



# Bayesian Inference for Multidimensional Scaling Representations with Psychologically Interpretable Metrics

Quentin F. Gronau<sup>1</sup> · Michael D. Lee<sup>2</sup>

Published online: 8 July 2020  
© The Author(s) 2020

## Abstract

Multidimensional scaling (MDS) models represent stimuli as points in a space consisting of a number of psychological dimensions, such that the distance between pairs of points corresponds to the dissimilarity between the stimuli. Two fundamental challenges in inferring MDS representations from data involve inferring the appropriate number of dimensions and the metric structure of the space used to measure distance. We approach both challenges as Bayesian model-selection problems. Treating MDS as a generative model, we define priors needed for model identifiability under metrics corresponding to psychologically separable and psychologically integral stimulus domains. We then apply a differential evolution Markov-chain Monte Carlo (DE-MCMC) method for parameter inference, and a Warp-III method for model selection. We apply these methods to five previous data sets, which collectively test the ability of the methods to infer an appropriate dimensionality and to infer whether stimuli are psychologically separable or integral. We demonstrate that our methods produce sensible results, but note a number of remaining technical challenges that need to be solved before the method can easily and generally be applied. We also note the theoretical promise of the generative modeling perspective, discussing new and extended models of MDS representation that could be developed.

**Keywords** Multidimensional scaling · Bayesian model selection · MDS dimensionality · Separable and integral stimuli · Warp-III bridge sampling

## Introduction

Multidimensional scaling (MDS) was developed in the 1950s in cognitive psychology as a statistical method for making inferences about human mental representations (Shepard 1957, 1962; Kruskal 1964). MDS models the similarities or psychological proximities between pairs of stimuli, representing each stimulus as a point in a multidimensional space, such that more similar stimuli are nearer each other. The core psychological motivation is that the similarities reflect the basic cognitive process of generalization. Generalization can be thought of as the ability to treat two stimuli as being the same, and has been argued to serve as a basis for the mental organization of

knowledge, and the capability of the mind to make adaptive predictions about properties and consequences (Shepard 1987). For these reasons, mental representations found via MDS methods have been and remain widely used in cognitive process models of identification, categorization, and decision making (e.g., Nosofsky 1992).

Soon after its development in cognitive psychology, however, MDS algorithms found application as a statistical method that produces a low-dimensional representation of a set of objects, based on a measure of the similarities between them. As a data-reduction or visualization method, MDS has been applied in the natural, biological, and human sciences, with application areas as diverse as representing the similarities of skulls in archaeology, the tastes of colas in marketing, and the voting patterns of senators in politics (e.g., Borg and Groenen 1997; Cox and Cox 1994; Schiffman et al. 1981).

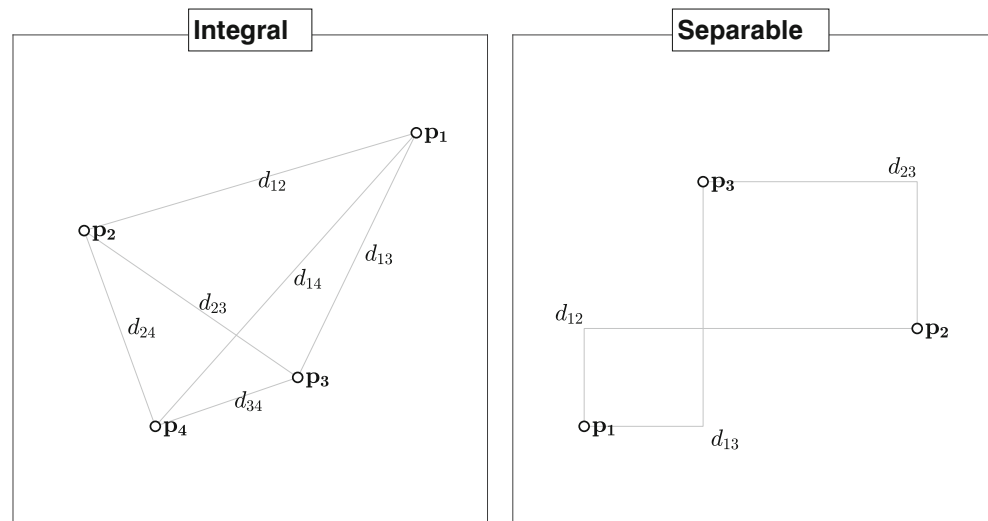
Whether viewed as a model of psychological representation or a data-reduction method, a foundational challenge in MDS modeling is determining the dimensionality  $M$  of the representational space. In his 1974 Presidential Address to the Psychometric Society, Roger Shepard identified six

✉ Quentin F. Gronau  
Quentin.F.Gronau@gmail.com

<sup>1</sup> Department of Psychological Methods, University of Amsterdam, 1012 WX Amsterdam, Netherlands

<sup>2</sup> Department of Cognitive Sciences, University of California Irvine, Irvine, CA 92697, USA

**Fig. 1** MDS representations with integral and separable metric structures



basic challenges for MDS, the third of which was “The problem of determining the proper number of dimensions for the coordinate embedding space” (Shepard 1974, p. 377). A number of methods for solving the problem of MDS dimensionality have been developed in both statistics and psychology. The most common approach is a scree test that aims to identify an “elbow” in the goodness-of-fit as dimensionality increases (Cox and Cox 1994; Kruskal 1964; Schiffman et al. 1981). Steyvers (2006) suggests the use of cross-validation methods, although this approach does not seem to be widely used.

Since choosing the correct dimensionality of an MDS is naturally regarded as a model-selection problem—that is, choosing between a one-dimensional versus two-dimensional versus three-dimensional representation, and so on—the statistically principled approach offered by Bayes factors should provide a solution (Kass and Raftery 1995). Along these lines, Lee (2001) implements an approach based on the Bayesian Information Criterion (BIC). The difference between BIC values for representations with different dimensionality provides a crude approximation to the Bayes factor. Oh and Raftery (2001) provide a different approach to approximation by computing the marginal likelihoods of different representations using plug-in point estimates for the stimulus locations. This is an approximation because the exact Bayes factor requires an integration across the stimulus location parameters. Oh (2012) developed a method based on spike-and-slab priors, in which the dimensionality is determined by the marginal posterior probabilities for each dimension that the coordinate locations are not zero for all stimuli.

From the perspective of MDS as psychological models however, none of these approaches qualifies as being principled and complete. The key issue is that the theory of mental representation developed by Shepard (1957, 1987,

1991) emphasizes the role that the metric structure of the space plays in capturing key psychological properties of the stimuli. In particular, the idea is that different metrics capture the theoretical and empirical distinction between separable and integral stimuli (Attneave 1950; Garner 1974). Separable stimuli are those for which the component dimensions can be attended to separately. An example is different shapes of different sizes, since it is possible for people to attend selectively to either shape or the size. Integral stimuli, by contrast, are those for which the component dimensions cannot be attended to independently. The standard example is color, since it is typically not possible for people to attend selectively to the underlying hue, saturation, and brightness components.

Figure 1 shows how different metric structures are used to represent integral and separable stimuli. In the left panel, there are four stimuli, represented by the points  $p_1, \dots, p_4$ . The pairwise distances between these points, such as  $d_{12}$  between the first point and the second point, are modeled using the Euclidean metric, and so correspond to standard straight lines. In the right panel, there are three stimuli, and the pairwise distances between them are modeled according to the city-block metric. Intuitively, this corresponds to comparing the stimuli on each underlying dimension independently, then adding those dissimilarities to get an overall measure of dissimilarity.

Admittedly, this account of integrality and separability is a theoretical and empirical caricature, and much more nuanced and detailed accounts are possible (Shepard 1991; Tversky and Gati 1982). The point is that psychological representations based on MDS need to make assumptions about the metric structure of the space, and use metrics other than the Euclidean metric. As Jäkel et al. (2008, p. 2) point out, from the origins of MDS as a psychological model, “There was no a priori reason to believe that mental

representations should be Euclidean.” Previous methods for determining the dimensionality of MDS representations using Bayesian model selection, however, have either been insensitive to the metric structure of the representation (Lee 2001) or have focused on the Euclidean metric (Oh 2012; Oh and Raftery 2001).

The use of non-Euclidean metrics raises another challenge, related to inferring MDS representations themselves. There is evidence that it can be computationally difficult to find multidimensional city-block MDS representations (Groenen et al. 1998; Hubert et al. 1992), as well as finding unidimensional MDS representations (Mair and Leeuw 2014). Given that these difficulties stem from basic geometric properties of the MDS representations, it seems likely they will continue to present an issue for Bayesian methods of inference.

Finally, there is the challenge of inferring the appropriate metric structure for an MDS representation. Shepard (1991) reviews the original statistical approach to this problem, which involved applying non-metric MDS algorithms for a large number of different metrics, and choosing the one with the best goodness-of-fit. As Lee (2008) pointed out, this approach neglects to account for the component of model complexity that arises from the functional form of parameter interaction (Pitt et al. 2006), which is often the only difference between MDS models using different psychologically interpretable metrics. Lee (2008) developed a Bayesian approach in which the possible metrics correspond to a parameter that is inferred jointly with the coordinate location parameters that represent the stimuli. Okada and Shigemasa (2010) developed and tested this approach further, and showed it is capable of recovering the correct metric in simulation studies. Both the Lee (2008) and Okada and Shigemasa (2010) methods, however, failed to resolve basic challenges in model identifiability that arise from treating the choice of metric structure as a parameter inference problem. It is possible these identifiability issues could be addressed by considering the choice as a model-selection problem, and restricting the set of possibilities to a few interpretable metrics.

