Political Analysis (2026), **0**, 1–18
doi:10.1017/pan.2026.10058

PA

ARTICLE

Euclidean Ideal Point Estimation from Roll-Call Data via Distance-Based Bipartite Network Models

Seungju Lee[1], In Kyun Kim[1,2], Jong Hee Park[2] and Ick Hoon Jin[1,3]

[1]Department of Statistics and Data Science, Yonsei University, Republic of Korea; [2]Department of Political Science and International Relations, Seoul National University, Republic of Korea; [3]Department of Applied Statistics, Yonsei University, Republic of Korea

Corresponding author: Jong Hee Park, Ick Hoon Jin; Email: jongheepark@snu.ac.kr, ijin@yonsei.ac.kr

(Received 12 December 2025 UTC; revised 22 May 2026 UTC; accepted 25 May 2026 UTC)

Abstract

Conventional ideal point models rely on Gaussian or quadratic utility functions that violate the triangle inequality, producing non-metric distances that complicate geometric interpretation and undermine clustering and dispersion-based analyses. We introduce a distance-based alternative that adapts the latent space item response model (LSIRM) to roll-call data, treating legislators and bills as nodes in a bipartite network jointly embedded in a Euclidean metric space. Through controlled simulations, Euclidean LSIRM consistently recovers latent coalition structure with superior cluster separation relative to existing methods. Applied to the 118th U.S. House, the model provides competitive predictive performance while yielding bill embeddings that clarify cross-cutting issue alignments. The results show that restoring metric structure to ideal point estimation provides a clearer and more coherent inference about party cohesion, factional divisions and multidimensional legislative behavior.

Keywords: legislative voting; ideal point estimation; metric spaces; bipartite networks; coalition structure

Edited by: Daniel J. Hopkins and Brandon M. Stewart.

1. Introduction

Ideal point estimation from roll-call voting data has become a foundational tool in legislative studies, enabling systematic analysis of partisan polarization, ideological trajectories and coalition dynamics (Clinton, Jackman, and Rivers 2004; McCarty 2016; Poole and Rosenthal 2007). The dominant methods—NOMINATE (Lewis *et al.* 2025; Poole and Rosenthal 2007) and Bayesian item response theory (BIRT; Clinton *et al.* 2004; Jackman 2024)—successfully recover ordinal ideological relationships and achieve high vote classification accuracy. Recent work expands these frameworks by incorporating domain-specific ideal-point models or multidimensional extensions (Binding and Stoetzer 2023; Lipman, Moser, and Rodriguez 2025; Marble and Tyler 2022; Moser, Rodríguez, and Lofland 2021), integrating text-based information (Kim, Londregan, and Ratkovic 2018) or accommodating both monotonic and non-monotonic response functions (Lei and Rodriguez 2025).

Yet despite these advances, existing models share a structural limitation: the utility functions underlying NOMINATE and BIRT do not induce proper metric distances. Their Gaussian and quadratic specifications violate the triangle inequality, the defining property of metric spaces. Although this issue is mathematically straightforward, its consequences for political analysis are substantive. When distances lack metric validity, dispersion statistics used to assess party cohesion become non-comparable, and clustering analyses may generate configurations in which A is "close" to B, B to C, yet A remains

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Society for Political Methodology.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/pan.2026.10058

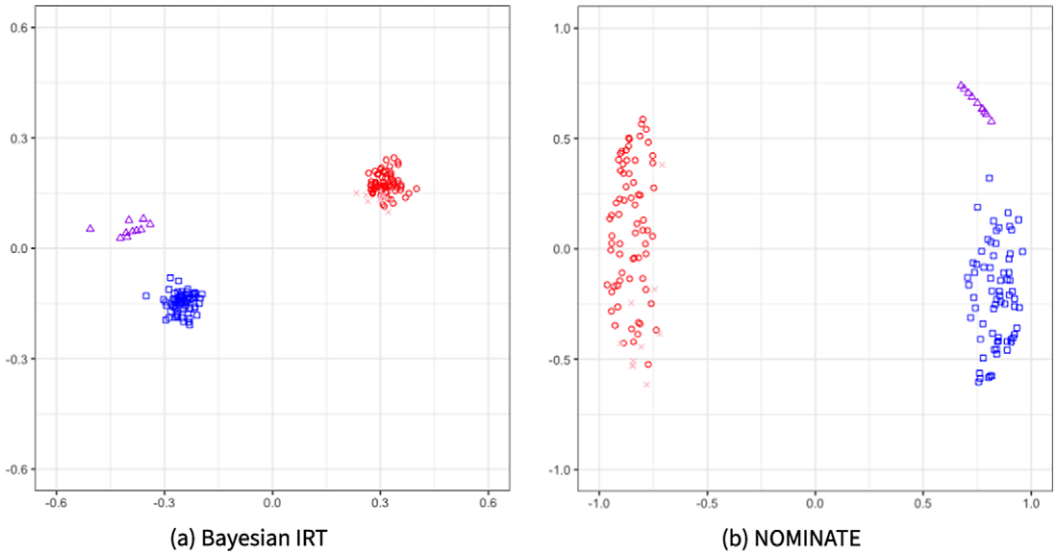


Figure 1. Coalition-cluster distortion in non-metric ideal point methods. Ground truth: *Four distinct coalitions* (marked by colors and shapes)—Party A majority (70, blue squares) and minority (10, purple triangles); Party B majority (70, red circles) and minority (10, pink crosses). The 400-bill agenda includes 200 partisan bills (party-line division) and 200 faction-specific bills (50 per faction). Despite this clear structure, both BIRT (a) and NOMINATE (b) compress the four factions into approximately two party blocs, with faction markers overlapping substantially within parties.

far from C. As contemporary legislatures exhibit growing intra-party heterogeneity (Clarke 2020; Harbridge-Yong, Volden, and Wiseman 2023), researchers increasingly rely on spatial distances to characterize factional structure, coalition tendencies or ideological spread—applications for which utility functions that do not induce metric distances ("non-metric utilities") provide an unstable foundation.

Recent work in the latent space literature highlights that different latent-space geometries can organize relational data in distinct ways. In particular, inner-product formulations capture directional similarity between latent vectors, whereas Euclidean distance models represent affinity through spatial proximity. These alternative geometries can lead to different representations of coalition structure even when the resulting models achieve similar predictive accuracy.

These geometric differences have practical consequences. Figure 1 illustrates this issue using simulated data with four distinct coalitions. Both NOMINATE and BIRT systematically compress factions into two party blobs. While both methods achieve high vote classification accuracy, the resulting spatial representations do not clearly recover the four underlying factions.

Recent work has recognized these limitations. Shin, Lim, and Park (2024) proposed an L_1-norm model using Manhattan distance: $U(x,y) = -\sum_k w_k|x_k - y_k|$, which satisfies the triangle inequality and provides a valid metric, where x and y denote the latent ideological positions of a legislator and a policy alternative, respectively, in the spatial model. However, Manhattan distance creates interpretive difficulties for spatial visualization: the distance between two points depends on their axis-aligned path rather than the direct geometric separation, complicating intuitive understanding of ideological proximity.

In this article, we propose using Euclidean distance (L_2 norm) $d(x,y) = \sqrt{\sum_k (x_k - y_k)^2}$, which provides both valid metric properties and intuitive geometric interpretability of spatial models. Euclidean distance satisfies the triangle inequality through the Cauchy–Schwarz inequality, provides the familiar notion of "straight-line" distance between points and admits natural geometric interpretation of clusters as spatially coherent regions.

https://doi.org/10.1017/pan.2026.10058

Our approach builds on recent network-based measures of political ideology. Barberá (2015) develops a Bayesian spatial following model for Twitter data, but its quadratic distance formulation ($-\|x_i - x_j\|^2$) inherits the same non-metric limitation. Similar work by Lo, Olivella, and Imai (2025) treats cosponsorship as a bipartite network and estimates latent positions via a mixed-membership blockmodel, where legislators and bills are represented as membership vectors on a simplex. These mixed-membership embeddings capture probabilistic affiliation with latent groups but do not retain the Euclidean distance structure assumed by spatial voting models. In contrast, we adopt the latent space item response model (LSIRM; Jeon *et al.* 2021), which treats roll-call data explicitly as a bipartite legislator–bill network and embeds both node types jointly in a Euclidean metric space. This symmetric formulation enables bills to serve as interpretive anchors for the latent dimensions while maintaining metric consistency for clustering and coalition analysis.

In the following, we provide theoretical and empirical evidence on how different latent-space geometries affect the interpretation of legislative voting behavior. We show that common ideal point utility specifications do not define metric distances, with implications specifically for distance-based interpretations. We then demonstrate that joint legislator–bill embedding aids substantive interpretation through bill locations, and that Euclidean LSIRM provides predictive performance competitive with BIRT while clarifying cross-cutting issue alignments.

Section 2 discusses the geometric properties of common ideal point utility functions and clarifies the role of metric distance in spatial interpretation. Section 3 introduces the bipartite network framework and Euclidean LSIRM specification. Section 4 presents simulations comparing Euclidean and inner-product formulations in recovering coalition structure. Section 5 applies the method to the 118th House, showing how bill embeddings serve as interpretive anchors for coalition structure. Section 6 concludes with implications and extensions.

