Polarization as Probabilistic Dependence

Luke Thorburn, Maria Polukarov and Carmine Ventre

Abstract

Many human communities are polarized to some degree, so it is important that preference and knowledge aggregation methods are robust to the presence of polarization. Evaluating this robustness requires a good measure of polarization, but existing metrics for quantifying the polarization of formal models of a population often diverge from intuition. In particular, many existing metrics imply that polarization is minimized only when there is perfect unanimity or homogeneity. Building on work from opinion dynamics, political science, and conflict studies, we argue that polarization is more accurately thought of as the degree of probabilistic dependence between the attributes of individuals in a population. We define a measure of *sortedness* which captures this intuition based on the information-theoretic concept of mutual information. We then show that this measure (i) allows for populations to be simultaneously diverse and minimally polarized, and (ii) is in a formal sense orthogonal to public opinion, meaning that polarization can, in principle, be reduced without disadvantaging existing political stakeholder groups.

1 Introduction

There is evidence of increasing political polarization in many societies [3], so it is important to understand how social choice functions behave in polarized contexts and to be able to design mechanisms that are robust to the presence of polarization. Valid measurement of polarization is also important in political science and in the context of emerging proposals to use algorithmic systems such as recommender systems on social media as a means of reducing polarization in society and promoting social cohesion [43].

Polarization can be conceptualized and measured in many distinct ways. Conceptually, polarization is often decomposed into issue polarization (disagreement over policy issues) and affective polarization (emotional distance between individuals). Such conceptualizations are usually measured with survey instruments such as feeling thermometers [21]. A parallel body of work focuses on formal definitions of polarization intended to represent the degree to which a formal model of affiliation or preference within a population is polarized [27, 5].

What is the opposite of polarization? It is presumably desirable for our measures of polarization to remain valid as polarization decreases. Particularly in settings where we are trying to "minimize polarization", it is important that the configurations of human relationships at which our measure of polarization is minimized are consistent with our values and ideals of a healthy public sphere. However, to minimize polarization according to most of the existing formal measures would correspond to complete uniformity, homogeneity, or consensus among a population (Section 2.1). This is inconsistent with freedom of expression and pluralism, and raises significant concerns about manipulation or influence [42].

An alternative understanding of polarization, common in literature on political science, is that of partisan sorting. In this conceptualization, polarization is not a property of individual people (a single person cannot be said to "be polarized"), nor is it a property of the distribution of positions on a single issue (the debate over, say, climate change action cannot in isolation be said to "be polarized"). Instead, polarization is a property of the dependency structure between the positions taken by individuals, or more generally between individual attributes that include opinions, preferences, identities, and demographics. The more these attributes tend to go together—and in particular, the greater the extent to which similar individuals "sort" themselves into the two major parties—the greater the degree of polarization. However, the measures used to quantify sorting are usually developed *ad hoc* for use with a particular dataset, and often only apply in the bivariate case (party affiliation, plus one additional attribute).

1.1 Contribution

In this paper, we formalize a measure of polarization called *sortedness* which characterizes polarization as the degree of probabilistic dependence between the attributes of individuals in a population. While political polarization as a type of sorting has been studied previously, to our knowledge this is the first time that these insights have been distilled into a general information-theoretic measure that is applicable regardless of the number of variables with which individuals are modelled.

We show that the opposite of polarization implied by the sortedness measure is not unanimity (as is the case with most existing measures), but perfect independence of the attributes of each individual. We then define formally what it means for a measure of polarization to be "orthogonal to public opinion", and prove that the sortedness measure satisfies this property. This is important because it means that, in principle, polarization can be reduced without disadvantaging existing political stakeholder groups.

1.2 Outline

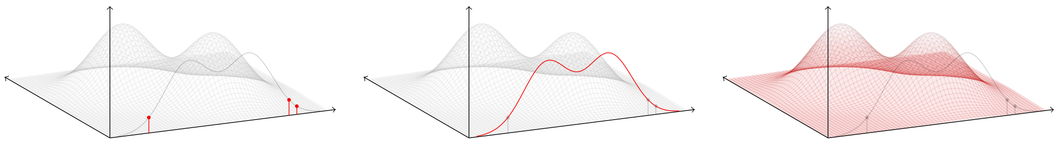
Section 2 contains a survey of existing polarization measures and related work, along with a brief introduction to relevant concepts from information theory. Our sortedness measure of polarization is defined in Section 3. In Section 4 we describe the extrema of the sortedness measure, formalize the notion of orthogonality and prove that sortedness satisfies this property. Section 5 contains a discussion of synergies with previous work, possible limitations of the sortedness measure, and open questions.

2 Background

2.1 Related Work

There are many existing measures of polarization in formal models of preference or opinion in a population. We maintain a living review of these at `https://bridging.systems/metrics/`, which at the time of writing contains 25 different measures [2, 4, 5, 8, 10, 11, 16, 18, 20, 19, 25, 26, 27, 30, 31, 39, 41, 45, 46, 58, 59]. These measures have variously been defined in the context of categorical, spacial, or graphical models; apply to individuals, sub-populations, or entire populations; may or may not require additional structure (such as clusters) to be known; and many are "unsafe" to optimize, in the sense that to minimize polarization according to these measures would mean creating perfect unanimity within the population.

Some works have proposed axioms that a polarization measure should satisfy, though primarily in the context of economic "polarization" or inequality, rather than political polarization. Of these, a series of papers [14, 13, 17, 28, 37, 44, 55] centering on the axioms proposed by Esteban and Ray [17] are the most prominent. However, for the most part these axioms apply only to univariate distributions, and some of these axioms are not widely accepted [1]. There are some proposals for multivariate and information theoretic measures of economic inequality [22, 33, 51, 52]. Most of these apply only to univariate distributions, or to attributes that are strictly positive, or are tailored to be measures of inequality, rather than polarization. The closest measure to ours is perhaps that of Gigliarano and Mosler



(a) "extremism" (b) "polarization" (c) "sorting"

Figure 1: Visual intuition for a model of a population as a joint probability distribution of the attributes of an individual selected uniformly at random. In this model, (a) "extremism" is a property of individual samples, (b) "polarization" is a property of a marginal distribution, and (c) "sorting" is a property of the dependency structure or copula.