Accordingly, the goals of this article are to examine the implementation of MDS models that use psychologically interpretable metrics, including both the Euclidean and a non-Euclidean metric, and explore the possibility of inferring the appropriate dimensionality and metric structure of these representations using Bayesian model-selection methods. The structure of the remainder of the article is as follows. In the next section, we define MDS models, and address the issue of model identifiability under different metrics. Consistent with previous literature, we argue that the city-block metric presents fundamental problems in making MDS representations identifiable. This leads to

the development of joint prior distributions on the stimulus location parameters for the Euclidean metric, and non-Euclidean metrics other than the city-block metric. With these priors established, we apply an approach to Bayesian inference using differential evolution Markov-chain Monte Carlo (DE-MCMC) computational sampling methods. The DE-MCMC method helps address the difficulties inherent in inferring MDS representations, which are especially evident in non-Euclidean cases. We then use the Warp-III bridge sampling method to approximate the marginal densities needed to determine Bayes factors. We apply the method to five previously studied data sets, differing in the type of stimuli and expected dimensionality of their MDS representation. For all five applications, the method makes sensible inferences about dimensionality, and produces interpretable stimulus representations. We conclude with a discussion of remaining statistical and computational challenges, and potential directions for refining and extending the approach.

## MDS Model Identifiability

### The Identifiability Problem

Formally, suppose there are  $N$  stimuli to be represented, based on observed proximity data from  $P$  participants, with  $d_{ijk}$  measuring the proximity between the  $i$ th and  $j$ th stimulus provided by the  $k$ th participant. We assume these observed proximities are normalized to lie between 0 and 1. The point representing the  $i$ th stimulus in a  $M$ -dimensional space is  $\mathbf{p}_i = (p_{i1}, \dots, p_{iM})$  and the distance between points  $\mathbf{p}_i$  and  $\mathbf{p}_j$  is measured by the Minkowski metric with metric parameter  $r$ , so that

$$\hat{d}_{ijk} = \left( \sum_{m=1}^M |p_{im} - p_{jm}|^r \right)^{1/r}. \quad (1)$$

The Minkowski metric has special cases of the city-block metric when  $r = 1$  and the Euclidean metric when  $r = 2$ . Values of  $r$  between 1 and 2 can potentially be interpreted as intermediate assumptions about the independence of stimulus dimensions between the end point of complete separability and complete integrality.

The goal of MDS is for the modeled distances  $\hat{d}_{ijk}$  to correspond to the observed proximities  $d_{ijk}$ . We use the probabilistic model

$$d_{ijk} \sim \text{Gaussian} \left( \hat{d}_{ijk}, \frac{1}{\sigma^2} \right), \quad (2)$$

where  $\sigma$  is the standard deviation with which the observed proximities are measured.<sup>1</sup> It is assumed to be the same for all of the proximities, and is given a prior

$$\sigma \sim \text{TruncatedGaussian}\left(0.15, \frac{1}{0.2^2}\right) T(0, \infty), \quad (3)$$

where the  $T(0, \infty)$  indicates the sampled value is truncated to be a positive real number. This is an informative prior (Lee and Vanpaemel 2018), consistent with previous data and modeling. Intuitively,  $\sigma$  corresponds to the average standard deviation of different individual ratings of the same pair of stimuli. Empirical estimates of this standard deviation in previous data tend to range from about 0.1 to about 0.2 (Lee 2001; Lee and Pope 2003).<sup>2</sup> Accordingly, the prior is centered on 0.15, but allows a wide range of possibilities.

We note that this MDS model does not incorporate individual differences. It is assumed that the same point  $p_i$  represents the  $i$ th stimulus for all participants. We also emphasize, however, that individual-level proximity data  $d_{ijk}$  are modeled, rather than averaged or aggregated data across participants. The problems inherent in averaging data have long been understood (Estes 1956), and have been studied in the specific cognitive modeling context provided by MDS representations (Lee and Pope 2003). Our approach is to require the same underlying MDS representation to provide an account of each individual proximity matrix.

To complete the generative model, a straightforward approach would be to give all of the coordinate locations for the representational points uniform priors  $p_{im} \sim \text{Uniform}(-1, 1)$ . These priors, however, made the model non-identifiable, because the distances between points are invariant under transformations (Borg and Groenen 1997, Ch. 2). The distances between points are preserved under translation, reflection, axes permutation (for non-Euclidean metrics), and rotation (for the Euclidean metric). A principled Bayesian approach for controlling these invariances to ensure model identifiability constrains the coordinate location parameters through a joint prior distribution that depends on the assumed metric.

## Previous Approaches

Existing MDS modeling methods that use Bayesian inference almost always rely on post-processing to address the issue of identifiability. The method developed by Lee (2008) post-processes posterior samples of the coordinate

location parameters to control for translation, reflection, and permutation. For example, to control for translation, the method zero centers every posterior sample of the sets of coordinate location. The Lee (2008) method does not control for rotation, which is problematic, because the method also attempts to infer the  $r$  metric parameter, and so the inferred representational space can have a Euclidean metric, which requires rotational invariance.

Most other methods, in contrast, assume the MDS space is Euclidean. The post-processing of the coordinate location parameters used by both Oh and Raftery (2001) and Oh (2012) assumes a Euclidean space and controls for translation, reflection, and rotation. Okada and Mayekawa (2018) extend the approach developed by Okada (2012), which relies on Procrustes analysis. Their post-processing uses a loss function to align posterior samples of the coordinate location, but again assumes a Euclidean space.

Besides the lack of flexibility in the nature of the distance metric, post-processing methods have the effect of implementing modeling assumptions without explicitly specifying those assumptions as part of the model. While this is often practical, it is theoretically inelegant, and contrary to the goals of generative modeling. Ideally, the constraints required for model identifiability should be part of the model itself. In the case of MDS models, these constraints are naturally imposed through the specification of a joint prior over the coordinate location parameters that addresses the transformational invariances, removes the need for post-processing, and makes bridge sampling feasible.

This generative approach is used by the “parameter fixing” method considered by Okada and Mayekawa (2018), who evaluate it as a contrast with the Procrustes methods that are their focus. Parameter fixing corresponds to setting a structured joint prior over the coordinate location parameters. Okada and Mayekawa (2018) define the appropriate prior for a Euclidean space using results provided by Bakker and Poole (2013), which were derived using an analytic method based on matrix properties.

Our goal is to extend this approach to include non-Euclidean representations. We start by considering one-dimensional MDS representations, before considering multidimensional representations in both Euclidean and non-Euclidean metric spaces. We take a geometric approach to identifying the required joint priors for invariance constraints, complementing the non-geometric approach of Bakker and Poole (2013) for the Euclidean metric.

## One-dimensional Representation

For a one-dimensional representation, all of the psychologically interpretable metrics we consider give the same

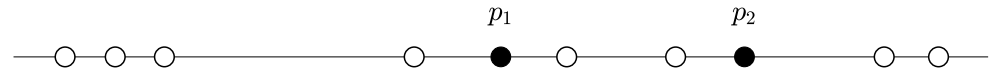
<sup>1</sup>We parameterize the Gaussian distribution in terms of mean and precision parameters, for consistency with the JAGS graphical modeling language.

<sup>2</sup>See also the data repository at <https://osf.io/ey9vp/>

**Fig. 2** Identification constraints for a one-dimensional representation

$$p_1 = 0$$

$$p_2 \geq 0$$



distances. The required constraints on the points are shown in Fig. 2, with one point fixed at the origin to control translation, and second point restricted to be positive to control reflection.

These constraints can be formalized by a joint prior with

$$\begin{aligned} p_1 &= 0 \\ p_2 &\sim \text{Uniform}(0, 1) \\ p_3, \dots, p_N &\sim \text{Uniform}(-1, 1). \end{aligned} \quad (4)$$

### Euclidean Multidimensional Representations

Figure 3 shows the constraints needed to identify Euclidean MDS representations in two and three dimensions. In the two-dimensional case, the first point  $p_1$  is fixed at the origin, to control translation, the second point  $p_2$  is constrained to the positive  $x$ -axis, to control reflection in the  $y$ -axis and rotation, and the third point  $p_3$  is constrained to have a positive  $y$ -value to control reflection in the  $x$ -axis. The same logic is applied in the three-dimensional case, with  $p_1$  controlling translation,  $p_2$  and  $p_3$  controlling reflection and rotation in successive axes, and  $p_4$  controlling the final reflection.

