but also on different topics that could benefit from the Brexit debate, mainly populist narratives. The two main topics that we uncovered are the Trump 2020 campaign, and perspectives of a further division in the EU.

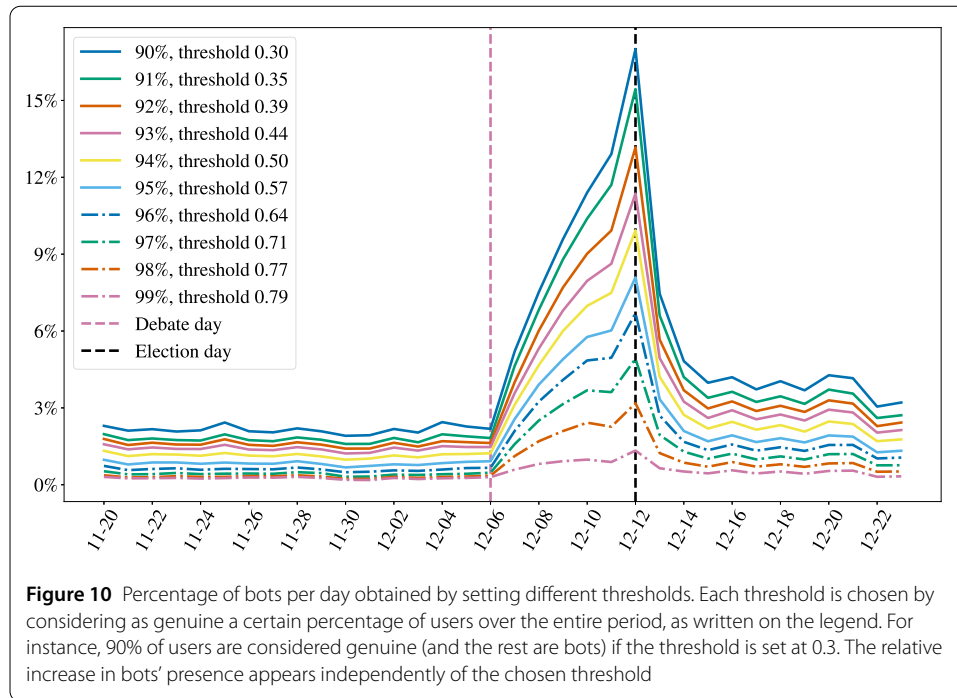
Throughout our analysis, we treated the categories of bots and suspended users separately, in pursuit of RO4. In doing so, we were able to find very different patterns from

the two types of agents. The topics most retweeted by both categories were quite similar, with a peculiar attention to populist topics and fake news websites. However, bots and suspended accounts are both injected in the conversation in mass, but not at the same time: bots increase radically in the days before the election and just after a pivotal debate, while the presence of suspended users grows rapidly after the election and is seemingly more linked to related populist topics or to Trump's propaganda. Moreover, the levels of activity of the two categories were very different: our research shows that the accounts that have been suspended by Twitter were much more active than the automated accounts that were not suspended. This evidence could mean that the bots whose main activity consists in boosting one profile or topic through their retweets are able to go under cover much more than noisier accounts.

In achieving our ROs, we found distinguishable traits of the behaviour of bots. Accounting for false positives, as well as the limitation due to the choice of a threshold, are common concerns in bot detection; nevertheless, the users that we are labeling as bots present behaviours that are very close to those found in other studies [5, 8, 19, 36, 37, 57], thus we are reasonably confident about our results. Consider also that our analysis of suspended users is independent of the bot detection procedure and it contributes deeply to the narrative of our results.

The aim of our study was to characterize the presence of automated accounts on the "Brexit" subject during UK national election and identify their role in the debate. Summarising, we have found a significant presence of automated accounts in the Brexit discussion during the UK elections of 2019, following the findings of Bastos and Mercea [29] and Howard and Kollanyi [30] that also found a large presence of bots during the Brexit referendum. The activity of bots also presents a steep, apparently coordinated, increase, and this result is robust to the choice of the threshold used in the Botometer score. This increase comes at a crucial time, after a widely popular TV debate between political candidates one week before the elections and the elections themselves. We further analysed the behaviour of the automated accounts by analysing several network features, such as the core-periphery positions of users and the similarities of bots in the retweet patterns via the analysis of different kind of networks. We found that the automated accounts are spread across all the discussions. Furthermore, the bots behave significantly differently with respect to the genuine users, and we found groups of bots retweeting hashtags not always perfectly centred on the Brexit discussion but relative to other populist narratives. Remarkably, our findings corroborate some other results present in the literature about bot behaviour [5, 8, 19], for the specific context of UK election campaign. As already noted above, the presence of foreign social bots is significant: they were relatively easy to identify, since they used pro-Trump hashtags, together with pro-Brexit ones. Such an observation suggests the presence of an attention of bots' masters to different events in various countries, as already observed in [9].

