American politics in 3D: measuring multidimensional issue alignment in social media using social graphs and text data

Pedro Ramaciotti[1,2,3*†], Duncan Cassells[2,3,4†], Zografoula Vagena[5†], Jean-Philippe Cointet[2†] and Michael Bailey[2,6†]

†Pedro Ramaciotti is main contributor. Duncan Cassells, Zografoula Vagena, Jean-Philippe Cointet and Michael Bailey have equal contributor.

*Correspondence:
pedro.
ramaciottimorales@sciencespo.fr

[1] CNRS, Complex Systems Institute of Paris Ile-de-France (ISC-PIF), Paris, France
[2] médialab, Sciences Po, Paris, France
[3] LPI, Université Paris Cité, Paris, France
[4] LIP6, Sorbonne Université – CNRS, Paris, France
[5] Data Intelligence Institute of Paris, Paris, France
[6] Georgetown University, Washington, DC, USA

Abstract

A growing number of social media studies in the U.S. rely on the characterization of the opinion of individual users, for example, as Democrat- or Republican-leaning, or in continuous scales ranging from most liberal to most conservative. Recent works have shown, however, that additional opinion dimensions, for instance measuring attitudes towards elites, institutions, or cultural change, are also relevant for understanding socio-informational phenomena on social platforms and in politics in general. The study of social networks in high-dimensional opinion spaces remains challenging in the US, both because of the relative dominance of a principal liberal-conservative dimension in observed phenomena, and because two-party political systems structure both the preferences of users and the tools to measure them. This article leverages graph embedding in multi-dimensional latent opinion spaces and text analysis to propose a method to identify additional opinion dimensions linked to cultural, policy, social, and ideological groups and preferences. Using Twitter social graph data we infer the political stance of nearly 2 million users connected to the political debate in the U.S. for several issue dimensions of public debate. We show that it is possible to identify several new dimensions structuring social graphs, non-aligned with the classic liberal-conservative dimension. We also show how the social graph is polarized to different degrees along these newfound dimensions, leveraging multi-modality measures in opinion space. These results shed a new light on ideal point estimation methods gaining attention in social media studies, showing that they cannot always assume to capture liberal-conservative divides in single-dimensional models.

Keywords: Social graphs, Graph embedding, Network homophily, Ideological scaling, Ideal point estimation, Polarization, Issue alignment

Introduction

The study of socio-political dysfunctions or disorders unfolding in digital social media and social networks (Benkler et al. 2018) has raised to prominence in the past decade, including studies of algorithmic bias (Bakshy et al. 2015), extremism (O'Callaghan et al. 2015), or echo chambers (Barberá et al. 2015). These studies hinge on assessments of the political positions or stances of online users. Bakshy et al. (2015), for example,

classified users and content on Facebook as Democrat- or Republican-leaning to analyze cross-cutting recommendations, and Barberá et al. (2015) positioned Twitter users on liberal-to-conservative continuous scales to investigate the so-called echo chambers. In several European countries, assessments of political positions require multiple dimensions (Bakker et al. 2012) to account for observed social choice data, from roll call voting (Cointet et al. 2021) to online social network activity (Ramaciotti Morales et al. 2021). In the United States, however, political positions typically are reduced to one-dimensional explanations, a natural result of the first-past-the-post electoral system that privileges two-party competition (Riker 1982) and the fact that opinions on economics, gun control, abortion, race and other issues are highly correlated (Poole and Rosenthal 1997) and increasingly polarized (Mason 2015).

Single-dimensional preferences in the United States are not necessarily inevitable, however. Certainly, the vast array of social experiences in the U.S. make it conceivable that not everyone falls simply into a one-dimensional cleavage. Views on trade have long been only weakly related to traditional ideological cleavages (Bailey 2001, 2003). And recently, populist and anti-elite sentiment does not always track with traditional left–right cleavages (Uscinski et al. 2021). (Ahler and Broockman (2018)) found, for example, that support for Donald Trump in 2016 was better predicted by conservatism on immigration and liberalism on taxes than it was by traditional left–right measures of ideology, suggesting that the policy underpinnings of one-dimensional ideological conflict in the U.S. have evolved in ways that may have reflected untapped off-dimensional preferences. A recent work has used data from the American National Election Studies to characterize several dimensions of polarization in American politics (Ojer et al. 2023). If a part of political competition can be understood through spatial political opinion models, off-dimensional axes of political competition that are relatively orthogonal to the main liberal-conservative axis, become important tools for understanding individuals near political positions that are most susceptible to preference swings.

This article builds on recent ideological scaling (Morales et al. 2020) and graph embedding methods for spatializing social graphs in multi-dimensional ideological spaces (Ramaciotti Morales et al. 2022). Exploiting graph embedding and text analysis methods, it proposes a methodology to identify new relevant political dimensions linked to cultural, policy, social, and ideological groups and preferences in social graphs. We apply the method to X/Twitter (hereinafter Twitter) social graph data of nearly two million users strongly connected to the online political debate in the US. We find that several opinion dimensions traditionally considered in social network analysis (*e.g.*, conservatism, gun control, patriotism, religion) are indeed strongly aligned, as most studies find. We also are able to quantitatively measure the relative alignment of these issues and, importantly, identify and compute positions of large numbers of users in emerging, quasi orthogonal dimensions that may reflect emerging lines of tension in politics. By placing U.S. social media participants in a multidimensional space that includes dimensions that are not highly correlated we are able to cast a new light on divisions within the U.S. political system. Issues not aligned with the main dimension distinguishing liberal from conservatives, and that are better captured by additional political dimensions in our sample, include attitudes towards cosmopolitanism or local or global views (Ramaciotti Morales et al. 2021), and attitudes towards liberal lifestyles or cultural change (Bakker et al. 2019).

One of the main results of this article is the measurement of alignment between the classic dimension retrieved using classic single-dimension ideal point estimation methods, and the dimensions best representing tensions that are often attributed to it: *e.g.*, party cleavage, ideological liberal-conservative divides, and candidate preferences.

We show theoretically and empirically that, because ideal point estimation models are invariant to rotations on ideal points, (1) single-dimension models cannot be taken *a priori* to capture any of these tensions, meaning that (2) they need *ex post* validation by different means, (3) that issues and divides attributed a priori to single-dimensional ideal point estimation models might not be completely aligned, and (4) that rotations of ideal points in retrieved political opinion space can produce improve ideological or political scales for separate issues, including one that are not highly aligned. After having identified spatial directions that best represent attitudes and ideologies, we then take interest in the degree to which these directions produce different polarized spatial arrangements or distributions of users. To measure this, we project the position of the users in our sample onto the different computed directions that best distinguish attitudes towards the analyzed issues. Using the new coordinates along these spatial directions, we apply measures of polarization developed in axiomatic theories to assess the degree to which these dimensions produce multimodal distributions. In a previous article (Ramaciotti Morales 2023) we laid out the principles of the method used here. In this extended version we provide a formal theoretical and methodological description, and we show how to leverage identified directions associated with political issues to provide a spatial semantic for the latent space.

This paper proceeds as follows. We begin by discussing the literature on political preference estimation ("Estimating political preferences in one and multiple dimensions" section) and then move on to explain the Twitter data that we will use in this study ("Social network data" section). Then, we present the latent space embedding procedure and the results produced using the selected fraction of the Twitter social graph, showing the distribution of users along dimensions of latent space ("Homophily network embedding in latent space" section). First, we propose an exploration of the dimensions of this multi-dimensional latent space based on words chosen by users in their Twitter profile bios ("Exploring political concepts in space using text profiles" section). This exploration will both point towards leads in linking the dimensions with political concepts, while highlighting the limits of this exercise, often employed in other social media research works of the literature. We will then propose a way of overcoming these limits by jointly exploiting graph embedding and text classification methods ("Discovering spatial directions of political tension" section). This allows us to propose several spatial directions within our multi-dimensional latent space that best capture positive and negative attitudes towards selected issues that are relevant in U.S. politics. This also allows us to quantify issue and ideology alignment in "Measuring issue alignment" section. In "Off-dimensional users" section we investigate how different types of users have diversely dispersed in our latent political opinion space, and in particular which type of users are the farthest from the main direction of political competition opposing liberals and conservatives. Using our newfound directions, we will finally assess the degree to which these dimensions represent polarizing tensions by measuring the degree

of multi-modality of the distributions of users of our sample along them ("Measuring polarization in spatial directions" section).

Estimating political preferences in one and multiple dimensions

Many researchers have used binary categorical classification of social media and network users counts, relying, *e.g.*, on self-reporting and surveys (Bakshy et al. 2015) or sophisticated methods using neural networks on heterogeneous graphs (Xiao et al. 2020).

One of the most prominent approaches to estimating preferences in the U.S. is Poole and Rosenthal's Nominal Three-Step Estimation (NOMINATE) method (Poole and Rosenthal 1985) which has been applied to measure congressional preferences based on their roll call votes. The NOMINATE method can estimate multiple dimensions but since the 1970s it strongly suggests that divisions in Congress are single-dimensional. The NOMINATE model assumes that legislators have unobservable ideal policy positions in n-dimensional space and vote for bills that are ideologically close to them in the unobservable space. Closeness is computed as distances based on positions estimated via an iterative maximum likelihood procedure. Clinton, Jackman and Rivers used a similar model and data to estimate preferences using MCMC Bayesian methods (Clinton et al. 2004).

These models have been extended to estimate preferences of survey respondents (Bafumi and Herron 2010). As with the legislative models, models using survey data in the U.S. suggest that preferences are largely—but not completely—one dimensional (Uscinski et al. 2021). Others have applied similar models to the Supreme Court (Bailey 2007; Martin and Quinn 2002; Lauderdale and Clark 2012), campaign contributions (Bonica 2018). Several papers discuss how to use EM algorithms to estimate these models efficiently (Imai et al. 2016; Peress 2022). Barbera (Barberá et al. 2015) extends the logic to network data by modeling connections as

$$P\left(i \to j | \alpha_i, \beta_j, \gamma, \vec{\phi}_i, \vec{\phi}_j\right) = \text{logit}^{-1}\left(\alpha_i + \beta_j - \gamma \|\vec{\phi}_i - \vec{\phi}_j\|^2\right), \tag{1}$$

using social media accounts of politicians—members of parliament (MP) in Barbera's work—and those of their followers. In these models, such as that of (1), the probability of observing user i following user j (*i.e.*, $i \to j$) depends on position and scale parameters α_i (activity of user in number of friends), β_j (popularity of MP in number of followers) and γ (sensitivity parameter), and, most importantly, on the distance between the unobservable position $\vec{\phi}_i$ and $\vec{\phi}_j$ of users i and j. Social choice data (*i.e.*, pairs $i \to j$), forming a social graph can then be used to infer position $\vec{\phi}_i$ for any user i. Applications of such models typically assume that one dimension is enough to retrieve the main social cleavage in the United States, namely the liberal-conservative one, and use social network data to compute the position of users in some liberal-conservative scale. On Twitter, for example, Barberá (2015), considered how users follow (or not) accounts of political figures, while on Facebook, Bond and Messing (Bond and Messing 2015) considered how users like pages of political figures. In both cases, they effectively apply ideology scaling or ideal point estimation techniques to explain how users provide signals of approval (following on Twitter or liking on Facebook) towards politicians, applying the same

principle previously used to explain how politicians provided signals of approval towards bills (*i.e.*, voting).

These works often rely on *ex post* validation using text cues to argue that the latent dimension reflects indeed on political positions of users. Multi-dimensional inference for $\vec{\phi}$ can be achieved in a computationally-tractable manner with Correspondence Analysis (Greenacre 2017) as it has been shown to approximate the inference of unobservable parameters of (1), both theoretically (Lowe 2008) and empirically (Barberá et al. 2015).