[22], who propose an information-theoretic measure of multivariate polarization designed to capture the idea that polarization is the "presence of groups which are internally homogeneous, externally heterogeneous, and of similar size". Their class of polarization measures thus consists of aggregations of three separate sub-measures that are intended to capture the three separate aspects of (their concept of) polarization.

There is a considerable body of work which articulates the view that polarization is a kind of population-level sorting, which is the version of polarization we formalize in this paper. (We note that existing literature varies as to whether sorting is considered a kind of polarization, as we treat it here, or a conceptually distinct concept, as illustrated in Figure 1.) Studies on conflict prediction have shown that the greater the extent to which ethnic identity in a population is cross-cut by religious identity, socioeconomic class or geographic region, the lower the probability of civil war onset [23, 47, 48]. These studies all use inherently bivariate measures of cross-cuttingness, or aggregations of pairwise bivariate measures. Studies on US politics have shown that partisan sorting—measured as the difference between the association of various political identities with support for each of the major parties—is increasing and predictive of emotional responses to political messaging [32, 34, 35]. Again in the US context, studies have shown increased sorting of beliefs into correlated clusters [12], and high levels of geographic sorting along party lines [6]. Finally, one simulation study has suggested that increasing exposure to a national or global news in digital media environments may increase partisan sorting, where sorting is measured using a formula that captures the proportion of attributes shared between agents [53].

2.2 Notation

This paper makes use of concepts from information theory (see Cover and Thomas [9] for a standard introductory text). Here, we briefly define some core concepts used in Sections 3 and 4.

Definition 1 (entropy of discrete random variables)**.** *Let $n \geq 1$ be an integer and $X_1, \dots, X_n$ be real-valued, discrete random variables which respectively take values in $\mathcal{X}_1, \dots, \mathcal{X}_n$ and have joint probability mass function $p(x_1, \dots, x_n) = \mathbf{P}(X_1 = x_1, \dots, X_n = x_n)$. If*

$$H(X_1, \dots, X_n) = -\sum_{x_1 \in \mathcal{X}_1} \cdots \sum_{x_n \in \mathcal{X}_n} p(x_1, \dots, x_n) \log p(x_1, \dots, x_n)$$

exists and is finite, it is called the joint entropy of $X_1, \dots, X_n$. In the case $n = 1$, it is simply the entropy.

Definition 2 (entropy of continuous random variables)**.** *Let $n \geq 1$ be an integer and $X_1, \dots, X_n$ be real-valued, continuous random variables which respectively take values in*

$\mathcal{X}_1, \ldots, \mathcal{X}_n$ *and have joint density function* $f(x_1, \ldots, x_n)$*. If*

$$H(X_1, \ldots, X_n) = -\int_{x_1 \in \mathcal{X}_1} \cdots \int_{x_n \in \mathcal{X}_n} f(x_1, \ldots, x_n) \log f(x_1, \ldots, x_n)\ dx_1 \ldots dx_n$$

exists and is finite, it is called the joint entropy of $X_1, \ldots, X_n$*. In the case* $n = 1$*, it is simply the entropy.*

Intuitively, entropy can be thought of as a measure of the uncertainty, information or surprise associated with a set of random variables. The two definitions above, for the discrete and continuous cases, are analogous. However, because they must be defined separately, we will throughout this paper often define, discuss or prove results separately in the discrete and continuous cases.

Definition 3 (Kullback-Leibler divergence)**.** *Let* p *and* q *be probability mass functions (in the discrete case) or two density functions (in the continuous case), characterizing distributions over a common sample space. Further, let* $X \sim p$*. If*

$$D_{\mathsf{KL}}(p \| q) = \mathbf{E}_p \left[\log \frac{p(X)}{q(X)} \right]$$

exists and is finite, it is called the Kullback-Leibler (KL) divergence between p *and* q*.*

Kullback-Leibler divergence is a measure of the distance between two distributions, though it is not symmetric so is not a true metric. The expectation in the definition above can be expanded as either a sum or integral in the discrete and continuous cases, respectively. In general, the two distributions being compared can be multivariate.

3 A Measure of Multivariate Sortedness

We begin with an example to build intuition. Consider a stylized world in which people have only two attributes, *politics* and *wealth*. Each of these attributes can take only two values. People are politically either `LEFT` or `RIGHT`, and either `RICH` or `POOR`. For an individual chosen uniformly at random, let the random variable X denote their politics and Y denote their wealth.

Let $p \in [0, 1]$ be the proportion of people that are politically `LEFT`, and $q \in [0, 1]$ be the proportion of people who are `RICH`. Figure 2 depicts three possible contingency tables, corresponding to three possible populations and three dependency structures between X and Y. Speaking intuitively, populations (a) and (b) are maximally polarized, according to the sorting view of polarization. Any two individuals will either be identical, or maximally different. In contrast, population (c) is maximally *un*polarized. The random variables X and Y are independent: knowing someone's politics provides no information about their wealth, and vice versa.

	`RICH`	`POOR`	
`LEFT`	$p = q$	0	p
`RIGHT`	0	$1-p = 1-q$	$1-p$
	q	$1-q$	1

(a)

	`RICH`	`POOR`	
`LEFT`	0	$p = 1-q$	p
`RIGHT`	$1-p = q$	0	$1-p$
	q	$1-q$	1

(b)

	`RICH`	`POOR`	
`LEFT`	pq	$p \cdot (1-q)$	p
`RIGHT`	$(1-p) \cdot q$	$(1-p) \cdot (1-q)$	$1-p$
	q	$1-q$	1

(c)

Figure 2: Dependency structures between politics and wealth in three example populations. Populations (a) and (b) are maximally polarized, population (c) is maximally *un*polarized.

3.1 Definition

To capture this intuition quantitatively we use mutual information, an information-theoretic measure of the degree of dependence between two random variables [9, Ch. 2 and 8].

Definition 4 (mutual information)**.** *Let X and Y be real-valued random variables (both discrete or both continuous), $P_{X,Y}$ be their joint distribution, P_X, P_Y their marginal distributions, and $P_X \otimes P_Y$ the product of their marginal distributions. If $H(X)$, $H(Y)$, $H(X,Y)$ all exist and are finite, then*

$$\begin{aligned} I(X;Y) &= D_{\mathsf{KL}}(P_{X,Y} \| P_X \otimes P_Y) \\ &= H(X) + H(Y) - H(X,Y) \end{aligned}$$

is called the mutual information between X and Y.