Due to the nature of our dataset, some questions remain unanswered. For instance, we do not know if the spike in bots' activity is due to political bots shifting their attention on Brexit, as that would require a more general stream of tweets. However, further similar analyses on the specialisation of bots on a certain subtopic of a general discussion are needed to investigate these questions. Regarding methodological limits, while the general approach of entropy-based null-models is well grounded, the effectiveness of the extraction of discursive communities, even if it has been validated *a posteriori* with really



good performances in different data sets [21, 51, 52, 60], has never been extensively tested. A comparison among different approaches is in preparation and suggests that the method implemented in the present manuscript is the one with the best performances [61].

### Appendix A: Choice of the CAP threshold in Botometer

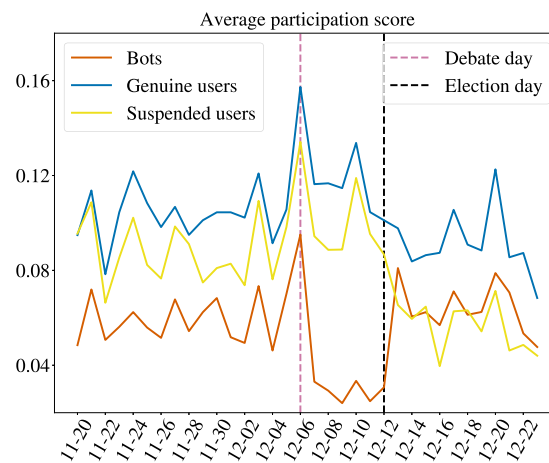
It is remarkable that the pattern observed in Fig. 3 is independent of the threshold chosen to label users as bots: in Fig. 10 we show that the relative percentage increase of bots happens with different thresholds. We compare the number of bots detected each day by selecting different thresholds in order to have in total from 1% (threshold 0.79) to 10% (0.3) of users labeled as bots. In our case, we chose a threshold that was already used in [37] and labels as bots about 7% of all users in the dataset.

### Appendix B: Average participation scores

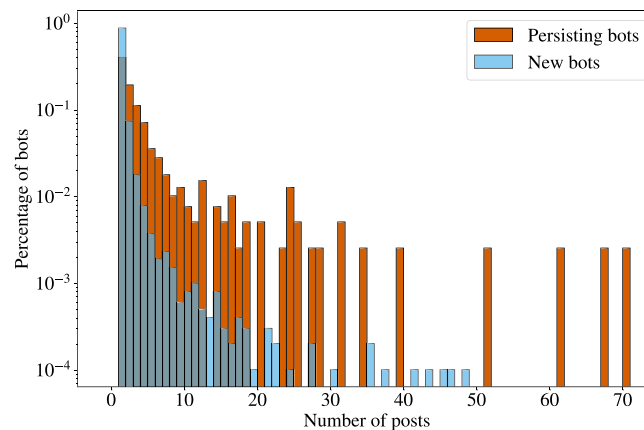
We computed the average participation score for the different categories of users over time (Fig. 11). The bots are almost always more focused on a single community than genuine users, and the difference increases in the week before the election. Suspended users behave similarly to the genuine users until the election day, after which they have scores more similar to the bots.

