Department of Statistics, The University of Chicago

Degree Paper
Master of Science in Statistics

Multidimensional Scaling Analysis of Legislative Roll Calls in Multi-Party Systems

ZHI RONG TAN

The University of Chicago

Advisor:

Lek-Heng Lim

Submitted July 2019

Abstract

Multidimensional Scaling (MDS) is a method for reducing high dimensional data to a low dimensional structure. In this paper, we apply Metric MDS to the legislative roll calls of three countries: United Kingdom, United States, and Canada, over a time period of roughly two decades, using a kernel method and dissimilarity function that has an analytical solution and is locally accurate. We propose a theoretical model that accounts for MDS embedding structures in a multi-party system where parties are well-separated, and explore the relationship between MDS and Spectral Clustering. Our results suggests the presence of two particular contemporary political phenomenon. First, there appears to be increasing inter-party polarization over time in all legislatures. Second, from the simulation of legislative behaviors, we argue that empirical results demonstrate intra-party fragmentation along a new issue space, indicating the insufficiencies of the traditional uni-dimensional spectrum.

Contents

1 Introduction

"One of the penalties for refusing to participate in politics is that you end up being governed by your inferiors."

– Plato

The geometrical representation of legislators and roll calls using dimensionality reduction methods is common in studying legislative voting patterns. One of such techniques is Multidimensional Scaling (MDS), which can mathematically model each legislator's voting patterns and ideologies. By reducing the high dimensional data to a low dimensional structure that most faithfully preserves the input data's inner products, MDS allows for the graphical interpretations of political positions and determination of the ideological rank order within a legislative body.

Beyond MDS, spectral methods have emerged as a powerful tool for nonlinear dimensionality reduction and manifold learning. These techniques find low dimensional sub-spaces in higher dimensional data from the top or bottom eigenvectors of specially constructed matrices. For example, Diaconis et al. 2008 applied spectral methods with multidimensional scaling on the 2005 United States (US) House of Representatives roll call votes to reveal horseshoes structures that are characteristic of kernel-based dimensionality reduction techniques.

The US two-party Congress is the world's most widely studied legislative institution. Yet, despite a wealth of existing literature, political trends in recent years are becoming more interesting but challenging to navigate. For instance, Doherty et al. 2018 described worsening partisanship as the ideological overlap between the two US parties is gradually vanishing. This inter-party polarization is further made complex by the rise of intra-party fragmentation. In the past, western political parties were built on a set of binary choices — big vs. small government, and conservative vs. liberal social norms. However, the set of binary choices may no longer be sufficient in explaining the range of political ideologies in long-established parties. Similarly, other western countries, such as Canada and the United Kingdom may be facing identical concerns. In the United Kingdom, the Conservatives are slowly fracturing under the weight of Brexit, while the Labor Party is ruled by the far left at the grassroots but is center-left in government, creating a dichotomy that renders the party unable to negotiate clear policies.

Are these political concerns just an anomaly, or will it become a new constant in the world? How are each country's legislative voting patterns different from one another, and what are the deviations over time? Can patterns in two-party political systems may be generalized to multi-party systems?

In order to better understand these questions and quantify modern political trends, we apply Multidimensional Scaling and kernel methods to the legislative bodies of countries such as the United States, United Kingdom, and Canada over the last two decades. We employed the same technique in Diaconis et al. 2008, having demonstrated its relevance by expanding

their proof for the validity of their method from a uni-dimensional spectrum to a multi-dimensional one.

We classify the MDS results for these three countries over 2 decades into groups of embedding patterns, presenting how MDS patterns are often a function of factors like inter-party polarization, cross-party coalition, intra-party unity, fragmentation, strength of party whips, and the effective number of parties. We also show that Diaconis et al. 2008's two-party model can be extended to a multi-party system, and discuss its relationship with spectral clustering. However, weaknesses of MDS on a multi-party legislature include its inability to capture the intra-party distribution of legislators' ideal points if parties are well-separated.

We are also able to relate the MDS patterns to contemporary political trends. By using the proportion of dissimilarities explained by the first eigenfunction as a measure of polarization, we discover increasing inter-party polarization over time for all three countries. In addition, in an attempt to account for the two-dimensional patterns observed in recent US House of Representatives, we perform a simulation of legislative voting behaviors by making assumptions about the intra-party distribution and dimensionality of legislators' ideal points. The simulation supports hypotheses of intra-party fragmentation along a new issue space in modern legislatures, indicating that the traditional uni-dimensional issue space, such as the left-right political spectrum, may no longer be sufficient.

The paper shall proceed as follows: Section 2 examines literature related to dimensionality reduction methods on legislative voting data. Section 3 outlines a rigorous justification of our methodology. Section 4 describes the data, while Section 5 briefly discusses our results. Section 6 to 8 detail a deeper analysis of our findings — the extension of insights from two-party to multi-party systems, and political trends in the last decade, including rising polarization and intra-party fragmentation.

2 Literature Review

The majority of dimensionality reduction techniques employed in the analysis of legislative voting records are based on the NOMINATE Scaling method, invented in the 1980s but still widely utilized today. New research similar to the NOMINATE method are also discussed, such as the Political DNA method, or Diaconis et al. (2008)'s mathematical analysis of horseshoe effects found from the multidimensional scaling results of the US Congress. We pay special attention to the latter since our paper is closely related to his analysis.

2.1 NOMINATE Scaling Method

Several approaches exist for scoring political ideologies from voting data, the most popular techniques being the family of NOMINATE methods, conceptualized by Poole and Rosenthal (1983). Before its invention, most ratings were constructed from a small set of divisive roll calls pertaining to specific issues, and excluded less divisive votes — the smaller sample of roll-calls hence artificially polarized congressional behaviors. On the other hand, NOMINATE uses almost all the recorded roll calls in its calculations.

NOMINATE represents legislators by points in a low-dimensional Euclidean space with their votes decided by maximizing a utility function. The simplest model assumes that legislators' views (defined as ideal points) lie on only one dimension — liberalism or conservatism, and have utility functions centered at their ideal points. When faced with the choice of $\{\text{Yea}, \text{Nay}\}$ for each roll call, they vote for the alternative closest to them on the dimension. Poole and Rosenthal (1983) further presumed that ideal points can be "reverse-engineered" from observing the roll call matrix, without taking into consideration the content of the votes themselves. Therefore, legislators have utility functions consisting of a deterministic component that is a function of the distance between the legislator's ideal point and the roll call outcome.

Further model improvements acknowledged the possibility of errors, and added a stochastic component to the utility functions to account for any poor judgement. Based on the legislator's ideal points and voting outcomes, we calculate the probabilities of observing the roll calls, and estimate their ideal points based on maximizing the joint probabilities of the votes.

The research further claimed that the entire congress up till the 1980s can be represented by at most two dimensions, also known as "issue spaces" — the first is the left-right spectrum, and the second measures intra-party unity, picking up regional differences, such as attitudes towards slavery, Bimetallism or civil rights. For example, the splitting of Democrats into Northern and Southern during the 1960s due to differing attitudes towards Civil Rights led to the need for a second dimension to adequately account for legislative voting patterns. Nonetheless, Poole and Rosenthal (1984) revealed that from the late 1960s till 1980s, race-related issues are no longer orthogonal to the liberal-conservative issue space, and the old split in the Democratic Party between North and South had largely disappeared.

Poole and Daniels (1985) improved the method to enable cross-chamber comparisons of legislators and roll calls across years, by using multidimensional unfolding in accordance with a spatial model to obtain a spatial configuration of the legislators. They introduced time as a new dimension, and the possibility for legislators to change their spatial positions from year to year. The final result is then derived from maximizing the following likelihood function, where t is the t-th congress, i is the i-th legislator, j the j-th roll call, and $l = \{0, 1\}$ for the roll call outcome. P represents the probability, a function of the legislators' estimated ideal points, while C is the observed outcome:

$$L = \prod_{t=1}^{\tau} \prod_{i=1}^{n_t} \prod_{j=1}^{p_t} \prod_{l=0}^{1} P_{ijl}^{C_{ijl}}. \tag{2.1}$$

Limitations of the NOMINATE method include its focus on two-party systems, and the possible complexity involved as the number and size of congress increases. As Equation (2.1) demonstrates, the method could be computationally demanding if party number or years increase.

2.2 Other Dimensionality Reduction Models

Since NOMINATE, further techniques have been developed to study legislative voting records, such as kernel methods in Multidimensional Scaling (described in Section 2.3), Principal Component Analysis by Leeuw (2011), Bayesian spatial voting models by Clinton et al. (2004), and graphical models by Guo et al. (2015), who developed a Markov model to characterize the dependence structures arising from voting records of legislators on different pre-labelled issues, such as defense, energy, and healthcare. By jointly estimating a collection of graphical models, Guo et al. described the internal networks of the US senate on critical legislative issues. Notably, Kim, Londregan, and Ratkovic (2015) also combined analysis of voting records with textual and speech data through Sparse Factor Analysis to determine legislators' ideal points and the number of policy dimensions.

More recently, Longo et al. (2018) introduced a new measure called "Political DNA" that is based on features selection and estimation of class posterior probabilities. By modelling the votes of the Italian Senate's XVII legislature as a mixture of random variables and utilizing sparse Principal Component Analysis, the authors selected the most relevant bills causing variances in legislators' political positions. This determined not only the separation of political positions between senators and the attachment to their own ideology, but also quantified the influence from other issue groups on each senator.

2.3 Horseshoe Effects in Kernel Methods

Diaconis et al. (2008) applied kernel methods with Multidimensional Scaling to a specific dataset – the 2005 US House of Representatives roll call votes. They first defined the

empirical distance between legislators as $d(l_i, l_j) = \frac{1}{p}\sum_{k=1}^{p}|v_{ik} - v_{jk}|$, where p is the total number of roll calls, i and j are individual legislators, and v_{ik} is the vote call result $= \{0.5, -0.5\}$ of the k-th roll call for the i-th legislator. To acknowledge that $d(l_i, l_j)$ is only accurate at small scales, they further defined the proximity $P(i, j) = 1 - \exp(-d(l_i, l_j))$ using a kernel method.

Next, P is doubly mean-centered and transformed into $S = -\frac{1}{2}HPH$, where $H = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$. The three-dimensional mapping of legislators revealed twin horseshoes that are characteristic of kernel projection techniques, and are orthogonal in the second and third component. According to Mardia, Kent, and Bibby (1979), such horseshoes result from ordered data in which only local inter-point distances can be accurately estimated.

To prove this is true, the paper proceeded with a theoretical model where within each party, the legislators are uniformly distributed, such that $1 \leq i, j \leq n$, $d(l_i, l_j) = |i/n - j/n|$, and $\tilde{P_{1n}}$ is derived using the same exponential kernel as above. To account for twins horseshoes instead of only a single horseshoe, the assumption that both parties are well-separated is made such that if i, j belongs to different parties, $P(i, j) = 1$. They then ended up with the proximity matrix

$$\tilde{P_{2n}} = \left[\begin{array}{c|c} \tilde{P_{1n}} & \mathbf{1} \\ \hline \mathbf{1} & \tilde{P_{1n}} \end{array}\right],$$

and $\tilde{S}$ is calculated as per above. To prove limiting results, Diaconis et al. (2008) further scaled $\tilde{S_{2n}} = \frac{1}{2n}S$, and find approximate eigenfunctions for S_{2n} by considering a kernel limit K of the matrix S_{2n} and then solving the corresponding integral equation

$$\int_0^1 K(x, y)f(y)\ dy = \lambda f(x).$$

The approximate eigenfunctions derived are graphed, and the theoretical twin horseshoes that result from the three-dimensional mapping match the empirical results exactly. Diaconis et al. (2008) hence mathematically proved from the MDS-produced twins horseshoe shape that legislators come from two separate populations (Democrat and Republican party), and within each population, the members are uniformly distributed on a regular grid in a unidimensional spectrum. The assumption of two separate populations is an indication that there already is a significant level of polarization in 2005, since for all legislators within a party, their distance to another legislator in the other party is the same. The model also suggests the MDS ordering or legislators should correspond to the political ideology within each party.

Diaconis et al.'s utilization of the exponential kernel also relates to other research on spectral methods. For example, kernel methods in Multidimensional Scaling provide a Spectral Clustering of legislators into k groups using the first k eigenvectors, as described by Shi and Malik (2000). This will be further explored in Section 6, when we extend the MDS technique to countries with more than two parties.

In addition, Leeuw (2007) generalized the exponential kernel used by Diaconis et al. as a function of the Kac–Murdock–Szegö matrix A, where $a_{ij} = \rho^{|i-j|}$, and in this case $\rho =$

$\exp(-1)$. Note that A is both totally positive and Toeplitz, the latter defined as a matrix in which each descending diagonal from left to right is constant. Leeuw proved that horseshoes will occur for matrices with these properties, and the double centering of a Toeplitz matrix also produces horseshoes.

As part of our methodology, we shall proceed with Multidimensional Scaling using the same exponential kernel that Diaconis et al. (2008) utilized, since it provides a simple analytical solution compared to other dimensionality reduction methods described above — further justifications are discussed in the next section. In addition to the objectives mentioned in Section 1, we are also interested to see if the horseshoe effect can be reproduced in recent years of the US Congress, or in the other legislative bodies, and if so, the implications of such observations.

3 Methodology

We explore the legislative voting records of four bodies in three countries — the US House of Representatives, US Senate, UK House of Commons, and Canadian House of Commons, from the last 15–25 years. We pick these countries as they are all western democracies that have at least two major parties and are English-speaking, enabling us to understand their voting records. Since the United Kingdom and Canada are both multi-party system, they allow for comparison with the United States' two-party system.

Before proceeding, let us first define some terminologies:

Definition 3.1. A political party is considered ***significant*** if it runs for national elections *and* has the capacity to obtain a sizeable portion of legislative seats.

Definition 3.2. A political party is a ***major party*** if it has the capacity to gain control of government. Major parties often fall on opposite end of the traditional ideological "left-right" spectrum. For example, the *Liberals* and *Conservatives* in Canada are major parties.

Definition 3.3. A political system is ***multi-party*** if there are more than two *significant* parties. For the purpose of this paper, a two-party system shall *not* be referred to as multi-party.

We apply Multidimensional Scaling to voting records generated by the legislators of the four bodies, with dissimilarity between legislators defined by their roll call votes. Suppose in each session, and for each body, there are N members and m roll calls. We will remove all r roll calls that are purely ceremonial or administrative, such as the first House call or election of speaker, using only $m - r = p$ roll calls. In addition, we further restrict our analysis to $n \leq N$ members that voted on at least a certain proportion of roll calls. Due to different voting policies, this proportion will depend on the specific country, in order to ensure our final count is not significantly diminished.

Then, we order our data in a $n \times p$ matrix V, with $v_{ij} \in \{+0.5, -0.5, 0\}$ which exactly correspond with the vote of "Yea", "Nay", and "Not Voting", where "Not Voting" could either imply the legislator was present but abstained from voting (due to a variety of reasons) or completely absent. We first begin by reviewing the theory behind Multidimensional Scaling.

3.1 Multidimensional Scaling Theory

Suppose we have a $n \times p$ matrix V with n observations $x_1, x_2, \ldots x_n \in \mathbb{R}^p$. Although MDS was designed to preserve inner products, it is now mainly motivated by the idea of preserving pairwise dissimilarities, where we want to find $y_1, y_2, \ldots y_n$ which lie in dimension $k < p$, that preserves the distance between these new observations as much as possible in p-dimensional space.