2. Metric properties of utility functions

2.1. Proximity voting theory and metric requirements

The theoretical foundation for spatial voting models rests on the proximity voting framework developed by Downs (1957) and formalized by Davis, Hinich, and Ordeshook (1970). Under this framework, each voter i possesses an ideal point x_i in a policy space and evaluates alternatives (y) by their distance from this ideal point. For a policy located at position y, the voter's utility follows:

$$U_i(y) = -d(x_i, y) + \epsilon_i, \tag{1}$$

where $d(\cdot,\cdot)$ represents a distance function and ϵ_i is a random utility component. The voter prefers alternative A over alternative B when $U_i(A) > U_i(B)$; absent stochastic shocks, this occurs when $d(x_i, A) < d(x_i, B)$, so A is closer to the voter's ideal point than B. With shocks, the comparison also depends on $\epsilon_{iA} - \epsilon_{iB}$: under $U_i(j) = -d(x_i, j) + \epsilon_{ij}$, $U_i(A) > U_i(B)$ holds when $d(x_i, A) - d(x_i, B) < \epsilon_{iA} - \epsilon_{iB}$.

The proximity voting framework implicitly assumes that $d(\cdot,\cdot)$ constitutes a proper distance metric. Three properties are essential for the theoretical coherence of this framework. First, symmetry requires $d(x,y) = d(y,x)$: the ideological distance from position x to position y equals the distance from y to x. Without symmetry, the model would imply that moving from a liberal to conservative position involves a different ideological journey than moving from conservative to liberal, contradicting the reciprocal nature of ideological proximity. Second, the triangle inequality requires $d(x,y) + d(y,z) \geq d(x,z)$ for all points: the direct distance between two positions cannot exceed the sum of distances through an intermediate position. Without this property, indirect ideological paths could be shorter than direct paths, violating basic intuitions about ideological space and enabling paradoxical configurations where three positions cannot be consistently arranged along ideological dimensions. Third, positive definiteness requires $d(x,y) > 0$ for $x \neq y$ and $d(x,x) = 0$: distinct positions must be separated by positive distance, while a position's distance from itself is zero. Without this property, distinct ideological positions could be indistinguishable, rendering the spatial representation meaningless.

https://doi.org/10.1017/pan.2026.10058

These metric properties are not merely mathematical conveniences—they are fundamental to the substantive interpretation of proximity voting. Consider the triangle inequality: if legislator A has an ideal point close to position x, and position x is close to position y, the triangle inequality guarantees that legislator A cannot be arbitrarily distant from position y. This transitive consistency underlies claims about coalition formation: if moderate Democrats are close to centrist positions, and centrist positions are close to moderate Republicans, then moderate Democrats and moderate Republicans must be relatively proximate, enabling bipartisan coalitions. Without the triangle inequality, such transitive reasoning fails—proximity becomes intransitive, and the spatial metaphor loses coherent meaning.

Given these metric requirements, we now examine the utility formulations commonly used in spatial voting models, focusing on whether quadratic and Gaussian utility preserve the triangle inequality required for a metric notion of ideological proximity. This discussion concerns the theoretical utility formulation; issues related to how ideal point models are represented and estimated in practice are addressed separately in Section 3.

2.2. The triangle inequality and its violation in conventional methods

A distance function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ defines a metric space if it satisfies four properties: (1) non-negativity: $d(x,y) \geq 0$ with equality if and only if $x = y$; (2) symmetry: $d(x,y) = d(y,x)$; (3) triangle inequality: $d(x,y) + d(y,z) \geq d(x,z)$ for all $x,y,z \in \mathcal{X}$; and (4) identity of indiscernibles. The triangle inequality is particularly important: it ensures that distances are additive, comparable and transitively consistent. Without this property, distance comparisons become difficult to interpret—for example, it becomes unclear whether one pair is "twice as distant" as another or whether distance-based clustering methods that assume metric structure are appropriate.

In terms of their theoretical specification, conventional spatial voting models typically employ utility formulations based on squared distance. In BIRT (Clinton *et al.* 2004), legislators maximize quadratic utility derived from proximity voting: for each roll-call j presenting a binary choice between "Yea" position ζ_j and "Nay" position ψ_j, legislator i with ideal point x_i votes "Yea" when

$$-\|x_i - \zeta_j\|^2 + \eta_{ij} > -\|x_i - \psi_j\|^2 + \nu_{ij}.$$

This formulation implies that utility decreases with the squared Euclidean distance between the legislator and policy positions. However, squared distance itself does not satisfy the triangle inequality and therefore does not define a metric distance. Consider $d_Q(x,y) = (x-y)^2$ in one dimension. The triangle inequality requires $(x-y)^2 + (y-z)^2 \geq (x-z)^2$. Expanding $(x-z)^2 = [(x-y) + (y-z)]^2 = (x-y)^2 + 2(x-y)(y-z) + (y-z)^2$, the inequality reduces to $2(x-y)(y-z) \leq 0$. This condition holds when $(x-y)$ and $(y-z)$ have opposite signs (or when one of them is zero), but fails when they have the same sign.

In NOMINATE (Poole and Rosenthal 2007), spatial voting is often motivated using Gaussian utility:

$$U_i(\zeta_j) = \beta \exp\{-\frac{1}{2}\sum_k w_k^2(x_{ik} - \zeta_{jk})^2\}.$$

The exponential transformation of squared distances alters the geometry implied by spatial proximity. To illustrate, consider the one-dimensional quantity $d_G(x,y) = -\exp\{-\frac{1}{2}(x-y)^2\}$. For three points $x = 3$, $y = 3$ and $z = -1$, we obtain

$$d_G(x,y) = -\exp\{0\} = -1, \quad d_G(y,z) = -\exp\{-8\}, \quad d_G(x,z) = -\exp\{-8\}.$$

The triangle inequality would require $d_G(x,y) + d_G(y,z) \geq d_G(x,z)$, but

$$-1 - \exp\{-8\} < -\exp\{-8\},$$

so the inequality is violated. Thus, the geometry implied by Gaussian utility does not correspond to a metric distance.

https://doi.org/10.1017/pan.2026.10058

Such configurations can arise when legislators are presented with policy alternatives that move policy in the same ideological direction but by different magnitudes. For example, suppose a legislator's ideal point lies near the status quo, and two amendments propose more moderate and more extreme versions of a policy change in the same direction. In this situation, the legislator, the moderate amendment and the more extreme amendment lie along a single ideological line but at increasing distances from one another. Under squared-distance comparisons, this configuration can violate the triangle inequality even in a one-dimensional policy space.

While such violations can arise even in one-dimensional settings, their practical importance depends on the structure of the voting environment. In highly structured left–right settings, where alternatives are consistently ordered along a single ideological dimension, conventional models can recover the dominant partisan structure reasonably well. However, violations become more substantively relevant when the voting environment departs from this simple ordering—for instance, when both alternatives lie on the same ideological side (e.g., two competing left-leaning amendments), when procedural votes place both positions far from most ideal points or when a unidimensional model is fit to an inherently multi-dimensional space.

When such configurations occur, the resulting non-metric geometry can complicate geometric interpretations of legislators' positions in the ideological space. For example, researchers often summarize party cohesion using the dispersion of legislators' positions within a party. If the underlying distance representation is non-metric, differences in dispersion may partly reflect how the utility formulation scales distances across the ideological space rather than only differences in voting cohesion. Similar considerations arise when clustering legislator positions to identify factions or when comparing ideological separations across policy dimensions. The concern we raise is structural rather than empirical: a non-metric distance function lacks the geometric guarantees—transitivity of separations, additivity through intermediate points and cross-dimensional comparability—that underlie distance-based inference, regardless of how often violations are realized in a given dataset.

2.3. Euclidean distance as a metric

Euclidean distance $d_E(x,y) = \sqrt{\sum_{k=1}^{K}(x_k - y_k)^2}$ satisfies the triangle inequality and defines a proper metric space. The proof follows directly from the Cauchy–Schwarz inequality: for any three points $x,y,z \in \mathbb{R}^K$, squaring $d_E(x,y) + d_E(y,z) \geq d_E(x,z)$ and expanding yields $d_E(x,y)d_E(y,z) \geq \sum_k (x_k - y_k)(y_k - z_k)$, which holds because $|\sum_k (x_k - y_k)(y_k - z_k)| \leq \sqrt{\sum_k (x_k - y_k)^2}\sqrt{\sum_k (y_k - z_k)^2} = d_E(x,y)d_E(y,z)$ by the Cauchy–Schwarz inequality.

This metric property has direct implications for legislative analysis. First, distances become cardinally interpretable: if $d_E(z_A,z_B) = 2d_E(z_C,z_D)$, legislators A and B are genuinely twice as ideologically separated as C and D in their voting behavior. Second, distances become comparable across dimensions: we can meaningfully ask whether the first or second dimension generates greater polarization by comparing distance statistics, because the metric structure allows these comparisons to be interpreted directly in terms of spatial separation in the latent space. Third, clustering methods that assume metric distances—k-means, hierarchical clustering and silhouette analysis—can be applied in a geometrically coherent way because pairwise distances satisfy the triangle inequality. For example, the silhouette coefficient measures cluster quality by comparing mean within-cluster distances to mean nearest-cluster distances, a computation that requires pairwise distances to be metrically consistent. Fourth, the geometric interpretation aligns with mathematical properties: points that appear close in two-dimensional visualizations genuinely vote similarly, and spatial neighborhoods correspond to actual voting coalitions rather than distortions induced by non-metric utility functions.