### Appendix C: Behaviour of incoming bots

It could be argued that the change of behaviour of the bots is not caused by the incoming automated accounts but rather by a change of behaviour of the ones that were already present in the dataset. Figure 12 suggests that this should not be the case. To explore this hypothesis, we separated the bots present on the 11th of December between the ones that were also present before the 6th of December, and the ones that were not. New bots are



**Figure 11** Average values of participation score by category. There is a sudden change of behaviour of bots after the election day in both the presence and participation score: in fact, the average values of the score are almost always higher for bots than for suspended users after the 12th of December. At the same time, the participation score of suspended accounts is always higher than average users before the election day, dropping to much lower levels after the elections



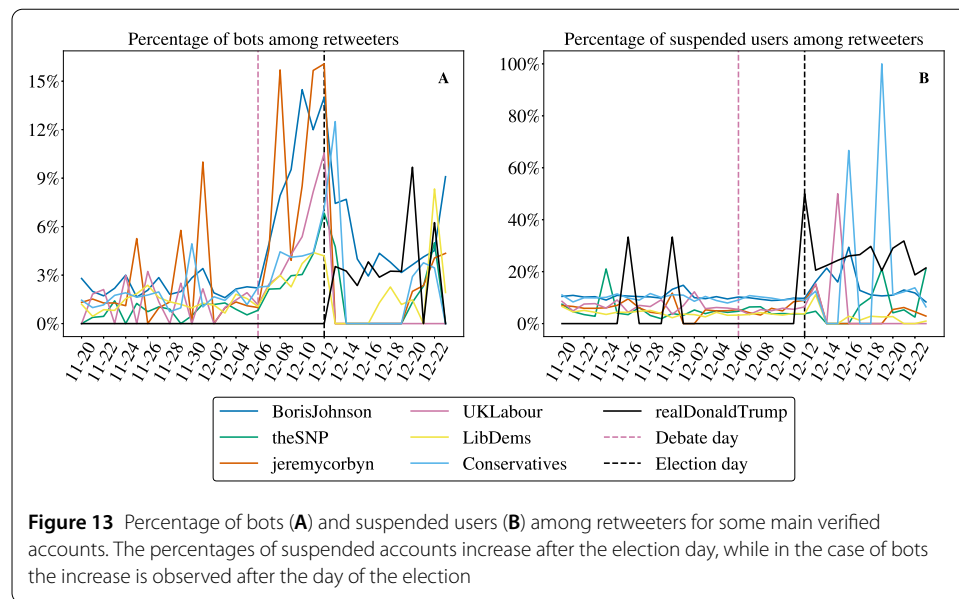
**Figure 12** Histogram of the number of posts made by bots. We took bots from the 11th of December, and the two histograms separate the activity of those that were already present in the discussion the 6th of December (persisting bots) from the new entries (new bots). Almost 40% of those of the 6th of December are still in the discussion, and their average activity is almost stable from the two dates. The old users' activity is much higher, so the effect of the less activity is only caused by the new bots' behaviour

clearly less active, usually limiting their activity to a single retweet, while the ones that were already in the discussion show a more variegated activity.

#### Appendix D: Automated retweeters of verified users

We also analysed what is the impact of bots and suspended users on the retweeters of some of the most popular verified accounts from various political parties. We found that the increase of bots is transversal across the users that we considered, while the suspended users' increment seems to be driven by mainly *realDonaldTrump*'s retweeters, and partially *BorisJohnson*'s retweeters (Fig. 13).





**Figure 13** Percentage of bots (A) and suspended users (B) among retweeters for some main verified accounts. The percentages of suspended accounts increase after the election day, while in the case of bots the increase is observed after the day of the election

#### Acknowledgements

Not applicable.

#### Funding

MB acknowledges support from the IMT School for Advanced Studies program Frontier Proposal Fellowship. FS acknowledges support from the IMT School for Advanced Studies PAI project Toffee. This work is supported by the European Union – Horizon 2020 Program under the scheme “INFRAIA-01-2018-2019 – Integrating Activities for Advanced Communities”, Grant Agreement n.871042, “SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics” ([www.sobigdata.eu](http://www.sobigdata.eu)).

#### Abbreviations

OSN, Online Social Networks; RO, Research Objective; CAP, Complete Automation Probability; BiCM, Bipartite Configuration Model; FDR, False Discovery Rate.

#### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. The list of all the tweets' identifiers relative to this study is available at <https://toffee.imtlucca.it/datasets>.

#### Declarations

##### Competing interests

The authors declare that they have no competing interests.

##### Authors' contributions

All authors designed the analysis, wrote and reviewed the manuscript. MB performed the analysis. All authors read and approved the final manuscript.

