



## Dynamic social media affiliations among UK politicians

Iain S. Weaver<sup>a,\*</sup>, Hywel Williams<sup>a</sup>, Iulia Cioroianu<sup>b</sup>, Matthew Williams<sup>a</sup>, Travis Coan<sup>b</sup>, Susan Banducci<sup>b</sup>

<sup>a</sup> School of Computer Science, University of Exeter, Exeter EX4 4QF, UK

<sup>b</sup> Department of Politics, University of Exeter, Exeter EX4 4RJ, UK

### ARTICLE INFO

#### Article history:

#### Keywords:

Twitter  
UK politics  
Dynamic networks  
Legislator networks  
Clustering

### ABSTRACT

Inter-personal affiliations and coalitions are an important part of politicians' behaviour, but are often difficult to observe. Since an increasing amount of political communication now occurs online, data from online interactions may offer a new toolkit to study ties between politicians; however, the methods by which robust insights can be derived from online data require further development, especially around the dynamics of political social networks. We develop a novel method for tracking the evolution of community structures, referred to as 'multiplex community affiliation clustering' (MCAC), and use it to study the online social networks of Members of Parliament (MPs) and Members of the European Parliament (MEPs) in the United Kingdom. Social interaction networks are derived from social media (Twitter) communication over an eventful 17-month period spanning the UK General Election in 2015 and the UK Referendum on membership of the European Union in 2016. We find that the social network structure linking MPs and MEPs evolves over time, with distinct communities forming and re-forming, driven by party affiliations and political events. Without including any information about time in our model, we nevertheless find that the evolving social network structure shows multiple persistent and recurring states of affiliation between politicians, which align with content states derived from topic analysis of tweet text. These findings show that the dominant state of partisan segregation can be challenged by major political events, ideology, and intra-party tension that transcend party affiliations.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Elected politicians do not represent their constituencies in a social vacuum. Members of Parliament (MPs) elected to the House of Commons typically serve in parliament for several years, during which time they necessarily form social connections with their colleagues. It is reasonable to expect that their social networks will have effects on their opinions and behaviour, and indeed there is a broad literature on the importance of social network structure for many aspects of human behaviour. Of particular relevance to politics and political behaviour, peer influence in online and offline social networks is known to affect opinions and attitudes (Bond et al., 2012; Centola, 2010; Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2008; Kramer et al., 2014; Muchnik et al., 2013; Salganik et al., 2006; Sunstein, 2007). Thus the pattern of who interacts with whom in political processes might be expected to give some explanatory power for actions such as voting, campaigning, and debating. A common finding in social networks of all types is the clustering of individuals that have similar attributes, known

as 'homophily' (McPherson et al., 2001), which can arise from selectivity in the formation and maintenance of network connections (preferential linking to alike others) as well as from peer influence (linked individuals becoming more alike) (Shalizi and Thomas, 2011). When homophilic interactions aggregate to create partisan groupings within social networks, also known as 'echo-chambers' (Adamic and Glance, 2005; Conover et al., 2012; Williams et al., 2015), the amplification of like-minded views (and exclusion of alternate views) can prevent effective debate and make cross-party consensus difficult (Sunstein, 2007).

Data for reconstructing politicians' social networks and community affiliation – apart from their party membership – is limited. Since we cannot realistically observe the social interactions of politicians, we need to rely on indirect observations and inference. Voting records of elected politicians have previously been used to identify individual ideological positions in some legislative bodies, for example, the US Congress and Italian Parliament (Dal Maso et al., 2014; Waugh et al., 2009). However, attempts to use roll calls to capture variation in legislator positions within parties in systems with a high degree of party discipline, such as the UK, have not been successful (Hix, 2010; Spirling and McLean, 2006), since

\* Corresponding author.

E-mail address: [i.s.weaver@exeter.ac.uk](mailto:i.s.weaver@exeter.ac.uk) (I.S. Weaver).

“... the positions of MPs are mainly determined by the positions of their parties” (Hix, 2010). This inability to use roll call votes to identify the positions of individual legislators generated a stream of papers that develop methods based on other types of data, such as Early Day Motions (Kellermann, 2012) or legislative speeches (Beauchamp et al., 2011; Lauderdale and Herzog, 2016). However, it is important to note the distinction between estimating legislator ideological positions and identifying the evolving structure of legislator social interactions, and evolving informal membership within peer communities. Our focus here is on capturing these informal affiliations using legislator social media interactions, with a specific aim of observing how they evolve over time in response to current events. Social interactions may indeed be strongly related to the underlying ideological positions of individual legislators, but testing this hypothesis is beyond the scope of this paper.

The widespread use of online social media by politicians (see Jungherr (2016) for a systematic review of the literature on the use of Twitter in electoral campaigns) provides a rich data resource for studying their interpersonal interactions. Previous studies on the use of Twitter by UK politicians have focused on: the characteristics which make candidates more likely to be active on the platform (Darren et al., 2010); differences in the way politicians use Twitter (Graham et al., 2013; Newman, 2010); and the content of their tweets (Baxter and Marcella, 2012; Theocharis et al., 2016). Despite of the importance of network-based approaches in understanding political phenomena (Lazer, 2011), no previous studies that we are aware of have looked at the networks of Twitter interactions between UK political elites. Outside the UK, a common finding across studies of politicians' social media interactions is that ideology plays a crucial role in the formation of Twitter links and communities. This finding holds across multiple countries and levels of elections. Boireau et al. (2015) analyse the network of Twitter interactions among candidates competing for federal, regional and European elections in Belgium, finding that interactions between candidates tend to be clustered around political ideology. Cherepnalkoski and Mozetic (2016) analyse the network of retweets among Members of the European Parliament (MEPs), showing that communities identified in social networks correspond to actual political groups. Other recent studies examine legislators' voting patterns, such as Cherepnalkoski et al. (2016), who study MEPs to show that there is considerable correlation between voting and retweeting patterns, and that the ideological left-to-right alignment of the political groups is reflected in the retweet network. Conversely, Cook (2016) analyses co-voting, bill co-sponsorship and diadic Twitter interactions between US Senators and Maine State Legislators, finding that Twitter ties are less partisan than voting and bill co-sponsorship.

A general challenge in use of social media to infer relationships amongst politicians is that the system under study is highly dynamic. Social media is a high-volume and high-velocity medium in which millions of utterances from millions of users form a complex web of communication. Gaining robust insights from such a dynamic complex system is challenging. A common approach in previous work is to identify a set of target users (typically elected politicians, in the field of political science) and then to aggregate content over some time period to form a single snapshot of their interactions. While this approach can be insightful, it ignores a key aspect of political discourse, which is the dynamics of communication and how the interactions between individual politicians are affected by events.

In this paper we make two unique contributions – one substantive and one methodological. First, we examine politician's networks over an eventful period spanning the 2015 General Election, in which the Conservative Party, led by David Cameron, won an unexpected parliamentary majority and the Scottish National Party, for the first time, gained a substantial share of MPs, and the

2016 Referendum on the UK's membership of the European Union, which saw a largely unforeseen majority in favour of withdrawal, so-called 'Brexit'. Outside of more formal interactions between legislators (votes, parliamentary questions, motions, debates), these networks of online interactions are important indications of less formal modes of party influence. Second, to study these networks a novel method of 'multiplex community affiliation clustering' (MCAC) is developed. This new method is necessary to both address challenges in the data and capture more accurately the dynamics of the interactions. Building on recent advances in multiplex network analysis, here we develop a novel method for tracking the evolution of community structure amongst a fixed set of nodes, referred to as 'multiplex community affiliation clustering' (MCAC). Although a thorough investigation of the sensitivity of the method to different implementations and algorithm choices would be valuable in establishing it as a robust analytical tool, this manuscript focuses on what we perceive as the least assumptive implementation, demonstrating its utility through application to the network of online interactions between MPs and MEPs across the study period.

The paper presents the first network analysis of online interactions between UK politicians, showing that analysis of online social networks can reveal party cohesion and patterns of affiliation between individual politicians in a case where more conventional methods (such as analysis of voting records) would fail. Importantly, the paper demonstrates that analysis of social media allows affiliations to be studied as a dynamic process, revealing major network restructuring and community formation to be driven by multiple factors, including major political events, ideology and party affiliation, concurrent with popular intuitions of the contemporary political landscape. In doing so, we can address two avenues of future research that Ringe et al. (2017) call for: new forms of data to study legislative networks and extending the empirical focus beyond the US case (p. 17).

Whilst not the focus of this paper, two sets of important theoretical questions can be addressed with an analysis of party communities over time. First, speaking to the literature on intra- and inter-party structures, we can examine the development of party factions and inter-party communities, and evaluate whether their formation is driven by current events or if they develop over a longer period of time in response to external pressure (e.g. shifts in public opinion). This question has motivated a body of research on intra-party coalitions (Sartori, 2005) but there have been recent calls for measures of intra-party factionalism to incorporate more dynamic elements (Boucek, 2009). Second, a study of party communities that incorporates both MPs and MEPs can address questions about the Europeanisation of national political parties (Hix and Goetz, 2000; Ladrech, 2002; Mair, 2008) and the extent to which MPs are responding to MEP communities or the other way around. Our analysis presents a method for detecting the community dynamics which makes it possible to answer the above questions.