There are a few ways to interpret this value. The mutual information $I(X;Y)$ can be thought of as the amount of information that X and Y share, or the degree to which knowing one of these variables reduces uncertainty about the other. Equivalently, the formulation in terms of Kullback-Leibler divergence suggests that $I(X;Y)$ is the additional cost of encoding X and Y using an encoding scheme optimized for independent random variables, when in reality they are at least partially dependent.

Mutual information (i) is always non-negative, (ii) equals zero if and only if X and Y are independent, and (iii) is bounded above by $H(X)+H(Y)-\max\{H(X), H(Y)\}$. Interpreted as a measure of polarization, $I(X;Y) = 0$ indicates no polarization, and

$$I(X;Y) = H(X) + H(Y) - \max\{H(X), H(Y)\}$$

indicates maximal polarization.

Mutual information appears to capture intuition for the sorting view of polarization in this simple model where people have only two attributes, but in reality people are much more high-dimensional. There are a number of proposed generalizations of mutual information to more than two random variables, including *interaction information* [36, 49], *total correlation* [56], and *dual total correlation* [24]. Of these, we believe total correlation is the easiest to interpret in the context of polarization, and the most computationally simple.

Definition 5 (total correlation)**.** *Let n be a positive integer, and $X_1, \ldots, X_n$ be real valued random variables (all discrete, or all continuous). If the entropies $H(X_1), \ldots, H(X_n)$ and $H(X_1, \ldots, H(X_n)$ all exist and are finite, then*

$$\begin{aligned} C(X_1; \ldots; X_n) &= D_{\mathsf{KL}}(P_{X_1,\ldots,X_n} \| P_{X_1} \otimes \cdots \otimes P_{X_n}) \\ &= \left[\sum_{i=1}^{n} H(X_i)\right] - H(X_1, \ldots, X_n) \end{aligned}$$

is called the total correlation of the variables $X_1, \ldots, X_n$.

Despite its name, total correlation is a measure of probabilistic dependence in general, not only linear correlation. Many of the properties for two-variable mutual correlation transfer analogously to the multivariate case. In particular, total correlation (i) is always non-negative, (ii) equals zero if and only if $X_1, \ldots, X_n$ are completely independent, and (iii) is bounded above by $[\sum_{i=1}^{n} H(X_i)] - \max\{H(X_1), \ldots, H(X_n)\}$. As a measure of polarization, $C(X_1, \ldots, X_n) = 0$ indicates no polarization, and

$$C(X_1, \ldots, X_n) = \left[\sum_{i=1}^{n} H(X_i)\right] - \max\{H(X_1), \ldots, H(X_n)\}$$

indicates maximal polarization.

Thus defined, the lower bound of total correlation is always zero, but the upper bound depends on the marginal distributions of $X_1, \ldots, X_n$. To measure polarization on the same interval regardless of the particular marginal distributions, we can revise the formula to give a value between 0 and 1.

Definition 6 (sortedness)**.** *Let n be a positive integer, and $X_1, \ldots, X_n$ be real valued random variables (all discrete, or all continuous). If the entropies $H(X_1), \ldots, H(X_n)$ and $H(X_1, \ldots, X_n)$ all exist and are finite, then*

$$S(X_1; \ldots; X_n) = 1 - \frac{H(X_1, \ldots, X_n) - \max\{H(X_1), \ldots, H(X_n)\}}{[\sum_{i=1}^{n} H(X_i)] - \max\{H(X_1), \ldots, H(X_n)\}}$$

we call the sortedness of the variables $X_1, \ldots, X_n$.

We propose the term *sortedness* for this specific measure to distinguish it from the qualitative notion of polarization and other formal measures. Sortedness equals 0 if and only if the variables are independent (that is, no sorting or polarization), and equals 1 if the variables are maximally dependent given the constraints imposed by the marginal distributions (that is, maximal sorting or polarization).

3.2 Application

We now briefly describe how sortedness can be calculated in the context of three common models of preferences, affinities, or affiliations in a population.

Categorical models In categorical or set-based models of a population, individuals are represented by the groups they are a part of. For example, these groups might include ethnic communities, political parties, religions, geographic neighbourhoods, friendship cliques, or employers. These groups can all overlap to varying degrees.

For the purposes of estimating sortedness, membership of these groups can be modelled as n Bernoulli random variables, where n is the total number of groups (of any variety). These discrete random variables are indicators for the event that an individual selected uniformly at random from the population belongs to each group. Intuitively, the sortedness of these indicator variables is an measure of the overall degree of overlap or redundancy among the set of groups considered, defined for an arbitrary number of groups.

Spacial models In spacial models of a population, individuals are represented as points in a metric space, representing a latent "opinion space" or "preference space". Alternatives over which individuals have preferences may also be represented as points in the same space, in which case the distance between an individual and each of the alternatives determines their ordinal preferences over the alternatives [15].

For the purposes of estimating sortedness, each of n dimensions in a spatial model can be interpreted as specifying an attribute of the individuals, and the coordinates of an individual chosen uniformly at random modelled as random variables $(X_1, \ldots, X_n)$. In empirical contexts or simulations, these distributions may be modelled as discrete (e.g., via binning) or continuous (e.g., via kernel density estimation).

Graphical models In graphical models of a population, individuals are represented as vertices in an (abstract, mathematical) graph. There may also be vertices that represent other entities, such as groups or alternatives with which individuals connect or interact. Optionally, these edges may be weighted.

For the purposes of estimating sortedness, the graph can be represented as an adjacency matrix, with one row per individual and one column for every vertex in the graph. The elements of the matrix represent the weight of the edge between the corresponding row and column vertices in the graph (which may be zero, if the edge doesn't exist). Depending on whether the edge weights are continuous or binary, the row vectors of this matrix can be interpreted as samples from a spatial or categorical model, as described above, and sortedness calculated in the same way.

4 Properties

In this section, we discuss two properties of sortedness as defined in Definition 6. These are (i) the joint distributions for which it is maximized and minimized, and (ii) the extent to which it can be said, as a measure of polarization, to be "orthogonal to public opinion".