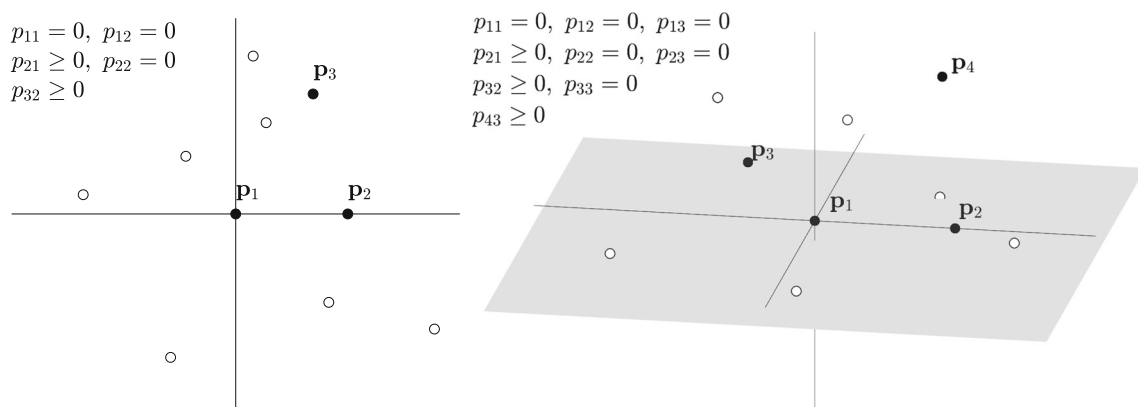
These are the first two cases of a general pattern, clear by induction, that applies to a  $M$ -dimensional representation, and corresponds to the matrix result provided by Bakker and Poole (2013). An intuitive presentation of the inductive pattern is shown below, where “0” denotes fixing a

coordinate location to zero, “+” denotes constraining it to be positive, and “ $\mathcal{R}$ ” denotes imposing no constraint.

	Dim 1	Dim 2	Dim 3	Dim 4		Dim $D$
$\boldsymbol{p}_1$	0	0	0	0	...	0
$\boldsymbol{p}_2$	+	0	0	0	...	0
$\boldsymbol{p}_3$	$\mathcal{R}$	+	0	0	...	0
$\boldsymbol{p}_4$	$\mathcal{R}$	$\mathcal{R}$	+	0	...	0
$\boldsymbol{p}_5$	$\mathcal{R}$	$\mathcal{R}$	$\mathcal{R}$	+	...	0

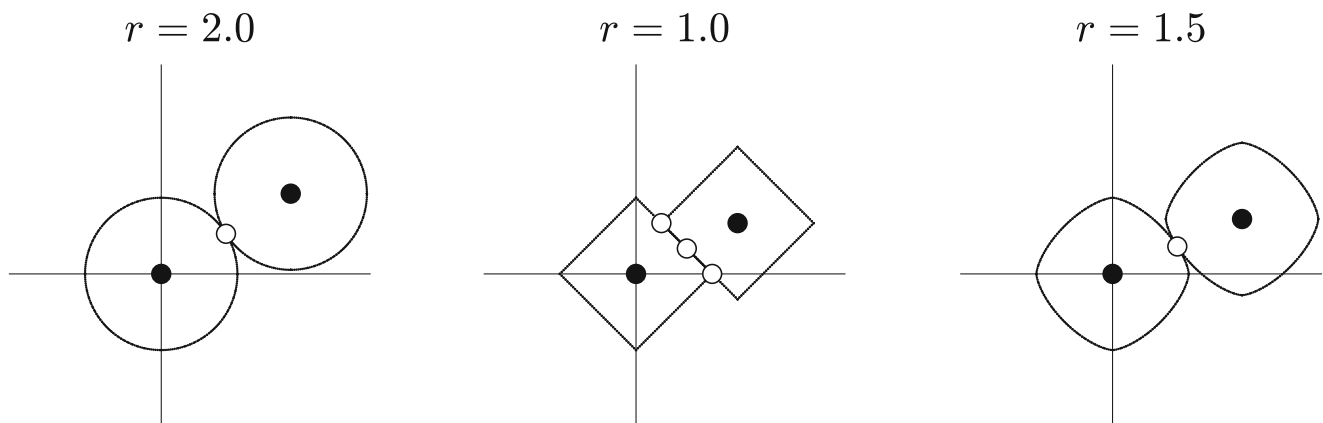
Formally, these constraints in  $D$  dimensions correspond to the joint prior

$$\begin{aligned} p_{11}, \dots, p_{1D} &= 0 \\ p_{21} &\sim \text{Uniform}(0, 1) \\ p_{22}, \dots, p_{2D} &= 0 \\ p_{31} &\sim \text{Uniform}(-1, 1) \\ p_{32} &\sim \text{Uniform}(0, 1) \\ p_{33}, \dots, p_{3D} &= 0 \\ p_{41}, p_{42} &\sim \text{Uniform}(-1, 1) \\ p_{43} &\sim \text{Uniform}(0, 1) \\ p_{44}, \dots, p_{4D} &= 0 \\ &\dots \end{aligned} \quad (5)$$



**Fig. 3** Identification constraints for Euclidean representations in two dimensions (left) and three dimensions (right)

## Non-Euclidean Multidimensional Representations



**Fig. 4** The nature of iso-distance curves and the identifiability of mid-points for the three Minkowski metrics corresponding to  $r = 2$  (Euclidean),  $r = 1$  (city-block), and  $r = 1.5$

### Non-Euclidean Multidimensional Representations

Finding constraints for invariance in non-Euclidean metrics is more complicated, and is especially difficult for the city-block metric. The basic geometric problem was noted as early as Arnold (1971), and discussed in Shepard's (1974) presidential address. A simple demonstration of the fundamental problem is provided by Fig. 4. The three panels correspond to Euclidean ( $r = 2$ ), city-block ( $r = 1$ ), and a general non-Euclidean ( $r = 1.5$ ) metric, and show unit iso-distance contours around the same two points in each metric, shown as black dots. These iso-distance contours are the “unit circles” of each metric, showing all the points in the space that are the same distance from the two points. For the Euclidean metric, these contours are familiar circles, and coincide at only one point, shown by the white dot. This means that there is a unique point in the space that is equally distant from the two points shown by black dots. In the context of an MDS representation, a stimulus that is equally different to both of the points can be uniquely identified.

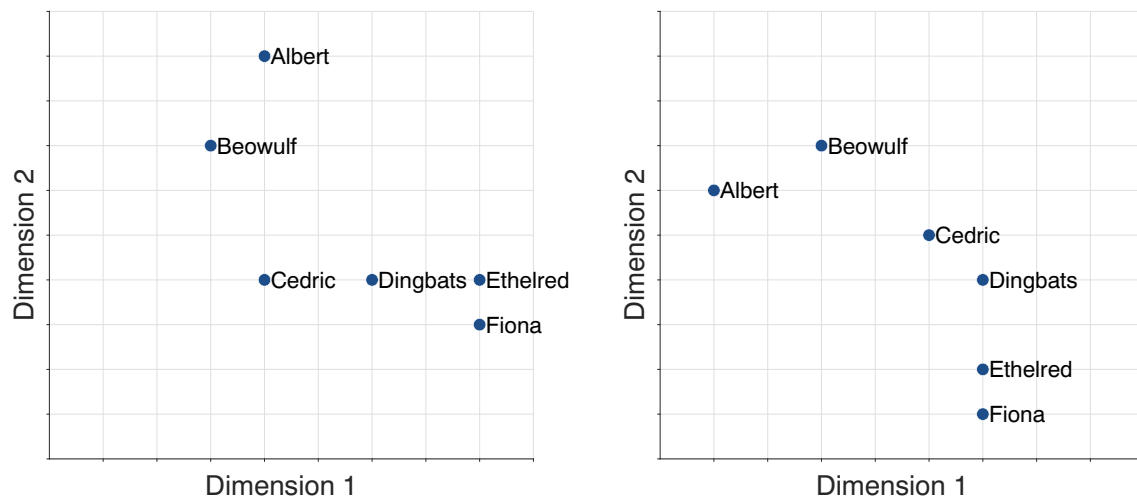
For the city-block case, however, the iso-distance contours are diamonds, and there are infinitely many points that are equally different. Three specific possibilities are shown by white dots, but clearly any point along the line where the iso-distance contours coincide is possible. In the context of an MDS representation, this means that there is a fundamental difficulty in identifying a stimulus that is equally different to both of the points. This basic problem is not, in general, solved by the introduction of additional stimuli that provide additional constraints. Indeed, the problem compounds for potential city-block representations with many stimuli. Bortz (1974, see, especially, Figures 2 and 3) provides compelling examples, and the same point is

emphasized in the seminal text by Borg and Groenen (1997, pp. 369–372).

Figure 5 provides a concrete example, based on the more general configuration examined by Borg and Groenen (1997, Figure 17.6). Each panel shows a representation of six fictitious people in terms of two underlying dimensions. The city-block distance between each pair of people is identical in both configurations. This means, of course, that this proximity matrix is equally consistent with both representations, and either could be inferred from the data. But, the two representations are substantively different, in non-trivial ways. The representations do not differ simply by changing the axes, and have basic structural differences. For example: Cedric, Dingbats, and Ethelred are co-linear in the first representation, but not in the second, where Dingbats, Ethelred and Fiona become co-linear; the ordering of Albert and Beowulf changes on both dimensions between the configurations; and so on. In fact, once the lack of invariance revealed by the Borg and Groenen (1997, Figure 17.6) analysis is understood, it is clear that many additional representations for the proximity between the six people could be constructed, supporting a wide range of different meaningful interpretations.