While there is little doubt that many preferences are well characterized by a single dimension in the United States, it may be unwise to ignore the possibility of multiple dimensions. First, not all issues map onto the one-dimensional policy space. For example, international trade policy has long been an issue that does not divide along conventional left–right lines as very progressive and very conservative people and politicians have often shared protectionist sentiments (Bailey 2001). And trade and related views toward globalization may not simply be an oddity, but may have played an important role in the recent emergence of Trump (Jensen et al. 2017) and the emergence of a conservatism focused on anti-trade, anti-immigrant and America first sentiment (Uscinski et al. 2021; Ahler and Broockman 2018). These views may relate to other important policies such as aid to Ukraine, as corners of the traditional left and the modern right have been more likely to praise Russia and raise concerns about supporting Ukraine (Campbell 2023). Historically, off-dimensional issues have been important. In the 1960s, race was off-dimensional as there were many Republicans and Democrats on both sides of the issue (Poole and Rosenthal 1997). In the 1970s, abortion was off-dimensional as there were many Republicans and Democrats on both sides of the issue (Adams 1997). Understanding off-dimensional issues holds importance for understanding possible reconfigurations of political competition. In a single-dimensional political liberal-conservative competition, from a proximity voting perspective (*i.e.*, voters casting preferences for political offers—candidates or parties—that are the closest to them (Downs 1957)), individuals that are susceptible to swing preferences lie at the frontier, equidistant from political offers. If political competition is structured along additional independent and orthogonal dimensions, swinging of political preference occurs in new regions of space
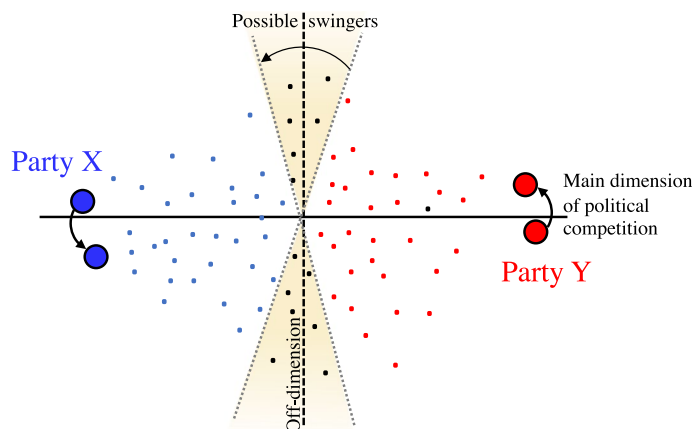

Fig. 1 Illustration of a multi-dimensional political competition setting showing how off-dimensional users are relatively more susceptible to swing preferences

characterized by these new dimensions, and that might be more sensible to changes of stance on the part of the political offer. Figure 1 illustrates such a setting in a two party system such as that of the U.S.

Just as issues may not map to the traditional left–right dimension, individuals may also not map easily into this single dimension. Broockman (2016) noted that many people have extreme views on specific policies but in a pattern that is poorly described by traditional left–right ideology. Fowler et al. (2022) found that about 20% of Americans "give a mix of liberal and conservative views that are not well described by the liberal-conservative dimension" but nonetheless are coherent. Such individuals constitute a non-trivial portion of the electorate, with their political importance magnified by the fact that they are more likely to be pivotal swing voters in hotly contested elections.

Understanding the nature of these off-dimensional issues and preferences may shed light on the dimensions that divide politics. Ideology is not a construct with fixed meaning; it evolves over time: it is, as Converse (1964) and Noel (2013) note, a question of what goes with what. During the presidency of George W. Bush some staunch conservatives, including President Bush and Fox personality Sean Hannity, sought to liberalize immigration policy. Such a position is almost unfathomable in today's conservative politics. At that time, privatizing Social Security and cutting Medicare were *de rigueur* for conservatives; such initiatives got less traction in the MAGA-version of modern conservatism.

Noel (2013) shows that ideology not only summarizes existing divisions in the United States, but also that ideological thinking can "organize policies and their proponents into coalitions that party leaders then seek to represent." For new thinking to matter, it needs to somehow differ from existing thinking in some way. One way that thinking can be new is to connect different policy positions in new ways. In practice, political competition might drive political figures and parties to compete and to present policy and ideological proposals to voters along off-dimensions: issue and ideological dimensions not aligning with the main liberal-conservative one. While the leading edge of this work is likely concentrated among intellectuals and political entrepreneurs, it also needs to filter out to a larger public if it is to be consequential. Social media is, therefore, a good venue for exploring new trends because the people who follow political actors are likely to be relatively motivated to explore new ideas. If a new way to connect policies or a new cluster of actors with off-dimensional preferences is proposed, this may be a sign of possible source of instability or change in the status quo one-dimensional paradigm.

There are two major challenges to estimating multi-dimensional models. First, they need to be estimated, something that can require identifying assumptions (Rivers 2003) and/or be computationally challenging. Greenacre shows that multi-dimensional versions of the model can be estimated in a computationally-tractable manner with Correspondence Analysis (Greenacre 2017). These models approximate the inference of unobservable parameters of (1). The second challenge with multi-dimensional models is that they need to be interpreted with care. Because (1) depends on unobservable parameters $\vec{\phi}$ through pairwise distances, their inference is invariant to isometric transformations. In particular rotation transformations mean that retrieved dimensions cannot be assured to be aligned with strong social cleavages that might be structuring political choices. This means that, in general, it cannot be assured that a single-dimensional

ideological scaling model will yield a political opinion scale completely aligned with some presumed main left–right or liberal-conservative dimension. Ideological scaling models need to test and validate how they relate to political concepts. In European settings, Ramaciotti Morales, Cointet and Muñoz-Zolotoochin use the position of referential users such as politicians of known political parties, and party positions in reference issue spaces (provided, *e.g.*, by political polls or surveys), to infer dimensions that align with issues of public political debate (Ramaciotti Morales et al. 2021). Using the position of several political parties, this fact has been leveraged in embedding large numbers of users in multi-dimensional space where dimensions stand for identifiable and separate political issues, not requiring *ex post* interpretation or validation (Ramaciotti Morales et al. 2022). These methods cannot be directly applied to the U.S. context because the two-party system does not allow for determining mappings from latent spaces produced by ideological scaling and spaces on which the two parties have been positioned along several dimensions.

This article proposes a two-step procedure for estimating multi-dimensional political preferences among U.S. Twitter users. First, we use Correspondence Analysis to estimate a multidimensional latent space in which users are arranged according to homophily in preference of MPs: users close in space follow similar sets of MPs on Twitter. Second, we use text descriptions written by users in their online profiles on Twitter constructing groups of referential users on more than a dozen possible issue cleavages. This allows us to estimate spatial directions within this latent space that can be associated with attitudes towards these issues. The goal is to better understand the dominant cleavage and to identify emergent opinions that are not highly correlated with the liberal-conservative dimension. This also allows us to evaluate the degree to which dimensions inferred by ideology scaling or ideal point estimation, often leveraged in literature, are aligned with main cleavages attributed to them: including party, candidate, or liberal-conservative ideological divides.

Social network data

To produce a sample of Twitter users that can be coherently positioned in multidimensional political spaces, we identify a population on the platform by their vicinity to political figures. Following multidimensional ideological scaling works in Europe (Ramaciotti Morales et al. 2021) and in the US (Barberá et al. 2015), we select a bipartite sub-graph of the Twitter social graph. To capture online social choices that might be revealing of several social and political preferences we take members of the US Congress as reference users. Our collection process was carried out in October 2020. We manually annotate the Twitter accounts of 550 members of the 116th United States Congress (looking for verified accounts corresponding to each congressperson), and collected their 17 952 824 followers (collection performed using Twitter's API in October 27th, 2020, see the Acknowledgements section for privacy-compliance information and references). To minimize the probability of followers being *bots* we follow criteria adopted by several studies (Ramaciotti Morales et al. 2021; Ramaciotti Morales and Muñoz Zolotoochin 2022; Morales et al. 2020; Ramaciotti Morales and Cointet 2021) and further identified followers with more than 25 followers (7 325 940), and users that have posted more at least 100 tweets (7 471 365). See Barberá (2015) for further details behind the

rationale for these parameters. This is done to identify users that are strongly connected to political debate, to limit the possibility of including users that follow an MP for reasons other than ideology or policy issues, and to ensure that users follow spatial preference models, we identify followers that follow at least three members of congress (3 846 925) (Barberá 2015). We select the 1 821 272 unique followers that satisfy all three conditions.

The next section describes how we produce a latent homophily space for this bipartite social graph. To establish reference points in latent space, we collect the text self-descriptions made by users in their Twitter profiles (also on October 27th, 2020). Out of 1 821 272 users, 1 442 716 had written any text entry in their Twitter profiles. This collection, performed in the days leading to the 2020 United States Presidential Election has the additional advantage of allowing us to investigate preferences for candidates.

Homophily network embedding in latent space

To identify dimensions that might be revealing of ideological or policy distinctions driving differences in how users follow MPs, we first produce a multi-dimensional space embedding in which these dimensions might emerge as spatial directions. For this, we take the bipartite social subgraph of the $m = 550$ members of congress and their $n = 1\ 821\ 272$ followers to produce an homophily embedding of the adjacency matrix to compute values $\vec{\phi}$ of (1). As described in "Estimating political preferences in one and multiple dimensions" section, this is achieved by computing the Correspondence Analysis of the adjacency matrix of this bipartite network, of which we will provide a summarized description (see Greenacre 2017 for further details). Formally, consider the adjacency matrix $A \in \{0,1\}^{m \times n}$ of the bipartite network, where $A_{ij} = 1$ if user i follows MP j, but has value $A_{ij} = 0$ if not. Now consider the marginal empirical discrete distributions $w_m = (1/a)A\mathbf{1}$ and $w_n = (1/a)\mathbf{1}^T A$, where $a = \sum_i \sum_j A_{ij}$ and $\mathbf{1}$ is a column vector of ones. Using the marginal distributions, we also consider diagonal matrices $W_m = \mathrm{diag}(1/\sqrt{w_m})$ and $W_n = \mathrm{diag}(1/\sqrt{w_n})$, and the standardized residuals matrix $S = (1/a)W_m(A - a\,w_m w_n)W_n$. If $S = U\Sigma V^T$ is the singular value decomposition of matrix S, the latent space coordinates of users are given by $F_m = W_m U\Sigma \in \mathbb{R}^{m \times \min(m,n)}$ for MPs, and $F_n = W_n V\Sigma \in \mathbb{R}^{n \times \min(m,n)}$ for their followers. More precisely, Correspondence Analysis approximates the Maximum Likelihood Estimation (MLE) of $\vec{\phi}_i$ and $\vec{\phi}_j$ in (1). Because several users follow the exact same set of MPs, it is admitted in this formulation that some users may share latent space coordinates. This is particularly true for combinations of MPs that have high visibility in the media. Coordinates F_m approximate MLE of $\vec{\phi}_i$ for followers and coordinates F_n approximate MLE of $\vec{\phi}_j$ for MPs. This is because it can be proven that the MLE expression for the $\vec{\phi}_i$ and $\vec{\phi}_j$ can be solved iteratively with a Markov Chain Monte Carlo Method, and that the coordinates computed with the Correspondence Analysis approximates the first iteration. See Lowe (2008, Section 7) for a proof of the approximation, and Barberá et al. (2015, Supplementary Material, Section 1) for empirical results using a bipartite Twitter network between MPs in the United States and their followers.

We consider the space in which MPs and followers have coordinates given by F_m and F_n. In this space, if singular values in Σ are ordered by magnitude, dimensions δ_p (for $p = 1, 2, \ldots$) are ranked according to the information they contain about choices

represented in the bipartite social graph, as measured by the inertia. The projection of positions $\vec{\phi}_j$ of MPs i and follower $\vec{\phi}_i$ along dimension δ_p of the latent space are then, correspondingly, $F_{n,j,p}$ and $F_{n,j,p}$. If singular values are ordered by magnitude, the inertia of each dimension provides an estimate of the relative importance of the dimensions in explaining the observed bipartite graph. The inertia of dimension δ_p is computed as $\epsilon_p = \sigma_p^2 / \sum_{k=1}^{\min(m,n)} \sigma_k^2$, where σ_p if the p-th singular value in Σ. To assess the contribution of each dimension to the explanation of observation A, we defined the incremental gain in inertia as $\tilde{\epsilon}_p = \epsilon_p - \epsilon_{p-1}$. Figure 2 shows the inertia of each dimension and their incremental gain, showing that at most the three first dimensions are relatively more informative than the rest. Figure 2 also shows the embedding positions of both, congressional members and followers, and the marginal density on these first three dimensions, estimated with kernel density estimation for the purposes of visualization. We compute party positions as the mean position of congressional members from the same party. As anticipated by previous works on Twitter in the U.S., the first and most explicative dimension, δ_1, stands qualitatively as a good candidate of scale of attitudes towards parties or liberal-conservative ideologies. Next sections will seek to quantify the degree to
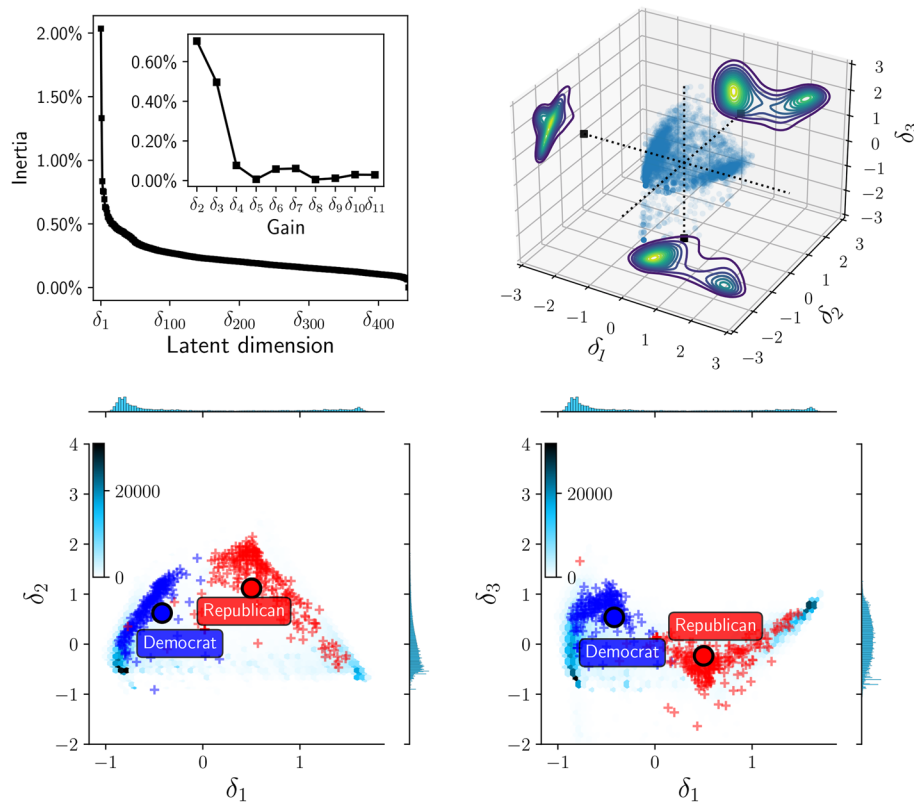