## 1. Summary of methods

The primary object of study is a directed multiplex network of social media (Twitter) interactions between UK politicians (Members of Parliament, MPs, and British Members of the European Parliament, MEPs). Although these two groups occupy distinct political institutions and cannot be directly compared (for example, by parliamentary voting patterns), we include both MPs and MEPs in the analysis for the following three reasons;

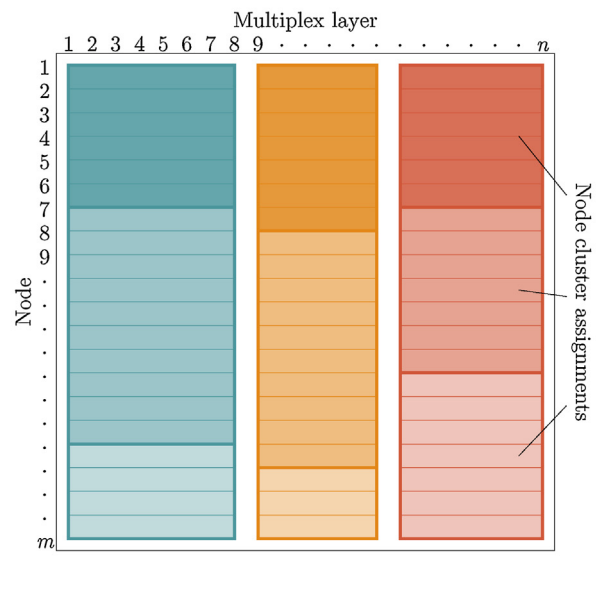
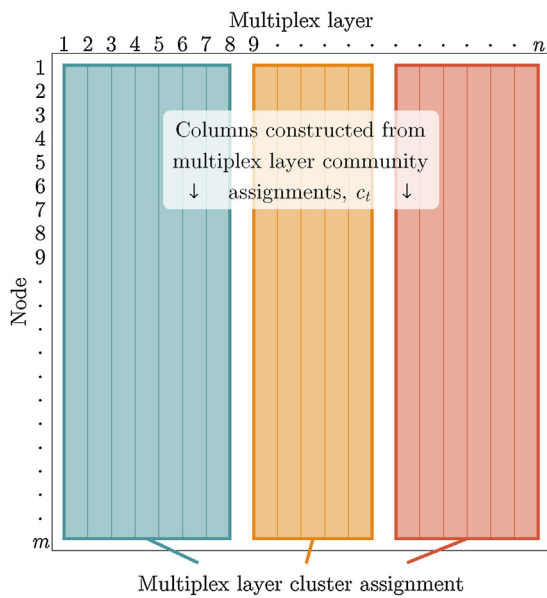
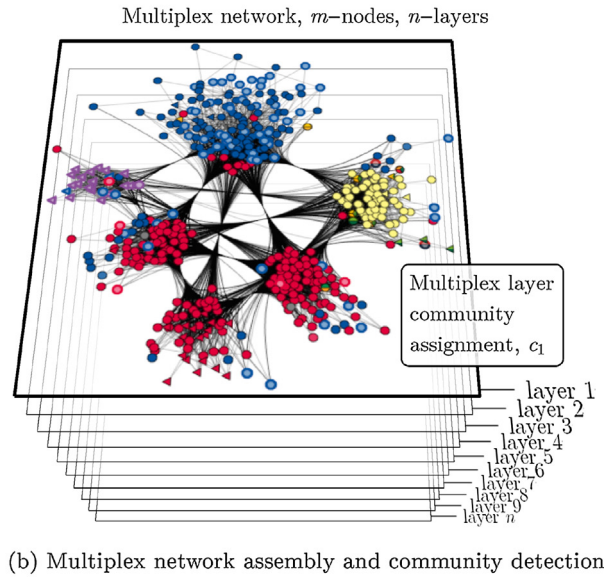
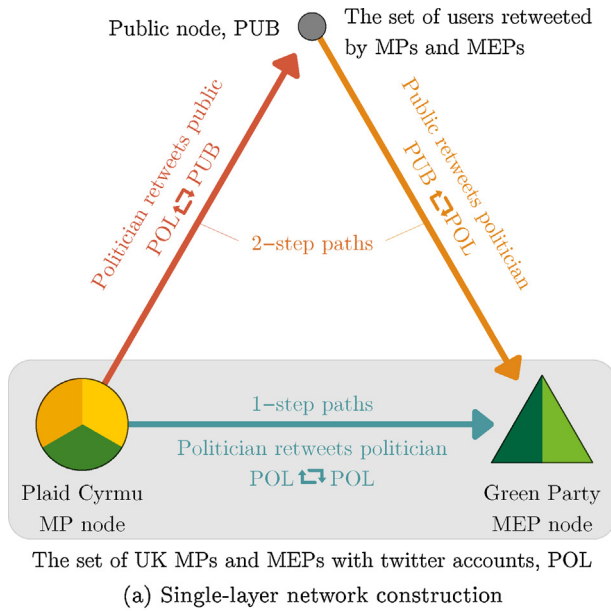
1. Preliminary analysis showed that both MPs and MEPs share strong connections in the realm of Twitter.

- Party factions important to the political developments during the selected time period are numerically underrepresented in the set of MPs. Specifically, the UK Independence Party (UKIP) had a prominent role in UK politics during the study period, despite winning just a single MP seat in 2015 General Election.
- Previous research found that discussions of EU topics take place mainly within countries, using local hashtags (Nulty et al., 2016), so the inclusion of MEPs gives insight into how multi-level govern-

nance influences national political parties and domestic political discourse (see Hix and Goetz, 2000; Ladrech, 2002 and others on the Europeanisation of national political parties).

Their inclusion is found to have no adverse effect on the clarity of results, serving only to strengthen the political narrative alongside the data.

The MP and MEP network is formed of links where a politician retweets an original tweet published by another UK politician.



**Fig. 1.** Summary of methods. The core methodology has four steps: (a) Single-layer network construction. The network is constructed from all one-step and two-step directed paths between MPs and MEPs in the set of politicians, POL. Two-step paths connect via intermediate nodes in the set of public users, PUB. In this example, the shape of each node denotes the institutional affiliation of each user (circles for an MP, triangles for an MEP and point for PUB), the colour denotes party affiliation, and the style denotes a politician’s stated position on the 2016 EU Referendum (closed for Remain, circle for Leave). A complete key is provided by Fig. 3. (b) Multiplex network assembly and community detection. Networks representing 7-day windows are assembled as the layers of a multiplex network. Community detection is applied to each layer to partition nodes into discrete communities. Singletons (here politicians which are not connected to any other politicians by two or less network steps) are placed in their own isolated communities. (c) Multiplex community affiliation clustering (MCAC). Node community affiliation vectors for each layer populate the columns of the multiplex community affiliation (MCA) matrix. The MCAC method groups layers into a number of ‘network states’ based on maximum similarity mapping of community affiliations between layers. A network state corresponds to a group of layers with similar community structure. (d) For each network state identified in the previous step, a further clustering is applied to the node community assignments (MCA matrix rows) and used to characterise each network state based on the similarity of node community affiliation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Retweets are chosen specifically for their clarity as agreement with the content of the original tweet, interpreted as a simple form of affiliation with the tweet's original author. Each layer of the multiplex is constructed from politician retweets aggregated into one-week intervals over an 85-week time period between 2014-12-22 and 2016-08-08, spanning the build-up and reception of both the 2015 General Election and 2016 EU Referendum. Single retweets result in weight-1 edges, while multiple retweets of one politician by another within an interval are combined into a single edge with summed weight. Importantly, we do not attempt any analysis of tweet *content*, such as sentiment analysis, except as a descriptive metric. The networks we analyse are constructed solely from the edges between the retweet source and target, grouped into temporal layers based on tweet time stamp.

While there was a large volume of Twitter activity by MPs and MEPs around the General Election and Referendum, there was a relative paucity of activity at some other times. To ensure sufficient data for creation of robust networks, we augment the direct first-order interactions (where one politician retweets another) with indirect second-order interactions (where a two-step retweet path connects two politicians via a non-politician intermediary). This extension captures the wider UK political landscape, including partisan and non-partisan media outlets, journalists, party members, and the politically engaged public. We define POL as the set of MP/MEP Twitter users and PUB as the set of public users who publish tweets which are retweeted by POL in the 1-week time span of a single multiplex layer, as in Fig. 1b. First-order network connections are defined  $POL \rightarrow POL$ , while second-order connections include both  $POL \rightarrow PUB$  and  $PUB \rightarrow POL$ . Fig. 2 shows the number of publicly available retweets collected from this scheme (Fig. 2a), and the numbers of edges and nodes (Fig. 2b and c respectively) in one-week layers of the resulting multiplex network.