4.1 Extrema

4.1.1 Minima

In part, we motivated sortedness as a measure of polarization with the fact that most existing formal measures of polarization are minimized only when the underlying population of individuals is completely unanimous or homogeneous. This significantly limits their validity: if minimizing polarization according to a given measure means removing all diversity from a population, we believe that measure cannot be capturing the structures of human relationships that we collectively care about. Most people would find the idea of eliminating all variation to "reduce polarization" deeply unethical.

To avoid this pitfall, we need a measure that implies an acceptable "opposite of polarization". In this respect, sortedness appears promising. The extreme opposite of polarization—as implied by the sortedness measure—is a world in which people's preferences or beliefs are probabilistically independent of one another.

Lemma 1 (minimizing sortedness)**.** $S(X_1, \dots, X_n) = 0$ *if and only if* $X_1, \dots, X_n$ *are independent.*

Proof. By definition, $S(X_1, \dots, X_n) = 0$ if and only if

$$H(X_1, \dots, X_n) = \sum_{i=1}^{n} H(X_i).$$

It is a standard result of information theory that this occurs if and only if $X_1, \dots, X_n$ are completely independent. See [9, Theorem 2.6.6] for the discrete case and [9, Theorem 8.6.2] for the continuous case. □

For any given marginal distributions of $X_1, \dots, X_n$, there is thus a unique joint distribution for which sortedness is minimized, corresponding to complete independence.

There are reasons why we would not want to advocate for such an extreme reality. The abstract ideal of perfectly independent beliefs would mean we have little in common with those in our families and communities, and perhaps formalizes a parochially Western value of individualism. Nonetheless, we think that as a goal it is considerably more acceptable than complete uniformity, and it is plausible that partially increasing the independence of beliefs would be a beneficial change in many conflict scenarios. The goal of independence is also consistent with formal results on epistemic democracy and wisdom of the crowds (which often work best under the ideal of independent voters [54]), and with emerging work on the benefits of *correlational discounting* in quadratic voting and related aggregation rules [57, 38].

4.1.2 Maxima

> All happy families resemble one another, but each unhappy family is unhappy in its own way. [50]

In contrast to the minimum point, which is unique for a specified set of marginal distributions, there are in general multiple joint distributions with the same marginals for which sortedness is maximized. All correspond to there being a high degree of dependency between the random variables $X_1, \ldots, X_n$, but the particular structure of this dependency may vary.

In the continuous case, one way in which sortedness can be maximized is for the random variables $X_1, \ldots, X_n$ to be comonotonic.

Lemma 2 (maximizing sortedness in the continuous case)**.** *Let $X_1, \ldots, X_n$ be continuous, real-valued random variables. If $X_1, \ldots, X_n$ are comonotonic, then $S(X_1, \ldots, X_n) = 1$.*

Proof. By definition, $S(X_1, \ldots, X_n) = 1$ if and only if

$$H(X_1, \ldots, X_n) = \max\{H(X_1), \ldots, H(X_n)\}.$$

Assume $X_1, \ldots, X_n$ are comonotonic. This means that every variable X_i is a monotonic transformation of each of the others, and knowing the value of any X_i absolutely determines the values of the others. Thus, each variable X_i contains the same amount of information (and hence has the same entropy), so $\max\{H(X_1), \ldots, H(X_n)\} = \max\{H(X_1)\} = H(X_1)$.

It remains to show that $H(X_1, \ldots, X_n) = H(X_1)$. Using the chain rule for differential entropy [9, Theorem 8.6.2], we see that

$$\begin{aligned} H(X_1, \ldots, X_n) &= \sum_{i=1}^{n} H(X_i \mid X_1, \ldots, X_{i-1}) \\ &= H(X_1) + \sum_{i=2}^{n} \underbrace{H(X_i \mid X_1, \ldots, X_{i-1})}_{\substack{=0\text{, because once you know} \\ \text{the value of } X_1 \text{you gain} \\ \text{no more information from} \\ \text{observing the other } X_i}} = H(X_1) \end{aligned}$$

□

The discrete case is more complicated. For discrete random variables $X_1, \ldots, X_n$ with fixed marginal distributions, there may in general not exist a joint distribution such that $S(X_1; \ldots; X_n) = 1$. To see this, consider $S(X; Y)$ where $X \sim \text{Bernoulli}(p)$ and $Y \sim \text{Bernoulli}(q)$ for some $p, q \in (0, 1)$ and $p > q$. Even in the case where X and Y are comonotonic, there will be three potential outcomes for the random vector (X, Y) (namely $(0, 0)$, $(1, 0)$, or $(1, 1)$), but only two potential outcomes for each of the marginal variables X and Y. Thus $H(X, Y) > \max\{H(X), H(Y)\}$, and $S(X; Y) < 1$.

A tight upper bound for sortedness $S(X_1; \ldots, X_n)$ in terms of the marginal distributions of discrete random variables $X_1, \ldots, X_n$ is beyond the scope of this paper. But the main takeaway is this: according to the sortedness measure, all perfectly unpolarized societies resemble one another, but each highly polarized society is polarized in its own way.

4.2 Orthogonality

Increasingly there are proposals for large-scale interventions to reduce political polarization, such as by modifying recommender systems in large news and social media platforms [43]. Because of ethical concerns regarding manipulation [42], it would ideally be possible reduce polarization without influencing people. It may also be easier to garner a mandate for such large-scale interventions if it could be guaranteed that the intervention would not disadvantage existing political stakeholder groups.

The idea of reducing polarization without influencing people *as individuals* is perhaps a non sequitur. However, we show below that there is a sense in which it is possible to reduce polarization without influencing people *collectively.* More precisely, if polarization is a property of the dependency structure between attributes of individuals—as in the case of sortedness—it is theoretically possible to reduce polarization without changing the marginal distribution of those attributes. If those attributes represent opinions, a measure of polarization could thus be said to be "orthogonal to public opinion".

This notion of orthogonality can be defined formally as follows.