At the same time, metric embeddings are not free from scaling artifacts. As in other low-dimensional embedding methods such as multidimensional scaling or spectral embeddings, geometric patterns such as curvature or horseshoe structures may arise when complex preference structures are approximated in a limited number of dimensions (e.g., Diaconis, Goel, and Holmes 2008). These limitations reflect the challenge of representing high-dimensional preference structures in a small number of spatial dimensions rather than a failure of the metric formulation itself.

https://doi.org/10.1017/pan.2026.10058

3. Euclidean distance model for bipartite legislative networks

3.1. Bipartite network framework

Roll-call voting data constitute a bipartite network linking legislators to bills through binary votes (excluding abstentions). Formally, let $G = (L,B,E)$ denote a bipartite graph, where $L = \{1,\dots,N\}$ represents legislators and $B = \{1,\dots,M\}$ represents bills. An edge between legislator i and bill j is represented by the adjacency variable y_{ij}, where

$$y_{ij} = \begin{cases} 1 & \text{if legislator } i \text{ votes "yea" on bill } j, \\ 0 & \text{otherwise.} \end{cases}$$

The resulting $N \times M$ adjacency matrix $Y = (y_{ij})$ corresponds directly to the roll-call voting matrix commonly analyzed in ideal point estimation. This representation treats roll-call voting as an affiliation network linking legislators to bills. Alternative network constructions are possible—for example, including status quo nodes or using a doubled network that treats "yea" and "nay" as separate edge types. We adopt the legislator–bill affiliation network because it corresponds directly to the observed roll-call matrix and enables joint embedding of legislators and bills in a shared latent policy space. In our formulation, nay votes are already incorporated through the complement of the modeled probability, so a doubled network would increase complexity without adding information.

Latent variable models provide a natural framework for representing relational data of this type. In the latent space approach to network modeling (Hoff 2007; Hoff, Raftery, and Handcock 2002), nodes are assigned positions in a low-dimensional space, and the probability of an edge depends on the relationship between their latent positions. Adapting this idea to the legislator–bill setting, let $z_i \in \mathbb{R}^K$ denote the latent position of legislator i and $w_j \in \mathbb{R}^K$ the latent position of bill j. The probability of a "Yea" vote depends on their spatial relationship:

$$g\left(\Pr(y_{ij} = 1 \mid z_i, w_j)\right) = \mu + \alpha(z_i, w_j), \tag{2}$$

where $g(\cdot)$ denotes a link function and $\alpha(z_i, w_j)$ specifies how latent positions determine voting propensity. Three canonical specifications exist. The latent class model (LCM; Airoldi *et al.* 2008; Holland, Laskey, and Leinhardt 1983; Miller, Jordan, and Griffiths 2009; Mørup, Schmidt, and Hansen 2011; Nowicki and Snijders 2001; Palla, Knowles, and Ghahramani 2012; Zhou 2015) and stochastic blockmodel (SBM; Airoldi *et al.* 2008; Daudin, Picard, and Robin 2008; Holland *et al.* 1983; Karrer and Newman 2011; Nowicki and Snijders 2001; Vu, Hunter, and Schweinberger 2013) set $\alpha(z_i, w_j) = m_{u_i, v_j}$ for discrete group memberships $u_i, v_j \in \{1,\dots,K\}$ and interaction matrix $M = (m_{k\ell})$, capturing stochastic equivalence where legislators in the same class vote identically. The latent eigenmodel (Hoff 2007; Hoff 2009; Hoff 2021; Li, Yeung, and Zhang 2011; Young and Scheinerman 2007) defines $\alpha(z_i, w_j) = z_i^\top \Lambda w_j$ for a diagonal matrix Λ, representing voting through weighted inner products that extend principal component analysis to binary data. The latent distance model (Handcock, Raftery, and Tantrum 2007; Hoff 2007; Hoff *et al.* 2002; Kemp *et al.* 2006; Krivitsky and Handcock 2008; Raftery *et al.* 2012; Schweinberger and Snijders 2003; Sewell and Chen 2015; Smith, Asta, and Calder 2019) specifies $\alpha(z_i, w_j) = -\gamma\|z_i - w_j\|$, where voting probability decreases monotonically with spatial separation, operationalizing homophily through spatial proximity.

Conventional ideal point methods, however, treat this bipartite structure asymmetrically: legislators receive interpretable spatial positions (ideal points), while bills serve primarily as instruments for scaling, with bill parameters (α_j, β_j) estimated but not embedded in the same space. In BIRT (Clinton *et al.* 2004), the probability of a "Yea" vote is typically written as

$$\operatorname{logit} P(y_{ij} = 1) = \alpha_j + \beta_j^\top x_i,$$

where x_i denotes the legislator's ideal point and (α_j, β_j) are bill parameters. Although this formulation is motivated by spatial proximity voting, the statistical representation of the model reduces to an interaction between legislator positions and bill loadings through the inner product $\beta_j^\top x_i$. As a result, bills are not represented as identifiable spatial locations comparable to legislators. The reparameterization from

https://doi.org/10.1017/pan.2026.10058

spatial bill positions (ζ_j, ψ_j) to discrimination parameters (α_j, β_j) is many-to-one—entirely different bill locations generate identical parameters—making bills difficult to interpret as spatial entities (Nakis *et al.* 2025; Shin *et al.* 2024). In addition, the model does not include a separate legislator-specific intercept capturing baseline voting propensity; instead, variation across legislators is represented solely through their spatial positions.

Jeon *et al.* (2021) introduced the LSIRM for educational testing, embedding respondents and items in a shared latent space. This symmetric framework applies naturally to roll-call voting data, where legislators and bills form the two modes of a bipartite system. It is also flexible with respect to the geometric operator governing interactions: Jeon *et al.* (2021) consider both an inner-product specification, analogous to the latent eigenmodel (Hoff 2007), and a distance-based specification. The former captures directional alignment through $z_i^\top w_j$, whereas the latter captures spatial proximity through Euclidean distance. In this article, we adopt the Euclidean specification.

Distance-based specifications have been applied to political networks in two contexts, though with different formulations and objectives than ours. Barberá (2015) employed a latent space model to estimate ideal points from Twitter following patterns, treating the data as a bipartite network where users choose whether to follow political actors. Users i and political actors j are embedded in a shared latent space, with following probability modeled as $P(y_{ij} = 1) = f(\phi - \gamma\|\theta_i - \beta_j\|^2)$, where $\|\theta_i - \beta_j\|^2$ represents squared Euclidean distance. However, as discussed in Section 2, this squared distance formulation does not satisfy the triangle inequality and therefore produces a non-metric spatial representation.

Taken together, these approaches reveal two limitations in the spatial analysis of roll-call voting. First, in conventional ideal point models, bills are treated primarily as scaling instruments with parameters that do not correspond to identifiable spatial locations. Second, distance-based specifications used in political network models typically rely on squared Euclidean distance, which—as discussed in Section 2—does not satisfy the triangle inequality and therefore produces a non-metric representation of ideological space. In the next section, we introduce a Euclidean latent space formulation for roll-call voting that embeds legislators and bills in a shared space using Euclidean distance and allows direct comparison with inner-product-based representations.

3.2. Euclidean distance specification

We specify voting probability through Euclidean distance in a bipartite network framework. For legislator i with latent position $z_i \in \mathbb{R}^K$ and bill j with latent position $w_j \in \mathbb{R}^K$:

$$\text{logit}(P(y_{ij} = 1 \mid \theta_i, \beta_j, \gamma, z_i, w_j)) = \theta_i + \beta_j - \gamma\|z_i - w_j\|, \tag{3}$$

where θ_i captures legislator i's baseline propensity to vote "Yea," β_j represents bill j's baseline popularity or controversy, and $\gamma > 0$ governs the strength of proximity effects. The Euclidean distance is $\|z_i - w_j\| = \sqrt{\sum_{k=1}^{K}(z_{ik} - w_{jk})^2}$. The probability of a "Yea" vote increases with legislator propensity θ_i and bill popularity β_j, but decreases with spatial distance $\|z_i - w_j\|$. This formulation directly implements proximity voting (Davis *et al.* 1970; Downs 1957): legislators support bills near their ideal points, with support declining monotonically as separation increases.

For comparison, an inner-product LSIRM can be written as

$$\text{logit}(P(y_{ij} = 1 \mid \theta_i, \beta_j, z_i, w_j)) = \theta_i + \beta_j + z_i^\top w_j, \tag{4}$$

which replaces the distance-based interaction with a bilinear form while preserving the same intercept structure.