This extension provides challenges beyond those inherent in generating multiplex networks with overlapping but variable sets of nodes. From observations of contemporary politics in the UK, our intuition is that political affiliations (and therefore the structure of the retweet network) over the time period chosen are multi-modal. That is, rather than being modeled by a single time-invariant probability distribution, we hypothesise that these networks are drawn from a number of different underlying structures. Under this assumption, we can leverage the tendencies of political groups to affiliate and form network communities to measure the similarity between multiplex layers in spite of their dissimilarity on the individual level.

To evaluate the extent of this similarity (or lack thereof), we introduce a novel method which we term 'multiplex community affiliation clustering' (MCAC). We begin by applying a community detection algorithm to each layer of the multiplex, illustrated by Fig. 1b; this produces a community label assignment for each politician in the network layer, whereby groups of densely interconnected nodes share the same assignment. The output of this step is used to populate the multiplex community affiliation matrix, where each column is a community affiliation vector corresponding to a multiplex layer, and each row is a politician's community assignment.

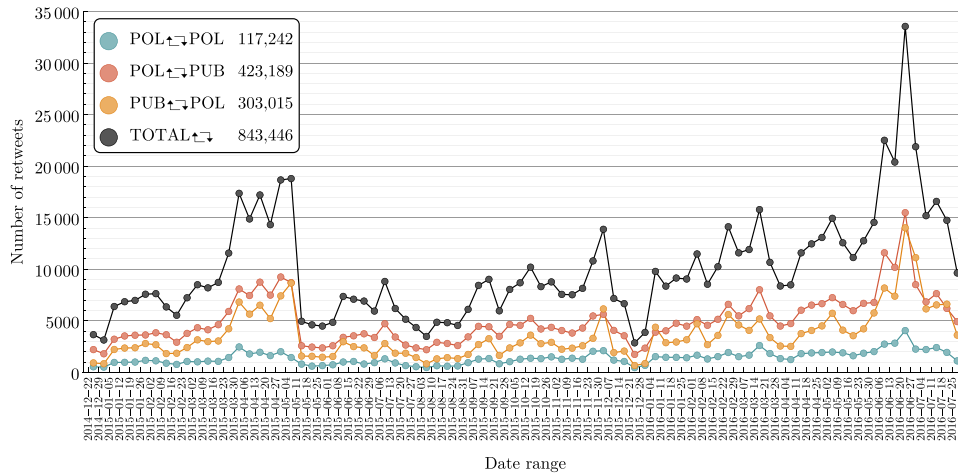
This matrix is subjected to two stages of clustering analysis. In the first instance, columns are clustered by their simple matching coefficient similarity, such that multiplex layers with similar community structure are clustered together. Having partitioned the matrix into clusters of similar columns, illustrated by Fig. 1c, we perform a second clustering step, this time on the matrix rows which encode the community affiliation of POL in each multiplex layer, illustrated by Fig. 1d. This produces clusters of MP/MEPs which tend to fall into the same graph communities and provides a method to characterise each cluster of multiplex layers.

In addition to the MCAC algorithm, we provide a simple topic model as a method to characterise the retweet content from which the network is constructed. We do not aim to produce a novel or particularly sophisticated topic model, but simply a coarse reference for the topics prevalent within the 1-week period corresponding to each multiplex layer. Briefly, the tweet content in each multiplex layer is merged into a single document, tokenised, and certain terms are removed; ubiquitous stop-words and tokens which occur in fewer than five documents. These data populate a matrix which is subsequently clustered so as to reveal textually similar documents. In our case, it provides insight into which multiplex layers share similar retweet content, independent of the resulting network structure.

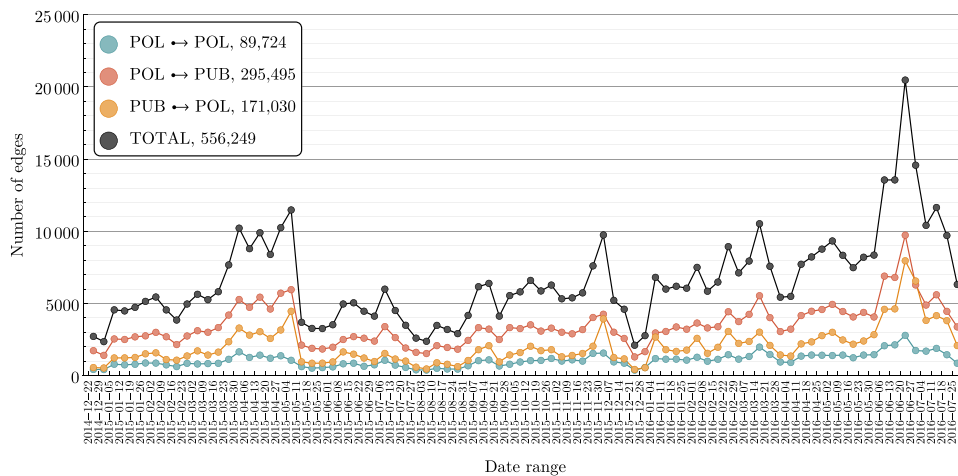
## 2. Results

One layer of the multiplex network is shown in Fig. 3, corresponding to the 2016-05-30 to 2016-06-06 window. The majority of MPs and MEPs are linked to a single giant network component encompassing 518 of 591 nodes in the set of UK politicians. Comparison of the network formed from direct retweet interactions between politicians (Fig. 3, Scheme 1, left) and the extended network including two-step paths via a non-politician intermediary (Fig. 3, Scheme 2, right) shows qualitative similarity. A measure of quantitative similarity between the network schemes is possible, but largely uninformative; the benefit of the extended network is that it captures more MPs and MEPs in the giant component, while preserving the patterns of interaction between politicians. This provides a dramatic improvement in consistency of network communities between different time periods, substantially improving the quality of subsequent clustering carried out in this section. The example network is drawn from the campaign period prior to the EU Referendum, where ideological division within the Conservative Party was strongly apparent and Conservative politicians were found on both sides in the debate. This is evident in Fig. 3, where Conservative Remain and Leave factions form distinct network communities; at this time, the Conservative Remain faction is actually more strongly connected to the Labour Party than to their Conservative Leave colleagues. There is significant intraparty linkage, with cross-party links accounting for 8% of the  $POL \rightarrow POL$  links in Fig. 3. It is important to note that party and institutional affiliations of politicians, as well as their stated EU referendum position, are used solely to style the network nodes in visualisations. They do not factor into the community detection or matrix clustering stages; The network layout and resulting structures in Fig. 3, along with all subsequent analysis, result entirely from politicians' retweeting behaviour, and not from content, sentiment or any political positions or opinions held.

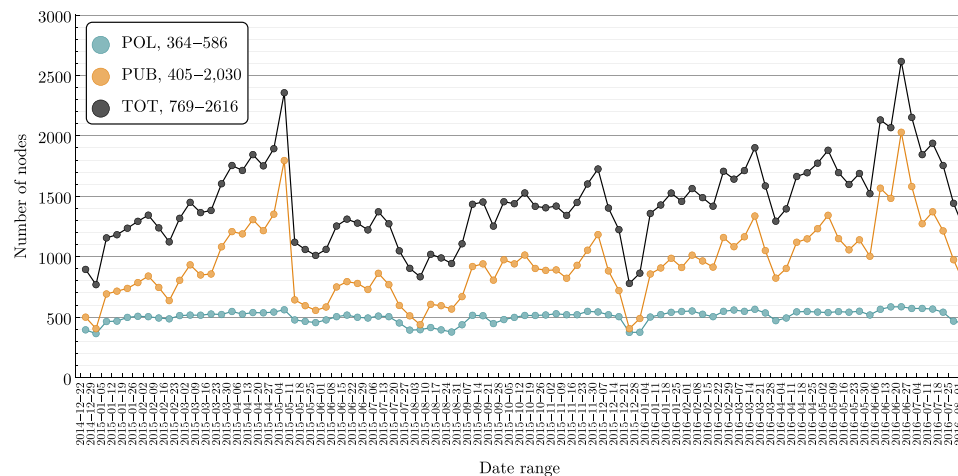
We find that the network layers typically fall into one of a few distinct states. Network analysis using the MCAC method shows multiple local optima in the silhouette index (data not shown) as the number of possible multiplex layer clusters increases. This is indicative of multiple hierarchical levels of clustering in the community affiliation patterns across the multiplex network. Focusing initially on high-level classification of network states, we take the first optimum of the silhouette index as our clustering scheme (*i.e.* we adopt the highest-level clustering of network states) which occurs at 4 network states. Fig. 4a summarises these states by clustering politicians on their community assignment vector within the network state. The frequency of politicians in different institutions and Brexit-vote categories is illustrated by the symbol sizes. Content analysis based on textual term frequencies shows a local maximum in the silhouette score at 6 topics, which are summarised by Fig. 4c, showing the top keywords as weighted by their TF-IDF scores summed across each content cluster. The time series of both retweet network and content states is shown in Fig. 4b.