A practical approach for identifying city-block representations, used by Nosofsky (1985), relies on determining the values of some stimuli on some dimensions, by means external to the MDS modeling. Ultimately, this strategy can solve the problem, if it is possible to find the values of every stimulus on every dimension. But, Fig. 5 suggests the strategy may not be effective in situations where the identification of just a few stimuli is possible. In both representations, Dingbats is at the same location,





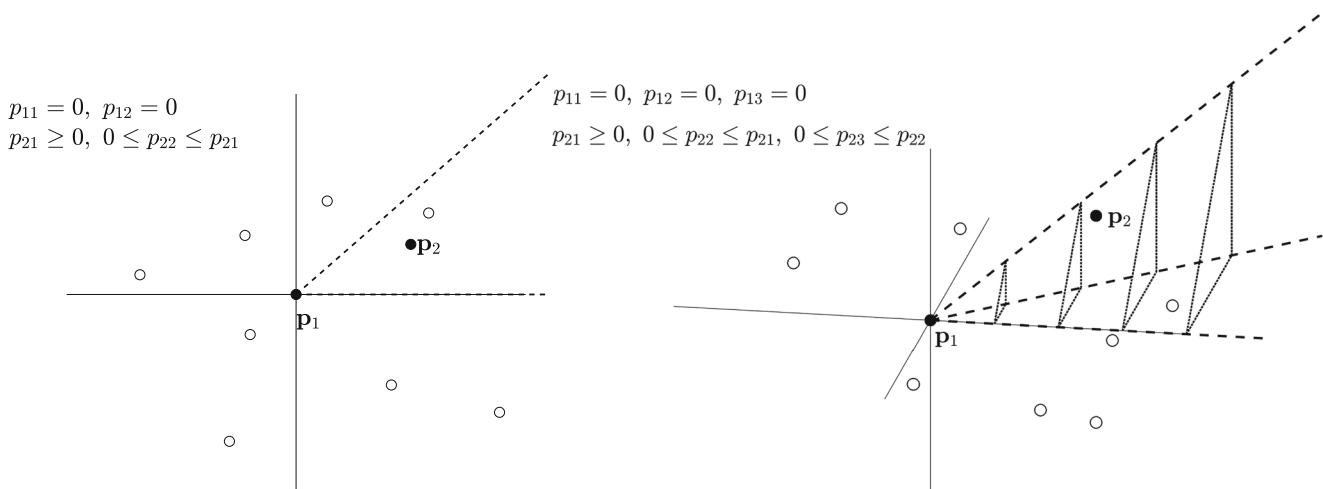
**Fig. 5** Two city-block representations of six fictitious people in terms of two dimensions. Both representations have identical proximity matrices

consistent with values on dimensions having been externally determined, yet the locations of the remaining stimuli are under-determined. In addition, if, for example, Albert was additionally identified as being located in the position shown in the first representation, that would constrain the inference about Beowulf and Cedric, but would not constrain Ethelred and Fiona, who could still be inferred to be at either of the possibilities shown in the two representations. Thus, while the addition of stimuli, or the identification of dimension values for some stimuli, may work in some specific circumstances, we do not believe either represents a general approach to making city-block MDS representations identifiable.

We do not know how to solve the problem of MDS model invariance for the city-block metric. As the right-most panel of Fig. 4 makes clear, however, the problem does not occur for Minkowski-metric parameters  $r > 1$ . For the  $r = 1.5$  metric, the iso-distance contours again

coincide at only one point. The asymmetry of these contours makes clear they do not have the rotational invariance of the Euclidean  $r = 2$  metric. In this way, general non-Euclidean metrics, such as  $r = 1.5$ , capture the psychological idea that the dimensions in an MDS representation have meaning and allow selective attention, while avoiding the degenerate lack of identifiability inherent in the city-block metric.

Figure 6 shows the constraints needed to identify these sort of non-Euclidean MDS representations in two and three dimensions. In the two-dimensional case, the first point  $p_1$  is once again fixed at the origin, to control translation, the second point  $p_2$  is constrained to the positive quadrant to control reflection. In addition, the constraint that  $p_{22} \leq p_{21}$  is imposed, requiring the value of the second stimulus on the y-axis not to be larger than its value on the x-axis. This constraint controls for axis permutation, preventing the two dimensions from being swapped, and so allocates a



**Fig. 6** Identification constraints for non-Euclidean representations in two dimensions (left) and three dimensions (right)

specific underlying stimulus dimension to each axis. The three-dimensional case extends this logic by requiring that the  $z$ -axis value of the second point be positive, to prevent reflection, and be less than the value of the second point on the  $y$ -axis, to prevent permutation.

These first two cases once again make clear a general pattern, in which the coordinate values of the second point are positive and order constrained.<sup>3</sup> Formally, the constraints for non-city-block but non-Euclidean  $D$  dimensions are

$$\begin{aligned} p_{11}, \dots, p_{1D} &= 0 \\ p_{21}, \dots, p_{2D} &\sim \text{Uniform}(0, 1) : p_{21} \geq \dots \geq p_{2D} \\ p_{31}, \dots, p_{3D} &\sim \text{Uniform}(-1, 1) \\ &\dots \end{aligned} \quad (6)$$

## Bayesian MDS Inference via DE-MCMC

When posterior samples for MDS models are obtained using conventional Markov-chain Monte Carlo algorithms (MCMC; e.g., Gamerman & Lopes, 2006), it can occur that chains get stuck in local maxima. In our experience, the reason is typically that the stimuli that are constrained are similar to each other. To prevent local maxima, we implemented a heuristic to order the stimuli in a way that those defining the constraints are dissimilar. We motivate and describe this heuristic in detail in Appendix 1. In addition, to improve sampling, we used the differential evolution Markov-chain Monte Carlo algorithm (DE-MCMC; e.g., Heathcote et al. in press; Turner et al. 2013) that helps to guide the chains to regions of high posterior density.

DE-MCMC is a population-based MCMC algorithm that generates efficient proposals via a population of interacting chains (Turner et al. 2013). One strength of the algorithm is that it works well for highly correlated target distributions. However, we used DE-MCMC primarily for the reason that the interacting chains can guide each other to regions of high posterior density which helps to avoid the issue of chains getting stuck in local maxima. Specifically, during burn-in, we used a *migration step* that remedies the problem of outlier chains in an effective manner (for details, see Turner et al. 2013, Appendix 2). We found that the combination of the ordering heuristic and DE-MCMC provides effective sampling consistently for the Euclidean metric, and is partially effective for non-Euclidean metrics.

<sup>3</sup>These order constraints can be imposed either in decreasing manner, as shown in Fig. 6 for easier visualization, or in an increasing manner, as they are in our code.

## Bayesian Model Comparison via Bridge Sampling

### Marginal Likelihood

Comparing MDS models with different dimensions and metrics via Bayes factors and posterior model probabilities requires the computation of the marginal likelihood for all of the models,  $\mathcal{M}_{m,r}$ , being considered where  $m$  denotes the dimensionality and  $r$  the metric. Let  $\mathbf{D}$  denote the observed data (i.e., the pairwise dissimilarity ratings  $d_{ijk}$ ) and  $\mathbf{P}$  denote the  $N \times m$  matrix with the latent stimulus coordinates for each stimulus. The marginal likelihood for model  $\mathcal{M}_{m,r}$  corresponds to the normalizing constant of the joint posterior distribution for  $\theta = (\mathbf{P}, \sigma)$ :

$$\begin{aligned} p(\mathbf{D} | \mathcal{M}_{m,r}) &= \int q(\theta | \mathbf{D}, \mathcal{M}_{m,r}) d\theta \\ &= \int \underbrace{\underbrace{p(\mathbf{D} | \mathbf{P}, \sigma, \mathcal{M}_{m,r})}_{\text{Likelihood}}}_{\text{Joint Prior on Stimulus Locations}} \underbrace{p(\mathbf{P} | \mathcal{M}_{m,r})}_{\text{Prior on Imprecision}} d\mathbf{P} d\sigma, \end{aligned} \quad (7)$$

where  $q(\theta | \mathbf{D}, \mathcal{M}_{m,r})$  denotes the unnormalized joint posterior density.

### Bridge Sampling

Since the marginal likelihood in Eq. 7 is not available analytically, we use Warp-III bridge sampling (Meng and Schilling 2002) to estimate this potentially high-dimensional integral. Bridge sampling (Meng and Wong 1996; for a recent tutorial, see Gronau et al. 2017) is based on the following identity:

$$p(\mathbf{D} | \mathcal{M}_{m,r}) = \frac{\mathbb{E}_{g(\theta)} [h(\theta) q(\theta | \mathbf{D}, \mathcal{M}_{m,r})]}{\mathbb{E}_{p(\theta | \mathbf{D}, \mathcal{M}_{m,r})} [h(\theta) g(\theta)]}, \quad (8)$$

where the numerator is an expected value with respect to a proposal distribution  $g(\theta)$ , the denominator is an expected value with respect to the parameter posterior distribution  $p(\theta | \mathbf{D}, \mathcal{M}_{m,r})$ , and  $h(\theta)$  is a function such that  $0 < \int h(\theta) p(\theta | \mathbf{D}, \mathcal{M}_{m,r}) g(\theta) d\theta < \infty$ . The bridge sampling estimate is obtained by sampling from the proposal distribution  $g(\theta)$  and the posterior distribution  $p(\theta | \mathbf{D}, \mathcal{M}_{m,r})$  to approximate the two expected values. Meng and Wong (1996) showed that the optimal choice for  $h(\theta)$  is given by

$$h_o(\theta) \propto [s_1 q(\theta | \mathbf{D}, \mathcal{M}_{m,r}) + s_2 p(\mathbf{D} | \mathcal{M}_{m,r}) g(\theta)]^{-1}, \quad (9)$$

where  $s_i = n_i / (n_1 + n_2)$ ,  $i \in \{1, 2\}$ ,  $n_1$  denotes the number of samples from the posterior  $p(\theta | \mathbf{D}, \mathcal{M}_{m,r})$ , and  $n_2$  denotes the number of samples from the proposal  $g(\theta)$ . The optimal choice for  $h(\theta)$  depends on the marginal



likelihood of interest. Therefore, in practice, the bridge sampling estimate is obtained via an iterative scheme, presented below, that updates an initial guess of the marginal likelihood until convergence.