Definition 7 (orthogonality)**.** *Let n be a positive integer, $X_1, \ldots, X_n$ be real-valued random variables with joint distribution $P \in \mathcal{P}$ and corresponding marginal distributions $P_1, \ldots, P_n$ (all discrete, or all continuous). The set $\mathcal{P}$ is the set of all possible joint distributions on n variables. Let $f : \mathcal{P} \to [0, 1]$ be a function, and $s = f(P)$ (the current value of the function). The function f satisfies the orthogonality axiom if, for any target value $s' \in [0, 1]$, there exists $P' \in \mathcal{P}$ such that (i) $f(P') = s'$ and (ii) P' has the same marginal distributions as P.*

The sortedness measure of polarization satisfies the orthogonality axiom in the continuous case, and a one-sided version of orthogonality in the discrete case. We prove these results below.

Theorem 1. *When applied to continuous random variables, the measure S satisfies the orthogonality axiom.*

Proof. Assume $X_1, \ldots, X_n$ are continuous random variables with specified (but arbitrary) marginal distributions, and $s \in [0, 1]$. We must show that there exists a joint distribution for $X_1, \ldots, X_n$ such that $S(X_1; \ldots; X_n) = s$. For brevity, we will write $S(P) := S(X_1; \ldots; X_n)$, where P is a joint distribution for $X_1, \ldots, X_n$.

The general approach is as follows. We will define a continuous family of distributions functions P_t for $t \in [0, 1]$ such that $S(P_0) = 0$ and $S(P_1) = 1$. In other words P_t is a joint distribution over $X_1, \ldots, X_n$ and, as a function of t, is able to continuously interpolate between maximally and minimally sorted distributions. By applying the intermediate value theorem, it then follows immediately that there exists some value of $t \in [0, 1]$ such that the joint distribution P_t satisfies $S(P_t) = s$.

It remains to construct this function P_t. To do so, we will use the notion of a copula from probability theory. Intuitively, copulas are a tool by which to decouple the dependency structure of a set of random variables from their marginal distributions. Concretely, a copula C is a multivariate cumulative distribution function for which all the marginal distributions are uniform on the interval $[0, 1]$. If $(U_1, \ldots, U_n) \sim C$, and $F_1, \ldots, F_n$ are arbitrary continuous, univariate distribution functions, then

$$(X_1, \ldots, X_n) = \left(F_1^{-1}(U_1), \ldots, F_n^{-1}(U_n)\right)$$

is a random vector that has the marginal distributions $F_1, \ldots, F_n$ and the dependency structure specified by C. The uniform random variables $U_1, \ldots, U_n$ can be interpreted as the quantiles of $X_1, \ldots, X_n$. See Nelsen [40] for a more comprehensive introduction to copula theory.

Using the results from Lemmas 1 and 2, we construct P_t as follows. Let

$$P_t(x_1, \ldots, x_n) = \begin{cases} C^{\perp}(F_1(x_1), \ldots, F_n(x_n)), & t = 0 \quad \text{(independent)} \\ \left[(1-t)C^{\perp} + tC^{\diagup}\right](F_1(x_1), \ldots, F_n(x_n)), & t \in (0, 1) \quad \text{(mixture)} \\ C^{\diagup}(F_1(x_1), \ldots, F_n(x_n)), & t = 1 \quad \text{(comonotonic)}, \end{cases}$$

where $C^{\perp}(u_1, \ldots, u_n) = \prod_{i=1}^{n} u_i$ is the independence copula and $C^{\nearrow}(u_1, \ldots, u_n) = \min\{u_1, \ldots, u_n\}$ is the comonotonicity copula. For $t \in (0, 1)$, the dependence structure of P_t is thus defined using a mixture of the independence and comonotonicity copulas (mixtures of copulas are still copulas). Finally, use P_t together with the intermediate value theorem, as described in the second paragraph of this proof, to see that there must be a value of t such that $S(P_t) = s$. □

As described in Section 4.1.2, if $X_1, \ldots, X_n$ are all discrete there may be some values of $s \in [0, 1]$ for which the marginal distributions $F_1, \ldots, F_n$ do not admit a dependency structure consistent with $S(X_1; \ldots; X_n) = s$. For this reason, sortedness does not (in general) fulfil the orthogonality axiom when applied to discrete random variables. However, it does satisfy a weaker notion of one-sided orthogonality.

Definition 8 (one-sided orthogonality)**.** *Let n be a positive integer, $X_1, \ldots, X_n$ be real-valued random variables with joint distribution $P \in \mathcal{P}$ and corresponding marginal distributions $P_1, \ldots, P_n$ (all discrete, or all continuous). Let $f : \mathcal{P} \to [0, 1]$ be a function, and $s = f(P)$ (the current value of the function). The function f satisfies the one-sided orthogonality axiom if, for any target value $s' \in [0, s]$, there exists $P' \in \mathcal{P}$ such that (i) $f(P') = s'$ and (ii) P' has the same marginal distributions as P.*

We note that one-sided orthogonality, rather than (full) orthogonality, may be "enough" in many cases. As with orthogonality, if a polarization measure is one-sided orthogonal, then we can in principle reduce polarization without influencing public opinion. The one-sidedness simply means it may not be possible to *increase* polarization with the same guarantee.

Theorem 2. *When applied to discrete random variables, the measure S satisfies the one-sided orthogonality axiom.*

Proof. The proof is very similar to that of Theorem 1. Assume $X_1, \ldots, X_n$ are discrete random variables with initial joint distribution P. Consider the marginal distributions to be fixed. Let $s = S(X_1, \ldots, X_n)$, the initial sortedness of the random variables, and $s' \in [1, s]$. Further, for $t \in [0, 1]$, define a family of distribution functions

$$P_t(x_1, \ldots, x_n) = \begin{cases} C^{\perp}(F_1(x_1), \ldots, F_n(x_n)), & t = 0 \quad \text{(independent)} \\ \left[(1-t)C^{\perp} + tC^0\right](F_1(x_1), \ldots, F_n(x_n)), & t \in (0, 1) \quad \text{(mixture)} \\ C^0(F_1(x_1), \ldots, F_n(x_n)), & t = 1 \quad \text{(initial)}, \end{cases}$$

where $C^{\perp}(u_1, \ldots, u_n) = \prod_{i=1}^{n} u_i$ is the independence copula and

$$C^0(u_1, \ldots, u_n) = F\left(F_1^{(-1)}(u_1), \ldots, F_n^{(-1)}(u_n)\right),$$

where $F_i^{(-1)}$ denotes the generalized inverse of F_i such that C^0 is a copula consistent with the initial joint distribution P. For $t \in (0, 1)$, the dependence structure of P_t is thus defined using a mixture of the independence copula and a copula consistent with the initial joint distribution (in the discrete case, there may be multiple such copulas). Finally, use P_t together with the intermediate value theorem to see that there must be a value of t such that $S(P_t) = s'$. □

5 Conclusions and Future Work

To design aggregation methods that are robust to the presence of polarization, it is important that we are able to measure it well. Distilling recent literature on polarization, we argue that viewing polarization as probabilistic dependence between the attributes of individuals would be consistent with many empirical results and capture most of the key intuitions about polarization emerging from political science and conflict studies.