(a) Retweets collected over the observation period



(b) Network edges resulting from collected tweets

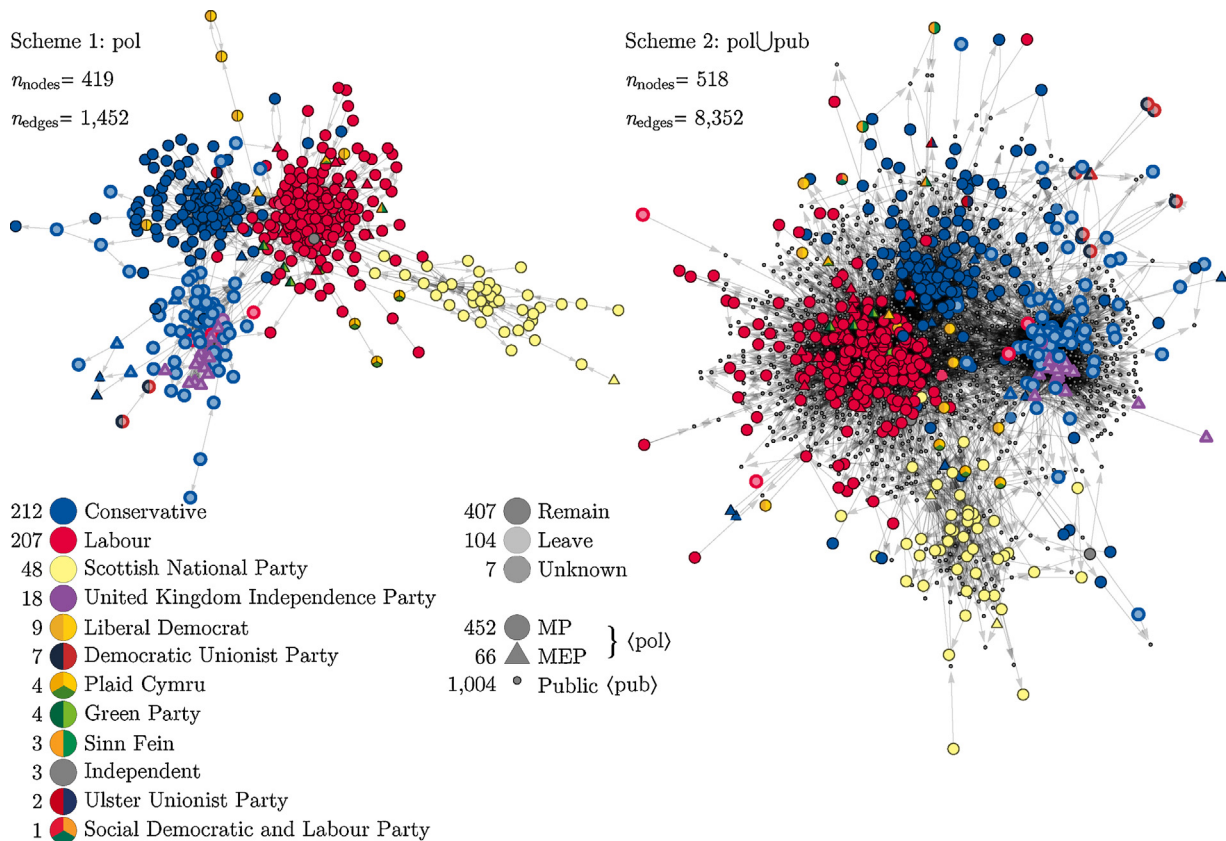


(c) Network nodes resulting from collected tweets

**Fig. 2.** Summary of Twitter data collecting and resulting multiplex metrics: (a) The volume of Twitter data collected over the time period spanning the 2015 UK General Election and 2016 EU Referendum. The period is divided into one-week intervals where tweets are further divided into categories by the identity of the user retweeting, and the author of the original Tweet. (b and c) Edges and nodes, respectively, resulting from the data collected, divided into categories by the identity of link ends and nodes. Instances where multiple retweets within a one-week window would produce identical edges instead iterate on edge weight such that not all retweets produce new edges.

One of our most striking findings is that without including any time component in our model, network states and content states are each strikingly persistent and recurrent over time. Adjacent lay-

ers of the multiplex network are highly likely to occupy the same state for both community affiliation and tweet content. This stability is perhaps surprising given the high-velocity and ephemeral



**Fig. 3.** Social networks of politicians show coherent communities. Nodes are styled according to their party affiliation, political institution (circles for MPs, triangles for MEPs and points for PUB) and stated position on the 2016 EU Referendum (closed circle for Remain, open circle for Leave). An example network based on retweet interactions (see Methods) over the 7-day period from 2016-05-30 to 2016-06-06. Nodes represent MPs/MEPs and node properties indicate party affiliation (colour), stated position on UK membership of the EU (closed – remain, open – leave), and their status as an MP (circle) or MEP (triangle). Scheme 1 (left) shows the network formed only of direct retweet interactions between politicians (members of the POL set). Scheme 2 (right) shows the network including indirect interactions via an intermediary from the PUB set. Supplementary nodes in the PUB set are drawn uniformly as small gray nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

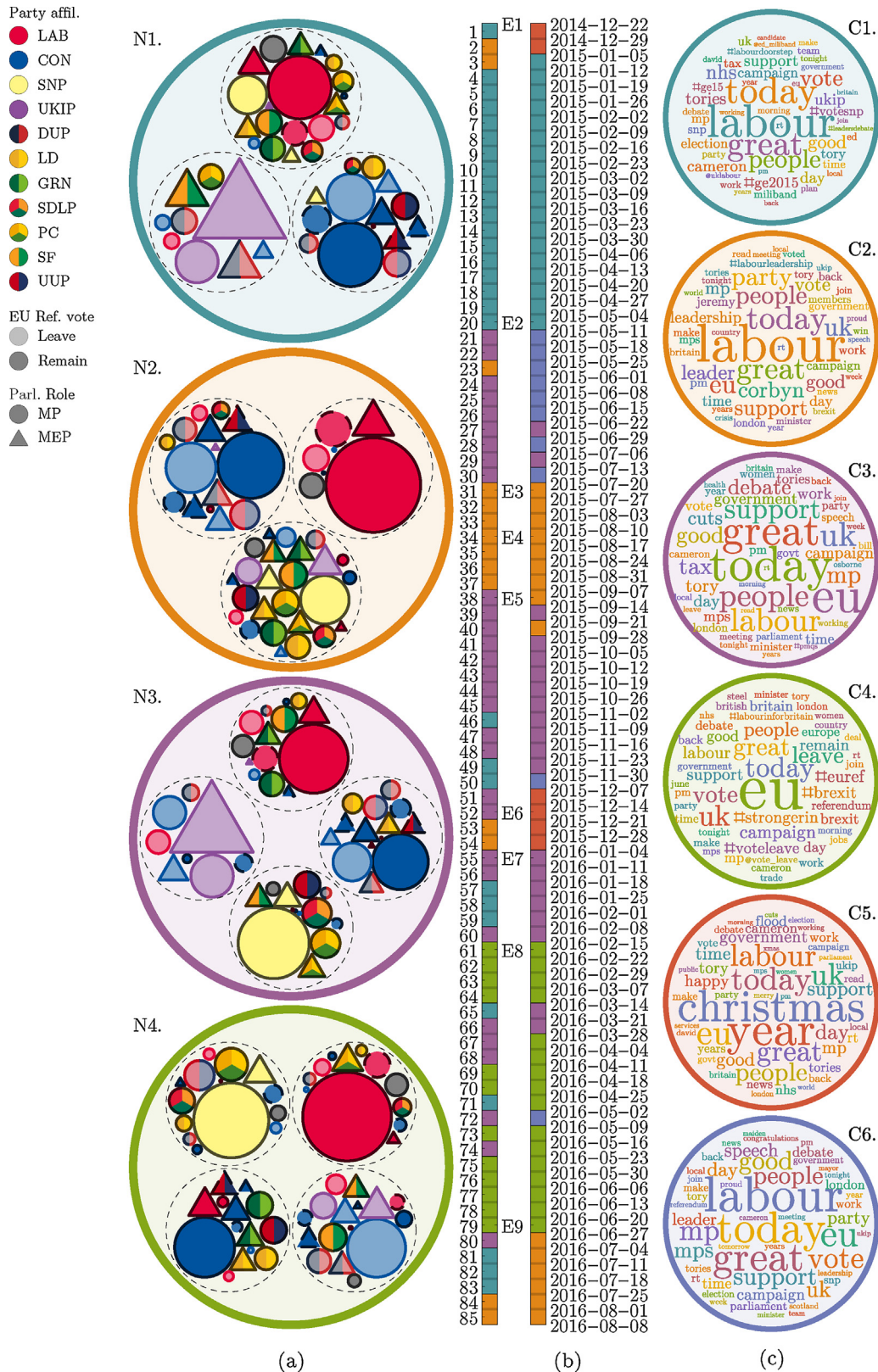
nature of social media communication. Temporal persistence of network and content states can be measured simply by computing the join count statistic between like-states; this is the number of points in the time series where the same state is encountered in succession. In an uncorrelated and unbiased random system with many intervals and  $n$  states, the expected join count is close to  $n^{-1}$ . The graph state and tweet content time series of Fig. 4b yield scores of 0.67 and 0.76 respectively, corresponding to a statistical significance of  $p < 10^{-9}$  when compared to the null expectation, implying a strong temporal component to system dynamics, as expected from a multistable dynamic network. Both network states and content states typically appear in several episodes over time. Furthermore the content and network states are highly dependent (Pearson chi-squared test of independence yields  $p < 10^{-9}$ ) indicating a strong correspondence between network state and the topics of discussion. Temporal correlation in either time series is itself remarkable; both the MCAC method and topic modeling are blind to the ordering of multiplex layers, yet their outputs display strong auto-correlation.