The variability of the bridge sampling estimate is governed not only by the number of samples but also, by the overlap between the proposal and the posterior distribution. To obtain estimates with low variability, it is therefore prudent to maximize the overlap between these two distributions. The Warp-III approach attempts to create a large overlap by fixing the proposal to a standard multivariate Gaussian distribution and then manipulating (i.e., “warping”) the posterior in a way that matches the first three moments of the two distributions.<sup>4</sup> Crucially, the warping procedure retains the normalizing constant of the posterior (i.e., the marginal likelihood of interest).

A prerequisite for the warping procedure is that all elements of the parameter vector are allowed to range across the entire real line. This can be achieved via a change-of-variables of the form  $\boldsymbol{\zeta} = f(\boldsymbol{\theta})$ , where  $f$  is a suitable<sup>5</sup> vector-valued function that transforms the constrained elements of  $\boldsymbol{\theta}$  so that all elements of  $\boldsymbol{\zeta}$  are unconstrained.<sup>6</sup> The Warp-III procedure is based on the following stochastic transformation of the unconstrained parameter vector  $\boldsymbol{\zeta}$ :

$$\boldsymbol{\eta} = b \mathbf{C}^{-1} (\boldsymbol{\zeta} - \boldsymbol{\mu}), \quad (10)$$

where  $b \sim \text{Bernoulli}(0.5)$  on  $\{-1, 1\}$ ,  $\boldsymbol{\mu}$  denotes the expected value vector of the posterior samples, and  $\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}^\top$  denotes the posterior covariance matrix (i.e.,  $\mathbf{C}$  is obtained via a Cholesky decomposition).

Figure 7 illustrates the warping approach for the univariate case. In the upper-left panel, the solid line corresponds to the standard Gaussian proposal distribution and the gray histogram depicts synthetic posterior samples. Subtracting the posterior mean from all samples matches the

first moment of the proposal and the posterior distribution, as shown in the upper-right panel. Dividing all samples by the posterior standard deviation matches the second moment of the two distributions, as shown in the lower-right panel. Finally, attaching a minus sign with probability 0.5 to the posterior samples achieves symmetry and thus matches the third moment of the proposal and the posterior distribution, as shown in the lower-left panel.

The Warp-III bridge sampling estimate based on  $h_o(\boldsymbol{\theta})$  is computed via an iterative scheme where the value of the estimate at iteration  $t$  is given by (for more details see, Gronau et al. 2019):

$$\hat{p}(\mathbf{D} \mid \mathcal{M}_{m,r})^{(t+1)} = \frac{\frac{1}{n_2} \sum_{i=1}^{n_2} \frac{l_{2,i}}{s_1 l_{2,i} + s_2 \hat{p}(\mathbf{D} \mid \mathcal{M}_{m,r})^{(t)}}}{\frac{1}{n_1} \sum_{j=1}^{n_1} \frac{1}{s_1 l_{1,j} + s_2 \hat{p}(\mathbf{D} \mid \mathcal{M}_{m,r})^{(t)}}}, \quad (11)$$

with

$$l_{1,j} = \frac{|\hat{\mathbf{C}}|}{2} \left[ q(2\hat{\boldsymbol{\mu}} - \boldsymbol{\zeta}_j^* \mid \mathbf{D}, \mathcal{M}_{m,r}) + q(\boldsymbol{\zeta}_j^* \mid \mathbf{D}, \mathcal{M}_{m,r}) \right], \quad (12)$$

and

$$l_{2,i} = \frac{|\hat{\mathbf{C}}|}{2} \left[ q(\hat{\boldsymbol{\mu}} - \hat{\mathbf{C}}\tilde{\boldsymbol{\eta}}_i \mid \mathbf{D}, \mathcal{M}_{m,r}) + q(\hat{\boldsymbol{\mu}} + \hat{\mathbf{C}}\tilde{\boldsymbol{\eta}}_i \mid \mathbf{D}, \mathcal{M}_{m,r}) \right]. \quad (13)$$

In Eqs. 12–13,  $q(\cdot \mid \mathbf{D}, \mathcal{M}_{m,r})$  denotes the unnormalized posterior density with respect to the unconstrained parameter vector  $\boldsymbol{\zeta}$ ,  $\{\boldsymbol{\zeta}_1^*, \boldsymbol{\zeta}_2^*, \dots, \boldsymbol{\zeta}_{n_1}^*\}$  denote  $n_1$  posterior samples, and  $\{\tilde{\boldsymbol{\eta}}_1, \tilde{\boldsymbol{\eta}}_2, \dots, \tilde{\boldsymbol{\eta}}_{n_2}\}$  denote  $n_2$  samples from the standard multivariate Gaussian proposal distribution. To compute the Warp-III estimate, one obtains  $2n_1$  posterior samples: the first half of these samples is used to approximate  $\boldsymbol{\mu}$  and  $\mathbf{C}$  with their sample versions  $\hat{\boldsymbol{\mu}}$  and  $\hat{\mathbf{C}}$ , the second half of the posterior samples is used in the iterative scheme (i.e., Eq. 11). We use the `bridgesampling` R package (Gronau et al. in press) to compute the bridge sampling estimate in Eq. 11.

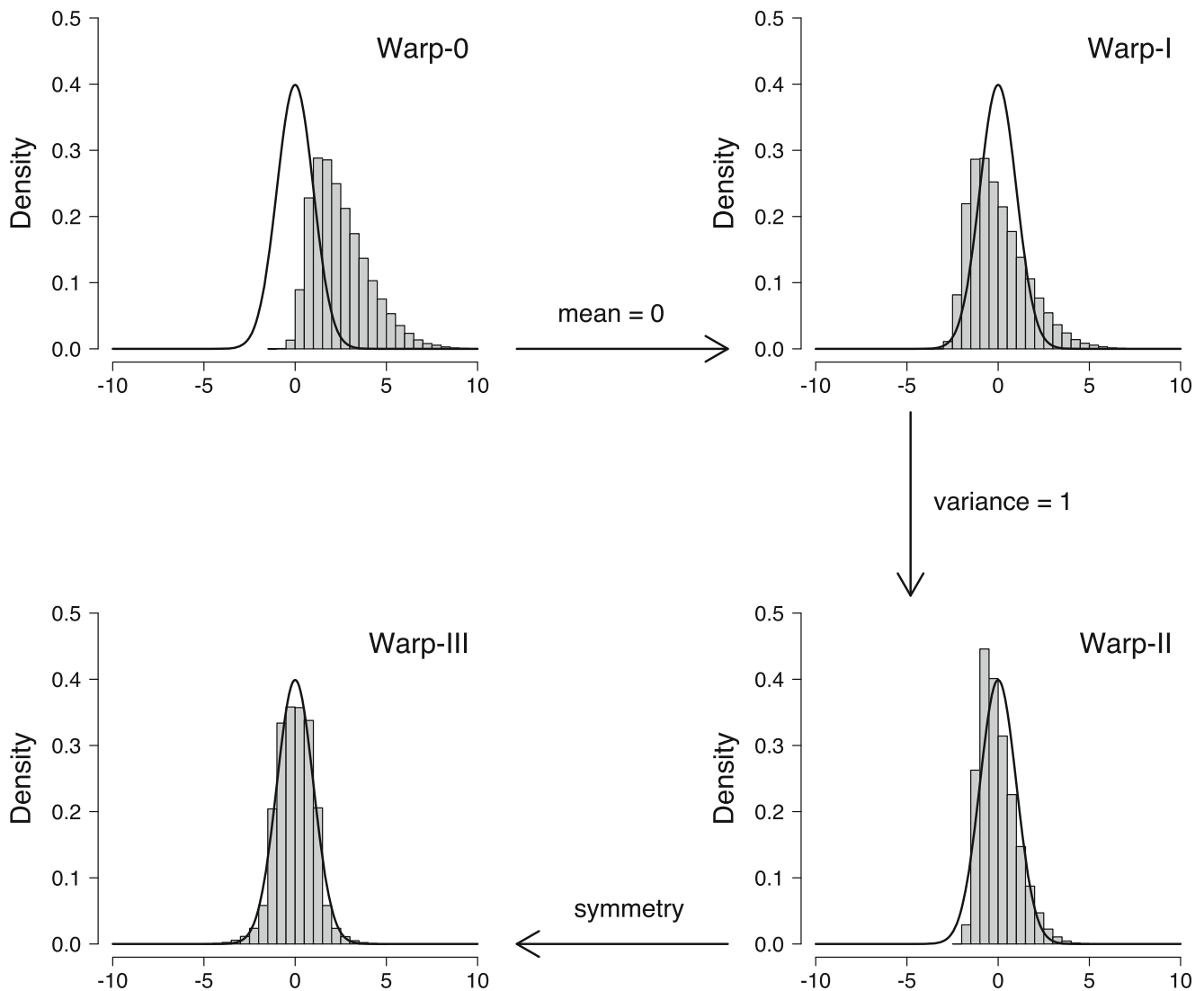
## Applications

In this section, we present applications of our method to five existing data sets. For each application, we describe the stimuli and the nature of the data, as well as make clear our expectations about the MDS representation that will be inferred. In particular, we state our expectations about both the dimensionality and metric structure of the representation whenever possible. The results we present are based on considering MDS models up to and beyond this expected dimensionality, so that the inference our method makes is clear. Where possible, we apply our method under the assumption that the metric space is both Euclidean ( $r = 2$ ) and non-Euclidean ( $r = 1.5$ ) so that an inference can also

<sup>4</sup>Note that other proposal distributions are conceivable. The only constraints are that the proposal has a zero mean vector, an identity covariance matrix, and exhibits no skewness.