Fig. 2 Multi-dimensional homophily embedding of the collected Twitter network. Dimensions ranked by inertia, and incremental gain of each dimension (top left). Scatter plot and estimated marginal densities for the position of users in the first three dimensions (top right). Density of followers and positions of members of congress colored by party (Democrat + and Republican + MPs), and party positions as mean position of MPs from a same party (bottom)

which δ_1 stands as an indicator of this concepts, and to clarify the conceptual issues captured by dimensions.

Because the probability of a topological observation in (1) is invariant to isometries over latent positions $\vec{\phi}$, the question remains whether isometric transformations (*e.g.*, rotations) might be able to improve the spatial distinctions between Democrat- and Republican-leaning followers. This means that, while it is the case that δ_1—the classic ideal point estimation dimension—is a good candidate for a liberal-conservative scale, we do not know if a rotation might improve the ability of a classifier to distinguish between Democrat- and Republican-leaning individuals. We know that δ_1 stands for a latent tension in choice of MPs, and we know that it is highly aligned with party cleavage, but we do not know if it is the best spatial direction for distinguishing these two groups. More broadly, it is not trivial to attribute an inductive meaning to what δ_2 and δ_3 might stand for, or to any other space direction for that matter.

Exploring political concepts in space using text profiles

In this section, we use the description text written by users in their Twitter profiles to explore the concepts associated with the dimensions of the homophily latent space computed in the previous one. This explorative analysis will both (1) suggest political concepts that might be associated with dimensions that order users according to attitudes, and (2) highlight the difficulties and the limits of producing text-based spatial interpretation in latent spaces. This explorative analysis is produced in three steps. First, we will distinguish user profiles by the sentiment they convey, as estimated using a pre-trained BERT base model for uncased words (Devlin et al. 2018), assigning to each profile text a sentiment from 1 (very negative) to 5 (very positive). We transformed texts into lower capitalization, and removed special character and emoji. We label text profiles as negative ($-$) if sentiment is equal to 1 or 2, as positive ($+$) if sentiment is equal to 4 or 5, and neutral (n) if sentiment value is equal to 3. We distinguish terms uttered in profiles with estimated positive, negative, and neutral sentiment. This is necessary to distinguish words that are bound to appear in expression of support or criticism, that sentiment might be able to capture. For example, we expect that term "liberal" will have different spatial properties according to whether it has been included in negative (*e.g.*,"don't vote for corrupt liberals!") and in positive statements (*e.g.*,"I am a proud liberal"). We distinguish the "liberal($-$)" (that appears in texts with negative sentiment) from "liberal($+$)" (appearing in texts with positive sentiment). Second, we consider salient terms in profiles and measure their semantic pertinence in order to focus only on the most relevant one. We automatically identify up to 2-grams contained in the text and which match a predefined grammatical pattern allowing us to gather noun phrases and adjectives. We then compute the C-value metric (Frantzi et al. 2000) of these terms to measure their unithood, that is, in the words of Kageura and Umino (1996): "the degree of strength or stability of syntagmatic combinations and collocations". Terms with the higher C-value are most likely to denote actual semantic units which may characterize user preferences. Third, we analyze the spatial distribution of the identified relevant and sentiment-specific terms. These three parts of the analysis are implemented as follows. First, we lemmatize the terms present in the texts. Then we distinguish them by the sentiment of the text in which they are present, and compute the C-value for each term. We then retain

the 2000 terms with the highest C-value, and compute their mean position along δ_1, δ_2 & δ_3, as the mean position of the texts in which they appear. Each text is a profile description, and thus has the position of the user that wrote it.

Having the mean position of the 2000 most important terms, we first examine the most extreme terms along each dimension. The 1st latent dimension δ_1 follows the expectation of distinguishing between liberals and conservatives. The most negative terms in δ_1 include "bidenharris(+)", "voteblue(+)", "bluewave(+)", "proud democrat"(+), "black lives matter(+)", or still "wearamask(−)" (often uttered with negative sentiment, or accompanied by critiques). The most positive terms in δ_1 include "maga kag(+)" (for "keep america great"), "maga patriot(+)", "president trump(+)", or "proud conservative(+)" and "conservative christian(+)". In contrast, dimension δ_2 does not immediately yield to interpretation by looking at extreme terms. Most negative terms in δ_2 include support for both Trump and Biden (e.g., "trump 2020(+)" and "biden harris(+)"), as well as terms associated with liberals (e.g., "resister(+)") and conservatives ("patriot american(+)", or "god fearing(+)"). This spatial shared trend between supporters of both candidates supports the idea of an underlying political notion orthogonal, or independent, of the main liberal-conservative divide. Most positive terms in δ_2 include many signaling the use of a collective or institutional voice, with less clearly marked liberal of conservative expressions: *e.g.*, "association(+)", "representing(+)", "twitter official account(+)". Most negative terms in δ_3 include terms of self-description: *e.g.*, locations (such as "Kentucky(+)", "Colorado(+)" or "Miami(+)"), words associated with occupations (such as "software(+)" or "actor(+)") or personal traits or hobbies ("obsessed(+)" or "games(+)"). Finally, most positive terms in δ_3 include terms of partisan conservative support: *e.g.*, "trump 2020(+)" or "maga patriot(+)". See "Appendix A" for a more detailed table of the most extreme terms by dimension.

While the 1st dimension seems to conform to expectations in the way the resulting terms are related to liberal-conservative and partisan divides, it is less clear what the most extreme terms say about the 2nd and 3rd dimension. Extreme terms might not necessarily provide good examples of the underlying political concepts that dimensions might be capturing. Instead, they could well be expressions regarding topics for which interest only develops in extremist users. Thus, a different exploratory approach consists of inspecting the skewness of the terms, measured as the skewness of the profile texts in which each term appears along a dimension. Skewness, as a measure of distributional asymmetry, measures whether a term is more used in the negative extreme positions, but with a long-tailed distribution towards the positive positions (very positive skewness), or if, for example, a term is more used in the positive extreme positions, but with a long-tailed distribution towards the negative positions (very negative skewness). Skewness tells us then whether a term is more frequently used as we move towards one extreme along one dimension. This is different from the mean positions of extreme terms, which might concern only a small niche position. We compute the skewness of each term and compare it to their mean position along each dimension (see Fig. 3). Skewness and position follow a clear and expected inverse relation for the 1st dimension: very negative terms are also positively skewed, while positive terms are also negatively skewed, following a tendency that is consistent along the whole range of δ_1. This suggests that term usage along this dimension reflects a continuous ideological tension, with people's
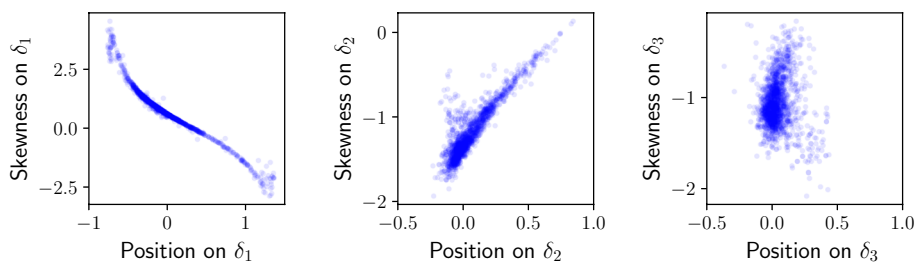
Fig. 3 Relation between skewness and mean position of terms along each spatial dimension. Position and skewness of a term are the mean position and skewness of the documents in which it appears

frequency of use of terms continuously changing across the spectrum subtended by this dimension. The same cannot be said of dimensions δ_2 and δ_3. Terms are generally negatively skewed along δ_2, with a clear relation between position and skewness: the more negative a term position is, the more negatively skewed the distribution of profiles on which it appears. Most negatively skewed terms along δ_2 include self-description of users referring to their families (e.g., "married(+)", "proud mother(+)"), expressions of personal attitudes and sentiments (*e.g.*, "love president(+)", "life to the fullest(+)", "love all(+)") or personal interests (*e.g.*, "love animals(+)", "rock(+)", "games(+)"). Terms are generally negatively skewed along δ_3, independent of the position. Most negatively skewed terms along δ_3 include expressions of partisan support (*e.g.*, "maga patriot(+)", "bidenharris2020(+)") and references to religion and family (*e.g.*, "god(+)", "god fearing(+)", "love god family(+)"). See Appendix B for a more detailed table of the most skewed terms by dimension.

These first exploratory results suggest that δ_2 might be related to individual vs collective or institutional perspective and attitudes, while δ_3 might be related to cultural or moral differences, but it is finally inconclusive. The difficulty in explaining underlying political notions attributable to dimensions beyond the first axis of political competition in social media in the U.S. has also been reported in other works with inconclusive results (Barberá and Rivero 2015). Given the fact that our sample is strongly connected to U.S. politics (in degree and distance with respect to political Twitter accounts), the presence of utterances of candidate preferences, and the format and length of text profiles, leaves little room for the emergence of other preferences that might characterize dimensions.

Discovering spatial directions of political tension

In this section we leverage a different strategy to attribute meaning to spatial dimensions. Instead of inspecting how terms are used along our three dimensions, we select terms that should be revealing of political tensions, and then estimate what is the spatial direction in our three-dimensional space along which this tension is best dichotomized. This strategy is inspired in recent works that show that, in latent multidimensional space for social graphs, dichotomous terms denoting sides in ideological or issue tensions (*e.g.*, people describing themselves as "left-wing" and "right-wing"), can be distinguished in latent space by linear classifiers (Ramaciotti Morales and Muñoz Zolotoochin 2022). In this strategy, we select pairs of groups of labels that might be revealing of political

tension or polarization, but considering a larger scope of possible tensions, beyond left–right divides. Following the example from Ramaciotti Morales and Muñoz Zolotoochin (2022) for terms "left" and "right", the goal is not to capture the diversity of ways in which users might signal left- or right-wing political affinities, but to select minimal pairs of groups of terms that will identify two groups of users that should be positioned in opposite sides of the latent space, revealing some spatial direction of political tension.