3.1.1 Euclidean Distances

For simplicity, suppose the dissimilarity is defined in terms of Euclidean distance — we form a distance matrix $D \in \mathbb{R}^{n\times n}$ where

$$d_{ij}(X) = \sqrt{(x_i^1 - x_j^1)^2 + (x_i^2 - x_j^2)^2 + \cdots + (x_i^p - x_j^p)^2}. \tag{3.1}$$

Since MDS focuses on preserving inner product, we must find the dot product matrix $S \in \mathbb{R}^{n\times n}$, $S = XX^T$ from the distance matrix. Note that

$$d_{i,j}^2 = (x_i - x_j)(x_i - x_j)^T = x_i x_i^T + x_j x_j^T - 2x_i x_j^T.$$

The above equation can be represented in matrix form as

$$D_2 = s\mathbf{1}^T + \mathbf{1}s^T - 2S,$$

where s is the $n\times$ 1 vector of diagonal entries of S, and one can find D_2 by squaring each element of D. This allows us to retrieve S from D_2 by double-centering it:

$$S = -\frac{1}{2}HD_2H,$$

$$H = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T.$$

Since for any matrix A, HAH will have row and column mean = 0, then

$$\begin{aligned} HD_2H &= Hs\mathbf{1}^TH + H\mathbf{1}s^TH - 2HSH \\ &= 0 + 0 - 2HSH \\ &= -2S. \end{aligned}$$

The last equality comes from the fact that the row sums of $S = 0$.

The problem of finding $Y \in \mathbb{R}^{n\times k}$ which preserves the inner products in a lower-dimensional space can be represented as a minimization problem, in which k is a given parameter:

$$Y = \underset{Y}{\operatorname{argmin}} \sum_{i,j} (d_{ij}^2 - \omega_{ij}^2), \tag{3.2}$$

where $\omega_{ij}^2(Y)$ is the squared distance of the new points y_i, y_j in the lower-dimensional space and similarly defined by Equation (3.1).

Young and Householder (1938) demonstrated that this can be treated as an eigenvalue problem, where they first decompose $S = U\Lambda U^T$ by the Spectral Theorem for Symmetric Matrices, in which Λ is diagonal and contains the eigenvalues of S, while U is orthogonal. Note that if the matrix is of Euclidean distance, then S is positive semi-definite. If S is positive semi-definite, then the eigenvalues are all non-negative.

Next, sort the eigenvalues and eigenvectors such that the diagonal entries of Λ are non-increasing, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_n$. It can be shown that Y may be obtained by taking the first k columns of U and scaling it by the square-root of Λ_{rr} for $1 \leq r \leq k$. Hence,

$$Y = U_k\Lambda_k^{1/2},$$

in which $U_k \in \mathbb{R}^{n\times k}$, and $\Lambda_k^{1/2} \in \mathbb{R}^{k\times k}$. This is known as the Classical Scaling solution.

3.1.2 Generalization of MDS

However, Euclidean distance may not be the best approach for all situations. The Euclidean distance squared matrix D_2 can also be represented with different forms of distances, or other types of "dissimilarities" Δ measuring differences between observations.

Regardless, the Classical Scaling solution remains relevant for finding points in a lower-dimensional Euclidean space whose inter-point distances approximate the orders of the new dissimilarity matrix Δ. We can run the same algorithm above with a non-Euclidean dissimilarity matrix, although there is no guarantee that all the eigenvalues shall be non-negative. As Critchley (1978) described,

Proposition 3.1. *Let Δ be a dissimilarity matrix, and k is fixed. Then the loss function as defined in Equation (3.2) is minimized over all possible matrix $\hat{Y} \in \mathbb{R}^{n\times k}$ when $\hat{Y}$ is given by $U_k\Lambda_k^{1/2}$, where $S = -\frac{1}{2}H\Delta H$, and S is decomposed to form $U\Lambda U^T$.*

Remark. Depending on the way Δ is defined, S may not be positive semi-definite, and some of the eigenvalues may be negative. If so, one can either proceed by simply ignoring the negative eigenvalues and their corresponding eigenvectors, or transforming the dissimilarity element — more information is available in Borg, Groenen, and Mair (2013).

In general, the dissimilarities Δ can be replaced in the loss function by a transformation $f(\delta)$, and the new points Y are found based on the minimization of this modified loss function, such that $\omega_{ij} \approx f(\delta_{ij})$. In this case, Equation (3.2) is transformed to:

$$Y = \underset{Y}{\operatorname{argmin}} \frac{\sum_{i,j}(f(\delta_{ij}) - \omega_{ij})^2}{\sum_{i,j}\omega_{ij}^2}. \tag{3.3}$$

If the transformation $f(.)$ is a continuous function, then this is known as Metric MDS, and the problem is typically solved using gradient-based iterative methods. If the transformation only preserves rank-order, then this is known as nonmetric MDS.

3.1.3 Kernel PCA and Metric MDS

Under the Kernel Principal Component Analysis (PCA) algorithm described by Schölkopf, Smola, and Müller (1998), an arbitrary mapping, $\Phi : \mathbb{R}^p \to \mathbb{R}^N$ where $N \geq p$, is first selected. One then creates the $N \times N$ kernel representing the inner product space of the new feature space,

$$K = k(x, y) = \Phi(x)^T\Phi(y) \in \mathbb{R}^{N\times N},$$

before forming S by double-centering K and obtaining the principal components via the eigenvalue decomposition of S. Kernel functions are normally chosen to be conditionally positive definite, so that the eigenvalues of the matrix K are non-negative.

It is possible for the transformation function of the dissimilarities in Metric MDS to be a kernel. Suppose we apply the same MDS scaling algorithm described in Proposition 3.1 to

the transformed dissimilarity matrix instead of using an iterative method. It is trivial to see that this procedure is rather similar to Kernel PCA.

Indeed, Williams (2000) demonstrated that Kernel PCA can be interpreted as a form of metric MDS when the kernel function is isotropic.

Definition 3.4. A kernel function is ***stationary*** if $K(x_i, x_j)$ depends only on the vector $x_i - x_j$. A stationary kernel function is ***isotropic*** if $K(x_i, x_j)$ depends only on (any definition of) distance $\|x_i - x_j\|$.

In this event, the desired configuration of points can be found via an eigen-problem instead of the iterative optimization of the loss function used in regular metric MDS algorithms.

Proposition 3.2. *Suppose for a metric MDS problem, the dissimilarities $Z = [\zeta_{ij}]_{i,j=1}^{n,n}$ are transformed using a isotropic kernel function, where $\Delta = K(Z)$.*
Then, the problem can be solved using a Classical Scaling solution. That is, given that the dimension k is fixed, the lower-dimensional representation in $\mathbb{R}^k$ is given by $U_k \Lambda_k^{1/2}$, where $S = -\frac{1}{2} H \Delta H$, and S is decomposed to form $U \Lambda U^T$.

3.1.4 Proportion of Dissimilarities

Suppose we solve a MDS problem using the classical scaling approach, where we treats the MDS as an eigenvalue problem. While $k \leq 3$ typically to allow for visualization, one can also determine k based on a trade-off between the proportion of dissimilarities captured in the lower-dimensional subspace and the size of k. According to Cox and Cox (2000), if S is positive semi-definite then the number of non-zero eigenvalues gives the number of dimensions required to completely capture the dissimilarities. Otherwise, the number of positive eigenvalues is the approximate number of dimensions.

Also note that a measure of the proportion of dissimilarities that is explained by the first k dimensions is

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{n} \lambda_j}.$$

If S is not positive semi-definite under a different definition of dissimilarity, then the above equation can be modified to

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{n} \mathbb{1}_{\lambda_j > 0} \lambda_j}.$$

Intuitively, if the first few components are able to capture a large proportion of the dissimilarities, this implies that the observations are well-separated.

3.2 Choice of Dissimilarity and Kernel Functions

Euclidean distance is not the most natural dissimilarity to use in our model, since for each dimension, the vote choices are discrete $v_{ik} \in \{+0.5, 0, -0.5\}$. Thus, we follow Diaconis et al. (2008)'s definition of empirical distance and dissimilarity. First, the empirical distance between two legislator l_i and l_j are defined by

$$\widehat{d}(i,j) = \frac{1}{p}\sum_{k=1}^{p}|v_{ik} - v_{jk}|, \tag{3.4}$$

where k is the index for the roll calls, and p is the total roll calls. This is almost equivalent to the proportion of roll calls that legislator i and j disagreed on, if not for the fact that a "no vote" is possible too.

3.2.1 Justification

As described in Section 2.1, the roll call votes are often a function of each legislator's ideal points, and the bill's nature. Although the ideal points and distances are unknown, they can be estimated from the roll call votes. Diaconis et al. (2008) presented a justification on why the empirical distance in Equation (3.4) is theoretically valid, but their proof assumes that legislators' ideal points are uni-dimensional. We generalize their model into one that is multi-dimensional, having previously discussed in Section 1 that a uni-dimensional issue space may sometimes be invalid, especially for legislatures with more than two parties.

Consider a theoretical model in which legislators' ideal points can be represented in a R-dimensional grid, where each dimension is of interval $I_r = [0, 1]$, and $R \ll p$. We define the theoretical distance between two legislators $d(i, j)$ by the Manhattan distance normalized such that $d(i, j) \in [0, 1]$:

$$d(i,j) = \frac{1}{R}\sum_{r=1}^{R}|l_{ir} - l_{jr}|.$$

Each dimension in the grid represents one issue space. We make the further assumption that for all R issue spaces, legislators' ideal points in one issue are independent of their ideal points in other issues — the issue spaces are orthogonal to one another.

Votes are also a function of the bills. Each bill is represented by a cutoff hyperplane $\in \mathbb{R}^{R-1}$ in the R-dimensional grid, which divides the legislators into yea or nay votes. Due to the assumption that all issue spaces are orthogonal to each other, it is reasonable to consider each bill k as a hyperplane only orthogonal in one dimension r, and parallel to other dimensions — it can be represented as a cutoff point in one of the R dimensions, $C_k \in I_r = [0, 1]_r$. Finally, assume that the cutoff points are independent w.r.t. one another, they are uniformly distributed across all R dimensions, and for each dimension r, uniformly distributed on $[0, 1]$.

Without loss of generality, let

$$v_{ik} = \begin{cases} +0.5 & \text{if } l_{ir} \leq C_k(r); \\ -0.5 & \text{if } l_{ir} > C_k(r). \end{cases}$$

Two legislators only take different views on a bill if a cutoff point divides them. Due to the uniform distribution of bills across all R dimensions and on each $[0,1]_r$ interval,

$$\mathbb{P}(v_{ik} \neq v_{jk}) = d(i,j).$$

Therefore, we can reduce Equation (3.4) to

$$\widehat{d}(i,j) = \frac{1}{p} \sum_{k=1}^{p} \mathbb{1}_{v_{ik} \neq v_{jk}}.$$

The rest of the proof are now similar to Diaconis et al.'s presentation. In particular, we will first discuss a lemma.

Lemma 3.1. *Given* $p \geq \log(n/\sqrt{\epsilon})/\epsilon^2$,

$$\mathbb{P}(|\widehat{d}_p(i,j) - d(i,j)| \leq \epsilon \text{ for all } 1 \leq i,j \leq n) \geq 1 - \epsilon.$$

Using Lemma 3.1, one can show that

$$\lim_{p \to \infty} \widehat{d}_p(i,j) = d(i,j) \quad a.s.$$

The asymptotic convergence justifies our choice of empirical distance.

We further presume that the information of roll call voting differences is only locally accurate, such that the distance can only be properly estimated when 2 legislators are nearby. We thereby define our dissimilarities with a kernel transformation

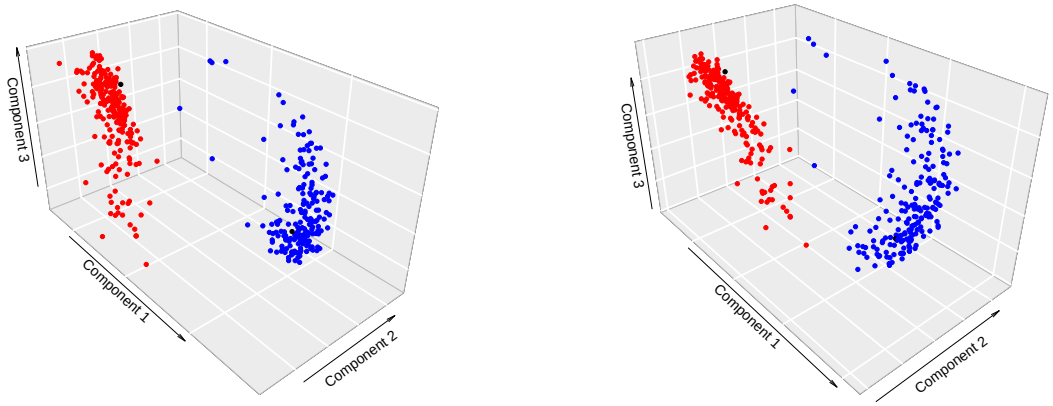
$$\delta(i,j) = 1 - \exp(-\theta \ \widehat{d}(i,j)), \tag{3.5}$$

where $\delta(i,j)$ are the elements of our dissimilarity matrix Δ_n, and θ is a hyperparameter. By letting $\theta = 1$, observe that $\delta(i,j) \approx \widehat{d}(i,j)$ when $\widehat{d}(i,j) \ll 1$.

This localization is commonly employed in dimensionality reduction algorithms, and allows the distance to not be sensitive to noises at large values of $d(l_i, l_j)$. As the number of votes two legislators disagree on increases, the marginal increase in dissimilarity decreases for every additional vote difference. By using this exponential kernel, any distance is down-weighted for points that are very far apart. Since our choice of kernel satisfies the requirements stated in Section 3.1.3, we will proceed with the algorithm to find the lower-dimensional representation as per Proposition 3.2.

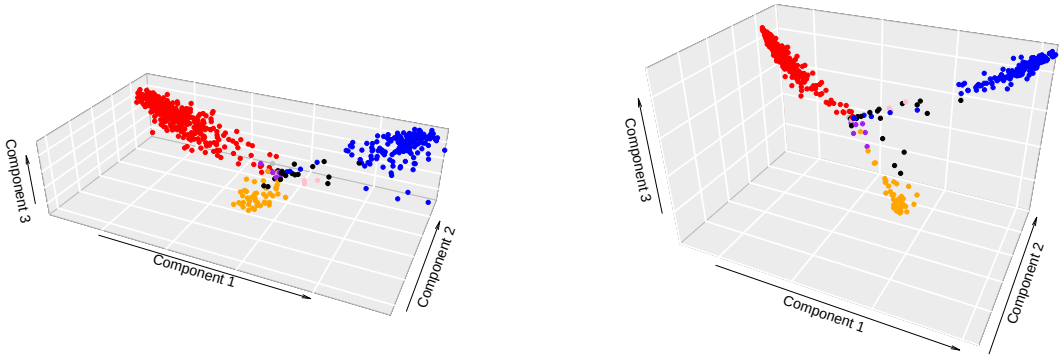
3.2.2 Comparison with Euclidean Distances

We also explore other definition of dissimilarities on the dataset. We observe that for most legislative bodies, the three-dimensional representation we obtained using a Euclidean distance-squared matrix (defined in Equation (3.1)) instead of the dissimilarity matrix described in Equation (3.5) are nearly similar. Often, the only notable differences are that compared to utilizing Euclidean distance, our dissimilarity matrix allows a clearer separation of observations, but points closer to each other are more spread out, as observed from Figure 1. In addition to the justifications described, the kernel use in our dissimilarity allows for a simpler interpretation of empirical distance as simply the proportion of voting roll calls that are different between two legislators, and relates to other analysis methods such as Spectral Clustering.