Characterisation of network states based on politician community memberships suggests that network states are reflective of ongoing political events and debates. Each network state is characterised by clustering of politicians within the multiplex layers belonging to that state, visualised by Fig. 4a, where symbols are sized in proportion to their fractional share of the cluster and the network whole, and coloured according to the party key given in Fig. 3. Labeling the four network states N1 to N4 in order of their appearance, we offer a simple interpretation of the patterns

observed in Table 1. Characterising content states by word frequencies suggests loose themes for online discussions making up each state. Tweet content within each content state is characterised by visualisation of the dominant words, weighted by their inverse-document-frequency in the content state (size is maximised by high frequency within content state and absence from other states), and shown in Fig. 4c. Some additional interpretation based on these content clusters along with thematic labels is given by Table 2. For reference, Fig. 4 also contains annotations for several periods of political significance, bounded by events E1 to E10 described by Table 3. While caution must be exercised (see below), we note that there appears to be some correlation between these events and transitions between states.

The social network produced by politicians' retweet activity leads to a small number of distinct states corresponding strongly to contemporary issues as summarised by the content of the tweets; who they retweet is related to what they retweet. From this we can see a number of intuitively sensible results:

- The 2015 General Election period is spanned almost entirely by network state N1 and content state C1. Throughout this period, we see predominantly left- and right-leaning parties cluster separately, and the extreme right parties, United Kingdom Independence Party and Democratic Unionist Party also form their own cluster. This result indicates that communication throughout the election period revolves not only around party affiliation but common issues or ideology which serve as talking-points for left- and right-leaning parties. The state recurs over the 2016-



**Fig. 4.** Analysis of the Twitter interaction networks and the textual tweet content finds both to have strong clustering indicative of distinct network and content states. Fig. 4b shows the time series of network states (left, N1 . . . 4) and tweet content states (right, C1 . . . 6) along with the corresponding state summaries in Figs. 4a and c respectively. In both instances, this conclusion is supported by network and content states having strong temporal clustering; a multiplex layer in one state is most likely to be followed by the same state than any other. Furthermore, there is a strong correlation between network and content states, indicating that the content of retweets is strongly tied to the interaction structure of MPs and MEPs. Note that the colours chosen to represent these states have been chosen as to maximise coincidence of colours between the two time lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 1**

Descriptions of political network community clusters. Cluster ID corresponds to the nomenclature of Fig. 4a.

Cluster ID	Description
N1.	Labour MPs cluster with the Scottish National Party and a number of smaller left-leaning parties including Green Party and Sinn Féin. Conservative MPs cluster with other predominantly right-wing parties, such as Ulster Unionist Party. Nodes affiliated with the United Kingdom Independence Party (UKIP) form their own (small) cluster along with another extreme-right group, the Democratic Unionist Party.
N2.	Labour MPs cluster with some aligned smaller parties. Conservative MPs cluster with some aligned smaller parties, including UKIP. Scottish National Party and the majority of other smaller parties.
N3.	Labour MPs cluster with some smaller left-wing parties. Scottish National Party nodes with Plaid Cymru. UKIP form their own cluster, joined only by a small number of Conservative nodes who would go on to vote “Leave” in the 2016 EU Referendum. Conservatives with other right-wing parties.
N4.	Scottish National Party MPs with Plaid Cymru. Labour MPs. Conservative “Remain”-voting MPs and MEPs, Labour “Remain”-voting MEPs and smaller parties. UKIP with Leave-voting Conservative MPs and MEPs, along with the majority of Leave-voting MPs and MEPs from smaller parties.

**Table 2**

Descriptions of retweet textual content clusters. Cluster ID corresponds to the nomenclature of Fig. 4c.

Cluster ID	Theme	Description
C1.	General Election	Many terms related to the 2015 General Election, including the dominant hashtag used to refer to the election ( <i>#ge2015</i> ), various voting/campaigning terms (e.g. <i>#votesnp</i> ), and the name of the Conservative leader during the election period ( <i>cameron</i> ).
C2.	Labour party leadership	Terms include <i>leader</i> , <i>leadership</i> , along with the popular candidate and present Labour leader, <i>Jeremy Corbyn</i> .
C3.	Economy	Financial talking points <i>tax</i> , <i>cuts</i> , <i>bill</i> along with the controversial Conservative treasurer <i>George Osborne</i> .
C4.	EU Referendum	This cluster surrounds several terms pertinent to the 2016 EU referendum, notably <i>eu</i> , <i>leave</i> and <i>remain</i> , along with hashtags promoted by the two opposing campaigns: <i>#voteleave</i> and <i>#strongerin</i> .
C5.	Festive Greetings	A relative paucity of partisan terms along with broad positivity and festive terms, <i>happy christmas</i> and <i>new year</i> .
C6.	No clear theme	No clear identity beyond a lack of the distinctive terms which differentiate other <i>labour</i> dominated clusters from it.

**Table 3**

Descriptions of political events throughout the study period. Event ID corresponds to the nomenclature of Fig. 4b.

Event ID	Date	Description
E1.	19 December 2014	Campaigning for the 2015 General Election officially begins.
E2.	7 May 2015	The day of the 2015 General Election and Ed Milliband's resignation (8 May 2015).
E3.	21 July 2015	UK Parliament enters summer recess period.
E4.	14 August 2015	Voting process for the Labour leadership election begins.
E5.	7 September 2015	Summer recess period ends and Labour leadership voting concludes (10 Sep 2015).
E6.	17 December 2015	UK Parliament enters Christmas recess period.
E7.	5 January 2016	Christmas recess period ends.
E8.	20 February 2016	The date of the 2016 EU Referendum is announced.
E9.	5 May 2016	Scottish Parliament, National Assembly for Wales and some local elections.
E10.	23 June 2016	EU Referendum concludes.

01-18 to 2016-02-08 period and most of July; it appears likely to surround issues which unites parties along ideological lines, such as a flurry of discussion and the vote to renew the UK's Trident nuclear programme in January 2016 and on 2016-07-18 respectively.

- The network states over Christmas and following parliamentary summer recess periods, 2015-07-21 to 2015-09-07 and 2016-07-21 to 2016-09-05, cluster together into the same network state, N2. These correspond very closely to content states C2 and C5. We choose to refer to this as a ‘rest-state’. However, this period also encompasses the Labour party's leadership elections in 2015 and 2016. Both contests and the debates around them have been criticised for not being able to set a clear and distinctive direction for the Labour party, which seems to be consistent with our classification both for the network and the content states. For this reason it is unsurprising that N2 is characterised by an almost completely pure cluster of Labour MPs; a result of a large volume of intraparty Twitter activity.
- Network state N3 follows on from the 2015 General Election (N1) and summer parliamentary recess (N2), corresponding with content states C6 and C3 respectively. The most obvious difference between the network states either side of the general election is the emergence of the Scottish National Party as their own cluster following their dominance of Scottish MP seats.
- The politically divisive referendum on the UK's continued participation in the EU produced a network state which dominated from mid-February, corresponding with David Cameron's announcement of the referendum date on 2016-02-20, to the end of June when the referendum took place, and corresponds to network state N4 and content state C4. As anticipated, this period saw MEPs affiliated with the United Kingdom Independence Party, who largely championed the vote to withdraw from the EU, forming consistent communities with the majority of MPs who expressed support for Britain's exit from the EU, particularly Conservative MPs who, over this period, were socially estranged from their peers. Also, politicians campaigning to Remain were well-



linked, including an unusual linkage between Labour MEPs and Conservative politicians.

Here we have attempted to enumerate only the coarsest features of Fig. 4. It is tempting to attribute transitions between network states to major political events; for example 2015-05-11 corresponds to multiplex layers either side of the end of the 2015 General Election (Event E2 – 2015-05-04), and a transition between network states N1 and N3. Following this, state N2 spans 2015-07-20 to 2015-09-07, exactly the layers across the summer recess of the House of Commons (events E3 and E5). However, there is clear danger in this type of attribution; disambiguating the signal produced by the clustering of qualitatively different states of interaction on the Twitter platform from the noise inherent in systems of relatively few actors is not trivial. For this reason, our interpretation is correctly focused on regions of relative invariance between major political events of broad interest.