We gave a specific information-theoretic formalization of this type of polarization, called sortedness, which is general enough to apply to several types of formal model and is able to capture polarization across an arbitrary number of individual attributes. The sorting view of polarization is compatible with several ideas about the nature of polarization. These include the claims that polarization is

- **a kind of fracturing or clustering** (one way in which variables can become more dependent is for their joint distribution to resemble a mixture of distributions with disjoint supports);
- **a reduction in the dimensionality of individual differences** [29] (one way in which variables can become more dependent is for them to become increasingly comonotonic or countermonotonic);
- **a bifurcation in the expected degree of agreement** (the greater the dependence between issue positions, the greater the extent to which you will be able to predict someone's position on every issue given their position on just one);
- **the politicization of previously non-political issues** [53] (one way for variables to become more dependent is for them to depend increasingly on party affiliation);

We showed that the "opposite of polarization" implied by sortedness is a situation in which all attributes are completely independent. We also defined what it means for a functional on a joint probability distribution to be orthogonal to the marginal distributions, and this is used to show that the sortedness measure is "orthogonal to public opinion"—meaning that polarization can in principle be reduced without disadvantaging existing political stakeholder groups.

For clarity, we emphasize that we are *not* advocating that the marginal distributions should stay fixed: both forced stasis and manipulated change are likely unethical. But we may want to design algorithmic systems (e.g. recommender systems [43]) that aim to reduce polarization without manipulating public opinion—or at least without having an incentive to manipulate [7]—and to do this we may need a way of measuring polarization that is orthogonal to public opinion.

We also note that sortedness, as with any measure, only measures polarization among the variables used to compute it. This makes it contingent on the variables included in the joint distribution, and studies that use sortedness to quantify "overall societal polarization" would need to justify this by including a sufficiently comprehensive set of variables.

Future Work There are many open questions which merit further study, the first of which are theoretical. For example, we would like a measure of polarization that is comparable across contexts. To what extent is this true of sortedness? Is it (or can it be modified to be) comparable across joint distributions with different marginal distributions, or different numbers of variables? Is it invariant to linear or, more generally, monotonic transformations of the marginal distributions? (E.g., we don't want our measure of polarization to depend on whether we encode left-right political affiliation from $[-1, 1]$ or $[1, -1]$.) It may also be useful to generalize the definition to all sets of random variables (not just those that are homogeneously continuous or discrete) using the most general definition of mutual information [9, Equation 8.54], and to characterize a general class of measures that satisfy our definition of orthogonality.

The second set of questions are more practical, relating to the use of this measure. How computationally complex is it to compute? How can we estimate it from incomplete or inferred models of a population? While conceptually appealing, to what extent is it psychologically and empirically valid? For example, the orthogonality result is valid if people "generate their attributes on the fly". This might effectively be true of some opinions, or if the population is large enough for this model to be a good approximation. But if the attributes or opinions of each individual are fixed, this introduces some discreteness to the problem, which may limit the degree to which orthogonality-like properties can be proven. Finally, while orthogonality holds in theory, to what extent is this achievable in practice? This paper has focused on the ontological question of how polarization should be defined, not the mechanics of how it evolves or can be influenced over time.

Acknowledgements

Luke Thorburn was supported by UK Research and Innovation [grant number EP/S023356/1], in the UKRI Centre for Doctoral Training in Safe and Trusted Artificial Intelligence (safeandtrustedai.org).

References

[1] Yoram Amiel, Frank Cowell, and Xavier Ramos. Poles apart? an analysis of the meaning of polarization. *Review of Income and Wealth*, 56(1):23–46, 2010.

[2] Samin Aref, Ly Dinh, Rezvaneh Rezapour, and Jana Diesner. Multilevel structural evaluation of signed directed social networks based on balance theory. *Scientific Reports*, 10(1):15228, Sep 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-71838-6. URL https://doi.org/10.1038/s41598-020-71838-6.

[3] Levi Boxell, Matthew Gentzkow, and Jesse M Shapiro. Cross-country trends in affective polarization. *Review of Economics and Statistics*, pages 1–60, 2022.

[4] Aaron Bramson, Patrick Grim, Daniel J. Singer, Steven Fisher, William Berger, Graham Sack, and Carissa Flocken. Disambiguation of social polarization concepts and measures. *The Journal of Mathematical Sociology*, 40(2):80–111, 2016. doi: 10.1080/0022250X.2016.1147443. URL https://doi.org/10.1080/0022250X.2016.1147443.

[5] Aaron Bramson, Patrick Grim, Daniel J. Singer, William J. Berger, Graham Sack, Steven Fisher, Carissa Flocken, and Bennett Holman. Understanding polarization: Meanings, measures, and model evaluation. *Philosophy of Science*, 84(1):115–159, 2017. doi: 10.1086/688938.

[6] Jacob R. Brown and Ryan D. Enos. The measurement of partisan sorting for 180 million voters. *Nature Human Behaviour*, 5(8):998–1008, Aug 2021. ISSN 2397-3374. doi: 10.1038/s41562-021-01066-z. URL https://doi.org/10.1038/s41562-021-01066-z.

[7] Micah D Carroll, Anca Dragan, Stuart Russell, and Dylan Hadfield-Menell. Estimating and penalizing induced preference shifts in recommender systems. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 2686–2708. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/carroll22a.html.

[8] Alessandro Cossard, Gianmarco De Francisci Morales, Kyriaki Kalimeri, Yelena Mejova, Daniela Paolotti, and Michele Starnini. Falling into the echo chamber: The italian vaccination debate on twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 14(1):130–140, May 2020. doi: 10.1609/icwsm.v14i1.7285. URL https://ojs.aaai.org/index.php/ICWSM/article/view/7285.