Let us illustrate this principle with a simple example based on party cleavages. Among the users of our sample embedded in the latent space, 7 895 use the word "republican" and 14 481 the word "democrat" in their profile without negative sentiment (so as to exclude utterances of criticism). While these terms do not capture the diversity of ways of expressing partisan support (with alternatives including, *e.g.*, "GOP voter"), we expect that the position of users on these two groups should reveal a spatial direction that is associated with party cleavage. To measure the degree to which δ_1, δ_2 or δ_3 might be good candidate directions for distinguishing these two groups, we fit a logistic regression model on each dimension based on these two classes. We then use the fitted logistic model as a binary classifier using a probability value equal to 0.5 as the threshold separating class regions. With this classifier, and looking at true and false positive and negative classifications, we can compute a precision, recall and F1-score metric. We use F1-score as a metric of the ability of a dimension to distinguish two classes. Figure 4 (left panel) shows these values and the distribution of these two groups along δ_1, δ_2 and δ_3. We observe that δ_1 is indeed the only dimension among the three to produce a meaningful distinction, with a F1 value of 0.815 for δ_1, but 0.318 and 0.0 for δ_2 and δ_3 respectively. This dimension, δ_1, is the traditional result of computing an ideological scaling, as done in Barberá (2015), Barberá et al. (2015), attributed in the literature with the concept of liberal-conservative political divide. While the described procedure allows for testing how dimensions distinguish pairs of groups, it does not readily tell us which spatial directions might best do so. Alternatively, instead of using a given dimension, we can fit a multivariate logistic regression model, and identify the direction perpendicular to the decision boundary surface (determined again with the 0.5 probability threshold). In the case of our three-dimensional model, the decision boundary will be a plane and the direction a three-dimensional vector (see in Fig. 4, right panel). This direction provides us with new coordinates (the projection over the vector of the direction) for users over the specific identified direction (direction $d_{\text{Dem-Rep}}$ in the case of Fig. 4). This discovered direction separating these two groups of users is well aligned with δ_1, but it does not produce an improvement in the F1-score. The established practice in ideological scaling in social media data in the U.S. is to suppose that a single-dimensional model (*i.e.*, δ_1) captures the main party cleavage. Yet, as this example shows, ideological scaling cannot rely on the a priori assumption that this will always be the case, especially in light of research suggesting a decline in left–right cleavages structuring collective choice (Grossman and Sauger 2019), as it is standard practice in many disciplines. Indeed in other national settings, left–right divides have been shown to be aligned to δ_2 and not to δ_1 (Ramaciotti Morales et al. 2021). This also stems from the fact that, in (1), the probability of a given topological observation is invariant to isometries in the positions of users and MPs in the latent space.
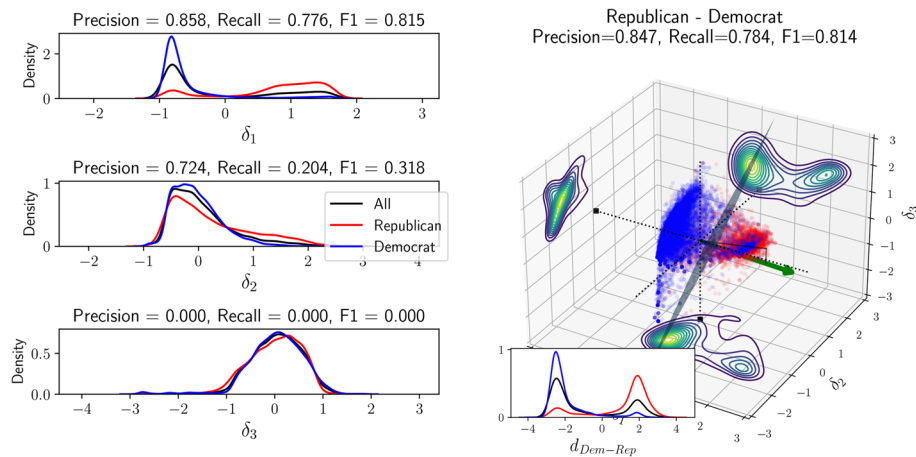
Fig. 4 Distribution Republican- or Democrat-leaning according to their Twitter text profile description, their distribution along the first 3 latent space dimensions, and the accuracy of logistic regression models fitted on each dimension (left). Conditional distributions and positions of labeled users in three-dimensions, and distributions along the direction perpendicular to the boundary of a multivariate logistic regression (right)

Table 1 Proposed issue partitions of users into minimal groups for mining spatial direction of political tension

Partition	Values	Users	Partition	Values	Users
Party	Republican	7895	Police	Pro-Police	1 811
	Democrat	14481		Anti-Police	1 686
Candidate	Trump sup.	50 159	Military	Pro-Military	17 721
	Biden sup.	10 229		Anti-Military	8 187
Ideology	Conservative	12 331	Patriotism	Patriot	27 907
	Liberal	8 917		Anti-Imperialism	120
Left/Right	Right-wing	8 462	Local/Global	Local focus	5 532
	Left-wing	3 423		Global focus	6 283
Black Lives Matter	Pro-BLM	13 350	Deepstate/rationalism	DeepState	486
	Anti-BLM	12 017		Rationalist	111
Black/Blue	BlackLivesMatter	16 690	Entrepreneur	Entrepreneur	13 209
	BlueLivesMatter	240		Anti-capitalist	1 287
Rural/Urban	Rural	2 174	Liberty/Justice	Justice	18 829
	Urban	3 382		Freedom	13 041
Religion	Christian	22 735	Pronouns	Anti-Incl. Pronous	1 195
	Atheist	1 081		Incl. Pronous.	66 218
Guns	Pro-Gun	605	International	International	5 957
	Anti-Gun	496		Globalization	84
Communism	Anti-Communism	1 160	Individual/institutional	Institutional	976
	Pro-communism	448		Individual	14 739
Liberal lifestyle	Pro-Liberal LifeStyle	9 659	Science	Science	14 795
	Anti-liberal LifeStyle	9 921		AntiExperts	59
Libertarian/welfare	Libertarian	3 045	Masks (COVID)	Pro-Mask	4 247
	Welfare	681		Anti-Mask	214

For each issue we identify two disjoint groups based on Twitter profile text descriptions

Following the previous example, we now set out to identify additional spatial directions associated with political tension. The purpose of this is threefold. First, we want to assess the degree to which δ_1 represents the main party and ideological cleavage, and what issues define it. Second, we want to measure issue alignment between different lines of tensions. Third, we want to leverage discovered directions of political tension in providing conceptual meaning to δ_1, δ_2 & δ_3. To propose pairs of groups of users that might be revealing of tensions, we surveyed issues reported by recent works in social media politics that grant special attention to the question of multi-dimensionality or emerging lines of tension (Baumann et al. 2020; Ramaciotti Morales et al. 2022; Uscinski et al. 2021). To characterize the first dimension, we identify pairs of users according to party, candidate, and ideological (liberal or conservative) preferences. We also include a number of issues well identified in the literature as usually aligned with the main cleavage: racial issues, gun policy, and religious principles. Finally, to explore possible directions of political tension, we include several issues from the literature proposed as tensions possibly not aligned with liberal-conservative divides: cleavages in regional politics (urban vs rural), the new cultural issue of communism in the US, political differences related to liberal "life-styles" (Bakker et al. 2019) (*e.g.*, homosexuality, feminism), attitudes on welfare state and libertarianism, on the military, on patriotism, on globalization and the internationalization of the economy, and on conspirationism and mistrust

Table 2 Groups of pairs of labeled users (according to criteria of Table 1), naming of the mined dimension perpendicular to the decision boundary of a multivariate logistic regression classification model, and the accuracy of the fitted model

Label 1	Label 2	Dimension	Precision	Recall	F1
Democrat	Republican	$d_{Dem-Rep}$	0.814	0.823	**0.819**
Trump sup.	Biden sup.	$d_{Trump-Biden}$	0.944	0.874	**0.908**
Liberal	Conservative	$d_{Ideology}$	0.877	0.848	**0.862**
Right	Left	$d_{left-right}$	0.615	0.484	0.542
Pro-BLM	Anti-BLM	d_{BLM}	0.488	0.205	0.288
BlackLM	BlueLM	$d_{BlueLives}$	0.687	0.925	**0.789**
Urban	Rural	d_{Urban}	0.569	0.655	**0.609**
Christian	Atheist	$d_{Christian}$	0.706	0.559	**0.624**
Pro-Gun	Anti-Guns	$d_{Pro-Gun}$	0.713	0.544	**0.617**
Pro-Communism	Anti-Communism	d_{Comm}	0.648	0.645	**0.647**
Liberal Life-Style	Anti-Lib. Life-Sty.	$d_{LifeStyle}$	0.537	0.867	**0.664**
Libertarian	Welfare	$d_{Welfare}$	0.695	0.720	**0.707**
Pro-Police	Anti-Police	d_{Police}	0.617	0.624	**0.620**
Pro-Military	Anti-Military	$d_{Military}$	0.599	0.474	0.529
Patriot	Anti-imperialism	$d_{Patriot}$	0.709	0.692	**0.700**
Global focus	Local focus	d_{Local}	0.547	0.676	**0.605**
DeepState	Rationalism	$d_{Deepstate}$	0.569	0.667	**0.614**
Anticapitalism	Entrepreneur	$d_{Entrepreneur}$	0.679	0.549	**0.607**
Justice	Liberty	$d_{Liberty}$	0.664	0.652	**0.658**
Pro-Incl. pronouns	Anti-incl. pronouns	$d_{Anti-Pronouns}$	0.630	0.710	**0.667**
Internationalization	Anti-globalization	$d_{International}$	0.731	0.583	**0.649**
Individual Views	Institutional Views	$d_{Individual}$	0.606	0.572	0.588
Pro-Mask	Anti-Mask	$d_{Pro-Mask}$	0.702	0.748	**0.724**
Science	Anti-Experts	$d_{Science}$	0.673	0.627	**0.649**

in institutions. Table 1 summarizes pairs of sets of users identified, specifying the name of the binary partition, the binary values, and the name of identified users. Users corresponding to each binary value are identified using the aforementioned approach based on minimal keywords. See Table 1 in Appendix C for a definition of the dictionary of terms used for the classification.

After identifying the binary groups of Table 1 we proceed to fit the best spatial direction that dichotomizes them, following the example from Fig. 4. We fit a multivariate logistic regression model for each group pair, and measure the classification accuracy of the model, reported in Table 2, highlighting in bold characters the cases with F1-score accuracy equal or greater to 0.6. When pairs are highly imbalanced (*e.g.*, for religious cleavages there are 22 735 identified "christian" users vs 1 081 "atheists"), we systematically sub-sample the majority group with a Near-Miss strategy (Mani and Zhang 2003). Figure 5 an example of labeled users, according to whether the express support for Biden or Trump, with the decision boundary and discovered orthogonal direction of the fitted multivariate decision model. This selection highlights the different qualities in the accuracy of the multivariate logistic regression classifier, corresponding to different strengths of cleavages for the pairs in each labeled group, under the assumption that the chosen criteria identify a relevant group of users.
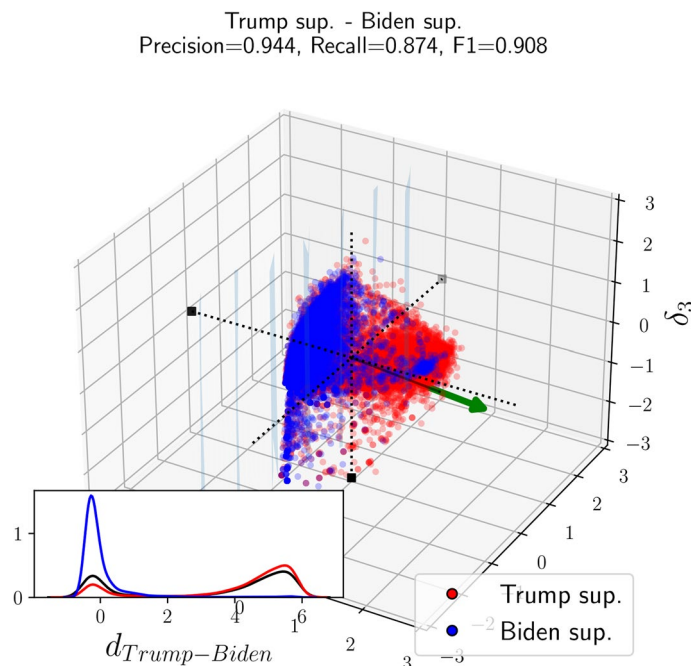
Fig. 5 Illustration of the discovery of spatial directions using pairs of groups of users identified with different issues. Users expressing support for Biden or Trump are shown in blue and red. The direction shown in the figure corresponds to the normal to the decision boundary of a multivariate logistic regression model trained to separate these two groups in the latent space computed using the follower graph. Precision, recall, and F1 metrics for this classification are provided for this model, as well as the density of position of these two groups in the spatial direction defined by the normal