In this analysis, we have by no means fully exhausted the potential of the multiplex. A fine-grained network analysis is possible, revealing sub-divisions within political parties. Fig. 5 shows the dendrogram of politicians in the network state N4, visualising the simple matching coefficient between node community affiliation vectors. Note that the hierarchical clustering displayed in a dendrogram is distinct from the k-means clustering methods employed previously. There is clear scope for an in-depth investigation and discussion around nestedness and fine detail in the social network of the UK parliament. Even the coarse analysis given here identifies a division within the Conservative Party nodes on Leave/Remain positions in the EU Referendum, and a nested structure within the Labour Party which is not trivially explained by Leave/Remain alignment. We speculate that the observed divisions reflect the ongoing Labour leadership battle and well-publicised internal conflicts which occurred at that time.

For comparison, we examined voting behaviour in the UK Parliament (data not shown) but this revealed minimal heterogeneity in party-affiliated MP voting records. While each vote has a small proportion of MPs who go against their party majority, there is no obvious clustering in the sense that there is no apparent division within the major parties. This is somewhat expected in legislatures with strong party discipline and voting unity, such as the UK Parliament where it is common practice for political parties to have an official designated as the party *whip*, whose purpose is to establish a party consensus for voting. The influence of party whips appears sufficient to promote uniform voting patterns and conceal any underlying ideological heterogeneity within parties. Other measures based on parliamentary activity in general, and roll call votes in particular, are not able to measure the underlying diversity which exists within political parties (Spirling and McLean, 2007). This makes voting records a poor measure for revealing social network structures for those legislatures, and comparison with social network modes is ultimately uninformative. We have also attempted comparison of MP clusters to the outcome of their constituency votes in the 2016 EU Referendum, finding no significant correlation.

### 3. Discussion

We have developed multiplex community affiliation clustering (MCAC) as a method for identifying groups of network nodes that consistently fall into the same network communities over time. We have applied this method to study the evolving patterns of affiliation between UK politicians as revealed by their interactions on social media. This analysis produces outputs that are coherent and reasonable based on our interpretation of the contemporary political narrative in the UK. Individual politicians are clustered

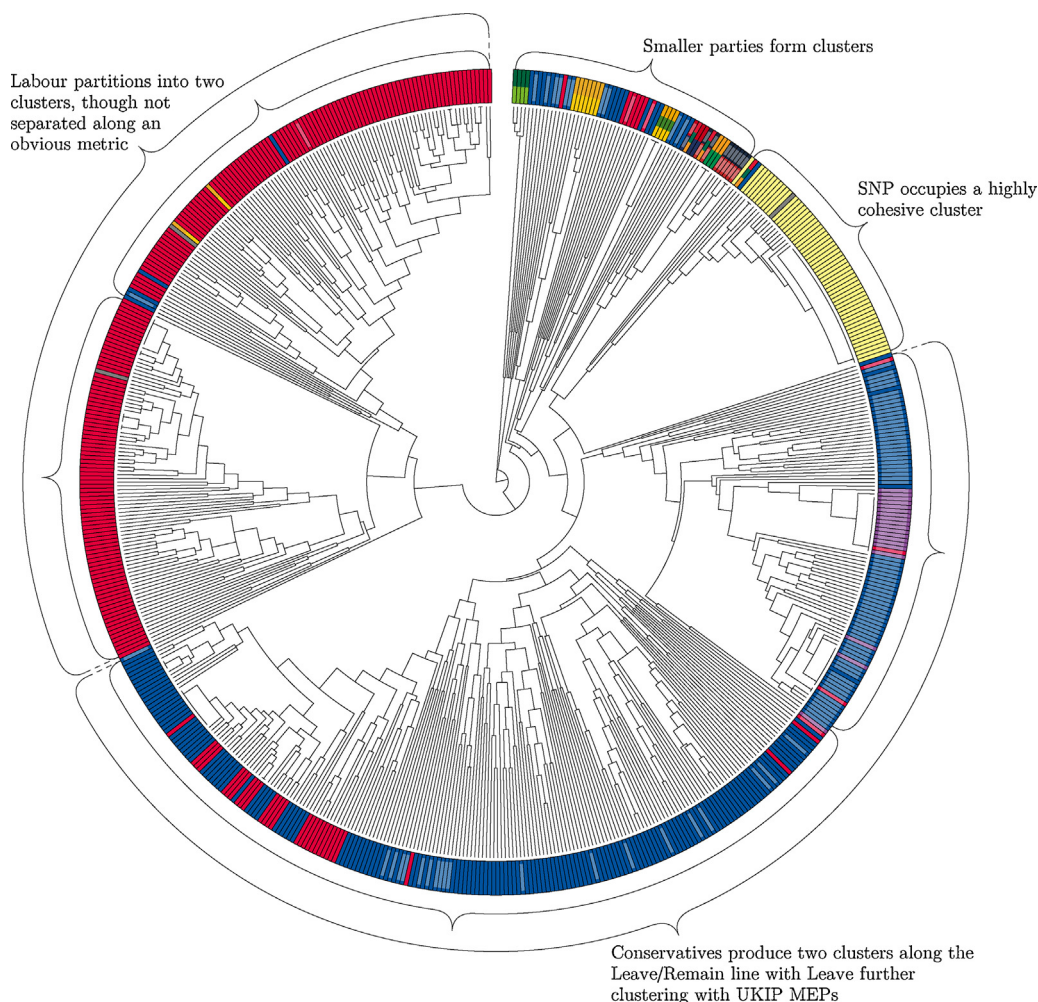
with members of the same party, while parties with similar ideologies are more likely to cluster together. The use of MCAC based on social media interactions reveals the existence of fine social structures within parties, in contrast to analysis based on voting records, whereby whipped voting conceals patterns below party level.

Outputs of this analysis show that political affiliations vary over time. The temporal dimension emerges from the analysis, despite the fact that we did not explicitly include time in our model. Over a 17-month period we find just four network “states” which are persistent and recurrent. We also find strong correspondence between social network structure and the topics of discussion; simple analysis of term frequencies in social media content reveals six content states which are well correlated with the network states. This suggests that the patterns of who interacts with whom are dependent both on ideology and also the current political and media landscape. Indeed, there is a clear analogue between the temporal evolution of network and content states and the popular narrative of political events during the study period. Politicians’ activity on social media changes around the General Election in 2015, the Brexit vote in 2016, and the festive periods around Christmas in 2015 and 2016. Substantial changes in network organisation and discussion topics occur at the commencement of campaign periods and again after the relevant votes have occurred.

In considering the potential to address important theoretical questions, our analyses and approach contribute to two areas of party factions and legislator behaviour. First, our results are important for understanding the dynamic evolution of intra and inter-party structures (see Boucek, 2009). Two features that are captured by our analysis are divisions within the two largest parties, Conservative and Labour. The Conservative party was subjected to significant internal stress during and following the 2016 Referendum campaign, with a significant proportion of the party campaigning on each side of the issue. The Labour party has undergone two highly publicised leadership elections following the General Election in 2015, causing substantial friction between center-left and hard-left factions within the party. In both cases, these divisions would not be visible through an analysis based purely on voting records. While these divisions are unlikely to surprise observers of UK politics, here we are able to identify them based purely on social media interactions with no prior knowledge of politicians or their opinions. Our multiplex clustering approach reveals intra-party divisions and complex structural changes in response to these unfolding political events. Second, the analysis of communities shows strong connections between MEPs and MPs with MEPs being prominent in online discussions.

Correspondence between the results of this analysis and the unfolding political narrative in the UK suggests that analysis of social media is a valid methodology for understanding aspects of UK politics. This is likely to accelerate as social media becomes increasingly embedded as a channel for communication in national political discourse. The richness of data provided by social media already permits multiple levels of analysis. Here we have focused on macroscopic analyses of network structure and discussion topics across all UK MPs and MEPs, over a relative long period containing two outstanding political events (2015 General Election and 2016 EU Referendum). However, we note that we could equally have focused more narrowly on particular ideological debates or smaller study populations. The hierarchical clustering of affiliations revealed in Fig. 5 shows multi-scale structure, revealing the broad division of individual politicians into parties, the within-party divisions created by the issues of the day (the Brexit vote in the data shown in Fig. 5), and finer-scale clusterings that are less obviously explained.

It is important to note that the MCAC network clustering alone recovered several key features of UK politics during the study period. Firstly, the temporal similarity of weekly network snap-



**Fig. 5.** Similarity of politicians according to their network community co-occurrences in network state N4. The figure shows the dendrogram of MP and MEP similarity across the snapshots that comprise this network state, where leaves are labeled with a colour scheme identical to that shown in Fig. 3. This reveals significant fine structure which is not visible in a coarse analysis comparing many network states; most significantly we can see the Labour dominated cluster to be split into two similarly sized components which one might expect from contemporary intraparty politics. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shots can largely be inferred by their network cluster similarity, without any temporal information being provided to the clustering method. This indicates a degree of temporal coherence and persistence in online interaction networks. Secondly, the clustering of individual politicians was able to accurately identify the ideological position of MPs and MEPs regarding the EU referendum. The Leave and Remain labels attached to individuals in Figs. 3 and 5 are added for clarity and were not used as input into the clustering process; instead, MCAC correctly separates politicians and predicts their position in this important debate. Similarly, the grouping of political parties in Fig. 4 is based purely on topological clustering from social media interactions, yet recovers ideological similarities and associations between parties. This ability of social network analysis to predict the ideological positions of individual politicians suggests a route towards predictive political science, whereby the likely voting intentions of politicians can be inferred from network position.