[9] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 2 edition, 2006.

[10] Sergio Currarini, Matthew O Jackson, and Paolo Pin. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045, 2009.

[11] Carlo Dal Maso, Gabriele Pompa, Michelangelo Puliga, Gianni Riotta, and Alessandro Chessa. Voting behavior, coalitions and government strength through a complex network analysis. *PloS one*, 9(12):e116046, 2014.

[12] Daniel DellaPosta. Pluralistic collapse: The "oil spill" model of mass opinion polarization. *American Sociological Review*, 85(3):pp. 507–536, 2020. ISSN 00031224, 19398271. URL https://www.jstor.org/stable/48595846.

[13] Jean-Yves Duclos and André-Marie Taptué. Chapter 5 - polarization. In Anthony B. Atkinson and François Bourguignon, editors, *Handbook of Income Distribution*, volume 2 of *Handbook of Income Distribution*, pages 301–358. Elsevier, 2015. doi: https://doi.org/10.1016/B978-0-444-59428-0.00006-0. URL https://www.sciencedirect.com/science/article/pii/B9780444594280000060.

[14] Jean-Yves Duclos, Joan Esteban, and Debraj Ray. Polarization: concepts, measurement, estimation. *Econometrica*, 72(6):1737–1772, 2004.

[15] Edith Elkind, Martin Lackner, and Dominik Peters. Structured preferences. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 10, pages 187–207. AI Access, (online), 2017.

[16] Hanif Emamgholizadeh, Milad Nourizade, Mir Saman Tajbakhsh, Mahdieh Hashminezhad, and Farzaneh Nasr Esfahani. A framework for quantifying controversy of social network debates using attributed networks: biased random walk (brw). *Social Network Analysis and Mining*, 10(1):90, Nov 2020. ISSN 1869-5469. doi: 10.1007/s13278-020-00703-1. URL https://doi.org/10.1007/s13278-020-00703-1.

[17] Joan-María Esteban and Debraj Ray. On the measurement of polarization. *Econometrica*, 62(4):819–851, 1994. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2951734.

[18] David Garcia, Adiya Abisheva, Simon Schweighofer, Uwe Serdült, and Frank Schweitzer. Ideological and temporal components of network polarization in online political participatory media. *Policy & internet*, 7(1):46–79, 2015.

[19] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying controversy in social media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, WSDM '16, page 33–42, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450337168. doi: 10.1145/2835776.2835792. URL https://doi.org/10.1145/2835776.2835792.

[20] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying controversy on social media. *Trans. Soc. Comput.*, 1(1), jan 2018. ISSN 2469-7818. doi: 10.1145/3140565. URL https://doi.org/10.1145/3140565.

[21] Noam Gidron, Lior Sheffer, and Guy Mor. Validating the feeling thermometer as a measure of partisan affect in multi-party systems. *Electoral Studies*, 80:102542, 2022. ISSN 0261-3794. doi: https://doi.org/10.1016/j.electstud.2022.102542. URL https://www.sciencedirect.com/science/article/pii/S0261379422000981.

[22] Chiara Gigliarano and Karl Mosler. Constructing indices of multivariate polarization. *The Journal of Economic Inequality*, 7(4):435–460, Dec 2009. ISSN 1573-8701. doi: 10.1007/s10888-008-9096-x. URL https://doi.org/10.1007/s10888-008-9096-x.

[23] Joshua R. Gubler and Joel Sawat Selway. Horizontal inequality, crosscutting cleavages, and civil war. *Journal of Conflict Resolution*, 56(2):206–232, 2012. doi: 10.1177/0022002711431416. URL https://doi.org/10.1177/0022002711431416.

[24] Te Sun Han. Nonnegative entropy measures of multivariate symmetric correlations. *Information and Control*, 36(2):133–156, 1978. ISSN 0019-9958. doi: https://doi.org/10.1016/S0019-9958(78)90275-9. URL https://www.sciencedirect.com/science/article/pii/S0019995878902759.

[25] Frank Harary. On the measurement of structural balance. *Behavioral Science*, 4(4): 316–323, 1959.

[26] Ruben Interian and Celso C. Ribeiro. An empirical investigation of network polarization. *Applied Mathematics and Computation*, 339:651–662, 2018. ISSN 0096-3003. doi: https://doi.org/10.1016/j.amc.2018.07.066. URL https://www.sciencedirect.com/science/article/pii/S0096300318306325.

[27] Ruben Interian, Ruslán G. Marzo, Isela Mendoza, and Celso C. Ribeiro. Network polarization, filter bubbles, and echo chambers: an annotated review of measures and reduction methods. *International Transactions in Operational Research*, oct 2022. doi: 10.1111/itor.13224. URL https://doi.org/10.1111%2Fitor.13224.

[28] Yoko Kawada, Yuta Nakamura, and Keita Sunada. A characterization of the esteban–ray polarization measures. *Economics Letters*, 169:35–37, 2018. ISSN 0165-1765. doi: https://doi.org/10.1016/j.econlet.2018.05.011. URL https://www.sciencedirect.com/science/article/pii/S0165176518301848.

[29] Mari Kawakatsu, Yphtach Lelkes, Simon A. Levin, and Corina E. Tarnita. Interindividual cooperation mediated by partisanship complicates madison's cure for "mischiefs of faction". *Proceedings of the National Academy of Sciences*, 118(50):e2102148118, 2021. doi: 10.1073/pnas.2102148118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2102148118.

[30] David Krackhardt and Robert N. Stern. Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, 51(2):123–140, 1988. ISSN 01902725. URL http://www.jstor.org/stable/2786835.

[31] Yphtach Lelkes. Mass Polarization: Manifestations and Measurements. *Public Opinion Quarterly*, 80(S1):392–410, 03 2016. ISSN 0033-362X. doi: 10.1093/poq/nfw005. URL https://doi.org/10.1093/poq/nfw005.

[32] Matthew Levendusky. *The Partisan Sort: How liberals became Democrats and conservatives became Republicans*. University of Chicago Press, 2009.

[33] Esfandiar Maasoumi. The measurement and decomposition of multi-dimensional inequality. *Econometrica*, 54(4):991–997, 1986. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1912849.