Measuring issue alignment

Having identified plausible spatial directions of political tension in the latent space spanned by dimensions δ_1, δ_2 & δ_3, we now address the question of the relation between these directions and our three dimensions. In particular, we seek to establish to which issues and ideologies are dimensions δ_1, δ_2 & δ_3 related, and to measure issue alignment in our three-dimensional latent space. In our new spatial directions, users can be projected to provide a measure of their attitudes towards a given issue. For example, direction $d_{Pro-Gun}$ captures positive and negative attitudes of users towards guns. In contrast, δ_1 is a proxy for party cleavages, but also for other positions on correlated issues (*e.g.*, racial or religious issues, see Fig. 5). By inspecting the alignment between different retrieved spatial directions we can identify and quantify issue alignment. Figure 6 shows the retrieved spatial directions of political tension (*i.e.*, with F1-score ≥ 0.6) and their pairwise angular distance. To measure this alignment we consider the minimal angle separating the lines containing the two given directions. This means that if two directions point in exactly opposite directions (*i.e.*, having an inner product value of -1 between the vectors normal to the decision boundary), their angular distance will be of 0°. Once all pairwise angular distances have been measured between these directions, we compute clusters of closely aligned directions using a Un-weighted Pair Group Method with Arithmetic (UPGMA) mean (Sokal 1958). More precisely, we compute a hierarchical cluster structure of the pairwise angular distance matrix. We then present the clusters that result cutting the dendrogram of the UPGMA hierarchical clustering at the first granularity level at which dimensions δ_1, δ_2 & δ_3 are separated into different clusters. While the granularity level of the cluster can be arbitrarily fixed, this prescribed threshold provides the closest issue directions associated with each dimension, and thus suggest meaning for the latent space dimensions. This procedure results in the identification of five groups or clusters of issue directions. We call these clusters *ideologies* in the sense that they are indicative of issue alignment as one of the main phenomena associated with polarization (Jost et al. 2022). This alignment is also reflective of ideology in the sense that individuals might be constrained to adopt preferences on certain issues by virtue of preferences that they have already adopted on others (Baldassarri and Gelman 2008). The five ideological clusters are: (1) a dominant ideology comprising party, candidate, and other stances correlated with δ_1, (2) an ideology separating people defining themselves using the words "local" and "global", (3) an ideology separating people that use inclusive pronouns, define themselves as using the word "international", or having positive mentions of sciences in opposition to people criticizing experts and inclusive pronouns, (4) an ideology separating those defining themselves using the words "welfare" and "libertarian", and (5) an ideology separating those with positive and negative mentions of issues relating to sexual diversity and feminism, and the use of the word "communism". This last cluster also includes attitudes towards wearing masks during the COVID19 pandemic. Five directions cannot be perfectly orthogonal in three-dimensions, but any two directions belonging to two different identified ideological clusters will display enough angular distance, so as to not be considered as highly aligned.
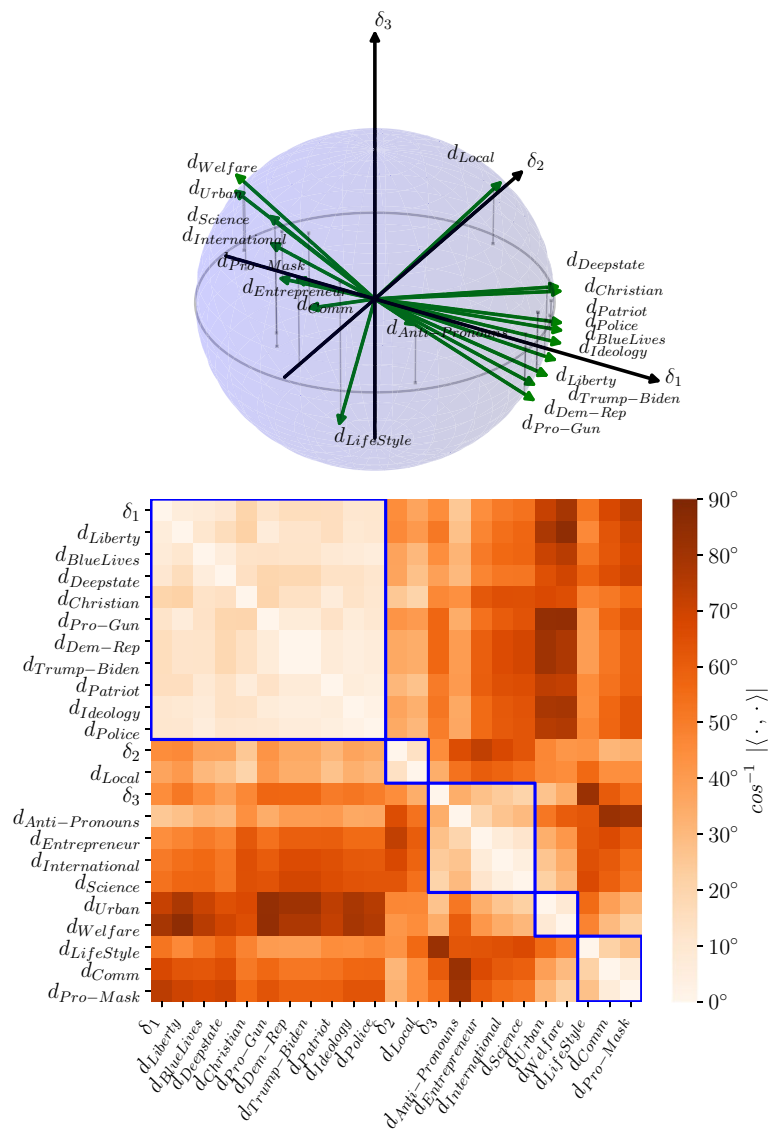
Fig. 6 Mined spatial direction of political tension linked to issues and ideologies (left) can be organized into five *ideologies* (in the sense of issue alignment polarization) computed with UPGMA clustering, shown in five groups in blue in the angular distance matrix (right)

Being able to disentangle issues in separate directions, enables us to conduct different investigations against the map positions of actors in now identifiable axes. Because we can also measure the position of reference users (politicians) in identified political tension directions, we can investigate intra-party diversity on separate issues: *e.g.*, of support for their presidential candidate, or attitudes towards welfare, religious diversity, or diversity of views on racial issues. Figure 7 shows, for example, that Republicans are more heterogeneous in their support for Donald Trump than the Democrats in their support for Joseph Biden, both the members of congress (in crosses in Fig. 7) and the followers (density shown in light blue in Fig. 7).
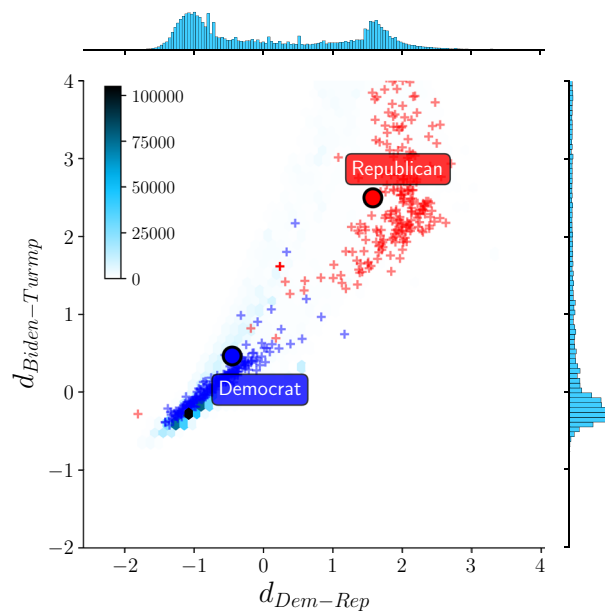
Fig. 7 Density of Twitter users and positions of members of congress along mined dimensions. The distribution of members of congress shows the intra-party diversity of stances towards presidential candidates
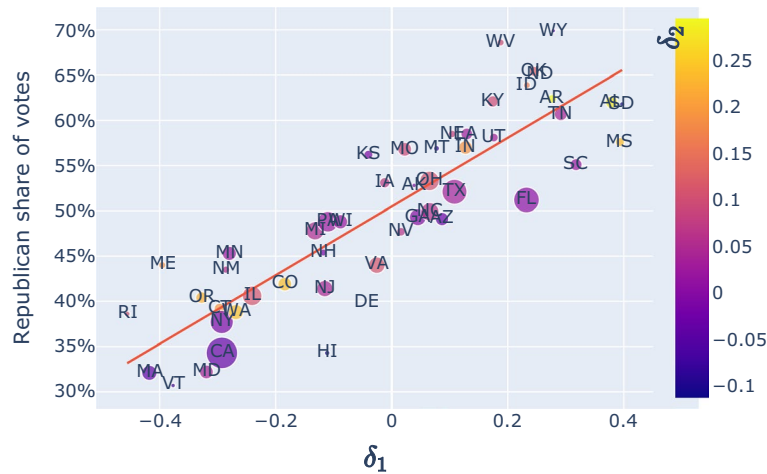

Fig. 8 Scatter plot showing the correlation between the average position of Twitter users on the first dimension of the ideological space and the proportion of votes obtained by the Republican Party during the 2020 presidential elections. Nodes' area scales with the size of the state. Each state is colored according to the average position of its users along the second dimension. The regression line is computed using OLS

Researchers have sought to further validate this type of Twitter ideology scaling using electoral results (Barberá and Rivero 2015). For the particular electoral outcome corresponding to the collection date of our dataset (October 2020), we propose a measure of validation using external data. To validate our dataset using electoral results we identify the geographical locations mentioned in texts of Twitter profiles (*e.g.*, "Dad of three, from Massachusetts"), to match users with states whenever possible. This allows us to identify the mean position of States along the first dimension

δ_1. We then compare the mean position of States computed with our dataset and the percentage of Republican voters.[1] The comparison shows a direct relation between the two quantities (see Fig. 8, with an adjusted R^2 value of 0.756. In comparison, dimensions δ_2 and δ_3 hold no relation with the electoral outcome (see color scale in Fig. 8 for δ_2), with adjusted R^2 values at 0.002 and 0.301 respectively.

Off-dimensional users

Having laid out several coherent arguments for the role of the first dimension δ_1 as the main dimension of political competition between liberals and conservatives, we seek to further characterize off-dimensional users: individuals whose position sits relatively distant to this dimension. This holds importance in political competition, as these off-dimensional individuals might be the most sensitive to change of stances on the part of parties and candidates (see Fig. 1). To characterize these individuals we use again the text of Twitter profiles and their positions in our latent space. To scout for possible text identifiers revealing the political identity of individuals, we select from the list of the 2 000 most explicative terms (according to their C-value) of "Exploring political concepts in space using text profiles" section all terms that speak to individual characteristics. These terms can be self-describing terms (*e.g.*, "christian", "gamer", "democrat", "artist", "teacher"), terms that convey criticism or opinion from a revealing stance (*e.g.*, "black lives matter", "blue lives matter", "imperialism", "woke"), or terms that identify preference, tastes, or that identify activities (*e.g.*, "yoga", "nature", "science", "tech"). We call these terms *labels*, of which we identified 172 among the first 2000. Next, we seek to determine how users that include these labels in their profiles are distant from δ_1 by measuring the eccentricity of the distribution of their use. Let us denote by Ω the region of three-dimensional latent space in which there are users present. For each label ℓ we consider the density $\rho_\ell(x)$ of users employing label ℓ at position $x \in \Omega$. We are interested in the eccentricity of $\rho_\ell(x)$ with respect to δ_1, which we measure as $r_{\delta_1}(x) = \min(\delta_1, x)$. Because we want to measure eccentricity
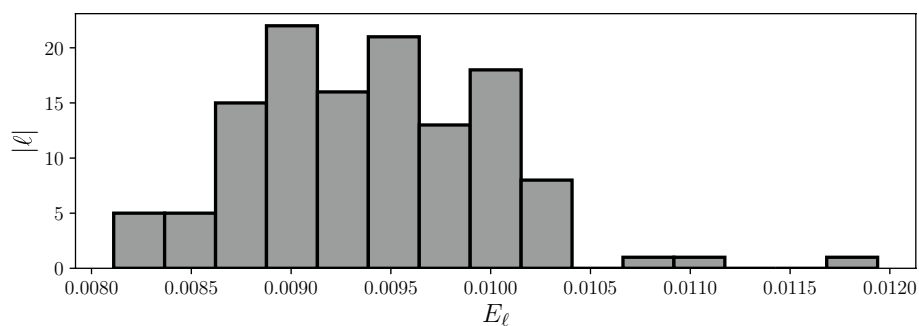

Fig. 9 Labels ℓ used by users in their Twitter profiles according to their eccentricity E_ℓ with respect to dimension δ_1 of the latent space

---

[1] Data downloaded from the Cook Political report https://www.cookpolitical.com/2020-national-popular-vote-tracker.

independently of the frequency with which different labels are used, we consider the normalized label density,

$$\hat{\rho}_\ell(x) = \frac{\rho_\ell(x)}{\int\limits_\Omega \rho_\ell(x)d\Omega}, \tag{2}$$

such that $\int\limits_\Omega \hat{\rho}_\ell d\Omega = 1$ for every label. We then measure the eccentricity E_ℓ of label ℓ with respect to δ_1 as:

$$E_\ell = \int\limits_\Omega r_{\delta_1}(x)\hat{\rho}_\ell(x)d\Omega. \tag{3}$$

We approximate E_ℓ by its Riemann integral dividing an arbitrary region encompassing all users $\Omega = [-3, 3]^3$ in 50 bins along each dimension and further restricting Ω to bins that contain at least 1000 users and labels that are used at least by 1000 users, so as to assure a robust estimation of ρ_ℓ as a proportion (changes in the arbitrary number of bins did not alter the ranking of most and least eccentric labels).