The baseline parameters θ_i and β_j play distinct roles from those in BIRT. In BIRT, bill positions are not separately identified such that α_j can be interpreted as a baseline level of support, and the model includes no legislator-specific intercept term. In our specification, β_j measures only the baseline tendency for legislators to support bill j averaging over all spatial locations—a pure "how popular is this bill" parameter separable from its ideological location w_j. Similarly, θ_i measures legislator i's general

https://doi.org/10.1017/pan.2026.10058

tendency to vote "Yea" independent of ideological proximity—some legislators vote "Yea" frequently (high θ_i) while others vote "Nay" frequently (low θ_i) for reasons orthogonal to spatial positioning. This separation between baseline propensities and spatial structure, which is similar to the degree correction in network science literature (Fortunato 2010), clarifies interpretation: if legislator i supports bill j, we can determine whether this reflects high baseline propensity (θ_i or β_j large), ideological proximity ($\|z_i - w_j\|$ small), or both.

The parameter γ quantifies the strength of spatial proximity relative to baseline propensities. Large γ indicates voting is primarily distance-determined—legislators vote almost exclusively for nearby bills regardless of baseline propensities—while small γ suggests the dominance of baseline propensities, a spatial structure playing a secondary role. This provides diagnostic information about legislative organization: chambers with strong, stable party or factional discipline exhibit larger γ (tight clustering with sharp spatial boundaries), whereas fragmented or weakly organized chambers exhibit smaller γ (diffuse positioning where baseline propensities matter more than ideological location). Our simulation studies demonstrate that $\hat{\gamma}$ responds predictably to known levels of coalition cohesion, increasing when within-group homogeneity strengthens and decreasing when independent voting attenuates the distance signal.

Inner-product models, including BIRT, measure affinity through directional alignment rather than spatial proximity, so geometric closeness does not necessarily imply a higher tie probability. As noted in the `latentnet` documentation, actors that are closer in a clustering sense need not have higher expected tie values (Krivitsky *et al.* 2009). More generally, as shown by Hoff (2007), inner-product formulations are representationally more flexible, which can be advantageous in networks characterized by directional equivalence or heterogeneous connectivity patterns.

By contrast, Euclidean distance models impose a stricter geometric interpretation in which spatial proximity directly governs interaction probability. This yields embeddings that are less sensitive to scale artifacts and limits the influence of unusually prominent units, such as procedural hub bills or omnibus measures. In homophilous settings such as legislative voting, the resulting geometry aligns clustering structure more closely with proximity in the latent space. Consistent with this intuition, Nakis *et al.* (2025) show that Euclidean representations can often achieve comparable structural fit with equal or lower latent dimensionality than inner-product formulations.

The bipartite embedding provides three interpretive advantages over conventional methods by locating legislators and bills in a shared latent space. Legislator-to-legislator distances reveal ideological similarity, bill-to-bill distances identify policy clustering and legislator–bill distances clarify coalition structure through spatial proximity, including the placement of cross-party bridge bills between cooperating factions. This joint embedding enables interpretation through bill locations as anchors: instead of asking "what does dimension 2 mean?", we can observe "which bills define dimension 2?" and infer substantive meaning from the specific legislation that separates high from low values.

3.3. Full conditionals and estimation

We estimate parameters via Markov chain Monte Carlo (MCMC) sampling with the Gibbs sampling and the Metropolis–Hastings algorithm with the following prior specifications:

$$\begin{aligned} \theta_i &\sim N(0, \sigma_\theta^2), \\ \beta_j &\sim N(0, \sigma_\beta^2), \\ \sigma_\theta^2 &\sim \text{Inv-Gamma}(a_\sigma, b_\sigma), \\ z_i &\sim \text{MVN}_K(\mathbf{0}, \mathbf{I}_K), \\ w_j &\sim \text{MVN}_K(\mathbf{0}, \mathbf{I}_K), \\ \log \gamma &\sim N(\mu_\gamma, \sigma_\gamma^2). \end{aligned} \tag{5}$$

https://doi.org/10.1017/pan.2026.10058

The overall scale of the latent space is regularized through the prior specification $z_i \sim \text{MVN}_K(\mathbf{0}, \mathbf{I}_K)$ and $w_j \sim \text{MVN}_K(\mathbf{0}, \mathbf{I}_K)$. Because the likelihood depends on the Euclidean distance $\|z_i - w_j\|$, it is invariant to joint rescaling of the latent positions and the distance coefficient γ. The prior therefore constrains the scale of the latent space, allowing γ to be interpreted as the strength of distance-based voting relative to this normalization.

Then, the joint posterior density of the model is given as

$$\begin{aligned}\pi(\boldsymbol{\beta}, \boldsymbol{\theta}, \gamma, \mathbf{Z}, \mathbf{W} \mid \mathbf{Y}) \propto & \prod_{i=1}^{N}\prod_{j=1}^{P} P(Y_{ij} = y_{ij} \mid \theta_i, \beta_j, \gamma, z_i, w_j) \\ & \times \pi(\gamma) \prod_{j=1}^{P} \pi(\beta_j) \prod_{i=1}^{N} \pi(\theta_i) \prod_{i=1}^{N} \pi(z_i) \prod_{j=1}^{P} \pi(w_j).\end{aligned} \tag{6}$$

The priors and the following posterior distribution can be expressed in Equations (5) and (6). This posterior kernel cannot be expressed with standard distribution, so the exact posterior density cannot be calculated and a Gibbs sampler is used to sample each parameter sequentially from their conditional density. Specifically, because each conditional kernel may not be expressed as a standard distribution form, Metropolis–Hastings within Gibbs sampler is used. Furthermore, for the generalized case with missing data, the imputed value for the non-respondent pair can be sampled with the logistic formula at the start of each Gibbs sampler step:

$$\pi(\theta_i) \propto \left[\prod_{i=1}^{N}\prod_{j=1}^{P} P(Y_{ij} = y_{ij} \mid \theta_i, \beta_j, \gamma, z_i, w_j)\right] \times \left[N_{\theta_i}(0, \sigma_\theta^2)\right]$$

$$\pi(\beta_j) \propto \left[\prod_{i=1}^{N}\prod_{j=1}^{P} P(Y_{ij} = y_{ij} \mid \theta_i, \beta_j, \gamma, z_i, w_j)\right] \times \left[N_{\beta_j}(0, \sigma_\beta^2)\right]$$

$$\pi(\log\gamma) \propto \left[\prod_{i=1}^{N}\prod_{j=1}^{P} P(Y_{ij} = y_{ij} \mid \theta_i, \beta_j, \gamma, z_i, w_j)\right] \times \left[N(\mu_\gamma, \sigma_\gamma^2)\right]$$

$$\pi(z_i) \propto \left[\prod_{i=1}^{N}\prod_{j=1}^{P} P(Y_{ij} = y_{ij} \mid \theta_i, \beta_j, \gamma, z_i, w_j)\right] \times \left[\text{MVN}_{K, z_i}(\mathbf{0}, \mathbf{I}_K)\right]$$

$$\pi(w_j) \propto \left[\prod_{i=1}^{N}\prod_{j=1}^{P} P(Y_{ij} = y_{ij} \mid \theta_i, \beta_j, \gamma, z_i, w_j)\right] \times \left[\text{MVN}_{K, w_j}(\mathbf{0}, \mathbf{I}_k)\right]$$

$$\pi(\sigma_\theta) \propto \text{Inv-Gamma}\left(\left(\frac{N}{2} + a_\sigma\right), \frac{1}{2}\sum_{i=1}^{N}\theta_i^2 + b_\sigma\right).$$

Like conventional ideal-point estimators, Euclidean LSIRM inherits rotational, reflective and translational indeterminacies: for any orthogonal matrix Q and vector c, the transformed configuration $(Qz_i + c, Qw_j + c)$ produces identical probabilities. Because we do not fix the positions of any legislators or bills, the absolute orientation and location of the latent space are not identified. Specifically, after obtaining posterior samples from the MCMC algorithm, we apply a Procrustes transformation (Gower 1975) to align each sampled configuration to a reference configuration selected as the draw with the highest posterior density (Friel *et al.* 2016). This post-processing removes rotational, reflective and translational indeterminacies while preserving pairwise distances among legislators and bills.[1]

[1] Recently, Shin *et al.* (2024) proposed an ℓ_1-norm-based Bayesian IRT model, which transforms the infinite rotational invariance of conventional methods into a more tractable signed perpendicular rotational invariance.

https://doi.org/10.1017/pan.2026.10058

4. Simulation studies

We designed three simulations to evaluate Euclidean LSIRM against two alternative formulations: BIRT, a widely used ideal-point model, and an inner-product LSIRM specification in which interactions between legislators and bills are modeled through the inner product of latent positions. To isolate the role of latent-space geometry, we constructed a comparison between Euclidean and inner-product LSIRM specifications that share the same baseline parameters, bill representation, scaling structure and identification procedure, differing only in the geometric form of interaction. Because BIRT also represents legislator–bill interactions through an inner-product structure between legislator ideal points and bill parameters, these two approaches provide useful contrasts to the Euclidean-distance formulation of LSIRM. The simulations progressively increase in structural complexity and address three questions:

1. *Cohesion gradient*: Does the proximity parameter γ respond appropriately to known levels of group cohesion?
2. *Cluster recovery*: Can the model recover multiple independent clusters in low-dimensional space?
3. *Party–faction structure*: Can the model distinguish factions within parties and cross-party coalitions?