[34] Lilliana Mason. "I Disrespectfully Agree": The Differential Effects of Partisan Sorting on Social and Issue Polarization. *American Journal of Political Science*, 59(1):128–145, 2015. doi: https://doi.org/10.1111/ajps.12089. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12089.

[35] Lilliana Mason. A Cross-Cutting Calm: How Social Sorting Drives Affective Polarization. *Public Opinion Quarterly*, 80(S1):351–377, 03 2016. ISSN 0033-362X. doi: 10.1093/poq/nfw001. URL https://doi.org/10.1093/poq/nfw001.

[36] W. McGill. Multivariate information transmission. *Transactions of the IRE Professional Group on Information Theory*, 4(4):93–111, 1954. doi: 10.1109/TIT.1954.1057469.

[37] Branko Milanovic. A new polarization measure and some applications. *Development Research Group, Word Bank*, 2000.

[38] Joel Miller, E Glen Weyl, and Leon Erichsen. Beyond collusion resistance: Leveraging social information for plural funding and voting. *Available at SSRN 4311507*, 2022.

[39] G. P. Nason and Robin Sibson. Measuring multimodality. *Statistics and Computing*, 2(3):153–160, Sep 1992. ISSN 1573-1375. doi: 10.1007/BF01891207. URL https://doi.org/10.1007/BF01891207.

[40] Roger B Nelsen. *An Introduction to Copulas (Second Edition)*. Springer Series in Statistics. Springer, 2006.

[41] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[42] Robert Noggle. The Ethics of Manipulation. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2022 edition, 2022.

[43] Aviv Ovadya and Luke Thorburn. Bridging Systems: Open Problems for Countering Destructive Divisiveness across Ranking, Recommenders, and Governance, 2023. URL https://arxiv.org/abs/2301.09976.

[44] Iñaki Permanyer. The conceptualization and measurement of social polarization. *The Journal of Economic Inequality*, 10(1):45–74, Mar 2012. ISSN 1573-8701. doi: 10.1007/s10888-010-9143-2. URL https://doi.org/10.1007/s10888-010-9143-2.

[45] Stephen D. Reese, Lou Rutigliano, Kideuk Hyun, and Jaekwan Jeong. Mapping the blogosphere: Professional and citizen-based media in the global news arena. *Journalism*, 8(3):235–261, 2007. doi: 10.1177/1464884907076459. URL https://doi.org/10.1177/1464884907076459.

[46] Anna Rumshisky, Mikhail Gronas, Peter Potash, Mikhail Dubov, Alexey Romanov, Saurabh Kulshreshtha, and Alex Gribov. Combining network and language indicators for tracking conflict intensity. In Giovanni Luca Ciampaglia, Afra Mashhadi, and Taha Yasseri, editors, *Social Informatics*, pages 391–404, Cham, 2017. Springer International Publishing. ISBN 978-3-319-67256-4.

[47] Joel Sawat Selway. Cross-cuttingness, cleavage structures and civil war onset. *British Journal of Political Science*, 41(1):111–138, 2011. ISSN 00071234, 14692112. URL http://www.jstor.org/stable/41241643.

[48] David Siroky and Michael Hechter. Ethnicity, class, and civil war: the role of hierarchy, segmentation, and cross-cutting cleavages. *Civil Wars*, 18(1):91–107, 2016. doi: 10.1080/13698249.2016.1145178. URL https://doi.org/10.1080/13698249.2016.1145178.

[49] Hu Kuo Ting. On the amount of information. *Theory of Probability & Its Applications*, 7(4):439–447, 1962. doi: 10.1137/1107041. URL https://doi.org/10.1137/1107041.

[50] Leo Tolstoy. *Anna Karenina*. Wordsworth Editions Limited, 1995. Translated by Aylmer Maude.

[51] Kai-Yuen Tsui. Multidimensional generalizations of the relative and absolute inequality indices: The atkinson-kolm-sen approach. *Journal of Economic Theory*, 67(1):251–265, 1995. ISSN 0022-0531. doi: https://doi.org/10.1006/jeth.1995.1073. URL https://www.sciencedirect.com/science/article/pii/S0022053185710733.

[52] Kai-yuen Tsui. Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation. *Social Choice and Welfare*, 16(1):145–157, Jan 1999. ISSN 1432-217X. doi: 10.1007/s003550050136. URL https://doi.org/10.1007/s003550050136.

[53] Petter Törnberg. How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences*, 119(42):e2207159119, 2022. doi: 10.1073/pnas.2207159119. URL https://www.pnas.org/doi/abs/10.1073/pnas.2207159119.

[54] Edward Vul and Harold Pashler. Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, 19(7):645–647, 2008.

[55] You-Qiang Wang and Kai-Yuen Tsui. Polarization orderings and new classes of polarization indices. *Journal of Public Economic Theory*, 2(3):349–363, 2000.

[56] Satosi Watanabe. Information theoretical analysis of multivariate correlation. *IBM Journal of Research and Development*, 4(1):66–82, 1960. doi: 10.1147/rd.41.0066.

[57] E Glen Weyl, Puja Ohlhaver, and Vitalik Buterin. Decentralized society: Finding web3's soul. *Available at SSRN 4105763*, 2022.

[58] Michael Wolfowicz, David Weisburd, and Badi Hasisi. Examining the interactive effects of the filter bubble and the echo chamber on radicalization. *Journal of Experimental Criminology*, 19(1):119–141, Mar 2023. ISSN 1572-8315. doi: 10.1007/s11292-021-09471-0. URL https://doi.org/10.1007/s11292-021-09471-0.

[59] Yan Zhang, A.J. Friend, Amanda L. Traud, Mason A. Porter, James H. Fowler, and Peter J. Mucha. Community structure in congressional cosponsorship networks. *Physica A: Statistical Mechanics and its Applications*, 387(7):1705–1712, 2008. ISSN 0378-4371. doi: https://doi.org/10.1016/j.physa.2007.11.004. URL https://www.sciencedirect.com/science/article/pii/S037843710701206X.

Luke Thorburn *(corresponding author)*
King's College London
London, UK
Email: luke.thorburn@kcl.ac.uk

Maria Polukarov
King's College London
London, UK
Email: maria.polukarov@kcl.ac.uk

Carmine Ventre
King's College London
London, UK
Email: carmine.ventre@kcl.ac.uk