By construction, labels with high eccentricity values will be those relatively more used by users that are geometrically distant from main dimension δ_1, while labels with low eccentricity will be those relatively more used by users geometrically close to δ_1. We compute values E_ℓ for our identified labels with which users define themselves and report those with extreme values. A handful of labels (see Fig. 9) display a relatively high eccentricity (E_ℓ): "non-profit" (0.0119), "federal" (0.011), "local" (0.0108), "state" (0.0103) "education" (0.0103), "farmer" (0.0103), "taxes" (0.0102), "islam" (0.0102). See Fig. 10 for a distribution of eccentricities. These labels refer to Twitter accounts that take institutional stance ("non-profit", "state", "federal"), but also accounts that define themselves with respect to "local" interests (*e.g.*, "your local historian", "interested in local politics", "science and technology, life and style, local news"). Most eccentric labels also include issues such as "education" (*e.g.*, "agricultural education teacher", "democratic nominee, fighter for workers, healthcare, education", "covering education and government in georgia"), and "taxes" ("paid taxes 45 years, tired of giving my money away", "the idiot pays taxes, the taxes that the dems are using to spend us into oblivion!"). Other defining labels include "farmers" (*e.g.*, "nature conservation is partnering with farmers and ranchers!", "corn farmer in georgia"), and "islam" ("I despise false teaching of islam", "anti-islamic fundamentalist and pro-democracy", "end racism and end islamophobia", "won't tolerate racism and islamophobia"). While seemingly diverse, these labels point towards accounts that take institutional stances in the political space, and that refer to issues rather than camps. Highly partisan labels are unsurprisingly the lowest eccentricity values. The 20 least eccentric labels are: "wear a mask", "progressive", "black lives matter", "liberal", "atheist", "vegetarian", "he/him", "she/her", "lgbt", "cat", "biden", "democracy", "pro-choice", "literature", "association". Many of these low eccentricity labels are often associated with liberal and progressive stances, with notable exceptions: "cat" and "literature". The comparison between labels with extremely high and low eccentricity points to issues on the attention of
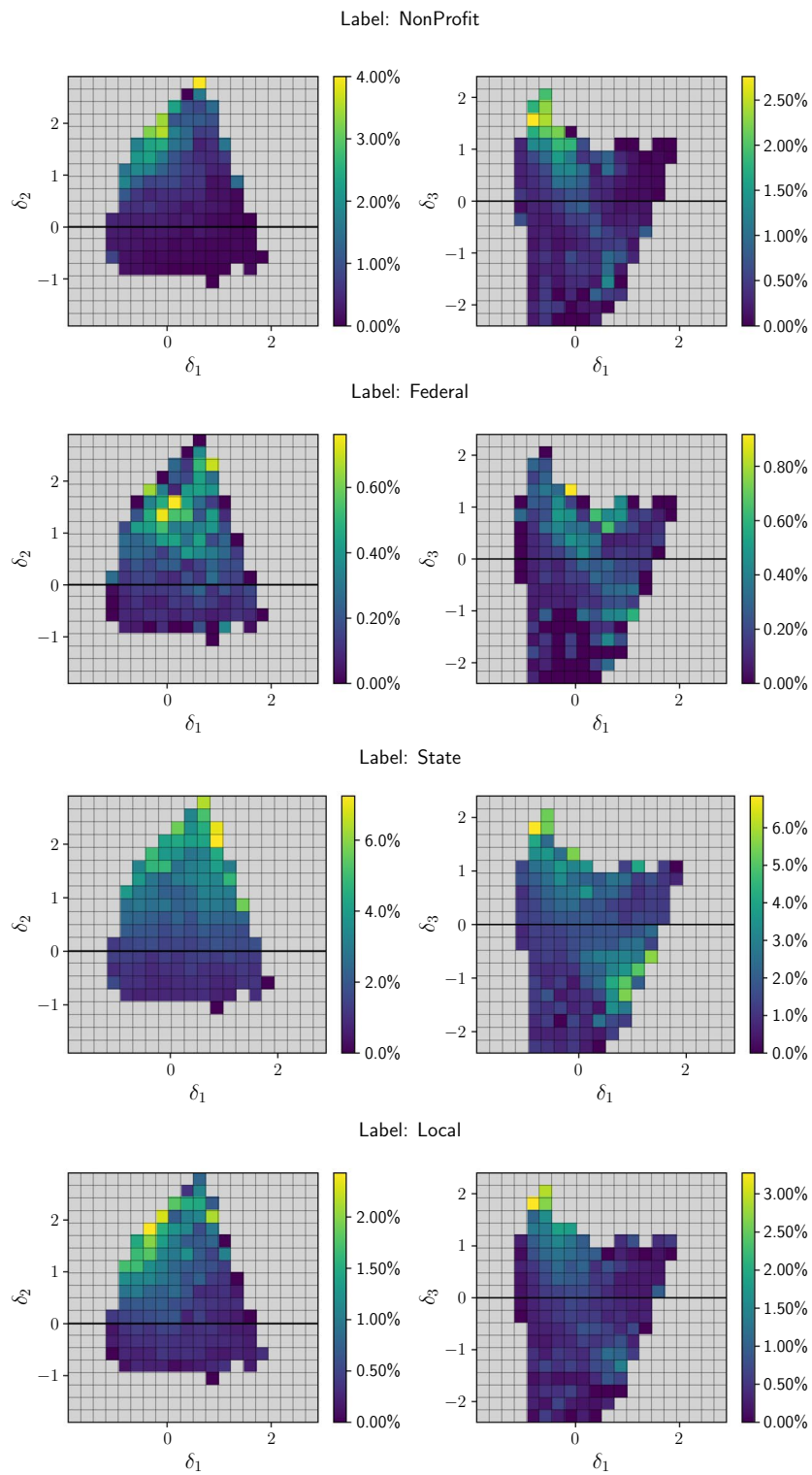
Fig. 10 Spatial density of the most eccentric labels used by users in their Twitter profiles with respect to the main dimension of political competition δ_1

institutional actors (as opposed to individual views), on issues that are comparatively closer to policy than to ideologies.

Measuring polarization in spatial directions

The dichotomous groups used to identify spatial directions of political tension in latent space do not allow us to say how polarized the distribution of our population is along these directions. This is because our choice of keywords is designed to identify users that are reliably in one or another of a public issue debate or ideological stance. In Fig. 4 (right panel), for example, two groups of users are identified (in blue and red curves): Democrat and Republican supporters. The spatial distribution of these two groups along the dimension they define (*i.e.*, $d_{Dem-Rep}$) is polarized according to several meanings often used in social polarization literature). On the one hand, members of each group are concentrated around distinguishable poles or positions in space. On the other hand, the distribution of users that belong to any of these two groups is clearly bimodal (black curve Fig. 4, right panel). See Bramson et al. (2016) for a comprehensive survey including these two conceptualizations of polarization. These distributions, however, do not tell us how polarized is the totality of users along $d_{Dem-Rep}$ (because our two groups do not include more subtle expressions of party support, *e.g.*, "hard to agree with dems on policy issues", neither do they capture users that simply do not utter party preferences in writing).

In order to assess polarization along identified spatial directions, and to compare it with how our binary groups identify directions, we compute two polarization metrics for each direction. First, we simply compute the binary *label spread* of binary labels; *e.g.*, for $d_{Dem-Rep}$, we compute the distance between the mean positions of users labeled Democrat and labeled Republican along the direction. Second, we compute a multi-modality metric of the distribution of the totality of users projected onto the direction. Our second metric is the Duclos–Esteban–Ray (DER) measure of polarization (Duclos et al. 2004), which captures two aspects of polarization that the authors term *alienation* and *identification*—analogous to affective and ideological polarization (Jost et al. 2022). For each spatial direction d, let x_i^d for $i = 1, \ldots, n$ be the positions of our $n = 1\,821\,272$ users projected onto d, and $\hat{f}_d$ the estimated density distribution. The DER metric is computed as:

$$P_\alpha(\hat{f}_d) = \int_d \int_d \hat{f}_d^{\alpha+1}(x)\hat{f}_d(y)|x - y|dxdy, \tag{4}$$

for $\alpha \in [1/4, 1]$, which we set at 0.5 (see Duclos et al. 2004, Section 3.2) for a discussion on the sensibility of the measure with respect to the choice of α). A sample based estimator for P_α is given by (see Section 4 of Duclos et al. 2004):

$$P_\alpha(\hat{F}) = n^{-1} \sum_{i=1}^{n} \hat{f}_d(x_i)^\alpha \hat{a}(x_i), \tag{5}$$

with $\hat{a}(x_i)$ given as

$$\hat{a}(x_i) = \hat{\mu} + x_i(n^{-1}(2i-1) - 1) - n^{-1}\left(2\sum_{j=1}^{i-1} x_j + x_i\right), \tag{6}$$

where $\hat{\mu}$ is the sample mean. We estimate $\hat{f}_d(\cdot)$ using kernel density estimation with bandwidth $h = 4.7n^{-0.5}\sigma\alpha^{0.1}$, with σ being the standard deviation (see Section 4.3 of Duclos et al. (2004) for the calculation of the optimal bandwidth). Figure 11 (top) compares these two polarization notions, showing the distribution of users labeled as Democrat and Republican supporters and the kernel density estimation of all users along the $d_{Dem-Rep}$ direction, with the corresponding DER polarization estimate (computed for the totality of users in our sample). Figure 11 (bottom) shows that our dichotomous binary labels define directions on which the separation of the means of the corresponding dichotomous groups are correlated with the polarization of the whole of users projected onto them. Binary labels identifying pairs of groups that are most distinguishable in space are also those that define spatial directions along which the whole of our sample is most bimodal. Some low polarization directions also have low label spread. This means that, for some dichotomous groups of users defining dimensions, the means
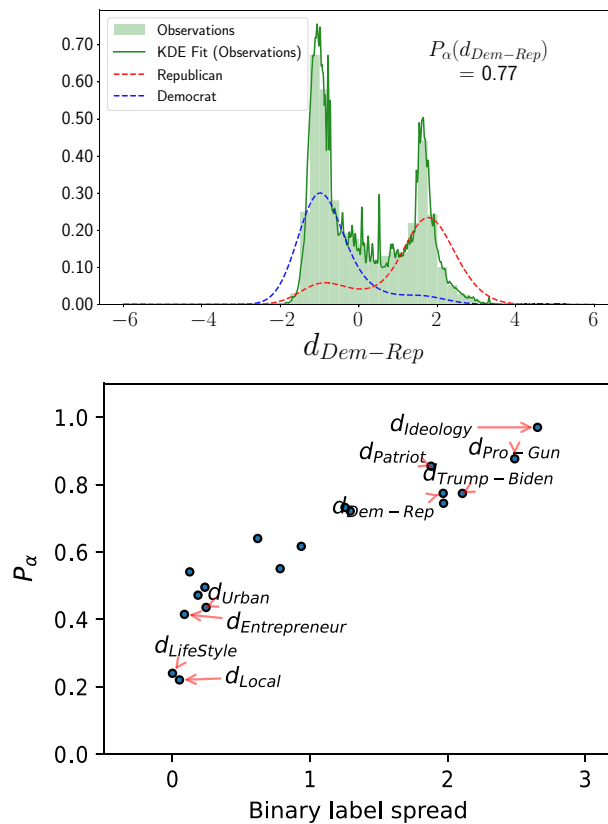

Fig. 11 Comparison of two polarization notions, showing the distribution of users labeled as Democrat and Republican supporters and the kernel density estimation of all users along the $d_{Dem-Rep}$ direction, with the corresponding DER polarization estimate P_α, computed for the totality of users in our sample (top figure). Comparison of DER and label spread metrics for computed spatial directions (bottom)

of both groups are similar due to outliers, all the while having boundaries separating enough members from both groups so as to achieve low enough false positives and false negatives, and sufficiently high F1-score (see Table 2).