All simulations used 100 replicated datasets. MCMC chains ran for 30,000 iterations with 5,000 burn-in and thinning factor 5. We evaluate performance using silhouette coefficients, and visual inspection of latent embeddings. For visualization, we display representative replications whose silhouette coefficients across the three models are jointly closest to their median values.

In the main text, we focus on the party–faction simulation, which is most directly connected to the article's goal of recovering within-party factions and cross-party coalitions. Additional results for the cohesion-gradient and cluster-recovery designs are reported in the Supplementary Material (Sections 1.1 and 1.2), while predictive accuracy and log-likelihood summaries across settings are reported in the Supplementary Material (Section 1.4).

The party–faction simulation considers two parties with internal factions, plus cross-party coalitions. The baseline structure consists of two parties, each containing two or three factions. We generated three types of bills: *partisan bills* supported by one party against the other, *faction bills* supported by a single faction and *bridge bills* supported by factions from opposing parties (Table 1).

4.1. Agenda sweep

Each party contains three factions (180 legislators total). We varied the share of partisan bills from 30% to 50%, with the remainder being faction-specific bills.

At both agenda compositions, BIRT and the inner-product LSIRM largely collapse factional structure, merging within-party factions into roughly three visible groups in the representative embedding (Figure 2). Euclidean LSIRM, by contrast, separates all six factions into distinct clusters. As the share of partisan bills falls from 50% to 30%, factional signals strengthen and the performance gap widens: Euclidean LSIRM's mean silhouette increases from 0.624 (SD = 0.081) to 0.723 (SD = 0.016), while

Table 1. Party–faction scenarios examined in the main text

Scenario	Manipulation	Key test
(i) Agenda sweep	Vary partisan vs. faction bill ratio	Faction recovery as partisan signal weakens
(ii) Cross-party coalitions	Add bridge bills	Representation of overlapping coalitions

Note: Additional noise-sweep results are reported in the Supplementary Material (Section 1.3.1).

https://doi.org/10.1017/pan.2026.10058

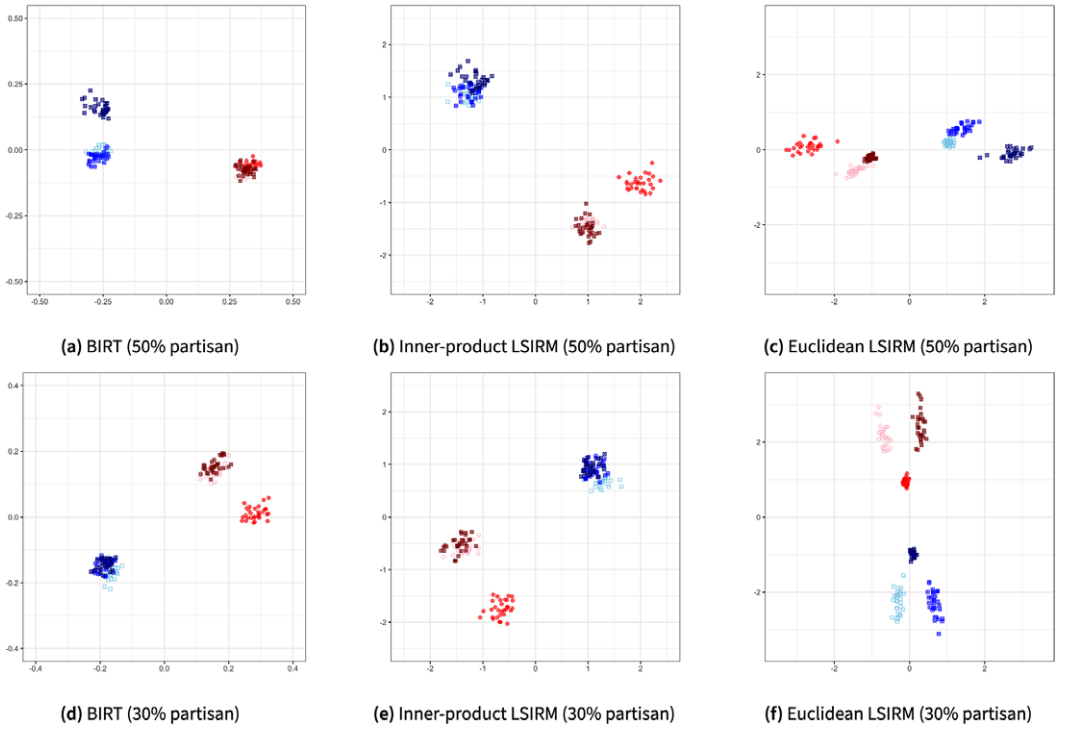


Figure 2. Faction recovery under shifting legislative agendas. Colors and markers denote six true factions: two parties, each containing three internal factions (180 legislators total, 30 per faction). Top row: 50% partisan bills; bottom row: 30% partisan bills. Euclidean LSIRM recovers all six factions as distinct clusters across both conditions, whereas BIRT and inner-product LSIRM collapse within-party factions into roughly three visible groups, failing to distinguish factional structure from party membership.

BIRT remains low (0.164– 0.172) and the inner-product LSIRM shows similarly weak separation (0.188–0.190).

4.2. Cross-party coalitions

We introduced unequal faction sizes (70 vs. 10 legislators per faction) and two types of bridge bills: *Bridge-A* supported by majority factions from both parties (L1 and C1), and *Bridge-B* supported by minority factions (L2 and C2). This design tests whether the model can represent overlapping coalitions that transcend party lines. Additional configurations and extended results, including simulations incorporating faction bills, appear in the Supplementary Material (Section 1.3.2).

In the majority consensus scenario (Figure 3), Euclidean LSIRM correctly draws the allied majority factions closer together near the inter-party boundary, with Bridge-A bills positioned between them. Both BIRT and the inner-product LSIRM fail to represent this structure—distances between allied cross-party legislators are no shorter than distances between ideological opponents, obscuring the coalition structure in the latent space.

In the ends-against-the-middle scenario (Duck-Mayr and Montgomery 2023), the two inner-product specifications—BIRT and the inner-product LSIRM—produce embeddings that stretch legislators along a pronounced secondary axis (Figure 4). This elongated geometry spreads factions along a single direction in the latent space.

The Euclidean LSIRM instead organizes legislators according to spatial proximity. As a result, the two minority factions are drawn toward one another while the majority blocs remain separated. This configuration places Bridge-B bills between the opposing minority factions and preserves interpretable distances between legislators and the bills that structure their voting behavior.

https://doi.org/10.1017/pan.2026.10058

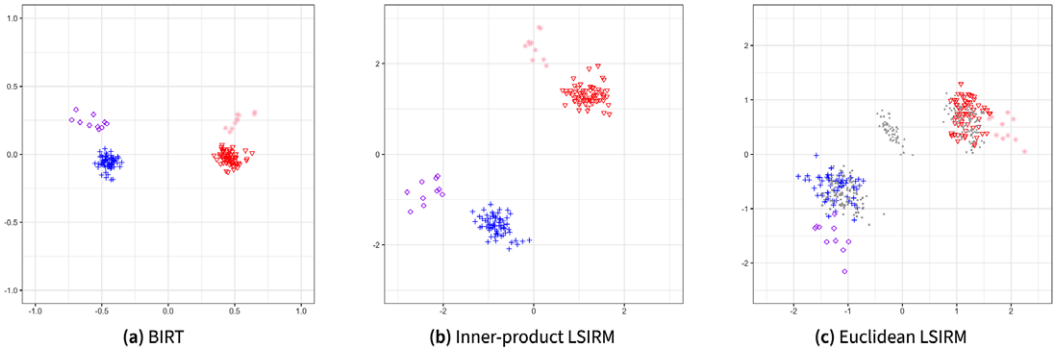


Figure 3. Cross-party majority coalition recovery. Two parties (Liberal: blue/purple; Conservative: red/pink) each split into a majority faction (L1, C1: 70 legislators) and a minority faction (L2, C2: 10 legislators). Bridge-A bills (gray triangles) are supported jointly by the two majority factions across party lines. Legislators are colored by party and shaped by faction (L1: blue squares, L2: purple triangles, C1: red circles and C2: pink crosses). Euclidean LSIRM correctly positions allied majority factions near the inter-party boundary with Bridge-A bills between them; BIRT and inner-product LSIRM fail to capture this cross-party coalition structure.

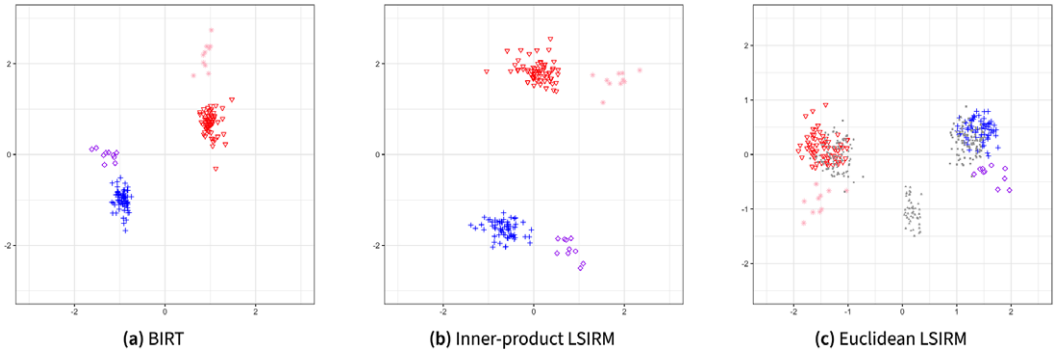


Figure 4. Cross-party minority coalition recovery: Ends-against-the-middle scenario. Bridge-B bills (gray triangles) are supported jointly by the two minority factions (L2, C2: 10 legislators each) across party lines, while majority factions (L1, C1: 70 legislators each) oppose. Legislators are colored by party and shaped by faction (L1: blue squares, L2: purple triangles, C1: red circles and C2: pink crosses). Euclidean LSIRM draws the minority factions toward each other with Bridge-B bills positioned near this pairing, preserving interpretable proximity between legislators and the bills structuring their coalition. BIRT and inner-product LSIRM instead stretch legislators along a pronounced secondary axis, obscuring the minority coalition.