(a) Euclidean Distance Matrix (b) Kernel Dissimilarity Matrix

Figure 1: MDS on 105th US House of Representatives



(a) Euclidean Distance Matrix (b) Kernel Dissimilarity Matrix

Figure 2: MDS on 52nd UK House of Commons

Nonetheless, this may not be the case in a legislature where the rate of absence is high, such as the United Kingdom — under the British parliamentary whip system, attendance for major party members is not compulsory for all roll calls. Figure 2 demonstrates how the MDS outputs are different using the Euclidean distance vs. kernel dissimilarity matrix for the 52nd UK House of Commons.

Suppose legislator i and j are close. If legislator i is present but legislator j is not, then the empirical dissimilarity between them increases by ≈ 0.5. However, the Euclidean squared distance increases only by $0.5^2 = 0.25$. Hence, our kernel dissimilarity method tends to amplify the different rates of voting participation within the major parties, which is captured in the second component of the MDS.

For consistency, we shall stick to the kernel dissimilarity method in our analysis for all countries.

3.3 Spectral Clustering

There is a relationship between MDS and Spectral Clustering, especially with our use of a kernel transformation. As Hu (2012) discussed, the core procedure in Spectral Clustering first employs spectral methods to find an embedding of points in a lower-dimensional Euclidean space, before a hard-cut algorithm like K-Means is applied to partition points in the embedded space into clusters.

The first step involves converting raw data input into a weighted graph adjacency matrix. In our case, the process of forming $-\frac{1}{2}H\Delta H$ converts the dissimilarity matrix into the appropriate pairwise inner product matrix which parallels the adjacency matrix. While modern spectral clustering methods may include further steps like forming the graph Laplacian matrix and normalization, these are flexible and not strictly specified. The final step involves eigen-decomposing the transformed version of the adjacency matrix, and obtaining the embedding coordinates in lower dimension using the sorted eigenvectors based on the eigenvalues' magnitude.

In general, the number of components in the lower-dimensional space should be equal to the expected number of clusters, although if the points are well-separated into the clusters, the components required may decrease — this can be verified by checking the rate of changes of the top few eigenvalues.

Since spectral clustering is not the focus of our paper, we will not go in depth. Yet, with this parallel, performing hard-cut clustering on our MDS lower-dimensional representation is akin to performing spectral clustering. How well-separated the representations are into different political parties (or even separation within parties) is a reflection of polarization or party split, and contributes to the richness of our analysis.

4 Data

Roll call data for the US Congress, the UK and Canadian Parliament are obtained from GovTrack and TheyWorkforYou, and scraped from the Canadian parliamentary website respectively. Due to data unavailability, our time range of analysis is not consistent across all countries.

4.0.1 Data Processing

As we are analyzing the data across time and countries, each dataset is defined to consist of a legislative session for a given legislative body. For example, a dataset may consist of all the roll calls during the 115th Congress from 2017-2018, and the voting choices made by legislators belonging to that particular congress. In the United States, the term length is fixed, since the midterms or presidential elections occur once every 2 years.

In the United Kingdom and most Commonwealth countries, elections are normally held within 4–5 years, although snap elections may happen before the 4–5 year mark in extraordinary cases, like a vote of no-confidence, and hence the legislative term length and dataset size vary accordingly.

Although Canada's elections are normally held every 4–5 years, Parliament only lasted 1–2 years during the 2000s due to multiple votes of no-confidence. Some terms only have one session, but others have multiple sessions. On average, each session is between one to two years. However, in recent terms, one session lasts for the entire term of five years. To ensure that the dataset sizes remain roughly similar, long sessions are divided up into smaller sub-sessions, sub-indexed by A, B, $C\ldots$ using specific years as a cutoff.

To recap,

Definition 4.1. A ***legislative term*** is the period of time in which a legislature meets, beginning after an election and ending just before the next election is held.

Definition 4.2. A ***legislative session*** is the period of time in which a legislature is convened for purpose of lawmaking, usually as one of the smaller divisions of the entire time between two elections.

A single dataset is an $n \times p$ matrix, where n is the number of valid legislators, and p is the total number of valid roll calls within that legislative session or sub-session. For each dataset, we remove all administrative and ceremonial roll calls to find p. All votes are converted into the $\{0.5, 0, -0.5\}$ scale, where the difference between a "yea" and "nay" vote $= 1$.

The criteria for removing legislators from our analysis because of missing votes differ across countries. A legislator is removed if he/she missed:

1. **United States**: ≥ 1 vote
2. **United Kingdom**: $\geq 33\%$ of all votes
3. **Canada**: $\geq 25\%$ of all votes

The criteria are different because in the United Kingdom and Canada, the stronger whip system implies that presence is not compulsory unless required by the party whip. Hence, the threshold of 33% and 25% is chosen such that the final number of legislators is not overwhelmingly reduced.

4.0.2 Influence of Party Whips

In most western democracies, the majority party and main opposition party have their own whips. Although US Congress often vote along party lines, its whip influence is weaker than in the UK system, and American legislators have more freedom to diverge from party line and a stronger mandate to support their constituents' position than that of the party.

This is not true of the United Kingdom and Canada, where MP (Member of Parliament, equivalent to Congressman) attendance is not compulsory. In legislatures with strong whip influence, MPs are often informed prior to the votes of the importance of their attendance. For example, a single-line whip implies that the vote is non-binding for attendance, while a three-line whip is a strict instruction to attend and vote according to the party's position. As discussed in Agozino (2015), whip effect varies from year to year too, and is a function of the party leader's strength and overall political climate.

In addition to the United States, United Kingdom and Canada, we explored other legislatures, such as Australia and New Zealand. However, the general attendance of these legislatures are poor. Since the original parliament size is already small, removal of legislators with insufficient attendance creates a dataset with a small n. Hence, we only chose to focus on the three countries in the paper.

In the following sections, we will describe our dataset in detail, giving a broad introduction to the political system in each country, the list of datasets and years involved, the n and p described in Section 3, the majority party for each session, and any significant events. The values of n, p are after data refinement.

4.1 United States

We look into both the House of Representatives and Senate from the 103rd to 115th Congress, from 1993 to 2018. For comparison purposes, we also briefly looked at the 84th to 86th, and the 116th (up till end-May 2019) House of Representatives.

The parties involved are the **Democratic Party (D)**, the **Republican Party (R)**, and **others (e.g. Independent)**. D* or R* in the table below implies that the party has control of both houses of Congress.

Congress	Year	n	p	Majority	Notable Events
84th	1955–1956	429	147	D*	Start of Civil Rights Movement
85th	1957–1958	420	191	D*	Record for longest filibuster
86th	1959–1960	424	178	D*	Greensboro sit-ins
———	———	———	———	———	———
103rd	1993–1994	422	1091	D*	WTC bombings
104th	1995–1996	425	1320	R*	Government shutdown
105th	1997–1998	423	1165	R*	Bill Clinton impeachment
106th	1999–2000	422	1208	R*	Bush vs. Gore litigation
107th	2001–2002	420	989	R	September 11th Attacks
108th	2003–2004	427	1217	R*	Iraq War
109th	2005–2006	420	1209	R*	Hurricane Katrina disaster
110th	2007–2008	420	1864	D*	Global Financial Crisis
111th	2009–2010	418	1646	D*	Affordable Care Act
112th	2011–2012	418	1601	R	Debt Limit Crisis
113th	2013–2014	420	1201	R	End of filibuster on executive branch
114th	2015–2016	405	1320	R*	N/A
115th	2017–2018	416	1206	R*	End of Supreme Court filibusters
116th	2019 (till May)	431	219	D	Federal shutdown

Table 1: Dataset Information for United States House of Representatives

Congress	Year	n	p	Majority
103rd	1993–1994	98	722	D*
104th	1995–1996	98	916	R*
105th	1997–1998	100	610	R*
106th	1999–2000	98	668	R*
107th	2001–2002	99	631	*Split*
108th	2003–2004	100	673	R*
109th	2005–2006	99	643	R*
110th	2007–2008	97	655	D*
111th	2009–2010	92	690	D*
112th	2011–2012	98	484	D
113th	2013–2014	97	655	D
114th	2015–2016	100	500	R*
115th	2017–2018	96	597	R*

Table 2: Dataset Information for United States Senate

4.2 United Kingdom

The House of Commons is the lower house of the UK Parliament, and is an elected body consisting of around 650 members known as Members of Parliament (MPs). Members are elected to represent constituencies by the first-past-the-post system. Despite the presence of the House of Lords, the Lords' power to reject legislation has been reduced to a delaying power since 1911. The Government is solely responsible to the House of Commons and the Prime Minister stays in office only if he/she maintains the support of a Commons majority. Parliament normally sits for a term of five years. However, an early general election can be brought about by a two-third majority, or by a vote of no confidence in the government.

We shall look into the 52nd to 57th (and current) Parliament's House of Commons, elected in 1997, 2001, 2005, 2010, 2015, and 2017 respectively.

The 2 major parties are:

1. **Labour Party (Lab)**
2. **Conservative Party (Con)**

Other significant parties include:

1. **Liberal Democrats (LDem)**: Positioned centred, internationalist and pro-European.
2. **Democratic Unionist Party (DUP)**: Described as right-wing, socially conservative, and favoring British identity.
3. **Scottish National Party (SNP)**: Supports Scottish independence within the European Union.

Remaining parties are **UK Independence Party (UKIP)**, **Independent Conservatives**, **Green Party**, **Social Democratic and Labour Party (SDLP)**, **Independent Labours**, and **others (e.g. PC, UUP, Alliance, Respect, UKUP, and Independent)**. Due to the large number of parties, the "Other Parties" column in Table 3 only includes parties with at least ten seats in Parliament.

Term	Year	n	p	Majority	Other Parties
52nd	Jun 1997 – May 2001	679	448	Lab	Con, LDem
53rd	Jun 2001 – Apr 2005	672	1246	Lab	Con, LDem
54th	May 2005 – Apr 2010	657	351	Lab	Con, LDem
55th	May 2010 – Mar 2015	664	1226	Con+LDem	Lab
56th	May 2015 – May 2017	654	467	Con	Lab, SNP
57th	Jun 2017 – (May 2019)	663	414	Con	Lab, SNP, LDem

Table 3: Dataset Information for United Kingdom House of Commons

4.3 Canada

The Canadian House of Commons is a democratically elected body whose members are known as Members of Parliament (MPs). There were 308 seats before 2015, and 338 after. Members are elected by simple plurality in each of the country's electoral districts, and serve for constitutionally limited terms of up to five years. However, terms have ended before their expiry. The lower of the two houses, the House of Commons in practice holds far more power than the upper house, the Senate. Although both Houses' approval are required for legislation to become law, the Senate very rarely rejects bills passed by the Common. Moreover, the Cabinet is responsible solely to the House of Commons.

We shall look into the 38th to 42nd Parliament's House of Commons, elected in 2004, 2005, 2008, 2011 and 2015 respectively.

The 2 major parties are the **Liberal Party (Lib)** and **Conservative Party (Con)**. Other significant parties include the **New Democratic Party (NDP)**: described as a social democratic party and seated left of the Liberal Party; and **Bloc Québécois (BQ)**: described as social democratic and separatist, it campaigns for the political secession of Quebec from Canada and is active only within the province.

In Table 4, the datasets are indexed under "Term", in the form (Term-Session (Subsession)) e.g. 41st-II (A). "Opposition" indexes the official opposition i.e. second largest party.

Term	Year	n	p	Majority	Opposition
38th	Oct 2004 – Nov 2005	295	190	Lib	Con
39th-I	Apr 2006 – Sep 2007	290	219	Con	Lib
39th-II	Oct 2007 – Sep 2008	247	161		
40th-II	Jan 2009 – Dec 2009	296	158	Con	Lib
40th-III	Mar 2010 – Mar 2011	290	204		
41st-I (A)	Jun 2011 – Sep 2013	292	455	Con	NDP
41st-I (B)		288	305		
41st-II (A)	Oct 2013 – Aug 2015	276	224		
41st-II (B)		278	243		
42nd (A)	Dec 2015 – (May 2019)	323	183	Lib	Con
42nd (B)		307	252		
42nd (C)		295	285		

Table 4: Dataset Information for Canadian House of Commons

5 Results

In this section, we briefly describe trends in our MDS results across time and country. Further discussion is presented in Section 6 to 8, where we provide a rigorous analysis, and relate our findings to contemporary issues.

The MDS result for each dataset is visualized in a three-dimensional plot, with each axis representing one component. The axes are proportional and accordingly to scale, to give perspective on the differences across components. For simpler visualization, we also made two-dimensional plots of pairs of components, and the proportion of dissimilarities each component accounts for.

Colors are added to indicate party affiliations. These colors are exactly similar to common party convention, and the exact labelling can be found in Section 4.

It is important to note that rotation and reflection do not matter in MDS — only the distance between observations is significant. Moreover, while one may provide an interpretation for each axis, these interpretations cannot be generalized across years, since the distribution of congressmen changes from term to term.

For multi-party systems, we also perform MDS on a subset containing only the top two major parties (often Conservatives and Liberals), to determine if we can reproduce the same structure as that of the US Congress.

Before proceeding, let us define some terms that will be used in the next few sections.

Definition 5.1. A ***party embedding*** is the structure formed by all the legislators/points from a single party, contained within the lower-dimensional MDS space.

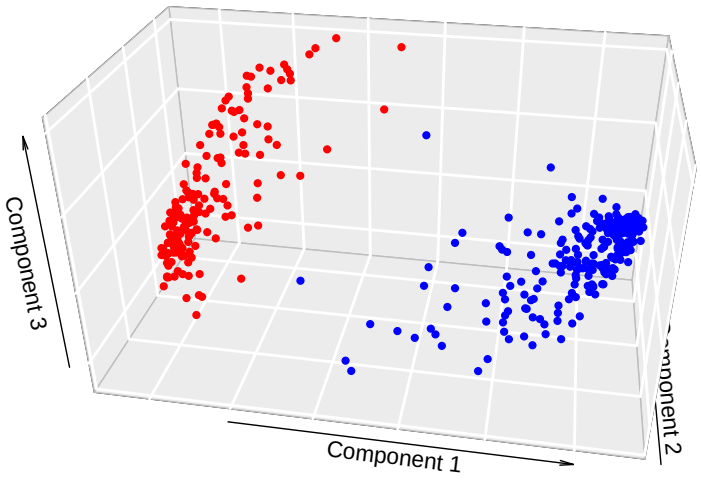
Definition 5.2. A party-embedding is ***one-dimensional*** if it can be roughly approximated by a curve within the three-dimensional MDS space. A party-embedding is ***two-dimensional*** if it can be roughly approximated by a surface in the three-dimensional MDS space.

Definition 5.3. Each j-th component of the Multidimensional Scaling results is also known technically as the j-th ***eigenfunction***, the value of which is given by the j-th eigenvectors multiplied by the square-root of the j-th eigenvalue $\sqrt{\lambda_j}v_j \in \mathbb{R}^{n\times 1}$.

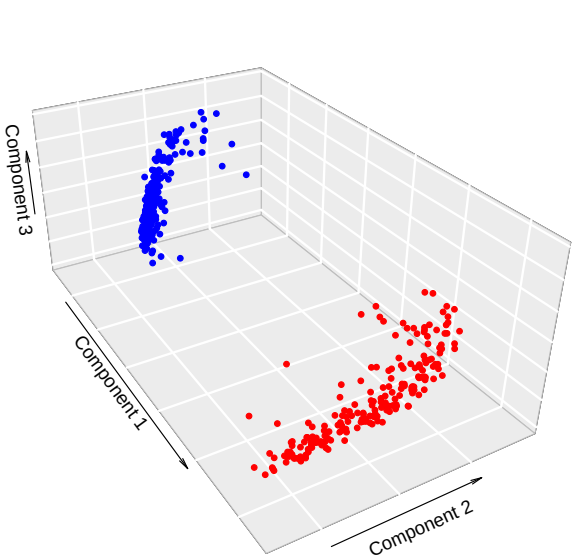
5.1 US House of Representatives

Recall from Section 3.1.4 that the proportion of dissimilarities can be measured from the eigenvalues. In the US House of Representatives, the first component is usually sufficient to explain most of the dissimilarities, which occurred normally between parties. However, since we want to capture intra-party dissimilarities in detail, the second and third component are also portrayed.