In summary, by analysing social media interactions of UK politicians we have shown that our analysis can reveal fine-grained social structures in contexts where other approaches (e.g. use of voting records) do not succeed. By analysing affiliation network structures over time, we have determined that network states fall into four broadly distinct and recurrent modes, each of which can be explained by reference to party ideology and ongoing political

events. Concurrent analysis of social media content helps to characterise the network states and ties them to political narratives. The MCAC method we developed has potential for future application to study political social structures across a range of scales. The wider methodology of social media analysis also offers the intriguing possibility of studying politicians within the context of a broader 'political society' including journalists, bloggers, and the engaged public.

## 4. Methods

### 4.1. Data collection

Our analysis is rooted in the social media interactions of two groups of UK politicians: the Members of Parliament (MPs) who won or retained their seat in the 2015 General Election (or a subsequent by-election held to fill an arising vacancy), and the Members of the European Parliament (MEPs) elected in the 2014 European Parliament election. Together these form a set of politicians we abbreviate POL. The observation period is 22nd December 2014 (the Monday following the start of campaigning for the 2015 UK General Election) and 8th August 2016 (approximately 6 weeks after the Referendum on membership of the EU). Along with the 650

members of parliament elected in the general election, there were four additional instances in which MPs were elected following resignation or death giving a total of 654 MPs and 73 MEPs in the study period. Of these, 591 MPs and 70 MEPs were found to have at least one active affiliated Twitter account at some point, giving a total of  $m = 661$  individual accounts which form the primary social media user population we study (POL). It is worth noting that this list is subject to change for a myriad of reasons; Twitter accounts can be deleted, suspended or protected, all of which prevent public access and tweets can be deleted en-mass at the will of the publisher. The data available at the time of collection therefore may not contain all Twitter activity from this time period, and subsequent studies may be reduced further.

A secondary user population was identified as the set of all non-politician accounts (*i.e.* non-members of POL) that were connected to members of POL by retweet activity (*i.e.* any other Twitter user account that was retweeted by an account in POL, during the study period). Inspection showed that this secondary user population consisted largely of journalists, commentators and other politically engaged members of the public; here we refer to this population as PUB. All relevant retweets from user accounts in POL and PUB were extracted from the Twitter API. The distribution of data collected from these user sets is displayed in Fig. 2a.

#### 4.2. Constructing politician social networks

Social networks were constructed from aggregated sets of retweets compiled for  $n = 85$  one-week intervals between 2014–12–22, and 2016–08–08. Each network was constructed by first creating a directed edge from each user in POL to any other user in POL that they retweeted in the corresponding one-week interval. That is, an edge from node  $x$  to node  $y$  is produced when a Twitter account affiliated with  $x$  retweets original content from  $y$  and is written  $x \rightarrow y$ . MPs and MEPs commonly have multiple affiliated Twitter accounts, such as an election campaign account. In such cases these are all attributed to a single user, rather than treated individually. Edges are weighted by frequency, so multiple retweets linking the same node pair increase the edge weight.

To increase network density and thereby improve the robustness and quality of community detection (see below), the initial set of edges based on retweets within POL is supplemented by additional edges representing all two-step paths that link members of POL via an intermediary member of PUB. That is, where a pair of politicians in POL are joined by a two-step retweet path in which the intermediate node is a non-politician member of PUB, both edges forming the two-step path are also included in the network (as illustrated by Fig. 1a). Put succinctly, the network for each week is constructed from *all* one-step and two-step directed paths which join nodes in POL.

The network construction process is summarised by the following steps for each one-week interval:

1. Identify *all* retweets made by users in POL. Those which are retweets of other users in POL produce 1-step directed paths between MPs and MEPs.
2. Those which are retweets of other users not in POL define PUB; the set of users retweeted by users in POL which are not contained in POL.
3. Identify *all* retweets of users in POL made by users in PUB. These produce the second step in 2-step paths between MPs and MEPs.
4. Generate directed, frequency-weighted edges from the aggregated set of retweets identified in steps 1 and 3 above to form the network.

This process was repeated to produce 85 distinct networks representing 1-week intervals in the study period, in which MPs

and MEPs are linked directly and indirectly via non-parliamentary nodes. It is convenient at this point to remove any degree-1 nodes in PUB; these users have been retweeted by PUB but do not mediate any 2-step paths between them. The network statistics which result from the collected retweet data are displayed in terms total edges and nodes in Fig. 2b and c respectively.

#### 4.3. Multiplex community affiliation clustering (MCAC)

The multiplex community affiliation clustering (MCAC) method is used to track the evolution of community structure over time. The 85 single-week networks are formed into a multiplex network with 85 layers representing each week. Then the MPs and MEPs in each network layer are partitioned into distinct communities. Of the many available community detection methods (Fortunato, 2010), simple modularity optimisation (Newman and Girvan, 2004) was found to be appropriate to the network scale and to produce intuitive network communities in the scope of this investigation. Several other algorithms were tested in preliminary work, with no significant difference in outcomes. For each layer, a community affiliation vector is produced which records the community label for each node in POL. The supplementary edges and nodes derived from PUB are included in the network for the purpose of community detection, since their inclusion gives higher quality partitioning. However, since we are not interested in the members of PUB beyond their intermediary role in relationships between members of POL, we do not need to record their community affiliation. This process is illustrated by Fig. 1b. The product of this analysis is a set of community affiliation vectors ( $c_t$ ) for each multiplex layer ( $t$ ). These vectors are used to populate the columns of a 'multiplex community affiliation' (MCA) matrix, shown in Fig. 1c. The MCA is then used as the input for cluster analysis.

Clustering of similar layers in the multiplex network begins with establishing a suitable measure of pairwise similarity between columns  $c_t$  in the MCA matrix. The output of the community detection algorithm is a set of arbitrary labels indicating the community to which each node is assigned in each layer. Determining the similarity in community structure between two network layers is not trivial as community labels are only meaningful in the context of a single layer and there are many possible label mappings between layers. For example, a naïve method would find the best-fit mapping of labels between layers using the simple matching coefficient of network nodes,  $SMC(c_s, c_t)$  to evaluate each possible mapping. However, the number of such mappings increases factorially with the larger of the label counts between the two snapshots, quickly becoming unreasonable to compute exhaustively (e.g. 10 labels would give over 3 million possible mappings). The problem can instead be posed as a form of the ubiquitous 'assignment problem'. By constructing the weighted bipartite network with edges  $e_{ij} = M_{ij}$  where  $i$  and  $j$  are the labels in snapshot community vectors  $c_s$  and  $c_t$  respectively, and  $M_{ij}$  is the total number of places where  $c_s$  and  $c_t$  have values  $i$  and  $j$  respectively, and finding the independent edge set of maximal total weight. The problem may now be solved (in polynomial time) to provide us a simple similarity measure with which to compare the community assignments between different network layers.

The final step in the MCAC process is to produce cluster assignments for multiplex layers, grouping those with a similar community structure under the same index. For this step we use the standard method of  $k$ -medoids; a local optimisation method which computes the distance between multiplex layers using a simple matching coefficient of politician community assignments described above, and iteratively merges layers into  $k$  distinct clusters. We choose  $k$  as the value which produces the first local maximum in the clustering silhouette score. If the data was unsuitable for partitioning into clusters, such as if it were drawn from

a continuous rather than categorical distribution, we would be unlikely to find a local optimum in the silhouette index and would not proceed with a cluster-based analysis. The output is a list of cluster assignments for MCA columns (corresponding to multiplex layers). As previously noted, the simple matching distance function, and  $k$ -medoids is one of many potential candidates for this type of clustering analysis. Several approaches were attempted, each producing broadly similar cluster assignments, with  $k$  between 3 and 5. Ultimately we have opted to present what we consider to be the simplest choice for description, implementation and interpretation.

#### 4.4. Clustering politicians by community co-occurrence.

Having grouped the multiplex network into layers with similar community structure, it remains to summarise these clusters in an intuitive way for human interpretation. For this purpose, we aim to highlight the differences in community structure by identifying which groups of nodes share a tendency for the same community affiliation by comparing their community affiliation vector over clustered layers of the multiplex, the matrix rows illustrated in Fig. 1d. Since we compare community labels within the same layers, there is no need for anything more complicated than the simple matching coefficient to measure similarity of the node community affiliation vector. The product of this process is that for each cluster of multiplex layers, we have a set of cluster assignments for each user (MCA rows) and can see the broad features of the community structure which characterise the clusters computed in the previous step. This is visualised in Fig. 4a by finding the frequency distribution of political affiliations within a user cluster, and drawing nodes whose style and size reflects this affiliation and prevalence within the cluster. As with MCAC, we evaluate the clustering quality by the silhouette index, selecting the number of clusters as that which produces the first local optimum in silhouette index. The methodological flow of MCAC is complete, and illustrated by Fig. 1.