Discussion and conclusions

This article argued that multidimensional preferences are interesting, even in the U.S. where preferences are overwhelmingly—and usefully—characterized as one-dimensional. Following traditional text-based analyses we illustrated the difficulty in proving multi-dimensional spatial models with inductive interpretation for dimensions. We then presented network embedding and NLP methods for estimating and interpreting multi-dimensional preferences in politically relevant ways. We applied the tools to the case of a political Twitter follower network around U.S. congressional members, identifying the main dominant cleavage, but also additional ones hypothesized as relevant by recent studies in social sciences (Uscinski et al. 2021).

We found that the main dimension is indeed aligned with traditional Democrat-Republican divides in the US. While not surprising, our results show that this should be verified, rather than assumed. In addition, having this measured and validated allows us to assess the degree of alignment between latent dimensions and different spatial directions of political tension. Standard practice in ideal point estimation consists of estimating position for a one-dimensional homophily model as in (1), to verify reliability in the way it positions users known to have liberal or conservative stances (*e.g.*, declaring themselves as progressives, sympathizers of the tea party, of black lives matters, or other groups), to then using this scale to analyze positions regarding other issues, such as attitudes towards abortion, immigration, racial issues, *etc.*. What our study suggests, both theoretically and empirically, is that the first dimension cannot always be expected to be a good indicator for liberal-conservative divides. Because the ideal point estimation is invariant to rotations, it is plausible that this old cleavage may lose importance in comparison to other divides in social media (as it has been observed in other countries Ramaciotti Morales et al. 2021). This can be caused by decline in the structuring power of this ideological divide (Grossman and Sauger 2019) (over collective choices revealed in digital traces), but also by the selection of particular online populations that might first be structured by other issues and ideologies (*e.g.*, politicized Twitter users, or users engaging a particular online debate). What our study also suggests, is that the first dimension of the latent space (*i.e.*, the scale of a one-dimensional ideal point estimation model) is not necessarily the best liberal-conservative scale retrievable in latent space, nor does it hold epistemic priority over other spatial directions. For example, consider a situation in which there are two closely aligned directions: (1) liberal-conservative and (2) pro- and anti-abortion stances. One common practice consists in computing a single-dimensional ideal point estimation model and validating adequate positioning of self-declared liberals and conservatives on opposite sides. We then might want to see how pro- and anti-abortion users are placed, leading us to some measurement of attitude polarization for this issue, for example. However, if, using our method, we retrieve a liberal-conservative axis that best separates self-declared liberals and conservatives, and if we inspect the positions of self-declared pro- and anti-abortion individuals projected

onto this axis, we might measure a different attitude polarization for abortion. If we are to grant epistemic precedence to a liberal-conservative axis on which to analyze other issues or ideologies, it might not be best captured by single-dimensional ideal point estimation models.

Our analysis also revealed several deviations from one-dimensional preferences. In particular, five ideologies, or bundled groups of polarization dimensions were identified. These groups of directions are not highly aligned between themselves, and represent new political tension dimensions that can be used in further studies. Further validation of these additional dimensions require additional data. One way of achieving this is by considering tweet streaming data from embedded users, or crossing Twitter identifiers with survey data on demographic, geographic, or voting characteristics of users. We were able to do so for the first and most determinant dimension of our latent space. We did this by identifying self-reported geographical positions of users, and comparing mean ideological stances per State with the fraction of Republican voters in the 2020 Presidential election. Acquiring these additional assurances about the main dimension of our latent space also allowed us to propose a new method for characterizing off-dimensional users, revealing that these users often adopt a less partisan and more institutional voice. Our results also suggest these off-dimensional users position themselves with regards to debates on issues (*e.g.*, taxes, education) rather than ideological camps (*e.g.*, liberals, progressives, atheists). The difficulty in obtaining new data with which to test the robustness of inferred ideological positions has regrettably increased with the change in access via the API of Twitter (now X) during the second quarter of 2023. While not impossible, the cost for conducting similar studies will become prohibitive for many research teams and will produce a steeper price on the volumes of data that, by virtue of abundance and diversity (*e.g.*, data on self-declared location, on interactions with other users, and uttered written expression) might provide a paths to proving robustness of this method.

This method, barring the new costs imposed for API access, also offers the possibility of developing new applications for explicitly measuring issue polarization as the alignment of bundled social cleavages, as well as a method for projecting large numbers of users onto space dimensions with explicit meaning in terms of the issues to which it measures positive and negative views.

This new possibility opens interesting paths for research, which we illustrated with a brief example. By measuring positions of Democrat and Republican congressional members on both a dimension of attitudes towards parties and towards candidate, this article showed that, when compared with Democrats, it may be proved that Republicans display higher heterogeneity in their support for their candidate. Beyond this example, many others could leverage these results and methods. In particular, having multidimensional distributions of political attitudes could be leveraged in the study of social mobilization (see for example Cointet et al. 2021; Ramaciotti Morales et al. 2021, 2022). Additionally, by leveraging information consumption practices and media diets, attitudinal positions could be attributed to news media articles and outlets, allowing for the study of diversity, or lack thereof, in information consumption patterns (Ramaciotti Morales et al. 2019; Morales et al. 2021). This, in turn, presents interesting possibilities for large-scale analysis of wide news and informational ecosystems (Cointet et al. 2021).

Table 3 Most extreme terms by dimension

1st dimension δ_1

Negative δ_1	Positive δ_1
bidenharris(+), bidenharris2020(+), bidenharris(−), voteblue(+), voteblue(−), resister(+), fbr(+), theresistance(+), fbr(−), bluewave(−), bluewave(+), biden harris(+), proud democrat(+), she/her(−), blacklivesmatter(−), theresistance(−), wearamask(−), resister(−), vote blue(−), black lives matter(+)	maga kag(+), maga patriot(+), kag(+), president trump(+), trump2020(+), trump 2020(+), maga kag(−), kag(−), trump maga(+), parler(+), trump supporter(+), conservative patriot (+), trump2020(−), maga(+), parler(−), americafirst(+), god family country(+), proud conservative(+), conservative christian(+)

2nd dimension δ_2

Negative δ_2	Positive δ_2
trump 2020(+), canadian(+), maga patriot(+), trump supporter(+), trump2020(+), patriot(+), resister(+), trump supporter(−), patriot american(+), resister(−), maga kag(+), trump2020(−), trump president(+), trump(+), love country(+), kag(+), god fearing(+), biden harris(+)	association(+), coalition(+), representing(+), organization(+), provides(+), advocacy(+), legislative(+), twitter official account(+), congressional(+), communities(+), senator(+), public affairs(+), official twitter(+), providing(+), account for(+), grassroots(+), families(+), promoting(+), programs(+), provide(+)

3rd dimension δ_3

Negative δ_3	Positive δ_3
que(+), miami(+), kentucky(+), bitcoin(+), gamer(+), republicans(+), los(+), florida(+), canadian(+), libertarian(−), software engineer(+), libertarian(+), actor(+), alabama(+), resister(−), colorado(+), games(+), dreamer(+), obsessed(+), software(+)	trump 2020(+), patriot maga(+), maga kag(+), president trump(+), kag(+), trump supporter(+), trump2020(+), kag(−), maga kag(−), maga(+), trump2020(−), trump maga(+), parler(+), president trump(−), maga(n), trump supporter(−), patriot(+)

We distinguish terms uttered in profiles with compute positive (+), negative (−), and neutral (n) sentiment. We report the 20 terms with most extreme positive and negative positions

Appendix A: mean position of terms of profiles in space

See "Exploring political concepts in space using text profiles" section for a description of the extraction of the sentiment-signed terms with most extreme means (Table 3).

While sentiment-signed terms are needed to discover spatial trends in otherwise highly used terms in both political extremes uttering both support and criticism, some terms related to candidate support still appear used with negative sentiment expressing support. The term "bidenharris(−)" is a clear example, which find instances on negative δ_1 such as: "political junkie ex gop coug fan pnw resistance fbr resist blacklivesmatter bidenharris", or "middle aged mom mba pursuer resister recovered evangelical overall pretty boring voteblue gocougs bidenharris2020 goawaytrumpandmaga". Similarly for "bluewave(−)" on negative δ_1: "married 31 yrs (this time) mother of 2 sons retired nurse democrat cincinnati reds fan impeachtrump muellertime bluewave2020", "theresistance bluewave2018 boycottnra". On the other side of political spectrum, for positive δ_1 we find cases such as "trump2020(−)" or "trump president(−)": 'just a regular guy husband father grandfather proud deplorable lifelong conservative supporter of Trump maga kag NRA member" or 'this Georgia wife mom & granny is a proud deplorable ! god bless president Trump". Several non-political keywords on these profile text bios have been changed to avoid the possibility of identification. Faced with the complexity of satire and negative sentiment for utterance of support, and mixed sentiments, our strategy aims to

distinguish this dimension as an additional one by differentiating positive and negative sentiment terms.

Table 4 Most skewed terms by dimension

1st dimension δ_1

Negative δ_1 (positive skewness)	Positive δ_1 (negative skewness)
biden harris(+), proud democrat(+), resist(+), biden-harris(+), resister(+), lgbtq(−), resistance(+), ally(−), blm(+), harris(+), blm(−), resist(−), resister(−), she(−), blacklivesmatter(−), resistance(−), black lives matter(+), feminist(+), black lives(+)	love president(+), maga(+), trump supporter(+), trump 2020(+), maga(n), nra(+), trump maga(+), president trump (−), trump president(+), god family country(+), maga(−), family country(+), maga kag(−), 2nd amendment(+), god country(+), trump supporter(−), america-first(+), god bless america(+), bless america(+), kag(−)

2nd dimension δ_2

Negative δ_2 (positive skewness)	Positive δ_2 (negative skewness)
No terms with highly positive skewness over δ_2. See Fig. 3.	canadian(+), world news(+), proud mother(+), love president(+), grandkids(+), happily married(+), cowboys(+), happily(+), love animals(+), american patriot(+), maga patriot(+), trump supporter(+), love all(+), fine(+), wife mother grandmother(+), huge(+), extraordinaire(+), games(+), rock(+), life to the fullest(+)

3rd dimension δ_3

Negative δ_3 (positive skewness)	Positive δ_3 (negative skewness)
No terms with positive skewness over δ_3. See Fig. 3.	love god family(+), patriot maga(+), parler(+), love country(+), god family(+), god fearing(+), fearing(+), bless(+), flag(+), god(+), god(+), trump president(−), patriots(+), god bless(+), parler(−), bidenharris2020(+), amendment 2nd(+), love god(+)

We distinguish terms uttered in profiles with compute positive (+), negative (−), and neutral (n) sentiment. We compute term skewness as the skewness of the positions of the profiles in which they appear, along dimensions δ_1, δ_2 & δ_3. We report the 20 terms with most extreme positive and negative skewness

Appendix B: skewness skewness of terms of profiles in space

See "Exploring political concepts in space using text profiles" section for a description of the extraction of the sentiment-signed terms with most extreme means (Table 4).

Appendix C: criteria defining binary labels

In addition to the keywords shown in Table 5, we rely on sentiment analysis of profile text to distinguish positive and negative mentions of keywords (using a pre-trained BERT base model for uncased words Devlin et al. 2018), assigning to each profile text

Table 5 Summary of the proposed issue partitions of users into minimal groups for mining spatial directions capable of classifying them

Partition	Values	Criteria	Users
Party	Republican	"republican(s)" AND ($\not=$)	7895
	Democrat	"democrat(s)" AND ($\not=$)	14481
Candidate	Trump sup.	("Trump" OR "MAGA") AND ($+$)	50 159
	Biden sup.	("Biden" OR "Build Back Better") AND ($+$)	10 229
Ideology	Conservative	"conservative" AND ($\not=$)	12331
	Liberal	"liberal" AND ($\not=$)	8917
Left/Right	Right-wing	"right" AND ($\not=$)	8462
	Left-wing	"left" AND ($\not=$)	3423
Black Lives Matter	Pro-BLM	("BLM" OR "black lives matter"] AND ($+$)	13350
	Anti-BLM	("BLM" OR "black lives matter") AND ($-$)	12017
Black/Blue	BlackLivesMatter	("BLM" OR "black lives matter"] AND ($-$)	16690
	BlueLivesMatter	"blue lives matter" AND ($+$)	240
Rural/Urban	Rural	(" rural" OR "small town" OR "heart land") AND ($\not=$)	2174
	Urban	" urban" AND ($\not=$)	3382
Religion	Christian	"christian" AND ($\not=$)	22735
	Atheist	"atheist" AND ($\not=$)	1081
Guns	Pro-Gun	"pro gun" OR "gunner" OR ("guns" AND ($\not=$))	605
	Anti-Gun	"anti gun" OR ("gun laws" OR "gun control" AND ($\not=$))	496
Communism	Anti-Communism	("communist" OR "communism") AND ($-$)	1160
	Pro-Communism	("communist" OR "communism") AND ($+$)	448
Liberal Lifestyle	Pro-Liberal LifeStyle	("gay" OR "feminist" OR "lgbt" OR "feminism") AND ($+$)	9659
	Anti-Liberal LifeStyle	("gay" OR "feminist" OR "lgbt" OR "feminism") AND ($-$)	9921
Libertarian/Welfare	Libertarian	"libertarian" AND ($\not=$)	3045
	Welfare	"welfare" AND ($\not=$)	681
Police	Pro-Police	"police" AND ($+$)	1811
	Anti-Police	("police" AND ($-$)) OR "defund the police" or "fuck the police"	1686
Military	Pro-Military	("army" OR "navy" OR "air force" OR "military") AND ($+$)	17721
	Anti-Military	("army" OR "navy" OR "air force" OR "military") AND ($-$)	8187
Patriotism	Patriot	"patriot"	27907
	Anti-Imperialism	"imperialism"	120
Local/Global	Local focus	"local" AND ($\not=$)	5532
	Global focus	"global" AND ($\not=$)	6283
Deepstate/Rationalism	DeepState	"deep state"	486
	Rationalist	"rationalist"	111
Entrepreneur	Entrepreneur	"entrepreneur"	13209
	Anti-capitalist	"capitalism"	1287
Liberty/Justice	Justice	"freedom" OR "liberty"	18829
	Freedom	"justice"	13041
Pronouns	Anti-Incl. Pronous	"pronouns" AND ($-$)	1195
	Incl. Pronous.	"he/him" OR "she/her" OR "they/them"	66218
International	International	"international" AND ($\not=$)	5957
	Globalization	"globalization"	84