<sup>5</sup>The function  $f$  needs to be one-to-one and its inverse  $f^{-1}$  needs to have a well-defined Jacobian.

<sup>6</sup>We use a function  $f$  that applies a log transformation to  $\sigma$  and (scaled) probit transformations to the non-zero elements of  $\mathbf{P}$ . The transformation for the ordered coordinates of the second stimulus for the non-Euclidean case is described in Appendix 2. Note that it is irrelevant whether the coordinates are ordered as decreasing, as shown in Fig. 6 for easier visualization, or increasing, as implemented in our code. The transformation described in the appendix assumes the latter. These transformations can be applied after having obtained posterior samples for  $\boldsymbol{\theta}$ . Furthermore, where necessary, the expressions are adjusted by the relevant Jacobian term  $|\det \mathcal{J}_{f^{-1}}(\boldsymbol{\zeta})|$ .



**Fig. 7** Illustration of the Warp-III procedure. The black solid line shows the standard Gaussian proposal distribution and the gray histogram shows synthetic posterior samples. Available at <https://tinyurl.com/y7owvsz3> under CC license <https://creativecommons.org/licenses/by/2.0/>

be made about the integrality or separability of the stimulus domain. For some applications, we were unable to generate samples with acceptable convergence for the  $r = 1.5$  metric. In those cases, we only report results assuming the  $r = 2$  metric.

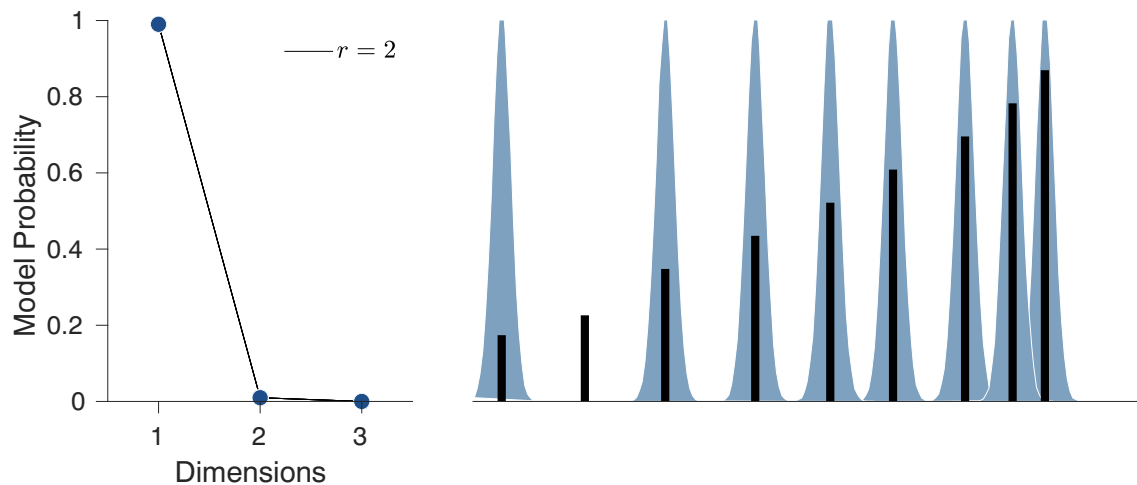
### Line Length

Our first application involves the similarity judgments between nine lines of equally increasing length provided by 27 participants, as reported in Cohen et al. (2001). We expect these stimuli to have a one-dimensional MDS representation, corresponding to line length. Because the Minkowski metrics are all equivalent in a one-dimensional space, we do not have any expectations about the metric structure. Thus, we applied our method to these data by

assuming a Euclidean metric.<sup>7</sup> As for all of our applications, we used 15 chains and 500 burn-in samples. During burn-in, the probability of a migration step was set to 0.05. After burn-in, migration was switched off, and the algorithm was run for 9000 iterations. We only retained every third sample so that we ended up with 3000 samples per chain for further use (i.e., a total of 45,000 samples collapsed across chains).

The left panel of Fig. 8 shows posterior model probabilities, assuming equal prior probabilities, for one-, two-, and three-dimensional MDS representations. To assess the stability of the posterior model probability estimates, we ran the Warp-III procedure five times based on new samples from the proposal distribution (we always used

<sup>7</sup>We note, however, for completeness that we had difficulty with convergence using the  $r = 1.5$  metric for these data.



**Fig. 8** Results for line-length similarity data from Cohen et al. (2001). The left panel shows the posterior model probabilities for one- through three-dimensional MDS representations. The right panel shows the

inferred one-dimensional representation with black lines showing the line stimuli at their inferred locations and blue histograms showing the marginal posterior distributions for these locations

the same set of posterior samples). These five repetitions are drawn as separate lines but, in this case, the results are so similar that they are visually indistinguishable. Because of the assumptions of equal prior probabilities, the ratio of any pair of posterior probabilities is naturally interpreted as a Bayes factor. The key result is that the expected one-dimensional representation is inferred, with a posterior probability near one.

The right panel of Fig. 8 shows the inferred one-dimensional MDS representation. The black lines show the stimuli in terms of their physical line lengths, located at the posterior mean of their location in the psychological space. The blue histograms show the marginal posterior distributions for each line stimulus. The MDS representation arranges the line stimuli in order of their length, but they are not evenly spaced, despite the lines increasing in constant physical increments. Instead, the psychological representation shows compression for the longer lines, consistent with basic psychophysics (Fechner 1966). This compression is large enough that the posterior distributions begin to overlap for the longest line stimuli.

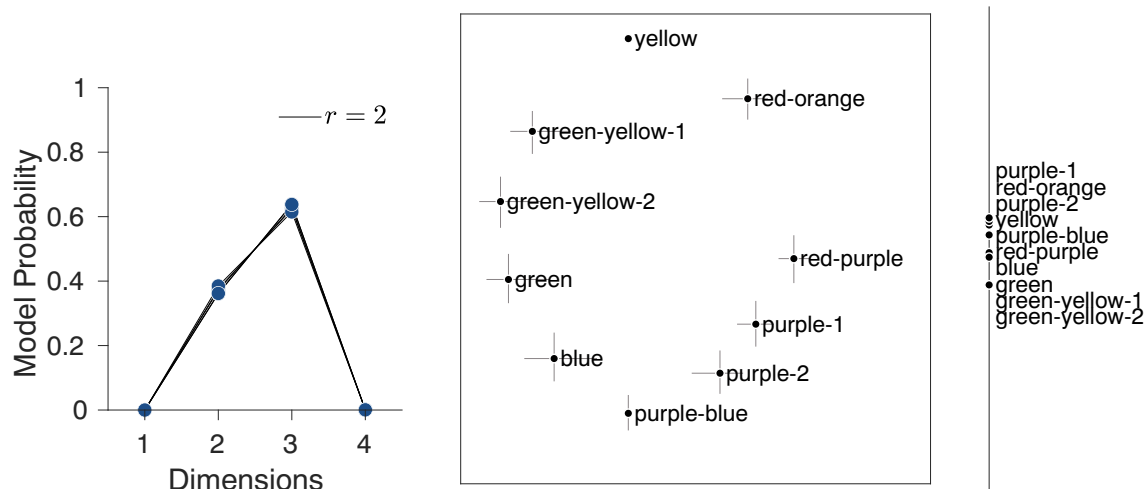
## Colors

Our second application considers classic data reported by Helm (1964), involving the similarities between ten colors. The experimental procedure involved trials in which participants were presented with physical tiles of three different colors, and moved one of the tiles to reflect their perceived overall similarity of the color of this tile to the colors of the other two tiles. Based on these responses, Helm (1964) calculated measures of pairwise similarities between the colors, and the resulting proximity data have previously

been considered in the MDS literature (e.g., Borg and Groenen 1997; Carroll and Wish 1974). We consider only the data from the ten participants with normal color vision.

We expect the MDS representation to use the Euclidean metric, consistent with the integral nature of the color stimulus domain. We also expect a two-dimensional representation, following the color circle found by previous MDS analyses of these and other color similarity data, such as the Shepard (1962) original MDS analysis of data reported by Ekman (1954).

Figure 9 shows the results of applying our method, assuming a Euclidean metric. This was a case in which we were unable to generate samples with acceptable convergence for the  $r = 1.5$  metric. For the Euclidean metric, there is uncertainty regarding the dimensionality, with a three-dimensional representation having probability a little over 0.6 and a two-dimensional representation having almost all of the remaining probability. The inferred three-dimensional representation is shown by pairing the first two dimensions as a two-dimensional plot in the center of Fig. 9, and showing the remaining third dimension separately to the right along an axis. Because of our ordering heuristic, the yellow and purple-blue stimuli were fixed at the origin and on the first axis. These assignments mean that the first two dimensions effectively represent the expected color circle that “bends” the visible physical spectrum from red to purple colors into a circle that reflects the psychological similarity between the end points. The third dimension, which we did not expect, could correspond to something like luminance, since low luminance purple-like colors are generally located at one end of the dimension and high luminance yellow-like colors are generally located at the other end.