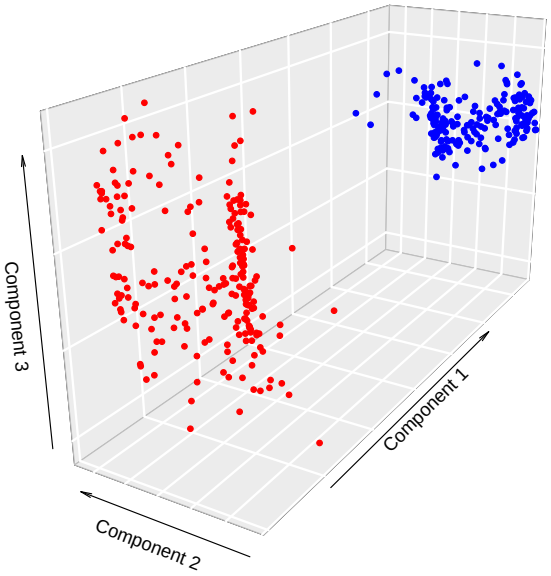
We observe three general patterns in the MDS of the 103rd to 116th House of Representatives.



(a) "Partial Horseshoes" — 103rd House of Representatives



(b) "Twins Horseshoes" — 110th House



(c) "Twins Surfaces" — 115th House

Figure 3: MDS on United States House of Representatives

1. **Twins Horseshoes**: Refer to Figure 3b.
   This pattern occurred from the 108th to 112th House. The pattern is observed as described in Diaconis et al. (2008), where each party is represented by a one-dimensional "horseshoe" at opposite ends of the first component. When visualized in the 2D plot of the second and third components, the two party embeddings are orthogonal to each other, and each embedding only spreads out in one component.

2. **Partial Horseshoes**: Refer to Figure 3a.
   This pattern occurred prior to the 108th House. Each party embedding is represented by a one-dimensional "partial horseshoe". For example, the Republican embedding

begins at one end of the first component and third component, but gradually moves towards the middle of the first component as it transverses along the third component, resembling an arc of a quarter-circle in the first-second component plane. The conjoining of both embeddings towards the middle of the first component represents less separation compared to the Twins Horseshoes. The two party embeddings are also orthogonal to each other in the second and third components.

3. **Twins Surfaces**: Refer to Figure 3c.
   This pattern occurs from the 113th House onward. The two embeddings are still well-separated in the first component, as per the "Twins Horseshoes" pattern. However, there is a loss of orthogonality in the second and third component. For both parties, the points are spread out in both axes, and the embeddings are two-dimensional instead of one-dimensional. Such patterns propose the invalidity of the assumption by Diaconis et al. (2008), that of uniform and independent one-dimensional distribution of ideologies for each party. Indeed, this suggests increasing fragmentation within both parties, especially for the Republican party with a larger embedding spread. Further discussion is presented in Section 8.

5.2 US Senate

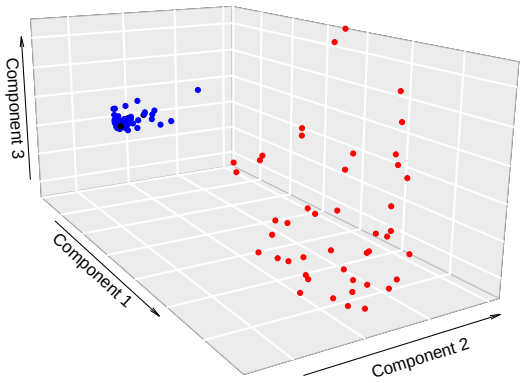
Within the US Senate, we witness less separation in the first component compared to the House of Representatives. The MDS results are more consistent across time too, since only a third of the Senate is up for re-election at the end of every session, as compared to the House. However, a central issue is the small number of observation ≤ 100, which create difficulties in trying to determine a pattern at times.

We observe two patterns in the MDS results, Most sessions are either one of these, or fall somewhere in between the two patterns.

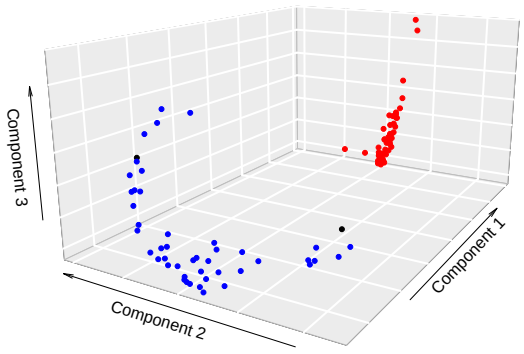
1. **Partial Horseshoes**: As previously discussed. E.g. 104th Senate.
2. **Point Representation**: Refer to Figure 4a.
   In several terms like the 108th and 113th Senate, we observed a pattern where either one of the party embedding is clustered into nearly a single point, at almost the extreme end of the first component. This is an indication of party unity, and normally occurs when that party is the majority with no need for compromise.

In the 115th Senate, we notice a noteworthy trend as portrayed in Figure 4b and 4c. The democratic party embedding is one-dimensional, and resembles a *partial horseshoe* taking a range of values in the first and second component. However, it *also* forms a horseshoe shape in the second and third component.

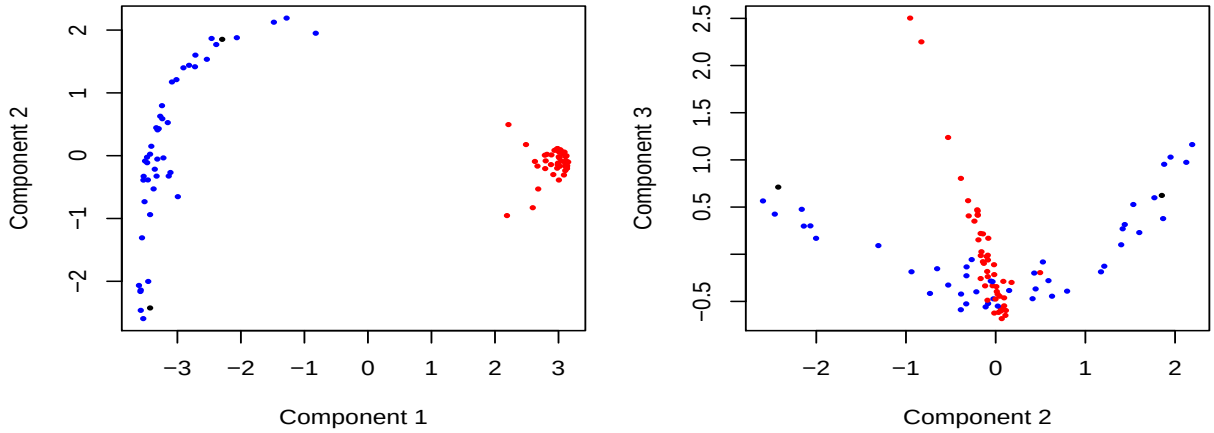
A closer look reveals that the ends of the embedding are *Bernie Sanders*, a democratic socialist, and *Joe Manchin*, a moderate conservative Democrat. Hence, it is possible that while the second component explains the variance of ideologies within the Democratic Party, the third component further determines how far a legislator deviates from the party agenda



(a) "Point Representation" — 113th Senate



(b) 115th Senate — 3D Plot



(c) 115th Senate — 2D Plot

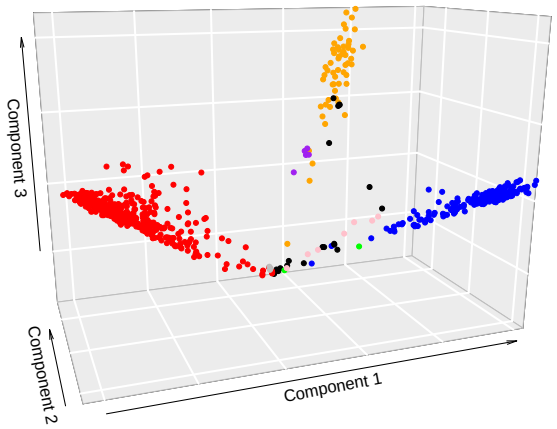
Figure 4: MDS on US Senate

for the Democratic (and Republican) party, since most legislators are clustered towards one end of the third component.

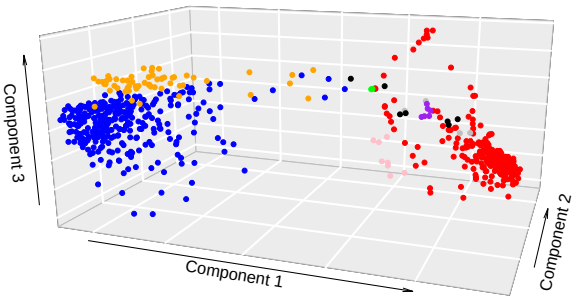
5.3 UK House of Commons

We point out three patterns from the MDS results.

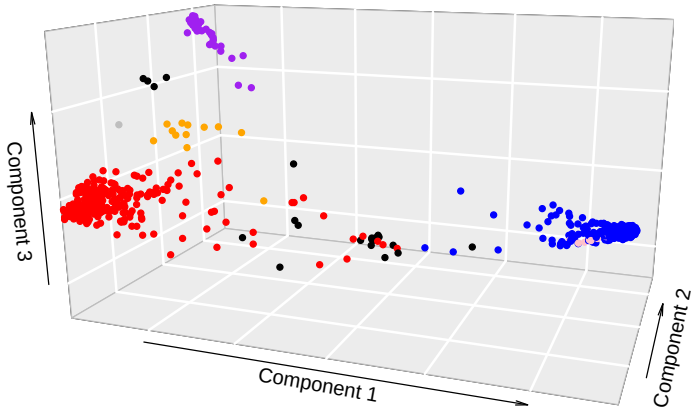
1. **Line Representation**: Refer to Figure 5a.
   This pattern is the most common across all UK parliamentary terms. Each party is represented by a one-dimensional embedding almost conjoining in the middle of the subspace. The first component captures the differences between the ruling party and the opposition, while the second component reveals some intra-party variance, and the attendance rate of the two dominant parties. The third component separates the



(a) "Line Representation" — 53rd House (2001)



(b) "Coalition" — 55th House (2010)



(c) "Post-Brexit" — 57th House (2017)

Figure 5: MDS on United Kingdom House of Commons

minority party, the Liberal Democrats, from the two major parties.
Notice that the embeddings are more structured and less spread out than that of the US Congress, and this is due to the stronger whip system preventing large deviation from party agenda.

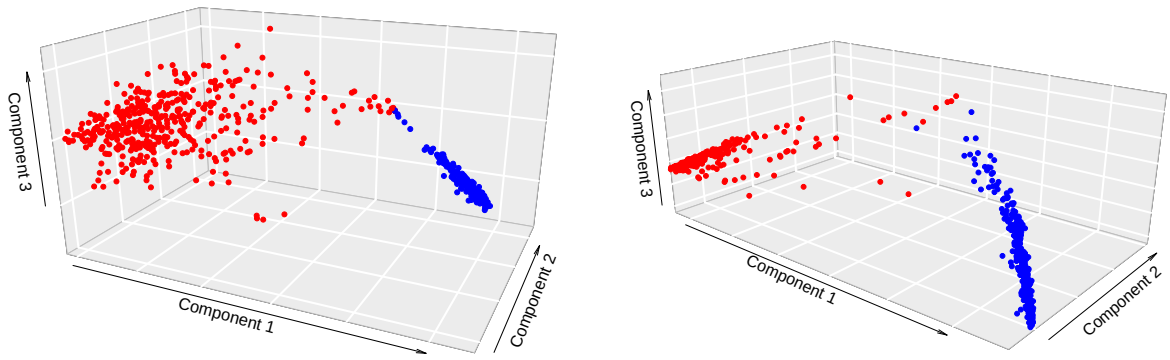
2. **Coalition**: Refer to Figure 5b.
   The plot of a coalition government between the LDem and Conservatives in the 55th House is slightly different from the previous terms. Observe that the LDem embedding lies just next to the Conservatives embedding. Although they are still distinct, the separation between them (captured in the third component) is negligible, compared to the "Line Representation" pattern.
   Indeed, this supports the strong effect of party affiliation and alliance on the voting patterns of legislators.

3. **Post-Brexit Government**: Refer to Figure 5c.
   From 2015 onwards, SNP took over LDem as the third largest party in parliament. Observe that while there are still some resemblance to the "Line Representation" pattern, more polarization is observed too, as there is no clear consensus between parties on the best approach to Brexit. Legislators are no longer uniformly distributed along their one-dimensional party embedding but have moved towards the extreme, and this is starting to resemble the pattern observed in Canadian Parliament, under Figure 7a.

5.3.1 UK Two-Party Dataset

We also performed MDS on the subset containing only legislators from the top two major parties (Conservative and Labour). We observe for all MDS plots, that the Labour Party is in general more fragmented and disunited than the Conservative Party, as discussed in Barone (2015).

In addition, we notice that the one-party-only MDS plot of the 57th Parliament reveals a twins horseshoe shape, where the party embeddings are well-separated in the first component, and orthogonal in the second and third component plane.



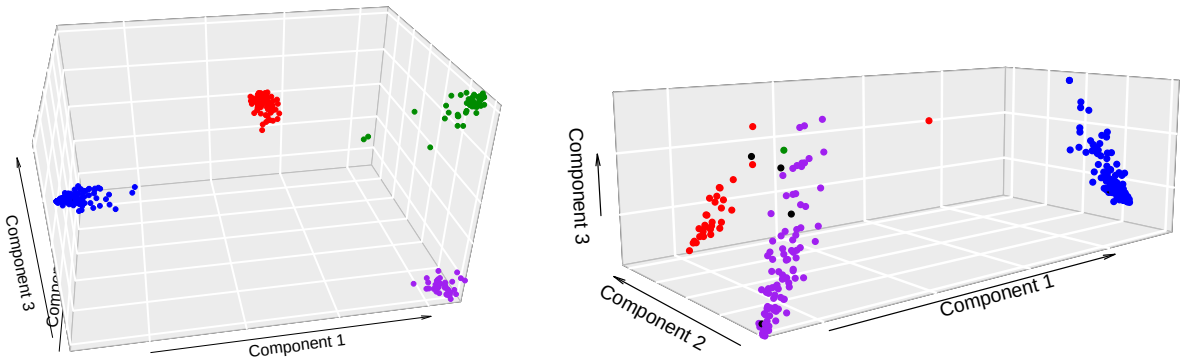
(a) 52nd House (1997) (b) 57th House (2017)

Figure 6: MDS on UK (Conservatives and Labour only)

5.4 Canadian House of Commons

Recall that Canada has four main parties — Liberal Party (Lib), Conservative Party (Con), New Democratic Party (NDP), and Bloc Québécois (BQ). There is not much variation in our MDS results across time, and we mainly obtain two patterns.