5. Application: 118th U.S. House of Representatives

5.1. Data and context

We apply Euclidean LSIRM to roll-call votes from the 118th U.S. House, obtained from Voteview (Lewis *et al.* 2025). Following standard practice, we excluded lopsided votes (overwhelming majorities) that provide little ideological information, yielding 1,225 informative roll-calls across 451 unique legislators (accounting for vacancies and replacements). This chamber provides a stringent test: extreme partisan polarization (the most conservative Democrat remains left of the most moderate Republican) yet well-documented intra-party factions—the progressive "Squad" (approximately 8 members) among Democrats and the House Freedom Caucus (approximately 22 members) among Republicans. These factions, though numerically small, exert disproportionate influence on leadership contests and agenda setting, making their accurate identification substantively important. We ran five independent MCMC chains and assessed convergence using overlaid trace plots and scale reduction diagnostics (see Section 2.1 of the Supplementary Material), which indicate satisfactory mixing and stationarity. As the latent space is invariant to rigid transformations, coordinates are not directly comparable across chains; visualizations therefore use posterior means from the first converged chain.

https://doi.org/10.1017/pan.2026.10058

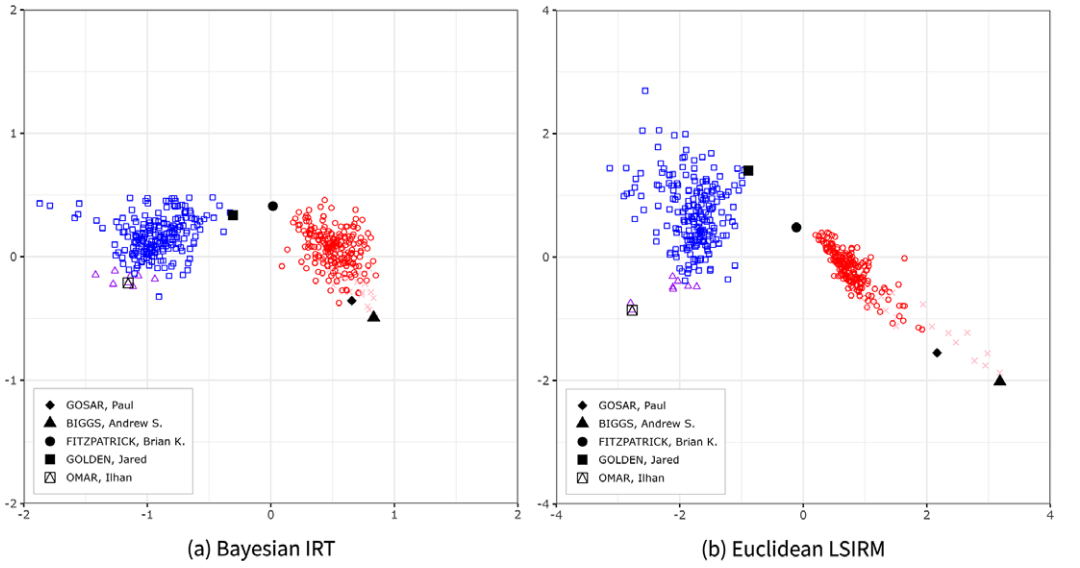


Figure 5. Two-dimensional legislator ideal point estimates for the 118th U.S. House. Legislators are plotted by party and factional affiliation: Mainstream Democrats (blue squares), "Squad" Democrats (purple triangles), Mainstream Republicans (red circles) and Freedom Caucus Republicans (pink crosses).

5.2. Legislator positions and faction recovery

Figure 5 displays two-dimensional embeddings from Euclidean LSIRM and BIRT. Both methods recover clear party separation on the primary (horizontal) dimension, with Democrats left and Republicans right. However, the secondary dimension reveals striking differences. Euclidean LSIRM differentiates factions: Freedom Caucus members (e.g., Andy Biggs and Paul Gosar) occupy far-right positions with distinctly negative second-dimension coordinates, separating them from mainstream Republicans, while Squad members (e.g., Ilhan Omar and Alexandria Ocasio-Cortez) appear below with slightly negative values, distinguishing them from other Democrats. The second dimension has substantial variance (SD = 0.69) in Euclidean LSIRM, indicating genuine factional structure. BIRT places faction members simply at extremes of the left–right continuum without meaningful secondary differentiation (SD of dimension 2 = 0.18 vs. 0.75 for dimension 1), effectively collapsing multi-dimensional coalition structure into one-dimensional ordering.

We evaluate predictive performance using classification accuracy (ACC), aggregate proportional reduction in error (APRE) and mean log posterior predictive density (LPPD). As shown in Table 2, Euclidean LSIRM and BIRT exhibit very similar in-sample predictive performance. LSIRM is slightly better on in-sample ACC, APRE and mean LPPD, but the differences are small. Out-of-sample performance is assessed using a 20% entry-wise holdout split. The held-out results also show that both models provide strong probabilistic predictions relative to an uninformative baseline. The Brier score measures the mean squared difference between predicted probabilities and actual binary outcomes, ranging from 0 to 1 with lower values indicating better probabilistic calibration; the uninformative baseline of always predicting $\hat{p} = 0.5$ yields a score of 0.25. Both models substantially improve upon this baseline, with Brier scores well below 0.25. These validation results indicate that Euclidean LSIRM remains predictively competitive with BIRT.

Beyond predictive performance, the Euclidean specification provides a useful representation of coalition structure through its joint metric embedding of legislators and bills. This has broader implications for ideal point estimation: legislatures where intra-party factionalism is consequential—such as those with multi-party systems, coalition governments or cross-cutting cleavages that do not map cleanly onto a single left–right axis—are precisely the settings where a non-metric distance formulation is most likely

https://doi.org/10.1017/pan.2026.10058

Table 2. In-sample and out-of-sample predictive performance for LSIRM and BIRT ($K = 2$)

	LSIRM	BIRT
In-sample		
Accuracy (ACC)	0.943	0.942
APRE	0.841	0.839
Mean LPPD	−0.154	−0.156
Out-of-sample		
Mean LPPD	−0.263	−0.165
Brier score	0.075	0.046

Note: Out-of-sample results are based on a 20% entry-wise holdout.

to obscure latent structure. In such contexts, recovering faction-level ideal points accurately is not only a methodological concern but a substantive one, as mismeasured positions can distort inferences about coalition formation, agenda influence and legislative bargaining.

Predictive gains from increasing dimensionality diminish beyond $K = 2$ for Euclidean LSIRM (see Table B2 in the Supplementary Material). Because improvements beyond two dimensions are modest and the substantive coalition structure is already clearly recovered in two dimensions, we adopt $K = 2$ for interpretation in the main analysis.

5.3. Bill embeddings as interpretive anchors

A key advantage of the bipartite network approach is that bill locations serve as interpretive anchors clarifying the substantive meaning of latent dimensions. Figure 6 displays the full legislator–bill embedding. Partisan bills cluster near their respective party centroids: Republican-supported bills (e.g., tax cuts and defense spending) locate in the far-right region near mainstream Republicans, while Democratic-supported bills (e.g., social programs and environmental regulation) cluster in the left region near mainstream Democrats. This validates that the first dimension captures the familiar liberal-conservative divide.

More revealing are bills that deviate from party centroids. War powers resolutions H.Con.Res. 21 (Syria withdrawal), H.Con.Res. 30 (Somalia withdrawal) and H.J.Res. 79 (terminating national emergency) position in the lower-left quadrant between Squad and Freedom Caucus locations (coordinates approximately $x \approx -0.5, y \approx -2$). These resolutions received support primarily from these two factions—progressive Democrats and libertarian Republicans—while being opposed by party establishments. Similarly, surveillance and privacy bills including H.R. 4755 (Privacy Enhancing Technology Research Act), H.R. 5403 (CBDC Anti-Surveillance State Act) locate far from the two extremist factions. These bill positions suggest the substantive meaning of the second dimension: it reflects an establishment-outsider cleavage where ideological extremes from both parties converge against centrists on issues of foreign military intervention and government surveillance—the "horseshoe" pattern where far-left and far-right align against the center (Duck-Mayr and Montgomery 2023).