**Fig. 9** Results for color similarity data from color-normal subjects reported by Helm (1964). The left panel shows the posterior probabilities for one- through four-dimensional MDS representations. The right panel shows the inferred three-dimensional representation, with two dimensions shown as a two-dimensional plot in the center, and the

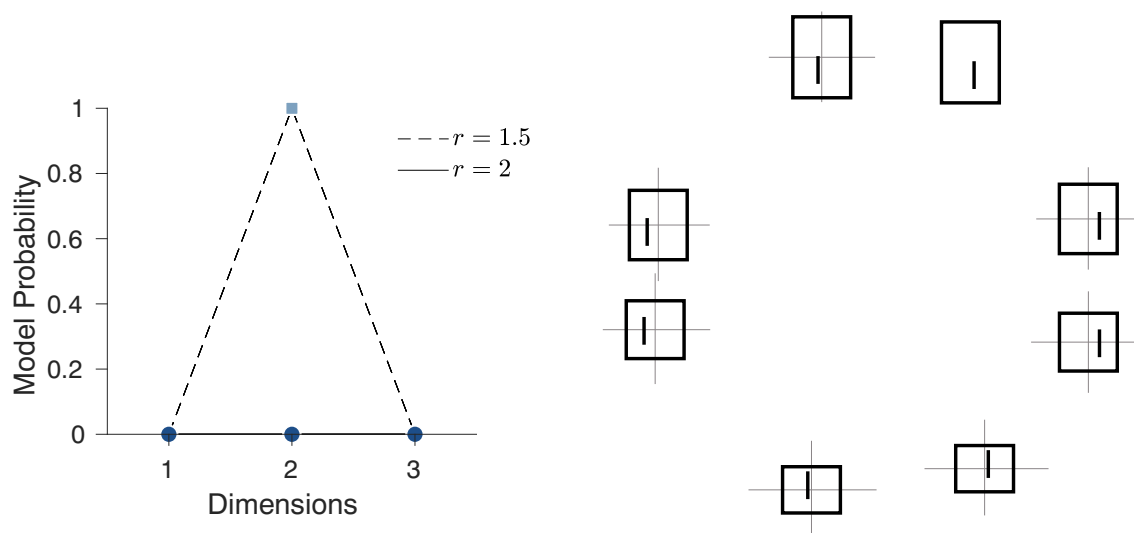
third dimension shown along an axis to the right. Circular markers and labels show the inferred locations of each stimulus and error bars show 95% credible intervals for the marginal posterior distribution for each dimension

### Rectangles with Line Segments

Our third application involves data reported by Kruschke (1993) involving the similarity between eight geometric stimuli. These stimuli consisted of rectangles with interior line segments, and varied in terms of the height of the rectangle and the horizontal location of the line segment. A total of 50 participants provided similarity ratings on a nine-point scale for all 28 stimulus pairs. Based on the original (Kruschke 1993) and subsequent (e.g., Lee 2001, 2008)

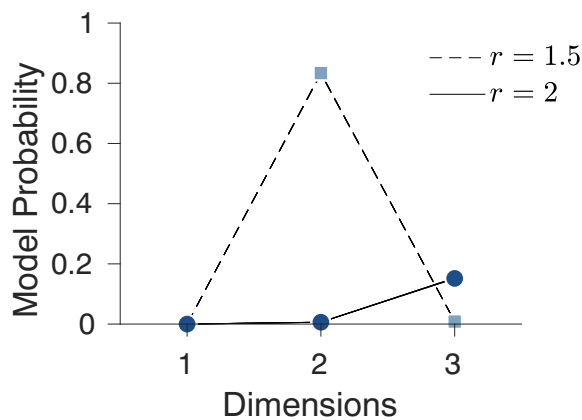
analyses of these data, we expect a two-dimensional MDS representation. We also expect the two stimulus dimensions to be psychologically separable.

Figure 10 shows the results of applying our method assuming both the  $r = 1.5$  and  $r = 2$  metrics. It is clear that a two-dimensional representation with the separable  $r = 1.5$  metric is inferred. It has essentially all of the posterior probability, with one- and three-dimensional  $r = 1.5$  representations, and all of the  $r = 2$  representations having essentially no posterior probability. The inferred

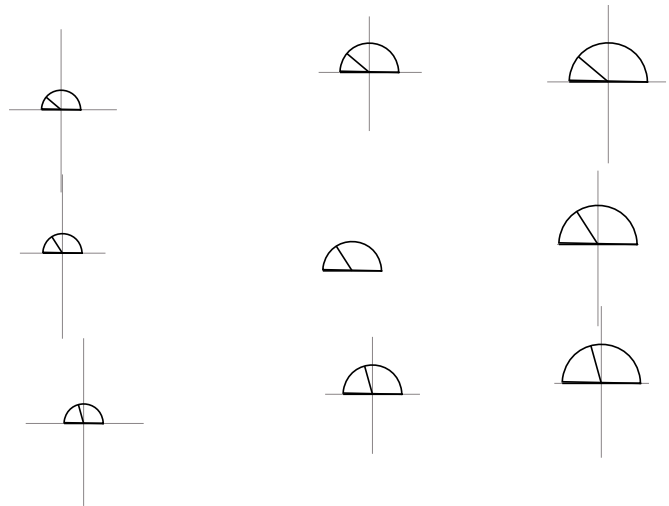


**Fig. 10** Results for rectangles with interior line segments data reported by Kruschke (1993). The left panel shows the posterior probabilities for one- through three-dimensional MDS representations, for both the Minkowski metrics with  $r = 1.5$  and  $r = 2$ . The right panel shows

the inferred two-dimensional representation. The stimuli are shown at their inferred locations and error bars show 95% credible intervals for the marginal posterior distribution for each dimension



**Fig. 11** Results for the Shepard circles data collected by Treat et al. (2001). The left panel shows the posterior probabilities for one- through three-dimensional MDS representations, for both the Minkowski metrics with  $r = 1.5$  and  $r = 2$ . The right panel shows



the inferred two-dimensional representation. The stimuli are shown at their inferred locations and error bars show 95% credible intervals for the marginal posterior distribution for each dimension

representation closely matches the ways in which the stimuli physically vary, with each psychological axis corresponding to an interpretable stimulus dimension. The horizontal axis corresponds to the position of the line segment and the vertical axis corresponds to the height of the rectangle.

### Shepard Circles

Our fourth application involves data collected by Treat et al. (2001), involving the similarity between nine geometric stimuli known as “Shepard circles.” These stimuli consist of a closed semi-circle with an interior ray from the center to the perimeter. The nine stimuli are constructed by exhaustively varying three different radius lengths and three different angles for the internal ray. As for the rectangles with line segments, we expect a separable two-dimensional MDS representation. For these stimuli, we expect the dimensions to correspond to the radius and angle dimensions.

Figure 11 shows the results of applying our method assuming both the  $r = 1.5$  and  $r = 2$  metrics.<sup>8</sup> It is clear, once again, that a two-dimensional representation with the separable  $r = 1.5$  metric is inferred. The inferred representation also again closely matches the ways in which the stimuli physically vary, with the horizontal axis corresponding to the radius of the semi-circle and the vertical axis corresponding to the angle of the ray.

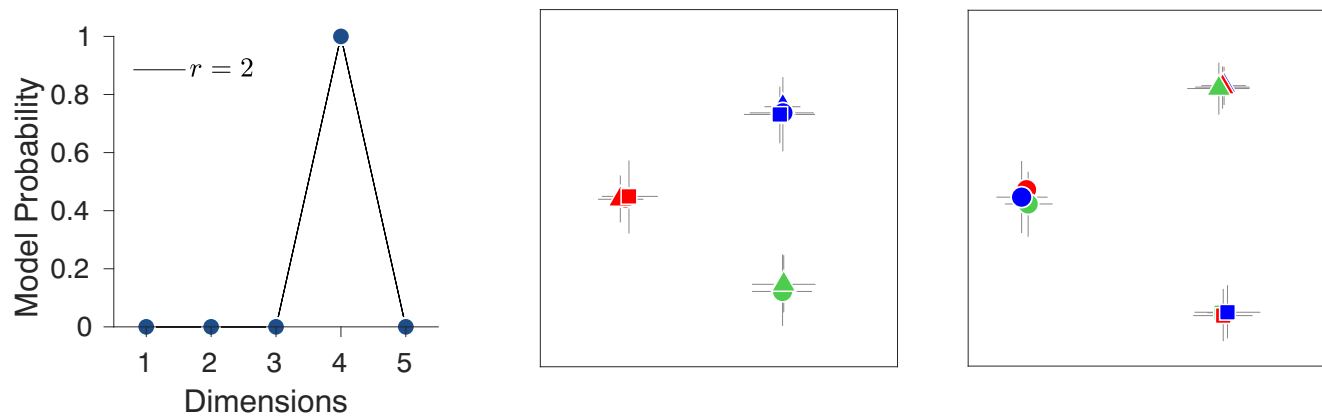
<sup>8</sup>For these stimuli, we did not have access to information about the precise physical values of the radius and angles, and so the depictions in Fig. 11 are approximate.

### Colored Shapes

Our final application considers similarity data for nine colored shape stimuli collected by Lee and Navarro (2002). The stimuli were circles, squares, and triangles that were colored red, green, and blue. The data were collected from 20 participants, each of whom rated the similarity of each pair of stimuli on a five-point scale.