1. **Clustered Representation**: Refer to Figure 7a.
   For most sessions, the MDS neatly separated each party in the three-dimensional space. Recall from Section 3.3 about the parallel between Metric MDS and Spectral Clustering. Here, the first component separates the ruling party from the opposition; the



(a) 40th-II Parliament

(b) 41st-II (B) Parliament

Figure 7: MDS on Canadian House of Commons

second component separates the two dominant parties from the smaller parties; the third component distinguishes the two smaller parties from each other. Since the first three eigenvalues are utilized for the proper separation of parties, the plot does not reveal the intra-party distribution very well. Note that the clear separation suggests that some degree of polarization exists within Canadian politics.

2. **Line Representation**: Refer to Figure 7b.
   The plot occurs only for the 41st Parliament, where there are only three significant parties, with NDP as the official opposition. With one less significant party, the third component now reveals the intra-party distribution for all three parties. This bears some resemblance to the "Line Representation" in Figure 5a, except the parties are more separated here than in the United Kingdom in 2001.

Methods to visualize the intra-party distribution are discussed in Section 6, where we attempt to generalize Diaconis et al. (2008)'s model to a multi-party system.

6 Generalization to Multi-Party Systems

Most literature revolves around the US two-party Congress, and little have focused on explaining the behaviors of multi-party legislatures with more than two parties. In this section, we attempt to create a model for multi-party legislatures by making assumptions about the distribution of legislators' ideal points within and across parties, and in the process understand the challenges faced in such an endeavor.

6.1 Interpretation of Eigenfunctions

As the MDS results from the US two-party Congress reveal, with increasing partisan polarization in the last two decades, the first component is responsible for party separation, while subsequent components explain the intra-party distribution of legislators. Recall that the eigenvalues are sorted. Thus, the first component always determines the largest proportion of dissimilarities in the observation, which in majority of the cases can be explained by the legislators' political party.

In our results, we observe a parallel between two-party and multi-party systems in terms of interpretations for eigenfunctions. Notably, one may consider each eigenfunction to primarily explain either:

1. **Separation of Parties**: If there are s significant parties that are well-separated, then around $s-1$ components shall determine the full separation of all clusters, as suggested in the Spectral Clustering method (Section 3.3).
2. **Intra-Party Distribution**: For example, a party embedding's horseshoe shape.

Analysis of the U.K.'s and Canada's MDS results suggests the following interpretations:

1. United Kingdom
   (a) **1st Component**: Separation between ruling party and non-majority parties
   (b) **2nd Component**: Intra-party variance (correlated with participation rates)
   (c) **3rd Component**: Separation of minor party from the 2 major parties
2. Canada
   (a) **1st Component**: Separation between ruling party and non-majority parties. May also be viewed as the left-right spectrum.
   (b) **2nd Component**: Separation between the major parties and non-major parties
   (c) **3rd Component**: Separation between the minor parties

It is not surprising that the biggest dissimilarities are between the majority party and other non-ruling parties, given that the majority party controls the legislative agenda and has to

face frequent opposition to the bills it puts forward. Similarly, the second biggest dissimilarities are between the major parties (e.g. Conservatives and Liberals), and non-major parties e.g. Liberal Democrats, since major parties are prepared to control the government and have a set of more traditional policies compared to others.

Unfortunately in this case, since the top few eigenfunctions typically account for party separation, it may be challenging to visualize the differences within a party in the three-dimensional MDS plot. We are originally able to perceive the intra-party distribution in the earlier UK parliaments. However, with increasing partisan polarization, majority of voting outcomes are simply based on party affiliation, and UK's MDS results began to converge towards one that looks similar to Canada's plot.

6.2 Model and Simulation

Diaconis et al. (2008) came up with a theoretical model to account for the "Twins Horseshoe" shape that appeared in the US House of Representatives. They made the following assumptions:

- The ideal points of all legislators within a common party is uniformly distributed on a one-dimensional interval $[0, 1]$. Hence, if legislators are ordered on that interval from $1, 2, \dots n$, the distance between legislator i and j is given by

$$\widehat{d}(i, j) = |i/n - j/n| \in [0, 1], \tag{6.1}$$

  where $1 \leq i, j \leq n$.

- The information is only locally accurate. Thus, the dissimilarity between legislator i and j is defined by $1 - \exp(-\widehat{d}(i, j))$, and we obtain the dissimilarity matrix

$$\Delta_n = \begin{bmatrix} 0 & 1 - e^{-1/n} & \dots & 1 - e^{-(n-1)/n} \\ 1 - e^{-1/n} & 0 & \dots & 1 - e^{-(n-2)/n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 - e^{-(n-1)/n} & 1 - e^{-(n-2)/n} & \dots & 0 \end{bmatrix}.$$

- Each party is well separated from each other such that the dissimilarity between any 2 legislators from different parties is 1. It is possible for the distance to be a constant 1 for all pairs of legislators from different parties, since we employed a localization method by using the kernel function in our analysis — if the vote differences are already large, any marginal change does not affect the dissimilarity. From this, we obtain

$$\Delta_{2n} = \left[\begin{array}{c|c} \Delta_n & \mathbf{1} \\ \hline \mathbf{1} & \Delta_n \end{array}\right] \in \mathbb{R}^{2n \times 2n},$$

  where $\mathbf{1}$ is the $n \times n$ matrix of 1.

We attempt to generalize this model to a system with more than 2 parties, before relating the model to our dataset. We chose to focus on modelling the Canadian Parliament since it has a larger number of significant parties (4) and less whip influence than the UK Parliament.

6.2.1 Extension from Two-Party Model

First, let us assume there are four parties of equal size n, the parties indexed by p, r, b, g, such that the whole legislature has $4n$ legislators. Suppose that within each party, the legislators are uniformly distributed on a one-dimensional space as described in Equation (6.1), and their dissimilarities are represented by Δ_n. The four parties are given the colors of purple, red, blue, and green, and are ordered accordingly in the dissimilarity matrix $\Delta_{1,4n}$.

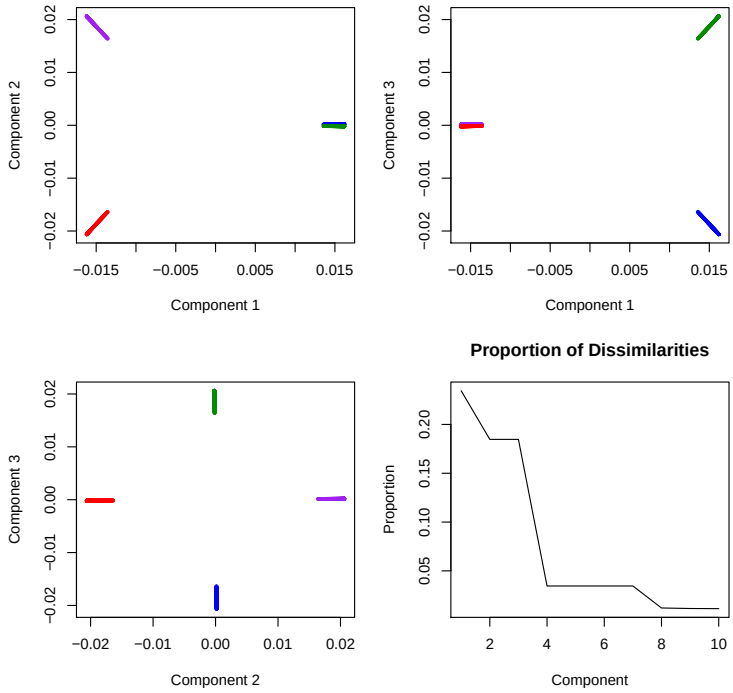


Figure 8: 2D Plots, MDS Simulation of Model 1

Next, assume that each party is a separate population, and they are located far enough such that the dissimilarity between any 2 legislator i and j, where $i \in$ Party a, and $j \in$ Party b is a constant c_{ab}. We further suppose that the dissimilarities between various parties vary such that c_{ab} is a function of Party a and b:

$$\kappa = \delta(p, r) = \delta(g, b) < \delta(p, g) = \delta(r, b) = \delta(p, b) = \delta(r, g) = \gamma.$$

We then obtain the theoretical dissimilarity matrix

$$\Delta_{1,4n} = \left[\begin{array}{c|c|c|c} \Delta_n & \boldsymbol{K} & \boldsymbol{\Gamma} & \boldsymbol{\Gamma} \\ \hline \boldsymbol{K} & \Delta_n & \boldsymbol{\Gamma} & \boldsymbol{\Gamma} \\ \hline \boldsymbol{\Gamma} & \boldsymbol{\Gamma} & \Delta_n & \boldsymbol{K} \\ \hline \boldsymbol{\Gamma} & \boldsymbol{\Gamma} & \boldsymbol{K} & \Delta_n \end{array}\right] \in \mathbb{R}^{4n \times 4n},$$

where $\boldsymbol{K}$ is the $n \times n$ matrix of κ, and $\boldsymbol{\Gamma}$ is the $n \times n$ matrix of γ. We find $S = -\frac{1}{2} H \Delta_{1,4n} H$, and solve the eigen-problem to determine the top 3 components of the MDS solution. In our

simulation, we let $4n = 1000$, $\kappa = 1$, and $\gamma = 1.1$. The results of Model 1 are presented in Figure 8.

From the plot, observe that the first eigenfunction separates Party p and Party r from Party g and Party b, since their dissimilarities of $\Gamma = 1.1$ are the largest. Next, the second and third component separates Party p from Party r, and Party b from Party g respectively. Notice that the second and third component explains the same proportion of dissimilarities, since the dissimilarity between each pair of parties are the same, at $K = 1$. Also note that given 4 significant parties, we require all first three components to be able to fully separate them, as discussed in Section 6.1.

Recall that in Diaconis et al. (2008), the second and third component explain the same proportion of dissimilarities too, where each component accounts for the intra-party distribution of only one party. In this four-party model, the fourth to seventh component also accounts for the intra-party distribution of each party. Given that within each party, for all 4 parties, the legislators are distributed uniformly, we noticed that the fourth to seventh component all explain the same proportion of dissimilarity.

6.2.2 Theoretical Models and Empirical Results

It is possible to produce a theoretical model which resembles some of our empirical MDS results, by modifying the inter-party dissimilarities. For our example, recall the MDS results of the 40th Canadian Parliament Session III, as depicted in Figure 9.

By varying the inter-party dissimilarities of our theoretical model, we are able to create a simulation with a MDS output that relates to the empirical data. First, let us list the observations in the order of **Con**, **Lib**, **BQ**, **NDP**, which refers to the Conservative Party, Liberal Party, Bloc Quebecois, and New Democratic Party respectively. Based on that particular year's parliamentary makeup, we let the simulated number of legislators from each party, $n_{con} = 450$, $n_{lib} = 250$, $n_{BQ} = 150$, $n_{NDP} = 150$, such that $\sum_s n_s = N = 1000$.

We continue to assume that the intra-party distribution is uniform on an uni-dimensional interval — note that in this situation, the intra-party distribution does not affect our MDS simulation significantly since we only plot the first 3 eigenfunctions — more discussion on this is available in Section 6.3.

Next, we assume the inter-party dissimilarities accordingly — As the majority party, the Conservatives' dissimilarities with other parties are the largest - but its dissimilarities with the other major party, the Liberals, are smaller than that with Bloc Quebecois, and New Democratic Party.

$$\delta(\text{Con, Lib}) = 1.2$$

$$\delta(\text{Con, BQ}) = \delta(\text{Con, NDP}) = 1.5$$

The dissimilarities between Bloc Quebecois and New Democratic Party, both minor parties, are the smallest at

$$\delta(\text{NDP, BQ}) = 1$$

$$\delta(\text{Lib, BQ}) = \delta(\text{Lib, NDP}) = 1.1$$

This forms a dissimilarity matrix of

$$\Delta_N = \left[\begin{array}{c|c|c|c} \Delta_{n_{Con}} & \mathbf{1.2} & \mathbf{1.5} & \mathbf{1.5} \\ \hline \mathbf{1.2} & \Delta_{n_{Lib}} & \mathbf{1.1} & \mathbf{1.1} \\ \hline \mathbf{1.5} & \mathbf{1.1} & \Delta_{n_{BQ}} & \mathbf{1} \\ \hline \mathbf{1.5} & \mathbf{1.1} & \mathbf{1} & \Delta_{n_{NDP}} \end{array}\right] \in \mathbb{R}^{N \times N}.$$

This implies that the Conservative Party is the furthest away from every other party, followed by the Liberal Party, while NDP and BQ have the smallest inter-party dissimilarity. We run the simulation and obtain the MDS results displayed in Figure 10.

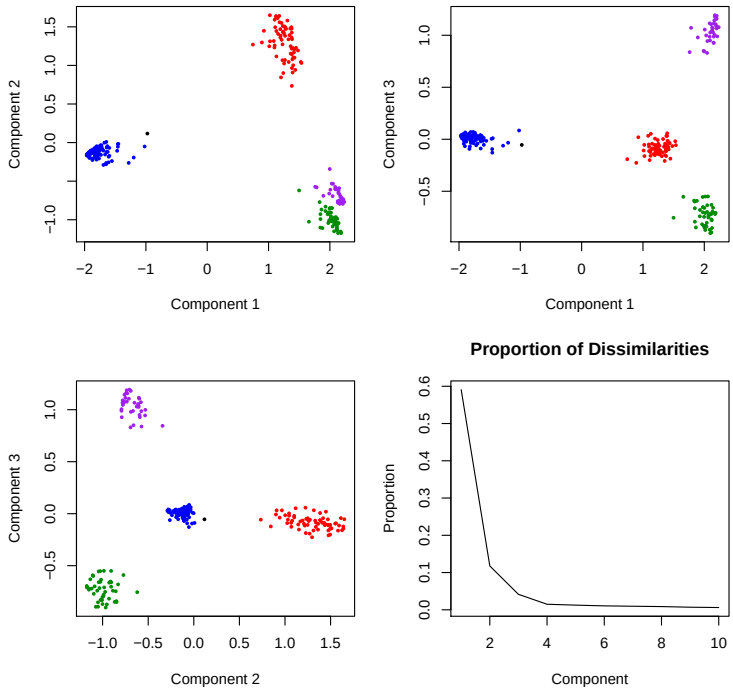


Figure 9: 2D Plot, MDS of Canadian 40th Parliament Session III

Note the resemblance with the empirical results from Figure 9, both for the party embeddings in lower-dimensional space, and the proportion of dissimilarities each component accounts for.

By modifying the inter-party dissimilarities in our theoretical models, we are able to simulate and create MDS results that roughly resemble our empirical results for other parliamentary sessions too, but their discussions are omitted for brevity.