##### Author details

<sup>1</sup>IMT School for Advanced Studies, Piazza S. Francesco 19, 55100 Lucca, Italy. <sup>2</sup>Mathematical Institute, University of Oxford, Woodstock Road, OX2 6GG Oxford, UK. <sup>3</sup>Institute for Applied Mathematics, National Research Council, Via dei Taurini 19, 00185 Rome, Italy.

#### Publisher's Note

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

Received: 30 July 2021 Accepted: 9 March 2022 Published online: 22 March 2022

#### References

- Dubois E, Blank G (2018) The echo chamber is overstated: the moderating effect of political interest and diverse media. *Inf Commun Soc* 21:729–745. <https://doi.org/10.1080/1369118X.2018.1428656>
- Valensise CM, Cinelli M, Nadini M, Galeazzi A, Peruzzi A, Etta G, Zollo F, Baronchelli A, Quattrocioni W (2021) Lack of evidence for correlation between COVID-19 infodemic and vaccine acceptance. <https://arxiv.org/abs/2107.07946v2>



3. Gallotti R, Pilati F, Sacco PL, De Domenico M Comment on "The COVID-19 infodemic does not affect vaccine acceptance". <https://doi.org/10.31219/OSFIO/M8J32>
4. Adamic LA, Glance N (2005) The political blogosphere and the 2004 U.S. election: divided they blog. In: Proceedings of the 3rd international workshop on link discovery, LinkKDD '05. ACM, New York, pp 36–43. ISBN 1595932151. <https://doi.org/10.1145/1134271.1134277>
5. Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M (2015) Fame for sale: efficient detection of fake Twitter followers. *Decis Support Syst* 80:56–71
6. Bessi A, Ferrara E (2016) Social bots distort the 2016 US presidential election online discussion. *First Monday* 21(11). <https://doi.org/10.5210/fm.v21i11.7090>
7. Del Vicario M, Vivaldo G, Bessi A, Zollo F, Scala A, Caldarelli G, Quattrociocchi W (2016) Echo chambers: emotional contagion and group polarization on Facebook. *Sci Rep*. <https://doi.org/10.1038/srep37825>
8. Ferrara E, Varol O, Davis C, Menczer F, Flammini A (2016) The rise of social bots. *Commun ACM* 59(7):96–104
9. Ferrara E (2017) Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday* 22. <https://doi.org/10.5210/FM.V22I8.8005>
10. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D, Schudson M, Sloman SA, Sunstein CR, Thorson EA, Watts DJ, Zittrain JL (2018) The science of fake news: addressing fake news requires a multidisciplinary effort. *Science* 359:1094–1096. <https://doi.org/10.1126/science.aao2998>
11. Garrett RK (2009) Echo chambers online?: politically motivated selective exposure among Internet news users. *J Comput-Mediat Commun* 14:265–285. <https://doi.org/10.1111/j.1083-6101.2009.01440.x>
12. Oshagan H (1996) Reference group influence on opinion expression. *Int J Public Opin Res* 8(4):335–354. <https://doi.org/10.1093/ijpor/8.4.335>
13. Artime O, d'Andrea V, Gallotti R, Sacco PL, De Domenico M (2020) Effectiveness of dismantling strategies on moderated vs. unmoderated online social platforms. *Sci Rep* 10:14392. <https://doi.org/10.1038/s41598-020-71231-3>
14. De Francisci Morales G, Monti C, Starnini M (2021) No echo in the chambers of political interactions on Reddit. *Sci Rep* 11:2818. <https://doi.org/10.1038/s41598-021-81531-x>
15. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M (2021) The echo chamber effect on social media. *Proc Natl Acad Sci USA* 118:e2023301118. <https://doi.org/10.1073/PNAS.2023301118>
16. Barberá P, Jost JT, Nagler J, Tucker JA, Bonneau R (2015) Tweeting from left to right: is online political communication more than an echo chamber? *Psychol Sci* 26(10):1531–1542. PMID: 26297377. <https://doi.org/10.1177/0956797615594620>
17. Puschmann C (2018) Beyond the bubble: assessing the diversity of political search results. *Dig Journal* 7:824–843. <https://doi.org/10.1080/21670811.2018.1539626>