The debt-limit deal H.R. 3746 positions near the center in LSIRM's embedding ($x \approx -0.4$, $y \approx 0.5$), reflecting its passage through a Democratic-anchored coalition with substantial Republican support—a bipartisan compromise occupying ideological middle ground. In contrast, BIRT cannot provide this interpretation because bills are not embedded spatially. The discrimination parameter β_j measures directional alignment but does not locate bills in the ideological space, precluding direct visualization of which bills define which dimensions. Without bill embeddings, conventional methods cannot directly show where such compromise legislation locates relative to partisan bills and extremist-coalition bills, obscuring the multi-dimensional coalition structure that characterizes contemporary congressional

https://doi.org/10.1017/pan.2026.10058

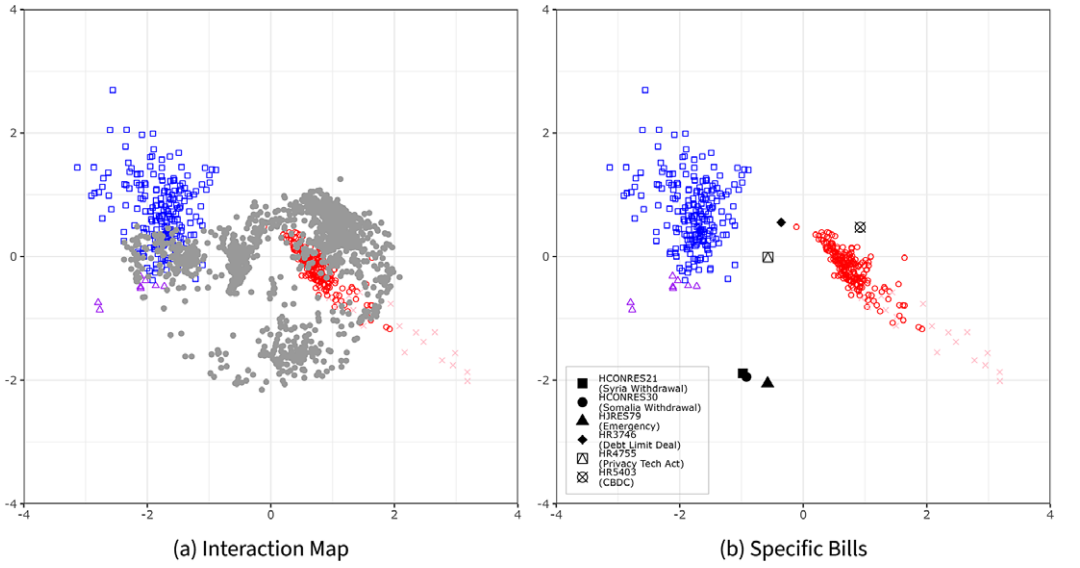


Figure 6. Two-dimensional legislator–bill ideal point estimates for the 118th U.S. House. Legislators are plotted by party and factional affiliation: Mainstream Democrats (blue squares), "Squad" Democrats (purple triangles), Mainstream Republicans (red circles) and Freedom Caucus Republicans (pink crosses). Bills are represented by gray markers.

politics. Additional analyses in the Supplementary Material (Section 2.4) visualize bill locations together with within-faction Yea ratios, providing a finer-grained view of how factional support patterns align with Euclidean LSIRM's spatial geometry.

6. Discussion

Conventional ideal point estimation methods employ utility functions that violate the triangle inequality, complicating meaningful interpretation of distance magnitudes and ratios. We show that Euclidean LSIRM provides a more interpretable representation for clustering and coalition analysis. In controlled simulations with known four-coalition structures, NOMINATE and BIRT compress factions into party blobs (silhouette coefficients: 0.778), while Euclidean LSIRM recovers the true structure (0.861). Application to the 118th House illustrates how joint bill embeddings clarify legislative coalition structure by placing war powers and surveillance bills between Squad and Freedom Caucus members, revealing an establishment-outsider dimension.

Beyond restoring metric validity, the Euclidean formulation offers a parsimonious and interpretable representation for homophilous roll-call settings. Prior work shows that Euclidean embeddings can often match the structural fit of inner-product formulations with equal or lower latent dimensionality (Nakis *et al.* 2025), while the dependence on pairwise distance rather than vector magnitude reduces sensitivity to scale artifacts or unusually influential units. Most importantly, metric coherence ensures that pairwise similarities satisfy the triangle inequality, supporting clustering, dispersion summaries and other distance-based interpretations in a way that non-metric formulations cannot.

These combined properties establish Euclidean LSIRM as particularly well suited for bipartite roll-call analysis when the goal extends beyond ordinal scaling. The bipartite network framework, by treating legislators and bills symmetrically in a shared metric space, enables bill locations to serve as interpretive anchors that define dimensional meaning and clarify which issues generate cross-cutting cleavages. This restores consistency between spatial voting theory, which assumes proximity preferences over metric distances, and empirical estimation, which should produce estimates that support distance-based inference.

https://doi.org/10.1017/pan.2026.10058

Three extensions merit development. First, dynamic specifications tracking coalition reconfiguration across sessions could model legislator positions as evolving while maintaining metric structure within periods, enabling analysis of how factional boundaries shift over time. Second, incorporating bill metadata (i.e., texts, issue codes and sponsors) through hierarchical priors could sharpen interpretation by partially pooling bill locations within policy domains while permitting cross-issue variation. Third, applications to multiparty systems could further exploit Euclidean distance's dimensional efficiency, potentially enabling more parsimonious representations of complex coalition structures in fragmented legislatures where conventional methods struggle with high dimensionality. A related extension would apply the framework to historical congressional sessions, such as those of the 1930s or 1960s, where cross-party factional structures are well documented and provide a richer setting for evaluating clustering-based representations. Each extension would require addressing identification and missing data issues particular to its context, but the fundamental advantages—dimensional efficiency, balanced influence and metric coherence—apply broadly across institutional settings where researchers seek to identify voting coalitions rather than merely scale legislators ordinally.

Acknowledgements. The authors thank the editor, associate editor and reviewers for their constructive comments. Kwangok Seo and Johan Lim provided valuable comments for the earlier version of this article.

Funding statement. This work was partially supported by the National Research Foundation of Korea (Grant Nos. NRF-2021S1A3A2A03088949, RS-2023-00217705 and RS-2024-00333701; Basic Science Research Program awarded to I.H.J.).

Data Availability Statement. Replication data are available in Harvard Dataverse: https://doi.org/10.7910/DVN/JHJEGT (Lee *et al.* 2026).

Author contributions. Conceptualization: S.L., I.K.K., J.H.P., and I.H.J.; Data curation: S.L. and I.K.K.; Formal analysis: S.L.; Methodology: S.L., J.H.P., and I.H.J.; Software: S.L.; Supervision: J.H.P. and I.H.J.; Validation: S.L.; Visualization: S.L.; Writing—original draft: S.L. and J.H.P.; Writing—review and editing: S.L., I.K.K., J.H.P., and I.H.J. All authors approved the final submitted draft.

Competing Interests. The authors declare none.

Ethical Standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Disclosure of use of AI tools. The authors used ChatGPT (OpenAI; standard conversational interface) to assist with language editing and proofreading during manuscript preparation and revision, and Codex (OpenAI) to assist with organizing and refactoring existing code during preparation of the replication materials. The original analysis code and statistical procedures were developed by the authors. The authors reviewed all AI-assisted outputs and are responsible for all content and accompanying replication materials.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/pan.2026.10058.

References

Airoldi, E. M., D. M. Blei, S. E. Fienberg, and E. P. Xing. 2008. "Mixed Membership Stochastic Blockmodels." *Journal of Machine Learning Research* 9 (65): 1981–2014. Accessed September 1, 2025. http://jmlr.org/papers/v9/airoldi08a.html.

Barberá, P. 2015. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23 (1): 76–91.

Binding, G., and L. F. Stoetzer. 2023. "Non-separable Preferences in the Statistical Analysis of Roll Call Votes." *Political Analysis* 31 (3): 352–365.

Clarke, A. J. 2020. "Party Sub-Brands and American Party Factions." *American Journal of Political Science* 64 (3): 452–470. https://doi.org/10.1111/ajps.12504

Clinton, J., S. Jackman, and D. Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98 (2): 355–370.

Daudin, J.-J., F. Picard, and S. Robin. 2008. "A Mixture Model for Random Graphs." *Statistics and Computing* 18 (2): 173–183. https://doi.org/10.1007/s11222-007-9046-7

Davis, O. A., M. J. Hinich, and P. C. Ordeshook. 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *American Political Science Review* 64 (2): 426–448.

https://doi.org/10.1017/pan.2026.10058

Diaconis, P., S. Goel, and S. Holmes. 2008. "Horsehoes in Multidimensional Scaling and Local Kernel Methods." *The Annals of Applied Statistics* 2. https://doi.org/10.1214/08-AOAS165

Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.

Duck-Mayr, J. B., and J. Montgomery. 2023. "Ends against the Middle: Measuring Latent Traits when Opposites Respond the Same Way for Antithetical Reasons." *Political Analysis* 31 (4): 606–625. https://doi.org/10.1017/pan.2022.33

Fortunato, S. 2010. "Community Detection in Graphs." *Physics Reports* 486 (3): 75–174.