#### 4.5. Characterising multiplex layers by tweet content

For purposes of comparison, we also apply a simple method for characterising the content of the retweets which produce our multiplex layers; those which join MPs and MEPs by at most two-step directed paths. For each multiplex layer, the textual content of all retweets is combined into a single document. Original tweets retweeted multiple times in a network also occur multiple times in the aggregated network; no tweets used in our MCAC method are omitted. These aggregated documents are first transformed into token-frequency distributions by splitting on non-alphabetic characters such as numbers and punctuation, and ignoring character case. Special characters with no obvious alphabetic mapping, such as UTF-16 characters encoding emojis are not tokenised, and omitted from the model. These token distributions are truncated at both ends by removing relatively rare tokens, those occurring in fewer than five documents, and overwhelmingly common words, often referred to as *stop-words*, which occur in more than half of the aggregated documents.

These document-token frequency-distributions are represented in matrix form by a term-frequency (TF) matrix, where each document is stored in a row, and each token as a column. The output of this process is a TF matrix with 85 rows and 71,926 columns, one for each week and token respectively. We provide additional power to tokens which co-occur in small numbers of documents (although above the threshold for removal by previous steps) by weighting each column by the inverse-document-frequency (IDF); that is the logarithmically scaled inverse of the fraction of documents in which the token occurs at least once. For example, a token appearing in 10 of our 85 documents would receive a weight of  $\log \frac{85}{10} \approx 2$  while

the weight of a token occurring in almost all documents would approach zero. Along similar lines to the previous analysis, this matrix is subject to  $k$ -medoids clustering with an appropriate distance function chosen as the cosine distance; the cosine of angular distance between vectors (which is independent of the magnitude of vectors, and therefore the differing size of documents).

More sophisticated methods exist for each step; tokenisation may be improved by considering  $n$ -grams, and there is a plethora of clustering algorithms which enjoy a great deal of attention from the topic-modelling community. However, our intent is not to produce a high-quality topic model, but to demonstrate a broad correspondence between the network state (which groups politicians are interacting via retweets) and content state (what common terms are being retweeted). We believe this is best demonstrated by comparison with a simple implementation of a relatively naive topic model as described.

#### Acknowledgements

This work was supported by the Economic and Social Research Council [ES/N012283/1]. The data used in this paper are available upon request from the lead author.

#### References

- Adamic, L.A., Glance, N., 2005. The political blogosphere and the 2004 us election: divided they blog. *Proceedings of the 3rd International Workshop on Link Discovery (ACM)*, 36–43.
- Baxter, G., Marcella, R., 2012. Does Scotland 'like' this? Social media use by political parties and candidates in Scotland during the 2010 UK General Election. *Libri* 62 (June(2)).
- Beauchamp, N., Laver, M., Nagler, J., Beck, N., Martin, A., Benoit, K., Lowe, W., Mead, L., Gordon, S., Stewart, B., Herzog, A., 2011. *Using Text to Scale Legislatures With Uninformative Voting*, pp. 10.
- Boireau, M., Gagliolo, M., Van Haute, E., Sudulich, M.L., 2015. The determinants and dynamics of Twitter-based interactions among candidates. In: *Papier préparé pour la journée d'étude*, June, pp. 2–3.
- Bond, Robert M., Fariss, Christopher J., Jones, Jason J., Kramer, Adam D.L., Marlow, Cameron, Settle, Jaime E., Fowler, James H., 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489 (September(7415)), 295–298.
- Boucek, F., 2009. Rethinking factionalism. *Party Polit.* 15, 1–31.
- Centola, D., 2010. The spread of behavior in an online social network experiment. *Science* 329 (5996), 1194–1197.
- Cherepnalkoski, D., Mozetic, I., 2016. Retweet networks of the European Parliament: evaluation of the community structure. *Appl. Netw. Sci.* 1 (June(1)), 2.
- Cherepnalkoski, D., Karpf, A., Mozetic, I., Grcar, M., 2016. Cohesion and coalition formation in the European parliament: roll-call votes and twitter activities. *PLOS ONE* 11 (November(11)), e0166586 arXiv:1608.04917.
- Christakis, N.A., Fowler, J.H., 2007. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* 357 (4), 370–379.
- Christakis, N.A., Fowler, J.H., 2008. The collective dynamics of smoking in a large social network. *N. Engl. J. Med.* 358 (21), 2249–2258.
- Conover, M.D., Gonçalves, B., Flammini, A., Menczer, F., 2012. Partisan asymmetries in online political activity. *EPJ Data Sci.* 1 (1), 1–19.
- Cook, J.M., 2016. Are American politicians as partisan online as they are offline? *Twitter networks in the U.S. senate and Maine state legislature. Policy Internet* 8 (March(1)), 55–71.
- Dal Maso, C., Pompa, G., Puliga, M., Riotta, G., Chessa, A., 2014. Voting behavior, coalitions and government strength through a complex network analysis. *PLOS ONE* 9 (December(12)), e116046.
- Darren, G., Lilleker, Nigel, A., Jackson, 2010. Towards a more participatory style of election campaigning: the impact of web 2.0 on the UK 2010 general election. *Policy Internet* 2 (August(3)), 69–98.
- Fortunato, S., 2010. Community detection in graphs. *Phys. Rep.* 486 (3), 75–174.
- Fowler, J.H., Christakis, N.A., 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham heart study. *Br. Med. J.* 337.
- Graham, T., Broersma, M., Hazelhoff, K., van't Haar, G., 2013. Between broadcasting political messages and interacting with voters. *Inf. Commun. Soc.* 16 (June(5)), 692–716.
- Hix, S., Goetz, K.H., 2000. *Europeanised Politics?: European Integration and National Political Systems*, vol. 23., pp. 10.
- Hix, S., 2010. Scaling the commons: using mps' left-right self-placement and voting divisions to map the British parliament, 1997–2005. *Annual Meeting of the American Political Science.*

- Jungherr, A., 2016. Twitter use in election campaigns: a systematic literature review. *J. Inf. Technol. Polit.* 13 (1), 72–91.
- Kellermann, M., 2012. Estimating ideal points in the British house of commons using early day motions. *Am. J. Polit. Sci.* 56 (3), 757–771.
- Kramer, A.D.I., Guillory, J.E., Hancock, J.T., 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci.* 111 (24), 8788–8790.
- Ladrech, R., 2002. Europeanization and political parties: towards a framework for analysis. *Party Polit.* 8 (4), 389–403.
- Lauderdale, B.E., Herzog, A., 2016. Measuring political positions from legislative speech. *Polit. Anal.* 24 (3), 374–394.
- Lazer, D., 2011. Networks in political science: back to the future. *PS: Polit. Sci. Polit.* 44 (January(1)), 61–68.
- Mair, P., 2008. Political parties and party systems. In: *Europeanization*. Springer, pp. 154–166.
- McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* 27, 415–444.
- Muchnik, L., Aral, S., Taylor, S.J., 2013. Social influence bias: a randomized experiment. *Science* 341 (6146), 647–651.
- Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. *Phys. Rev. E* 69 (February), 026113.
- Newman, N., 2010. #UKelection2010, Mainstream Media and the Role of the Internet. Reuters Institute for the Study of Journalism.
- Nulty, P., Theocharis, Y., Popa, S.A., Parnet, O., Benoit, K., 2016. Social media and political communication in the 2014 elections to the European parliament. *Elect. Stud.* 44 (Supplement C), 429–444.
- Ringe, N., Victor, J.N., Cho, W.T., 2017. *Legislative Networks*.
- Salganik, M.J., Dodds, P.S., Watts, D.J., 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311 (5762), 854–856.
- Sartori, G., 2005. *Parties and Party Systems: A Framework for Analysis*. ECPR Press.
- Shalizi, C.R., Thomas, A.C., 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* 40 (2), 211–239.
- Spirling, A., McLean, I., 2006. The rights and wrongs of roll calls. *Gov. Oppos.* 41 (4), 581–588.
- Spirling, A., McLean, I., 2007. UK OC OK? Interpreting optimal classification scores for the U.K. house of commons. *Polit. Anal.* 15 (1), 85.
- Sunstein, C.R., 2007. *Republic.com 2.0*. Princeton University Press, Princeton, NJ, USA.
- Theocharis, Y., Barbera, P., Fazekas, Z., Popa, S.A., Parnet, O., 2016. A bad workman blames his tweets: the consequences of citizens' uncivil twitter use when interacting with party candidates. *J. Commun.* 66 (6), 1007–1031.
- Waugh, A.S., Pei, L., Fowler, J.H., Mucha, P.J., Porter, M.A., 2009. Party Polarization in Congress: A Network Science Approach, arXiv:0907.3509 [physics], July.
- Williams, H.T.P., McMurray, J.R., Kurz, T., Lambert, H.F., 2015. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Glob. Environ. Change* 32, 126–138.