18. Dubois E, Minaeian S, Paquet-Labelle A, Beaudry S (2020) Who to trust on social media: how opinion leaders and seekers avoid disinformation and echo chambers. *Soc Media Soc* 6. <https://doi.org/10.1177/2056305120913993>
19. Stella M, Ferrara E, De Domenico M (2018) Bots increase exposure to negative and inflammatory content in online social systems. *Proc Natl Acad Sci USA* 115(49):12435–12440
20. Broniatowski DA, Jamison AM, Qi S, Alkulaib L, Chen T, Benton A, Quinn SC, Dredze M (2018) Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Publ Health* 108(10):1378–1384
21. Caldarelli G, De Nicola R, Del Vigna F, Petrocchi M, Saracco F (2020) The role of bot squads in the political propaganda on Twitter. *Commun Phys* 3(1):1–15
22. Shao C, Ciampaglia GL, Varol O, Yang KC, Flammini A, Menczer F (2018) The spread of low-credibility content by social bots. *Nat Commun*. <https://doi.org/10.1038/s41467-018-06930-7>
23. Pastor-Galindo J, Zago M, Nespoli P, Bernal SL, Celdrán AH, Pérez MG, Ruipérez-Valiente JA, Pérez GM, Mármol FG (2020) Spotting political social bots in Twitter: a use case of the 2019 Spanish general election. *IEEE Trans Netw Serv Manag* 17(4):2156–2170
24. Rossi S, Rossi M, Upreti B, Liu Y (2020) Detecting political bots on Twitter during the 2019 Finnish parliamentary election. In: Proceedings of the 53rd Hawaii international conference on system sciences
25. Lee K, Caverlee J, Webb S (2010) Uncovering social spammers: social honeypots+ machine learning. In: Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval, pp 435–442
26. Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M (2017) Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling. *IEEE Trans Dependable Secure Comput* 15(4):561–576
27. Mehrotra A, Sarreddy M, Singh S (2016) Detection of fake Twitter followers using graph centrality measures. In: 2016 2nd international conference on contemporary computing and informatics (IC3I), pp 499–504. IEEE
28. Jia J, Wang B, Gong NZ (2017) Random walk based fake account detection in online social networks. In: 2017 47th annual IEEE/IFIP international conference on dependable systems and networks (DSN), pp 273–284. IEEE
29. Bastos MT, Mercea D (2019) The Brexit botnet and user-generated hyperpartisan news. *Soc Sci Comput Rev* 37(1):38–54. <https://doi.org/10.1177/0894439317734157>
30. Howard PN, Kollanyi B (2016) Bots, #strongerin, and #Brexit: computational propaganda during the UK-EU referendum. Available at SSRN 2798311
31. Bradshaw S, Howard P (2018) How does junk news spread so quickly across social media? Algorithms, advertising and exposure in public life. White Paper, Oxford Internet Institute
32. Davis CA, Varol O, Ferrara E, Flammini A, Menczer F (2016) Botornot: a system to evaluate social bots. In: Proceedings of the 25th international conference companion on world wide web, pp 273–274
33. Sayyadiharikandeh M, Varol O, Yang K-C, Flammini A, Menczer F (2020) Detection of novel social bots by ensembles of specialized classifiers. In: Proceedings of the 29th ACM international conference on information & knowledge management. ACM, New York. <https://doi.org/10.1145/3340531.3412698>
34. De Nicola R, Petrocchi M, Pratelli M (2021) On the efficacy of old features for the detection of new bots. *Inf Process Manag* 58(6):102685. <https://doi.org/10.1016/j.ipm.2021.102685>
35. Rauchfleisch A, Kaiser J (2020) The false positive problem of automatic bot detection in social science research. *PLoS ONE* 15(10):e0241045
36. Varol O, Ferrara E, Davis C, Menczer F, Flammini A (2017) Online human-bot interactions: detection, estimation, and characterization. In: Proceedings of the international AAAI conference on web and social media, vol 11