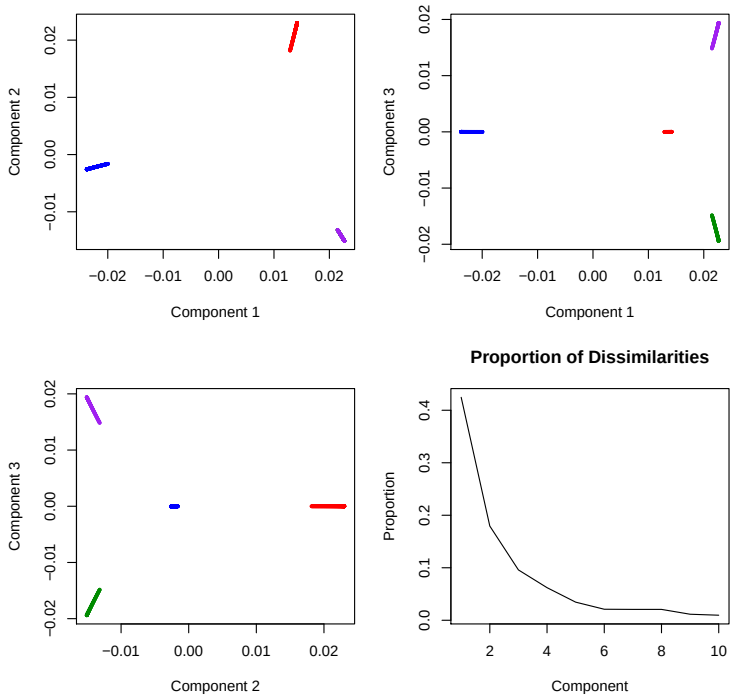


Figure 10: 2D Plot, MDS Simulation of Model 2

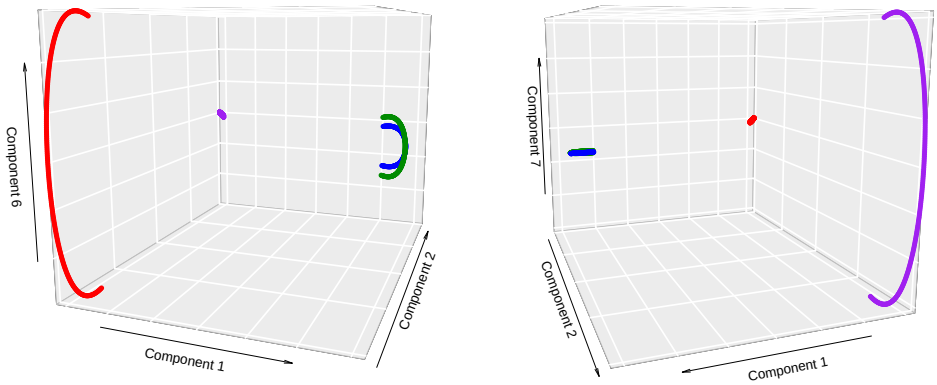
6.3 Capturing Intra-Party Variances

If a legislature has more than three significant parties that are well-separated in a polarized political climate, then the three-dimensional MDS plot may only reveal the inter-party splits, and not the intra-party variances, as demonstrated by the sharp drop in the proportion of dissimilarities after the third component for all the models.

Nonetheless, it is possible to depict the intra-party distribution of legislators if we utilize the fourth to seventh eigenfunction. Refer back to Model 1 described in Section 6.2.1. Due to the assumption that the legislators within a party are uniformly distributed, we are able to reproduce the horseshoe shape if we made a plot of the first, second, and either the fourth, fifth, sixth or seventh component, where each choice reveals the horseshoe shape for only one party, as depicted in Figure 11.

Appearance of the horseshoe shape in our theoretical model proposes that even if the three-dimensional plot reveals each party being in a small cluster, it is possible to obtain and determine the intra-party variance through other eigenfunctions, if these intra-party variances are significant enough.

In our informal work, we explored three-dimensional plots containing eigenfunctions beyond the third. Interesting results include a party embedding being further clustered into sub-structures, or the appearances of horseshoe structures, both which appeared in the MDS plot of the Canadian 42nd Parliament as shown in Figure 12, and also the 41st Parliament



(a) Model 1: Component 1,2,6 (b) Model 1: Component 1,2,7

Figure 11: Model 1 Simulation MDS Plot

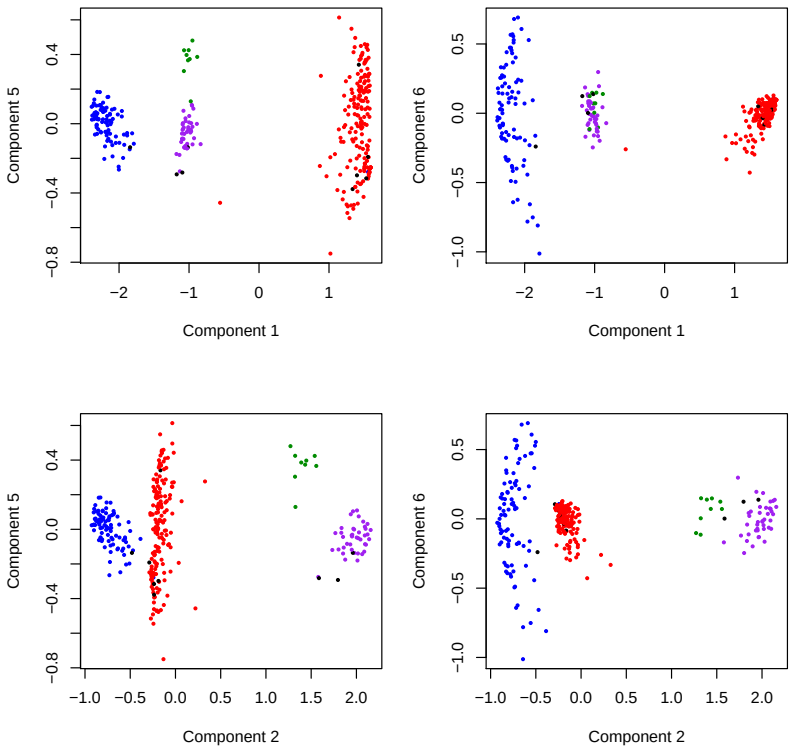


Figure 12: 2D MDS Plot, 42nd Parliament (A)

Session I under a different ruling party. However, spotting structural patterns in the fourth to seventh component is challenging as these dimensions no longer make up a large proportion of dissimilarities compared to the first to third component, and thus are subjected to more variations. The investigation of eigenfunctions after the third is not within the scope of this paper, but may be a future extension.

7 Political Polarization

Multiple indexes exist for measuring the level of polarization within a legislature. For example, NOMINATE Scale's first dimension is a good measure of polarization on the left-right spectrum, while Azzimonti (2013)'s index quantifies polarization monthly by determining the frequency of newspaper articles reporting political disagreement about government policy.

7.1 Measuring Polarization via Multidimensional Scaling

Our MDS results also enable us to determine the political polarization of a legislature across time. First, we consider the ***Proportion of Dissimilarities accounted for by the First Component***. Recall that the first eigenfunction separates the majority party from other non-ruling parties, and is often indicative of a legislator's position within the traditional left-right spectrum. If this proportion is large in our MDS results, it implies that most of the dissimilarities between members in the legislature are based on party affiliations.

Next, we consider the ***Differences in Values of the First Component***. For this, we measure the difference in the first eigenfunction value between either the **median**, **20th percentile**, or **90th percentile** of the majority party, and the non-ruling party. We intend for the 90th percentile to measure the extreme members, and the 20th percentile the moderates.

We expect the 90th-percentile differences (termed ***Extreme Differences***) to change at a slower rate across time, since extreme positions are not likely to deviate much. On the other hand, if the 20th-percentile differences (termed ***Moderate Differences***) increase, this implies more polarization since there are large dissimilarities even between the moderate legislators of both parties. If the range between the *Moderate Differences* and *Extreme Differences* decreases, it also indicates polarization where parties are moving towards the extreme in their voting records.

7.2 Polarization in the United States

We shall first analyze political polarization in two-party systems, specifically the US Congress. To provide a stronger contrast, we begin with the MDS results from the 84th to 86th House of Representatives in the 1950s. The 2D MDS Plot and Proportion of Dissimilarities by Components of the 85th Congress are displayed in Figure 13.

Observe from the figure, that based on the strong overlap between Republicans and Democrats in the first and second components, no clear separation can be made from just one dimension alone; one requires both components to create a dividing plane that separates the two parties, suggesting little polarization. Likewise, the "Proportion of Dissimilarities" plot reveals a very gentle drop from the first to the second component, indicating the second component is almost as important as the first in explaining the dissimilarities. Indeed, from the 84th to 86th Congress, the ratio of the proportion explained by the first to the second component

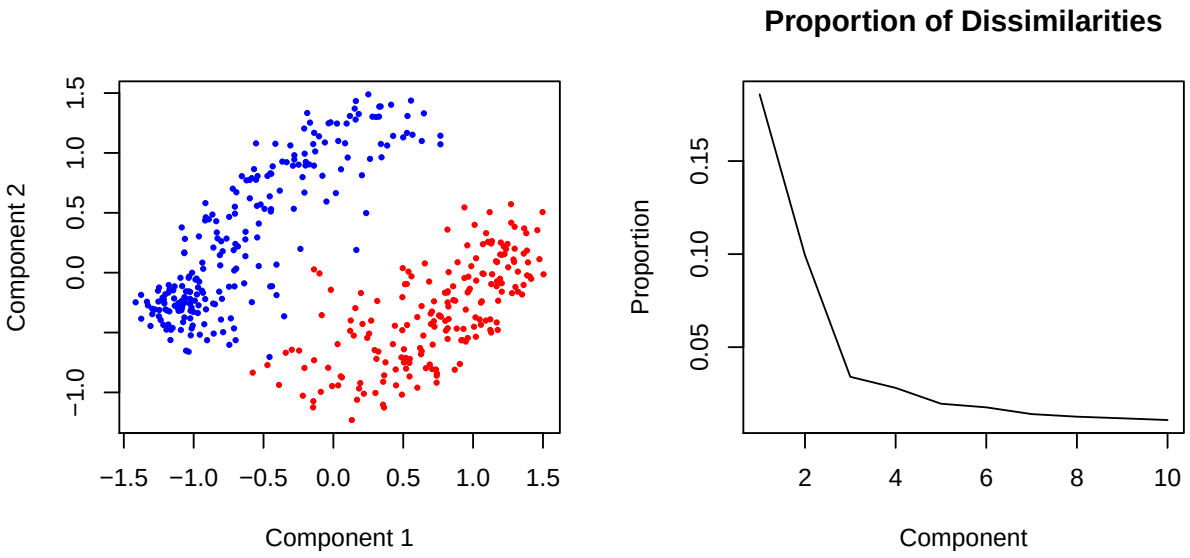


Figure 13: 2D MDS Plots, 85th House of Representatives

(0.19/0.08, 0.18/0.1, and 0.24/0.1 respectively) is just around 2, compared with a ratio of over 10 in recent congresses.

Lamentably, our MDS results reveal an increasingly divided political arena. Since the 103rd Congress, the overlapping between both parties in the first component gradually decreased until the 2 parties can be perfectly separated in solely the first component with a cutoff point — refer to Figure 16c.

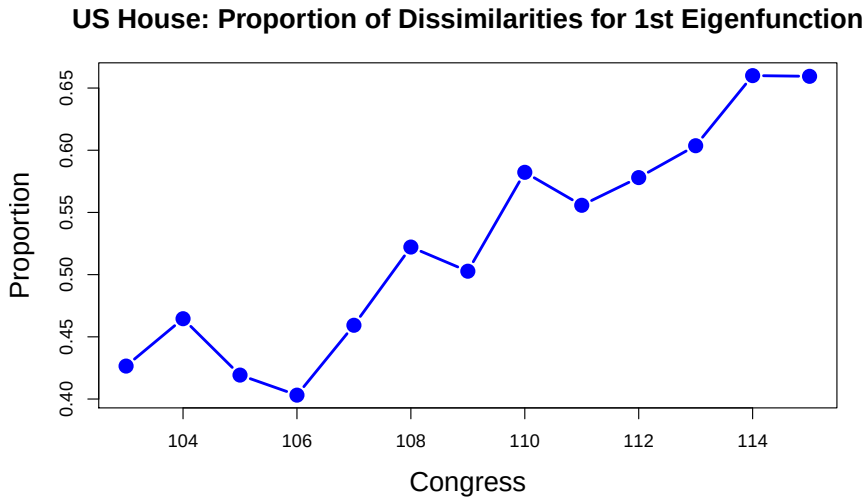


Figure 14: Proportion of Dissimilarities explained by First Component over Time

Indeed, observe from Figure 14 that the First Component takes up a larger proportion of

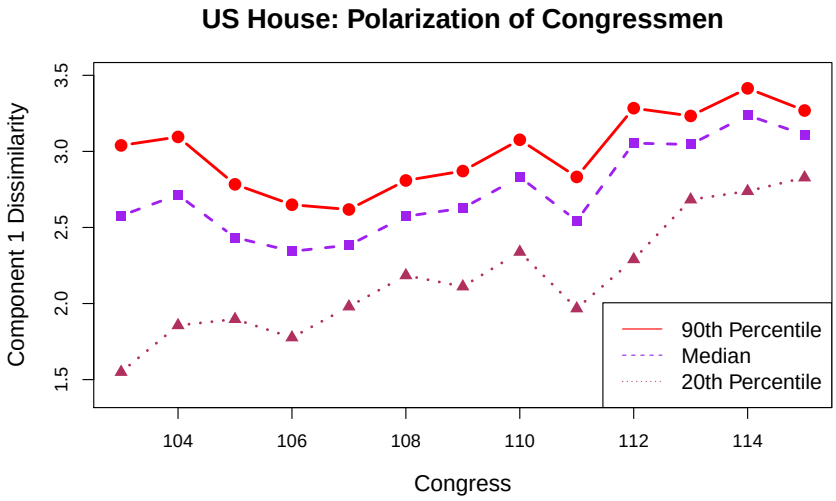
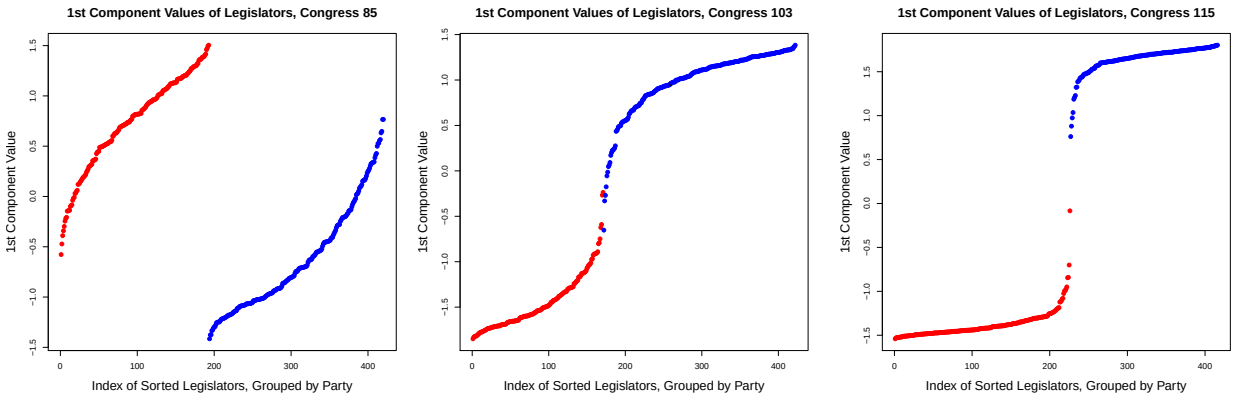


Figure 15: Party Differences in Values of First Component, over Time

dissimilarities with each congress. In addition, from Figure 15, the Moderate, Median, and Extreme differences are also increasing with time, demonstrating trends of growing party polarization over the years, as justified in Section 7.1.

Trends of polarization are well-supported even in our MDS structures. In the 1990s Congress, we observe the "Partial Horseshoes" structure in our MDS, where their bending towards the middle indicates overlapping ideologies, but Diaconis et al. (2008)'s twin horseshoe shapes support a legislature with two separate populations.

To further support the phenomenon, for selected Congress (85th, 103rd, 111th, and 115th), we divide the legislature into each respective parties, before sorting the legislators based on their first eigenfunction. The plots of the first eigenfunction against the ranked legislators for these selected congresses are available in Figure 16.