Friel, N., R. Rastelli, J. Wyse, and A. E. Raftery. 2016. "Interlocking Directorates in Irish Companies Using a Latent Space Model for Bipartite Networks." *Proceedings of the National Academy of Sciences of the United States of America* 113: 6629–6634.

Gower, J. C. 1975. Generalized Procrustes Analysis [in en]. *Psychometrika* 40 (1): 33–51. https://doi.org/10.1007/BF02291478

Handcock, M. S., A. E. Raftery, and J. M. Tantrum. 2007. "Model-Based Clustering for Social Network." *Journal of the Royal Statistical Society, Series A* 170: 301–354.

Harbridge-Yong, L., C. Volden, and A. Wiseman. 2023. "The Bipartisan Path to Effective Lawmaking." *The Journal of Politics* 85: 1048–1063. https://doi.org/10.1086/723805

Hoff, P., A. Raftery, and M. S. Handcock. 2002. "Latent Space Approaches to Social Network Analysis." *Journal of the American Statistical Association* 97 (460): 1090–1098.

Hoff, P. 2007. "Modeling Homophily and Stochastic Equivalence in Symmetric Relational Data." In *Advances in Neural Information Processing Systems*, edited by J. Platt, D. Koller, Y. Singer, and S. Roweis, Vol. 20, Red Hook, NY: Curran Associates, Inc.

Hoff, P. 2009. Multiplicative Latent Factor Models for Description and Prediction of Social Networks [in en]. *Computational & Mathematical Organization Theory* 15 (4): 261–272. https://doi.org/10.1007/s10588-008-9040-4

Hoff, P.. 2021. Additive and Multiplicative Effects Network Models. *Statistical Science* 36 (1): pp. 34–50. Accessed September 12, 2025. https://www.jstor.org/stable/26997946.

Holland, P. W., K. B. Laskey, and S. Leinhardt. 1983. "Stochastic Blockmodels: First Steps." *Social Networks* 5 (2): 109–137. https://doi.org/10.1016/0378-8733(83)90021-7

Jackman, S. 2024. "pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory." R package version 1.5.9, Sydney, Australia: University of Sydney. https://github.com/atahk/pscl/.

Jeon, M., I. H. Jin, M. Schweinberger, and S. Baugh. 2021. "Mapping Unobserved Item-Respondent Interactions: A Latent Space Item Response Model with Interaction Map." *Psychometrika* 86 (2): 378–403.

Karrer, B., and M. E. J. Newman. 2011. "Stochastic Blockmodels and Community Structure in Networks." *Physical Review E* 83 (1): 16107. https://doi.org/10.1103/PhysRevE.83.016107

Kemp, C., J. B. Tenenbaum, T. L. Griffiths, T. Yamada, and N. Ueda. 2006. Learning Systems of Concepts with an Infinite Relational Model. In *Proceedings of the 21st National Conference on Artificial Intelligence AAAI'06*, edited by Y. Gil and R. J. Mooney, Vol. 1, 381–388. Menlo Park, CA: AAAI Press. Accessed August 31, 2025.

Kim, I. S., J. Londregan, and M. Ratkovic. 2018. "Estimating Spatial Preferences from Votes and Text." *Political Analysis* 26 (2): 210–229.

Krivitsky, P. N., M. S. Handcock, A. E. Raftery, and P. D. Hoff. 2009. "Representing Degree Distributions, Clustering, and Homophily in Social Networks with Latent Cluster Random Network Models." *Social Networks* 31: 204–213.

Krivitsky, P. N., and M. S. Handcock. 2008. "Fitting Position Latent Cluster Models for Social Networks with Latentnet." *Journal of Statistical Software* 24 (5). https://doi.org/10.18637/jss.v024.i05

Lee, S., I. K. Kim, J. H. Park, and I. H. Jin. 2026. "Replication Data for: Euclidean Ideal Point Estimation from Roll-Call Data via Distance-Based Bipartite Network Models." Harvard Dataverse, V1. https://doi.org/10.7910/DVN/JHJEGT

Lei, R., and A. Rodriguez. 2025. "A Novel Class of Unfolding Models for Binary Preference Data." *Political Analysis* 33 (1): 32–48.

Lewis, J. B., K. Poole, H. Rosenthal, A. Boche, A. Rudkin, and L. Sonnet. 2025. "Voteview: Congressional Roll-Call Votes Database." Data, https://voteview.com.

Li, W.-J., D.-Y. Yeung, and Z. Zhang. 2011. "Generalized Latent Factor Models for Social Network Analysis." In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence IJCAI'11*, edited by T. Walsh, Vol. 2, 1705–1710. Menlo Park, CA: AAAI Press. Accessed August 31, 2025.

Lipman, E., S. Moser, and A. Rodriguez. 2026. "Explaining Differences in Voting Patterns across Voting Domains Using Hierarchical Bayesian Models." *Political Analysis* 34 (2): 237–257.

Lo, A., S. Olivella, and K. Imai. 2026. "A Statistical Model of Bipartite Networks: Application to Cosponsorship in the United States Senate." *Political Analysis* 34 (3): 451–470.

Marble, W., and M. Tyler. 2022. "The Structure of Political Choices: Distinguishing between Constraint and Multidimensionality." *Political Analysis* 30 (3): 328–345.

McCarty, N. 2016. "In Defense of DW-NOMINATE [in en]." *Studies in American Political Development* 30 (2): 172–184. https://doi.org/10.1017/S0898588X16000110

Miller, K., M. Jordan, and T. Griffiths. 2009. "Nonparametric Latent Feature Models for Link Prediction." In *Advances in Neural Information Processing Systems*, edited by Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, Vol. 22. Red Hook, NY: Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2009/file/437d7d1d97917cd627a34a6a0fb41136-Paper.pdf.

https://doi.org/10.1017/pan.2026.10058

Mørup, M., M. N. Schmidt, and L. K. Hansen. 2011. "Infinite Multiple Membership Relational Modeling for Complex Networks." In *2011 IEEE International Workshop on Machine Learning for Signal Processing*, 1–6. https://doi.org/10.1109/MLSP.2011.6064546

Moser, S., A. Rodríguez, and C. L. Lofland. 2021. "Multiple Ideal Points: Revealed Preferences in Different Domains." *Political Analysis* 29 (2): 139–166.

Nakis, N., N. R. Holm, A. L. Fiehn, and M. Mørup. 2025. "How Low Can You Go? Searching for the Intrinsic Dimensionality of Complex Networks Using Metric Node Embeddings." In *International Conference on Learning Representations (ICLR 2025)*, edited by Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu.

Nowicki, K., and T. A. B. Snijders. 2001. "Estimation and Prediction for Stochastic Blockstructures." *Journal of the American Statistical Association* 96 (455): 1077–1087. Accessed September 1, 2025. https://www.jstor.org/stable/2670253.

Palla, K., D. A. Knowles, and Z. Ghahramani. 2012. "An Infinite Latent Attribute Model for Network Data." In *Proceedings of the 29th International Conference on International Conference on Machine Learning, ICML'12*, edited by J. Langford and J. Pineau, 395–402. Edinburgh: Omnipress.

Poole, K. T., and H. Rosenthal. 2007. *Ideology and Congress*. New York: Routledge.

Raftery, A. E., X. Niu, P. D. Hoff, and K. Y. Yeung. 2012. "Fast Inference for the Latent Space Network Model Using a Case-Control Approximate Likelihood." *Journal of Computational and Graphical Statistics* 21 (4): 909–919.

Schweinberger, M., and T. A. B. Snijders. 2003. "Settings in Social Networks: A Measurement Model." *Sociological Methodology* 33: 307–341.

Sewell, D. K., and Y. Chen. 2015. "Latent Space Models for Dynamic Networks." *Journal of the American Statistical Association* 110: 1646–1657.

Shin, S., J. Lim, and J. H. Park. 2024. "*l*1-Based Bayesian Ideal Point Model for Multidimensional Politics." *Journal of the American Statistical Association* 120 (550): 631–644. https://doi.org/10.1080/01621459.2024.2425461

Smith, A. L., D. M. Asta, and C. A. Calder. 2019. "The Geometry of Continuous Latent Space Models for Network Data." *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 34 (3): 428.

Vu, D. Q., D. R. Hunter, and M. Schweinberger. 2013. "Model-Based Clustering of Large Networks." *The Annals of Applied Statistics* 7 (2): 1010–1039. https://doi.org/10.1214/12-AOAS617

Young, S. J., and E. R. Scheinerman. 2007. "Random Dot Product Graph Models for Social Networks [in en]." In *Algorithms and Models for the Web-Graph*, edited by A. Bonato and F. R. K. Chung, 138–149. Berlin: Springer. https://doi.org/10.1007/978-3-540-77004-611

Zhou, M. 2015. "Infinite Edge Partition Models for Overlapping Community Detection and Link Prediction." In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, edited by G. Lebanon and S. V. N. Vishwanathan, Vol. 38, 1135–1143. San Diego, CA: PMLR.

Cite this article: Lee, S., Kim, I. K., Park, J. H. and Jin, I. H. (2026). Euclidean Ideal Point Estimation from Roll-Call Data via Distance-Based Bipartite Network Models. *Political Analysis*, 1–18. https://doi.org/10.1017/pan.2026.10058

https://doi.org/10.1017/pan.2026.10058