37. Ferrara E (2020) # covid-19 on twitter: bots, conspiracies, and social media activism. arXiv preprint, [arXiv:2004.09531](https://arxiv.org/abs/2004.09531)
38. Chowdhury FA, Allen L, Yousuf M, Mueen A (2020) On Twitter purge: a retrospective analysis of suspended users. In: Companion proceedings of the web conference 2020, pp 371–378
39. Saracco F, Straka MJ, Di Clemente R, Gabrielli A, Caldarelli G, Squartini T (2017) Inferring monopartite projections of bipartite networks: an entropy-based approach. *New J Phys* 19(5):053022. [http://arxiv.org/abs/1607.02481](https://arxiv.org/abs/1607.02481). <https://doi.org/10.1088/1367-2630/aa6b38>
40. Vallarano N, Bruno M, Marchese E, Trapani G, Saracco F, Squartini T, Cimini G, Zanon M (2021) Fast and scalable likelihood maximization for exponential random graph models with local constraints. *Sci Rep* 11(1):15227
41. Saracco F, Di Clemente R, Gabrielli A, Squartini T (2015) Randomizing bipartite networks: the case of the world trade web. *Sci Rep* 5:10595. <https://doi.org/10.1038/srep10595>
42. Park J, Newman MEJ (2004) Statistical mechanics of networks. *Phys Rev E* 70:066117. <https://doi.org/10.1103/PhysRevE.70.066117>
43. Garlaschelli D, Loffredo MI (2008) Maximum likelihood: extracting unbiased information from complex networks. *Phys Rev E* 78:015101. <https://doi.org/10.1103/PhysRevE.78.015101>
44. Squartini T, Garlaschelli D (2011) Analytical maximum-likelihood method to detect patterns in real networks. *New J Phys* 13(8):083001. <https://doi.org/10.1088/1367-2630/13/8/083001>
45. Fronczak A (2014) Exponential random graph models. In: Encyclopedia of social network analysis and mining, pp 500–517. <https://doi.org/10.1007/978-1-4614-6170-8>
46. Cimini G, Squartini T, Saracco F, Garlaschelli D, Gabrielli A, Caldarelli G (2018) The statistical physics of real-world networks. *Nat Rev Phys* 1(1):58–71. [http://arxiv.org/abs/1810.05095](https://arxiv.org/abs/1810.05095). <https://doi.org/10.1038/s42254-018-0002-6>
47. Deheuvels P, Puri ML, Ralescu SS (1989) Asymptotic expansions for sums of nonidentically distributed Bernoulli random variables. *J Multivar Anal* 28(2):282–303. [https://doi.org/10.1016/0047-259X\(89\)90111-5](https://doi.org/10.1016/0047-259X(89)90111-5)
48. Volkova AY (1996) A refinement of the central limit theorem for sums of independent random indicators. *Theory Probab Appl* 40(4):791–794. <https://doi.org/10.1137/1140093>
49. Hong Y (2013) On computing the distribution function for the Poisson binomial distribution. *Comput Stat Data Anal* 59:41–51. <https://doi.org/10.1016/j.csda.2012.10.006>
50. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300
51. Becatti C, Caldarelli G, Lambiotte R, Saracco F (2019) Extracting significant signal of news consumption from social networks: the case of Twitter in Italian political elections. *Palgrave Commun* 5(1):1–16
52. Caldarelli G, De Nicola R, Petrocchi M, Pratelli M, Saracco F (2021) Flow of online misinformation during the peak of the Covid-19 pandemic in Italy. *EPJ Data Sci* 10:34. <https://doi.org/10.1140/EPJDS/S13688-021-00289-4>
53. Raghavan UN, Albert R, Kumara S (2007) Near linear time algorithm to detect community structures in large-scale networks. *Phys Rev E* 76(3):036106
54. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* 2008(10):P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
55. Fortunato S (2010) Community detection in graphs. *Phys Rep* 486:75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>
56. Guimerà R, Nunes Amaral LA (2005) Functional cartography of complex metabolic networks. *Nature* 433(7028):895–900. <https://doi.org/10.1038/nature03288>
57. González-Bailón S, De Domenico M (2021) Bots are less central than verified accounts during contentious political events. *Proc Natl Acad Sci USA* 118(11):e2013443118
58. Stella M (2021) Cognitive network science for understanding online social cognitions: a brief review. *Top Cogn Sci*. <https://doi.org/10.1111/tops.12551>
59. Lewis B, Marwick AE (2017) Media manipulation and disinformation online. Data & Society Research Institute. <https://datasociety.net/library/media-manipulation-and-disinfo-online/>
60. Radicioni T, Squartini T, Pavan E, Saracco F (2021) Networked partisanship and framing: a socio-semantic network analysis of the Italian debate on migration. *PLoS ONE* 16:e0256705. <https://doi.org/10.1371/JOURNAL.PONE.0256705>
61. Guarino S, Mastrostefano E, Saracco F (2022) Discursive community detection on Twitter. In preparation

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)