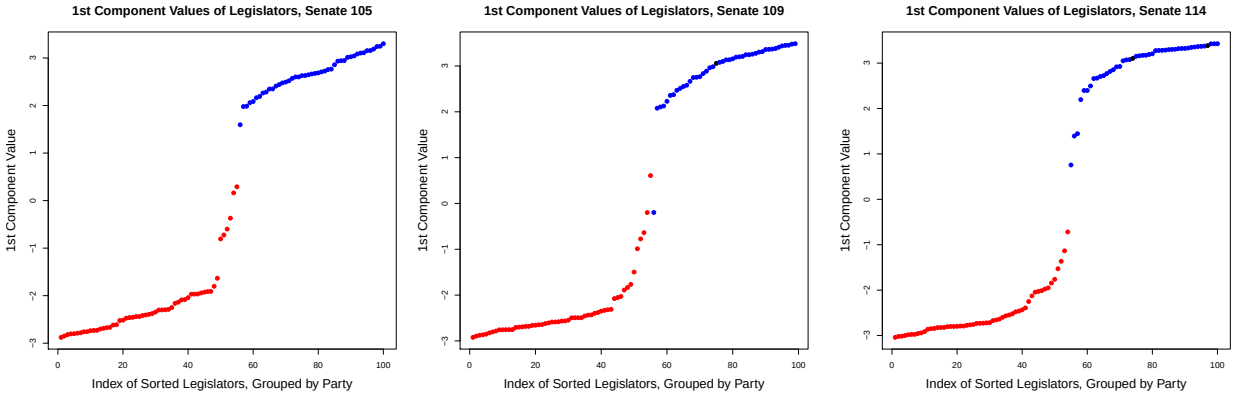


(a) 85th House (b) 103rd House (c) 115th House

Figure 16: First Eigenfunction Value against Ranked Legislators (House)

Observe that in the earlier congresses, the range of values the legislators took is very large for both parties, and there are significant overlaps between the two parties. However, in recent congresses, there are less legislators in the middle, and each party takes a smaller range of values in the first eigenfunction, indicating members from the same party are more likely to vote alike, at least on the spectrum dividing both parties. Notice the huge separation between both parties in the 115th Congress. The gradient of the points are nearly constant until it approaches near infinity in the middle, when there is a switch in party. In past congresses with less polarization, the gradient of the points changes at a more constant rate compared to that of the 115th Congress.

We compare the same plots of the first eigenfunction against the ranked legislators for the 105th, 109th, and 114th Senate, as displayed in Figure 17.



(a) 105th Senate (b) 109th Senate (c) 114th Senate

Figure 17: First Eigenfunction Value against Ranked Legislators (Senate)

Recognize that the range of values for the first eigenfunctions are more consistent over time compared to the House of Representatives, with less degree of polarization than the house in more recent congresses.

As Ellenberg (2001) discussed, legislatures become more polarized not when individual politicians adopt extreme views, but when they are unseated by more extreme politicians — polarization is an effect of replacement, not conversion. Since only one-third of the senate changes every election, this allows for more consistency across congresses and a slower rate of polarization.

7.3 Polarization in Multi-Party Systems

Through the Multidimensional Scaling analysis, we similarly discover the trend of growing polarization over time in both United Kingdom and Canada.

7.3.1 United Kingdom

Like before, we plot the "Proportion of Dissimilarities" and "'Party Differences by First Component Values", as described in Section 7.1. Notice that the proportions are strictly increasing over time; the indicators "Moderate Differences" and "Extreme Differences" are increasing too, indicating growing polarization in the UK Parliament since 1997.

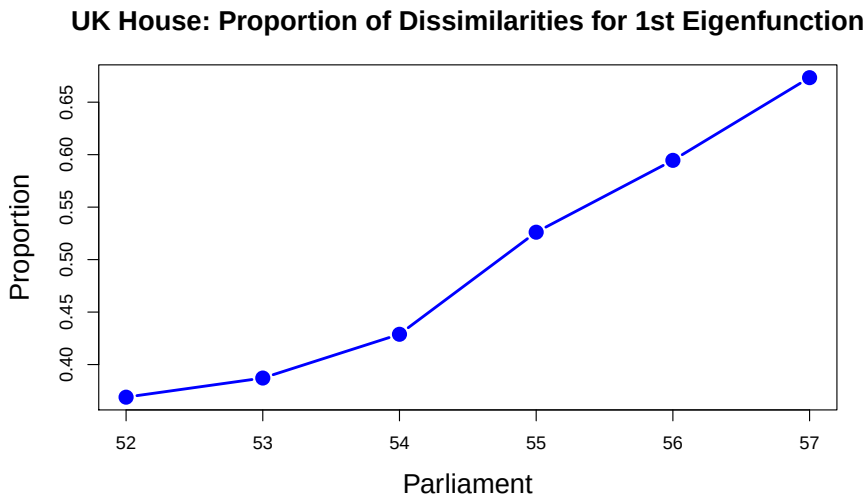


Figure 18: UK Proportion of Dissimilarities explained by First Component over Time

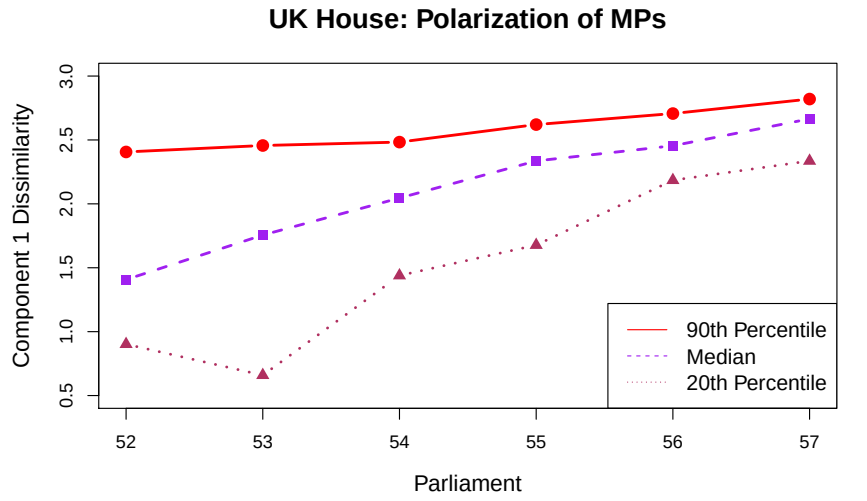
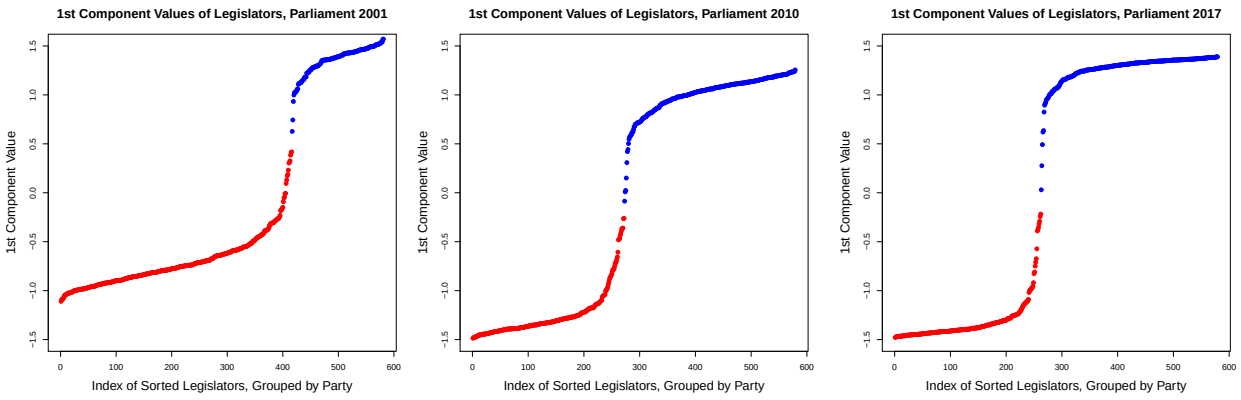


Figure 19: UK Party Differences in Values of First Component

We display the first eigenfunction against the ranked legislators for the 53rd, 55th and 57th parliament in Figure 20, considering only legislators from the *Labour* and *Conservative Party* to enable a clearer distinction. We noticed the same trends as that of plots from the US House of Representatives, of increasing differences between legislators from each party over time,



(a) 53rd Parliament (b) 55th Parliament (c) 57th Parliament

Figure 20: First Eigenfunction Value against Ranked Legislators (UK)

implying rising polarization in the United Kingdom. This is because the two major parties have been bitterly divided since Brexit, and demands from regional parties in Scotland and Northern Ireland since their rise have made consensus harder to build.

7.3.2 Canada

We repeat the same analysis performed earlier for the Canadian Parliament, with results in Figure 21 and 22.

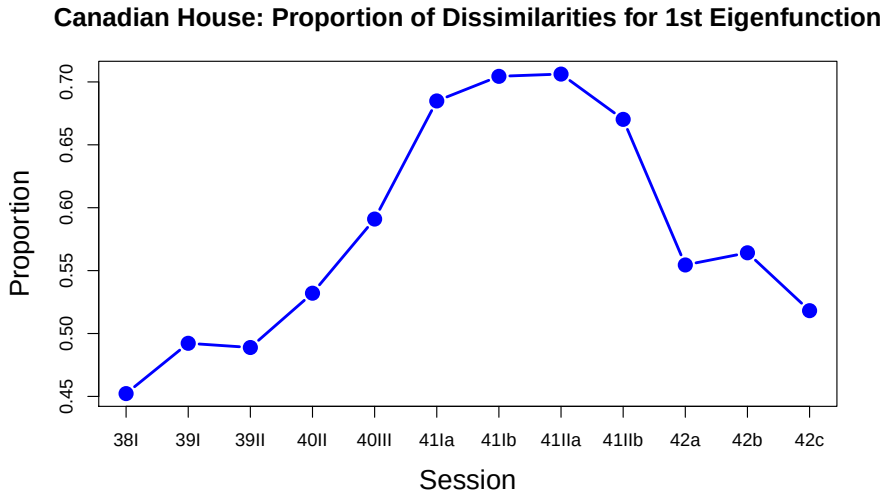


Figure 21: Canada Proportion of Dissimilarities by First Component over Time

The same trends of polarization in the US House of Representatives and UK House of Commons can similarly be observed in the Canadian House of Commons, at least up till the 41st Parliament. Then, there is a sudden reversal in the most recent 42nd Parliament

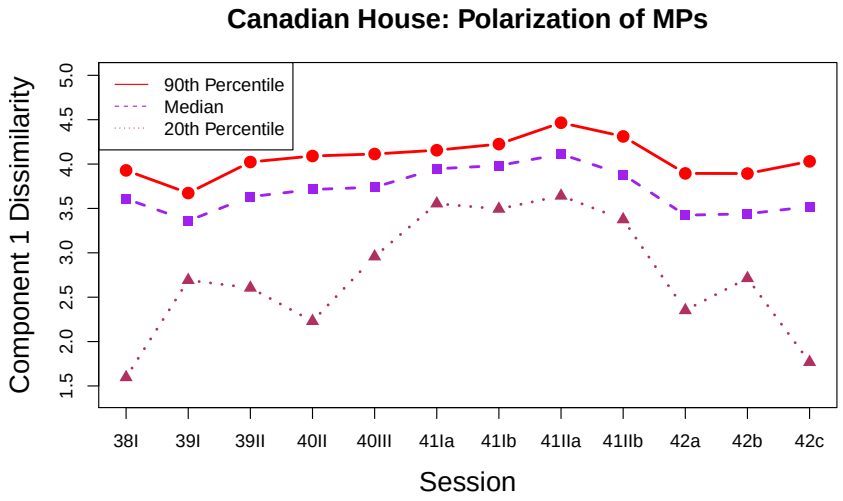


Figure 22: Canada Differences in Values of First Component

(controlled by the *Liberal Party*), where we noticed a drop in the proportion of dissimilarities by the first component, and a decrease in the "Moderate Differences".

Accounting for these patterns requires an understanding of Canadian Politics. The Conservatives and Liberals traditionally occupied the center, while the New Democratic Party occupied the left of the political spectrum. However, since the 2000s, the Conservatives were moving gradually towards the right. The 41st Parliament was seen as the culmination of this trend, as Canada witnessed the virtual disappearance of the Bloc Québécois, the rise of New Democratic Party as the second largest party, and the collapse of the Liberal Party into the third party. As Clarkson (2011) described, these events reflect the long-term decline of a centrist Liberal party, a reorientation of right-leaning voters clustering in the Conservative Party, and the left-leaning voters towards the New Democratic Party.

While the moderate middle is said to have disappeared in the 41st Parliament, the Liberal Party made a comeback in the 4nd Parliamentary elections, as detailed by Ives (2015). The control of government by a (more) centrist party reduces the parliamentary polarization, although how long-term this reversal will be remains to be seen.

8 Intra-Party Fragmentation

Contemporary political discussions frequently examined the divisions within political parties. Brooks (2016) mentioned how intra-party fragmentation has been driven by the weakening of traditional identities that are foundations of party structure, as other factors such as openness to globalization, nationality, and ethnicity become more relevant.

For instance, as Kamarck and Podkul (2018) described, the US Republican Party is now strongly divided into the (socially) Conservative, the Tea Party, and the Libertarian, whereas the Democratic Party is divided into the Progressive, Establishment, and the Moderates. Likewise, in a post-Brexit world, modern political opinions can no longer be described exclusively in terms of "left" and "right" spectrum. Bloodworth (2017) suggested that an "open vs. closed" issue space may better divide political ideologies of Britons than "left vs. right".

Our MDS results shall shed light on the types of fragmentation within parties. For example, will intra-party fragmentation lead to party split and the rise of a multi-party system? Friedman (2019) hypothesized that the United States may become a four-party system in the near future, given that many old binary choices no longer line up with challenges to citizens in an age of globalization, technology and climate change. A closer analysis of different types of intra-party fragmentation will provide useful knowledge about the future of the United States' political system.

8.1 Legislature Simulation: A Model

We refer specifically to the MDS three-dimensional plot of the US 113th to 115th House of Representatives. The plots demonstrate evidence of intra-party fragmentation — instead of being one-dimensional, both party embeddings are two-dimensional, termed "Twins Surfaces" in Section 5.1.

Attempts were made to reproduce the "Twins Surfaces" pattern by holding certain assumptions about the inter-legislator distances and creating a theoretical dissimilarity matrix, as we did in Section 6.2. However, informal work failed to reveal any findings.

Therefore, we performed a more fundamental simulation — making assumptions about the distribution of legislators' ideal points on a theoretical R-dimensional grid, simulating the roll call votes, and performing MDS on the simulated roll call votes and observing the patterns in the MDS three-dimensional plot. We discover that certain assumptions about the R-dimensional grid enable us to reproduce the same MDS patterns.

8.1.1 Assumptions and Procedure

For our simulation, we shall fix n, the number of legislators in the system, and p, the number of roll calls or bills being voted on. The outcome of each roll call is either $\{+0.5,\ -0.5\}$. For simplicity, we shall simulate only a two-party system, in which the seats are eventually divided between the two parties such that each party has $\frac{n}{2}$ seats.