Table 5 (continued)

Partition	Values	Criteria	Users
Individual/Institutional	Institutional	"official account"	976
	Individual	"opinions are my own" OR "own opinions" OR "views are mine"	14739
Science	Science	"science" AND $(\nearrow)$	14795
	AntiExperts	"experts"	59
Masks (COVID)	Pro-Mask	"wear a mask" OR "wearamask"	4247
	Anti-Mask	"mask mandate' OR "unmask" OR "stop wearing mask" OR "no mask" "anti.?mask' "mask hater" OR "burn your mask' OR "masks don't work" OR "masks off"	214

For each issue we identify two disjoint groups defined by queries of the Twitter profile text descriptions, including keywords (case insensitive, here all written in lowercase), and sentiment: positive (+), negative (−), and non-negative ($\nearrow$)

a sentiment from 1 (very negative) to 5 (very positive). We label text profiles as negative (−) if sentiment is equal to 1, and as positive (+) if sentiment is equal to 5. In Table 5 we also distinguish users whose profiles are not negative ($\nearrow$). This is needed, for example, to identify users that might use the word "republican" in their profiles, but in order to utter criticism (*e.g.*, "I hate republicans!").

Availability of data and materials
The analyses presented in this article include political preferences data of individuals. Due to privacy concerns, no data leading to identification is released with this article. A partial anonymized release can be found at https://github.com/pedroramaciotti/CORG in the form of a tutorial, along with reproducibility for some of the cases treated in "Discovering spatial directions of political tension" section. The code used for producing the computations of the homophily latent space of "Homophily network embedding in latent space" section is available at https://github.com/pedroramaciotti/LINATE.

Declarations

Competing interest
The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

References
Adams GD (1997) Abortion: evidence of an issue evolution. Am J Polit Sci 41(3):718–737

Ahler DJ, Broockman DE (2018) The delegate paradox: why polarized politicians can represent citizens best. J Polit 80(4):1117–1133. https://doi.org/10.1086/698755

Bafumi J, Herron MC (2010) Leapfrog representation and extremism: a study of American voters and their members in congress. Am Polit Sci Rev 104(3):519–542

Bailey MA (2001) Ideal point estimation with a small number of votes: a random effects approach. Polit Anal 9(3):192–210

Bailey MA (2003) The politics of the difficult: the role of public opinion in early cold war aid and trade policies. Legis Stud Q 28(2):147–178

Bailey MA (2007) Comparable preference estimates across time and institutions for the court, congress and presidency. Am J Polit Sci 51(3):433–448

Bakker R, Jolly S, Polk J (2012) Complexity in the European party space: exploring dimensionality with experts. Eur Union Polit 13(2):219–245

Bakker R, Hooghe L, Jolly S, Marks G, Polk J, Rovny J, Steenbergen M, Vachudova MA (2019) Chapel hill expert survey. Chapel Hill. www.chesdata.eu

Bakshy E, Messing S, Adamic LA (2015) Exposure to ideologically diverse news and opinion on facebook. Science 348(6239):1130–1132

Baldassarri D, Gelman A (2008) Partisans without constraint: political polarization and trends in American public opinion. Am J Sociol 114(2):408–446

Barberá P (2015) Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. Polit Anal 23:76–91

Barberá P, Rivero G (2015) Understanding the political representativeness of twitter users. Soc Sci Comput Rev 33(6):712–729

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

Baumann F, Lorenz-Spreen P, Sokolov IM, Starnini M (2020) Modeling echo chambers and polarization dynamics in social networks. Phys Rev Lett 124(4):048301

Benkler Y, Faris R, Roberts H (2018) Network propaganda: manipulation, disinformation, and radicalization in American politics. Oxford University Press, Oxford

Bond R, Messing S (2015) Quantifying social media's political space: estimating ideology from publicly revealed preferences on facebook. Am Polit Sci Rev 109:62–78

Bonica A (2018) Inferring roll-call scores from campaign contributions using supervised machine learning. Am J Polit Sci 62(4):830–848

Bramson A, Grim P, Singer DJ, Fisher S, Berger W, Sack G, Flocken C (2016) Disambiguation of social polarization concepts and measures. J Math Sociol 40(2):80–111

Broockman DE (2016) Approaches to studying policy representation. Legis Stud Q 41(1):181–215. https://doi.org/10.1111/lsq.12110

Campbell D (2023) Russia and the us press: the article the cjr didn't publish. bylinetimes.com

Clinton J, Jackman S, Rivers D (2004) The statistical analysis of roll call data. Am Polit Sci Rev 98(2):355–370

Cointet J-P, Cardon D, Mogoutov A, Ooghe-Tabanou B, Plique G, Ramaciotti Morales P (2021) Uncovering the structure of the French media ecosystem. arXiv preprint arXiv:2107.12073

Cointet J-P, Ramaciotti Morales P, Cardon D, Froio C, Mogoutov A, Ooghe-Tabanou B, Plique G (2021) What colours are the yellow vests? An ideological scaling of facebook groups. Statistique et Société

Converse PE (1964) The nature of belief systems in mass publics. In: Apter DE (ed) Ideology and discontent. University of Michigan Press, Michigan

Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805

Downs A (1957) An economic theory of political action in a democracy. J Polit Econ 65(2):135–150

Duclos J-Y, Esteban J, Ray D (2004) Polarization: concepts, measurement, estimation. Econometrica 72(6):1737–1772

Fowler A, Hill SJ, Lewis JB, Tausanovitch C, Vavreck L, Warshaw C (2022) Moderates. Am Polit Sci Rev. https://doi.org/10.1017/S0003055422000818

Frantzi K, Ananiadou S, Mima H (2000) Automatic recognition of multi-word terms: the c-value/nc-value method. Int J Digit Libr 3:115–130

Greenacre M (2017) Correspondence analysis in practice. CRC Press, Boca Raton

Grossman E, Sauger N (2019) Economic internationalization and the decline of the left–right dimension. Party Polit 25:36–49

Imai K, Lo J, Olmsted J (2016) Fast estimation of ideal points with massive data. Am Polit Sci Rev 110(4):631–656

Jensen J, Quinn D, Weymouth S (2017) Winners and losers in international trade: the effects on us presidential voting. Int Organ 71:1–35. https://doi.org/10.1017/S0020818317000194

Jost JT, Baldassarri DS, Druckman JN (2022) Cognitive-motivational mechanisms of political polarization in social-communicative contexts. Nat Rev Psychol 1(10):560–576

Kageura K, Umino B (1996) Methods of automatic term recognition: a review. Terminol Int J Theor Appl Issues Spec Commun 3(2):259–289

Lauderdale BE, Clark T (2012) The supreme court's many median justices. Am Polit Sci Rev 106(4):847–866

Lowe W (2008) Understanding wordscores. Polit Anal 16:356–371

Mani I, Zhang I (2003) knn approach to unbalanced data distributions: a case study involving information extraction. In: Proceedings of workshop on learning from imbalanced datasets, vol 126. ICML, pp 1–7

Martin A, Quinn K (2002) Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999. Polit Anal 10(2):134–153

Mason L (2015) "i disrespectfully agree": the differential effects of partisan sorting on social and issue polarization. Am J Polit Sci 59(1):128–145

Morales PR, Cointet J-P, Laborde J (2020) Your most telling friends: propagating latent ideological features on twitter using neighborhood coherence. In: 2020 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM). IEEE, pp 217–221

Morales PR, Lamarche-Perrin R, Fournier-S'Niehotta R, Poulain R, Tabourier L, Tarissan F (2021) Measuring diversity in heterogeneous information networks. Theor Comput Sci 859:80–115

Noel H (2013) Political ideologies and political parties in America. Cambridge University Press, Cambridge

O'Callaghan D, Greene D, Conway M, Carthy J, Cunningham P (2015) Down the (white) rabbit hole: the extreme right and online recommender systems. Soc Sci Comput Rev 33(4):459–478

Ojer J, Cárcamo D, Pastor-Satorras R, Starnini M (2023) Charting multidimensional ideological polarization across demographic groups in the united states. arXiv preprint arXiv:2311.06096

Peress M (2022) Large-scale ideal point estimation. Polit Anal 30(3):346–363

Poole KT, Rosenthal H (1985) A spatial model for legislative roll call analysis. Am J Polit Sci 357–384

Poole K, Rosenthal H (1997) Congress: a political-economic history of roll call voting. Oxford University Press, Oxford

Ramaciotti Morales P, Cointet J-P, Froio C (2022) Posters and protesters. J Comput Soc Sci 5:119–1157

Ramaciotti Morales P (2023) Multidimensional online American politics: mining emergent social cleavages in social graphs. In: Complex networks and their applications XI: proceedings of the eleventh international conference on complex networks and their applications: COMPLEX NETWORKS 2022–Volume 1. Springer, pp 176–189

Ramaciotti Morales P, Cointet J-P, Benbouzid B, Cardon D, Froio C, Metin OF, Ooghe B, Plique G (2021) Atlas multi-plateformes d'un mouvement social: Le cas des gilets jaunes. Statistique et Société

Ramaciotti Morales P, Cointet J-P, Muñoz Zolotoochin G (2021) Unfolding the dimensionality structure of social networks in ideological embeddings. In: Proceedings of the 2021 IEEE/ACM international conference on advances in social networks analysis and mining, pp 333–338

Ramaciotti Morales P, Cointet J-P, Muñoz Zolotoochin G, Fernández Peralta A, Iñiguez G, Pournaki A (2022) Inferring attitudinal spaces in social networks

Ramaciotti Morales P, Cointet J-P (2021) Auditing the effect of social network recommendations on polarization in geometrical ideological spaces. In: 15th acm conference on recommender systems, RecSys' 21

Ramaciotti Morales P, Muñoz Zolotoochin, G (2022) Measuring the accuracy of social network ideological embeddings using language models. In: 2022 international conference on information technology & systems (ICITS22)

Ramaciotti Morales P, Tabourier L, Ung S, Prieur C (2019) Role of the website structure in the diversity of browsing behaviors. In: Proceedings of the 30th ACM conference on hypertext and social media, pp 133–142

Riker WH (1982) The two-party system and duverger's law: an essay on the history of political science. Am Polit Sci Rev 76(4):753–766

Rivers D (2003) Identification of multidimensional ItemResponse models. Stanford University, Stanford

Sokal RR (1958) A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38:1409–1438

Uscinski JE, Enders AM, Seelig MI, Klofstad CA, Funchion JR, Everett C, Wuchty S, Premaratne K, Murthi MN (2021) American politics in two dimensions: partisan and ideological identities versus anti-establishment orientations. Am J Polit Sci 65(4):877–895. https://doi.org/10.1111/ajps.12616

Uscinski JE, Enders AM, Seelig MI, Klofstad CA, Funchion JR, Everett C, Wuchty S, Premaratne K, Murthi MN (2021) American politics in two dimensions: partisan and ideological identities versus anti-establishment orientations. Am J Polit Sci 65(4):877–895

Xiao Z, Song W, Xu H, Ren Z, Sun Y (2020) Timme: Twitter ideology-detection via multi-task multi-relational embedding. In: Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery and data mining, pp 2258–2268

Publisher's Note