We shall make the same assumptions as we did in Section 3.2.1:

- Each legislator i is represented by his/her ideal point, l_i. The ideal point exists in a R-dimensional grid, $[0,1]^R$.
- If $R > 1$, each dimension r in the R-dimensional grid is independent with respect to another dimension r'. Each dimension r represents an uni-dimensional issue space that is orthogonal to all other uni-dimensional issue spaces r'.
- The ideal points follow a distribution with c.d.f. $\mathcal{F}_\psi$ in $\mathbb{R}^R$ that is dependent on the party ψ. The component distributions of $\mathcal{F}_\psi$ are independent.
  - For example, if $R = 2$, then the first dimension may be the left-right traditional spectrum, while the second dimension measures attitudes towards NATO/European Union.
  - In this example, we can employ a *Normal Mixture* distribution in the first dimension, where each mixture component represents one party. In the second dimension, we can employ a *uniform* distribution on $[0,1]$. Each dimension of the ideal point are independently drawn.
- Every bill k can be represented by C_{rk}, where $r \in \{1,\ldots,R\}$ is a dimension, and $C_{rk} \in [0,1]_r$ is a cutoff point in that dimension r.
- The cutoff points for all p bills are uniformly distributed on the R dimensions, and for each dimension r, uniformly distributed on $[0,1]$. This uniformity is fixed for ALL simulations.
- During every bill k, a legislator's vote is determined by his *post-error* ideal point, $\tilde{l_{ir}} = l_{ir} + \epsilon \sim N(0,\sigma^2) \in \mathbb{R}$. If $\tilde{i_{ir}} > C_{kr}$, then his vote, $v_{ik} = +0.5$, otherwise $v_{ik} = -0.5$. The error ϵ is to account for the fact that legislators may sometimes be muddle-headed or be distorted by other incentives.
  - For example, for each bill, we first draw it from a multinomial distribution with parameter $p_r = 1/R$ for all $i \in 1,\ldots,R$, to determine which dimension r the bill is in.
  - Once we obtained r, we draw the cutoff point C_{kr} from $[0,1]$.
  - For each legislator i, we retrieve the r-th component of his/her ideal point l_i, find $\tilde{l_{ir}}$, and determine the vote results.

We shall fix some parameters for our simulation. For all simulations, $n = 500$, $p = 1000$, $\sigma = 0.05$; we only vary R, and the distributions of ideal points for parties.

8.2 Simulation Results and Discussion

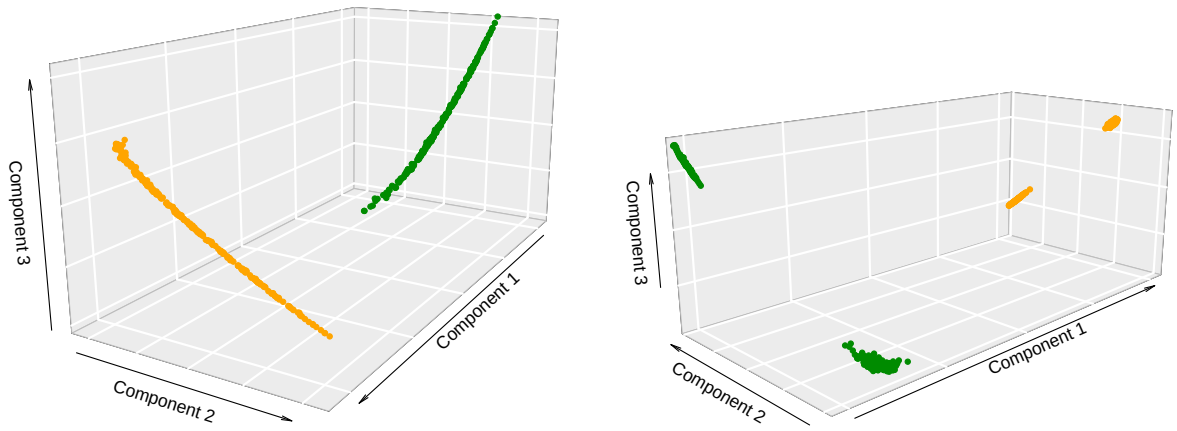
8.2.1 One-Dimensional Issue Grid

We try to obtain the various party embedding observed in the empirical results via the theoretical simulations, using the simplest model where $R = 1$. This is the case where all legislators' ideal points can be represented on a uni-dimensional spectrum, viewed traditionally as the left-right spectrum. To reflect party arbitrariness, the party a and b are colored orange and green in the plot.

By changing certain parameters, we are able to obtain the respective results:

Twins Horseshoe Pattern: MDS Plot depicted in Figure 23a. We choose a Normal distribution for the legislators' ideal points to reflect polarization on the left-right spectrum.

1. $R = 1$
2. Party a distribution $f_a = \mathrm{N}(0.1, 0.05^2)$
3. Party b distribution $f_b = \mathrm{N}(0.9, 0.05^2)$



(a) Twins Horseshoe

(b) Clustered Representation

Figure 23: Party Embeddings in MDS Results of Simulated Legislatures

Next, we consider the event where there is intra-party fragmentation on the uni-dimensional issue grid — we simulate this split by using a bi-modal distribution for each party, the Normal Mixture distribution. We obtain a MDS pattern resembling the "Clustered Representation" pattern observed in Canada's empirical results.

Clustered Representation Pattern: MDS Plot depicted in Figure 23b. Suppose party fragmentation occurs on the same uni-dimensional spectrum.

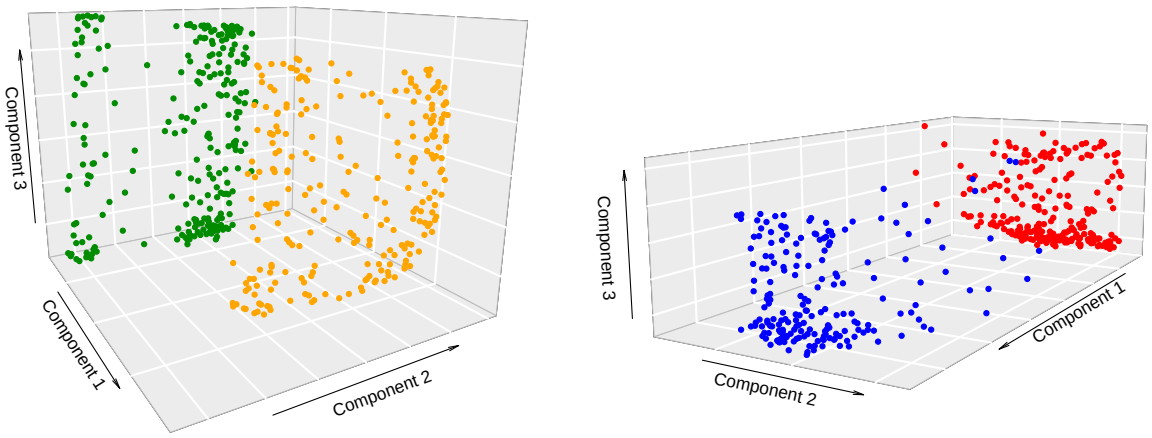
1. $R = 1$
2. Party a distribution $f_a = \pi_s \mathrm{N}(\mu_s, 0.01^2)$
   - $s = \{1, 2\}$, $\pi_1 = \pi_2 = 0.5$, $\mu_1 = 0.1$, $\mu_2 = 0.3$
3. Party b distribution $f_b = \pi_t \mathrm{N}(\mu_t, 0.01^2)$
   - $t = \{3, 4\}$, $\pi_3 = \pi_4 = 0.5$, $\mu_3 = 0.7$, $\mu_3 = 0.9$

Note that the resemblance to Canada's results suggests how party fragmentation on the uni-dimensional issue grid may lead to the rise of new political parties. Indeed, recall from Section 7.3.2 that in Canada, Conservative Party lies to the right of the political spectrum, New Democratic Party, and Bloc Québécois to the left, and Liberal Party in the center. Our simulation demonstrates the possibility for an intra-party split to occur on the uni-dimensional issue grid — in such cases, the legislature could take after a polarized multi-party system.

8.2.2 Multi-Dimensional Issue Grid

We attempted to simulate the "Twins Surfaces" embedding pattern observed in the 113th to 115th United States House of Representatives, but are unable to do so using just an uni-dimensional issue grid. We only obtain a simulation that resembles the pattern when we let $R = 3$.

Twins Surfaces Pattern: MDS Plot depicted in Figure 24a. Suppose party polarization occurs on the issue grid's first dimension labelled I, while intra-party fragmentation occurs on the second and third dimension labelled II, III.



(a) Simulation (b) Empirical MDS (113th US House)

Figure 24: Twins Surface Embedding in Simulated and Empirical Results

1. $R = 3$
2. Party a Dimension I distribution $f_{a,I} = \mathrm{N}(0.1, 0.05^2)$
3. Party b Dimension I distribution $f_{b,I} = \mathrm{N}(0.9, 0.05^2)$
4. Party a and b Dimension II distribution $f_{a,b,II} = \mathrm{unif}(0, 1)$
5. Party a and b Dimension III distribution $f_{a,b,III} = \pi_q \mathrm{N}(\mu_q, 0.15^2)$
    - $q = \{5, 6\}$, $\pi_6 = \pi_6 = 0.5$, $\mu_5 = 0.25$, $\mu_6 = 0.75$

8.2.3 Discussion

Notice that the MDS results of our simulation with a multi-dimensional issue grid closely resembles the empirical MDS results in recent US House of Representatives. Hence, this suggests that fragmentation in the House of Representatives has occurred in recent years, and this fragmentation happens in new issue spaces that are *not* the traditional left-right spectrum, but possibly in issue areas such as technology policy and immigration.

We acknowledge: (i) that our simulated model and the MDS output is not an *if and only if* relationship, and (ii) the possibility of the "Twins Surfaces" pattern being produced from an issue grid with less than 3 dimensions. Nonetheless, it is very likely that $R \neq 1$, since in our informal work, all simulations where $R = 1$ produce party embeddings that are either clustered or one-dimensional.

The simulation therefore supports hypotheses that contemporary legislatures can longer be simply represented with a one-dimensional left-right spectrum. This had happened before — Ellenberg (2001) discussed how the arrival of an issue so divisive, such as slavery in the 1800s, or civil rights in the 1960s, can stretch the political union along a new axis and strain traditional political bonds. Once again, one-dimensional voting is breaking down.

> *"Democracy does not demand uniformity. Our founders quarreled then compromised, and they expected us to do the same. They understood that democracy requires solidarity — the idea that for all our outward differences, we rise or fall as one."*
>
> – Barack Obama

9 Conclusion

Our paper applies Multidimensional Scaling (MDS) as a dimensionality reduction technique to the legislative roll calls of three countries: United Kingdom, United States, and Canada, over a time period of roughly two decades. We first review the dissimilarity function and kernel method employed by Diaconis et al. (2008) on the 2005 United States House of Representatives roll calls, expanding their proof for the validity of the empirical distance from a uni-dimensional voting model to a multi-dimensional one. Next, we proceed with utilizing their dissimilarity function to provide an analytical solution that is also locally accurate.

The paper continues our discussion by presenting a qualitative survey of embedding patterns obtained from the MDS results. Then, we extend Diaconis et al. (2008)'s theoretical model from a two-party to a multi-party system, relate the multi-party model to our empirical results, and, in the process, discuss MDS's relationship with spectral clustering. In addition, we perform a deeper analysis of our results and connect them to contemporary political trends — (i) By conceptualizing MDS as a polarization index, we discover increasing inter-party polarization over time for all three countries; and (ii) By simulating legislative behaviors and relating the simulation to our MDS results, we suggest intra-party fragmentation along a new issue space, indicating the growing insufficiency of a traditional uni-dimensional issue space.

References

Agozino, Biko. 2015. "The Whip in the House: Rituals of Social Control in Parliament and in Society." *Social Crimonol* 3 (118).

Azzimonti, Marina. 2013. "The Political Polarization Index." *FRB of Philadelphia Working Paper* 13 (41).

Barone, Michael. 2015. "America's politics is polarized, but Britain's is fragmented." *Washington Examiner.* Accessed May 24, 2019. https://www.washingtonexaminer.com/americas-politics-is-polarized-but-britains-is-fragmented-2563735.

Bloodworth, James. 2017. "Will the politics of 'open and closed' replace 'left versus right'?" *International Business Times.* Accessed June 5, 2019. https://www.ibtimes.co.uk/will-politics-open-closed-replace-left-versus-right-1606328.

Borg, Ingwer, Patrick J.F. Groenen, and Patrick Mair. 2013. "Applied Multidimensional Scaling." *SpringerBriefs in Statistics.*

Brooks, David. 2016. "The Coming Political Realignment." *New York Times.* Accessed June 5, 2019. https://www.nytimes.com/2016/07/01/opinion/the-coming-political-realignment.html.

Clarkson, Stephen. 2011. "Has the Centre Vanished? The past and future of the middle ground in Canadian politics." *Literary Review of Canada.* Accessed July 1, 2019. https://reviewcanada.ca/magazine/2011/10/has-the-centre-vanished/.

Clinton, J, et al. 2004. "The statistical analysis of roll call data." *American Political Science Review* 98:355–370.

Cox, T.F., and M.A.A. Cox. 2000. "Multidimensional Scaling." *Chapman and Hall, London.*

Critchley, F. 1978. "Multidimensional scaling: A short critique and a new method." *COMPSTAT 1978.*

Diaconis, Persi, et al. 2008. "Horseshoes in multidimensional scaling and local kernel methods." *Ann. Appl. Stat.* 81 (3): 777–807.

Doherty, Carroll, et al. 2018. "Wide Gender Gap, Growing Educational Divide in Voters' Party Identification." *Pew Research Center.*

Ellenberg, Jordan. 2001. "Mapping Congress's growing polarization." *Slate.* Accessed May 26, 2019. https://slate.com/human-interest/2001/12/mapping-congress-growing-polarization.html.

Friedman, Thomas. 2019. "Is America Becoming a Four-Party State?" *The New York Times.* Accessed June 28, 2019. https://www.nytimes.com/2019/02/19/opinion/four-party-america-politics.html.

Guo, Jian, et al. 2015. "Estimating Heterogeneous Graphical Models for Discrete Data with an Application to Roll Call Voting." *Ann Appl Stat* 9 (2): 821–848.

Hu, Pili. 2012. "Spectral Clustering Survey."

Ives, Andrew. 2015. "The Canadian general election of 2015: The Liberal victory marks a swing back to the center in Canadian politics." *IdeAs. Idées d'Amériques* 6.

Kamarck, Elaine, and Alexander Podkul. 2018. "2018 Primaries Project:Internal divisions." *Brookings Intuition.* Accessed June 17, 2019. `https://www.brookings.edu/research/the-2018-primaries-project-what-are-the-internal-divisions-within-each-party/`.

Kim, In Song, John Londregan, and Marc Ratkovic. 2015. "Voting, Speechmaking, and the Dimensions of Conflict in the US Senate." *Asian Political Methodology Conference.*

Leeuw, Jan de. 2007. "A Horseshoe for Multidimensional Scaling." *UCLA: Department of Statistics.*

———. 2011. "Principal Component Analysis of Senate Voting Patterns." *UCLA: Department of Statistics.*

Longo, Antonio, et al. 2018. "Learning Political DNA in the Italian Senate." *CoRR* abs / 1812.07940.

Mardia, K., J. Kent, and J. Bibby. 1979. "Multiariate Analysis." *Academic Press, NY.*

Poole, Keith T., and Steven Daniels. 1985. "Ideology, Party, and Voting in the U.S. Congress, 1959-1980." *The American Political Science Review* 79 (2): 373–399.

Poole, Keith T., and Howard Rosenthal. 1983. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29 (2): 357–384.

———. 1984. "The Polarization of American Politics." *Journal of Politics* 46 (2): 1061–1079.

Schölkopf, Bernhard, Alexander Smola, and Klaus-Robert Müller. 1998. "Nonlinear Component Analysis as a Kernel Eigenvalue Problem." *Neural Computation* 10 (5): 1299–1319.

Shi, Jianbo, and Jitendra Malik. 2000. "Normalized cuts and image segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (8): 888–905.

Williams, Christopher K.I. 2000. "On a connection between kernel PCA and metric multidimensional scaling." *NIPS:* 675–681.

Young, G., and A.S. Householder. 1938. "Discussion of a set of points in terms of their mutual distances." *Psychometrika* 3 (1): 19–22.