Following the previous analysis in Lee and Navarro (2002), we expect a four-dimensional representation. This representation is best understood as being the product of a pair of two-dimensional representations, with one representing the similarities between the shapes, and the other representing the similarities between the colors. There are only three shapes and three colors, and neither set of three has a natural ordering. Instead, the circle, square, and triangle are all approximately equally different from one another, and the same is true of the red, green, and blue colors. These equal similarities are naturally represented by two-dimensional approximately equilateral triangles. The four-dimensional representation we expect is simply the independent combination of these two two-dimensional subspaces.

Our expectations for the metric structure of the MDS representations are less straightforward. Theoretically, the interaction between the shape and color dimensions is a classic example of a separable relationship. The metric structure within the color subspace, however, is theoretically integral, as for the previous application. Countering these theoretical expectations is the fact that there are only three values for the color and shape dimensions present in the



**Fig. 12** Results for colored shapes data reported by Lee and Navarro (2002). The left panel shows the posterior probabilities for one-through five-dimensional MDS representations for the Euclidean metric. The middle and right panels show the inferred four-dimensional

representation, with two dimensions shown in each panel. The colored shapes show the inferred locations of each stimulus and error bars show 95% credible intervals for the marginal posterior distribution for each dimension

stimulus set. The corresponding approximately equilateral triangles could be equally well accommodated by any of the Minkowski metrics we are considering. Thus, from a statistical perspective—without regard to the theory of separable and integral stimuli—we expect the simplest metric to be inferred. Since all metrics should be able to fit the data, the one with the smallest functional form complexity should be preferred.

We found that this was a third case in which we were unable to generate samples with acceptable convergence for the  $r = 1.5$  metric. Accordingly, Fig. 12 shows the results of applying our method assuming the Euclidean metric. A four-dimensional representation is clearly favored. This representation is shown in terms of two two-dimensional subspaces, and has the expected structure. The middle panel of Fig. 12 shows a subspace that captures the similarity relationships between the red, green, and blue colors. The right panel shows a subspace that captures the similarity relationships between the circle, square, and triangle shapes. These subspaces were found using an orthogonal Procrustes method (Borg and Groenen 1997, p. 162). In particular, we solved for the orthogonal transformation matrix that most closely mapped the inferred coordinate locations to the expected representational structure, defined as the product of two subspaces each with an equilateral triangle configuration.

## Discussion

Collectively, the five applications demonstrate that our method is able to make reasonable inferences about MDS representations. The inferred number of dimensions, and the inferred stimulus locations, generally matched theoretical

expectations, with the exception of the color application. In addition, where inferences about whether a Euclidean or non-Euclidean metric structure were made, they matched theoretical expectations. It is interesting to note that all of the applications for which non-Euclidean metrics made inference difficult involved stimulus domains for which the expectation was that the Euclidean metric was appropriate.

We also think that the five applications serve to demonstrate the usefulness of our approach to determining dimensionality and metric structure. Our approach is to treat these determinations as Bayesian model-selection problems and use Bayesian posterior probabilities to make inferences. Complete Bayes posterior probabilities have not been used in this way previously to determine either dimensionality or metric structure, and our introduction of the Warp-III method to solve the difficult computational approximation problems involved represents progress on these long-standing challenges in MDS modeling.

Despite this progress, we think the greatest contribution of the current work is to highlight fundamental challenges in MDS models of mental representation, and suggest new avenues for theoretical development. The challenges largely stem from our insistence on fully Bayesian inference, which has enormous advantages in terms of reaching complete, coherent, and principled conclusions, but also raises technical hurdles. The opportunities largely stem from our adoption of a generative modeling approach (Lee 2018). In particular, we think there are many remaining possibilities relating to the use of different metrics in MDS representations, and that there is an opportunity to extend the generative approach to develop more complete cognitive process models for inferring MDS representations. We conclude by discussing some of these challenges and opportunities.



## Technical Challenges

Developing a generative MDS model in a Bayesian setting required the key issue of identifiability and invariance to be solved in terms of prior information, rather than more heuristically through post-processing. We used an existing solution to this challenge for the Euclidean metric, and proposed a solution for psychologically interpretable non-Euclidean metrics with  $1 < r < 2$ . We also highlighted, however, the fundamental intractability of MDS representations using the city-block metric. This intractability has been documented before (Bortz 1974; Frank 2006, Figure 5.4; Shepard 1974, Figure 11), but has not prevented the use of MDS representations inferred based on the city-block metric in the cognitive modeling literature (e.g., Kruschke 1993; Lee and Wetzels 2010).

Our current approach to determining the appropriate metric treats this inference as a model-selection problem, and only considers the possibilities  $r = 1.5$  and  $r = 2$ . Allowing for other metrics is theoretically interesting, but computationally difficult. One obvious cost is the need to generate posterior probabilities across a larger set of candidate models. But it also seems likely that some models will be difficult to make inferences about. We tried our DE-MCMC approach for  $r = 1.1$  on a number of data sets, and were not able to achieve satisfactory convergence. Furthermore, as explained above, for a few of the applications, we were also not able to achieve satisfactory convergence for  $r = 1.5$ . These challenging cases involved stimulus domains for which the expectation was that the Euclidean metric was appropriate, which leads to a speculative suggestion that failure is related to model mis-specification. This is a potential example of a general aspect of Bayesian model comparison that can be computationally challenging: in order to rule out models that are likely mis-specified, one needs to be able to infer them well enough that they can be part of the model comparison. Although we believe that DE-MCMC is a powerful sampling algorithm which substantially helps alleviate the issue of non-converging chains, future research should explore different sampling algorithms that may perform better, particularly for non-Euclidean metrics.

Collectively, these technical challenges mean that our approach cannot currently be applied to large naturalistic stimulus domains. For example, Nosofsky et al. (2018) consider MDS representations based on sparse matrices of pairwise similarity judgments for a set of 360 images of rocks, and Hebart et al. (2020) report extensive crowd-sourced triadic comparison similarity data for 1854 images of real-world objects. Being able to determine the dimensionality, metric structure, and psychological representations of MDS representations of these domains using the Bayesian framework would potentially offer deep

insight into how people represent the real-world stimuli. The successful applications we presented—in which there were clear expectations about dimensionality, metric, and representational structure—provide a basis for believing the Bayesian framework can provide this insight to situations where, because there are no clear theoretical expectations, answers must be inferred from data, if and when the computational technical hurdles are overcome.

## Other Representations

We did not consider Minkowski metrics with  $r < 1$ . This possibility has been proposed as a way of representing stimulus domains in which the component dimensions compete for attention (Shepard 1987, 1991; Tversky and Gati 1982). The identifiability constraints for this metric present an open research challenge, and it is not clear how well DE-MCMC sampling methods will perform in inferring representations.

There is also the possibility of moving beyond the Minkowski family of metrics. In his presidential address, Shepard (1974, Figure 11) presented a taxonomy of metric spaces, each of which makes different fundamental representational assumptions that could be appropriate for at least some stimulus domains. There has been relatively little work in exploring these possibilities. Lindman and Caelli (1978) investigated MDS representations using Riemannian spaces with constant curvature, and Cox and Cox (1991) presented compelling applications for a special case of this approach involving MDS representations on a sphere.

A new idea raised by our application to the colored shape stimuli involves the possibility of different metric structures within the same representation. These stimuli involved two sorts of stimulus dimensions: those representing color, which are usually considered to be integral, and those representing qualitatively different shapes, which seems more separable. Certainly the interaction between the color dimensions and the shape dimensions would be expected to be separable, since it seems likely people can selectively attend to either the color or the shape of a stimulus, depending upon the cognitive context. This suggests a generalization of the MDS models in which each pair of dimensions is associated with a metric.

Finally, there are alternative representational models, which do not assume stimuli are represented by values on dimensions, that can compete with or complement MDS models. These alternatives include feature-based representations (Tversky 1977), such as those found by additive clustering and related methods (Shepard and Arabie 1979) and special cases such as tree-based models (Cortier 1996; Shepard 1980). One attraction of the Warp-III approach we used is that it could estimate Bayes factors between fundamentally different sorts of representations—such as

comparing dimensional and featural representations—since it operates directly on posterior samples for each model applied independently to the data. Even further, Navarro and Lee (2003) proposed a hybrid model of stimulus representation that combined both dimensions and features, and it would be conceptually elegant to choose between all of the candidate models, with various combinations of dimensions and features, using our methods. Navarro and Lee (2003) used an approximate analytic approach for this purpose, which would be significantly improved by an approach based on Bayes factors.

### MDS Cognitive Process Models

Our modeling approach is generative, but is based on an extremely simple cognitive model. In essence, we assume that all participants have the same MDS representation, and produce dissimilarity judgments for pairs of stimuli that directly reflect the distances between those stimuli in the representation. It is likely that much better generative models can be developed by considering more realistic processing assumptions, and especially by including individual differences.

One example, involving the line-length application, was presented in a preliminary form by Lee (2014). A simple plot of the raw behavioral data suggests that one of the 27 participants appears to have reversed the scale that was used to judge similarity. This means that their judgments contaminate the inference of the MDS representation. Lee (2014) used a simple latent-mixture model extension of the basic MDS generative model, in which either the scale was used correctly or reversed. One participant was inferred to have reversed the scale, as expected. Perhaps more importantly, however, the resulting inference about the one-dimensional MDS representation was shown to have less uncertainty than the one shown in Fig. 8. In this way, the introduction of individual differences in the cognitive process of similarity judgment helped decontaminate the inference about the representation of stimuli.

The same basic generative approach could support much more general cognitive process modeling using MDS representations. The hierarchical, latent mixture, and common cause model structures advocated by Lee (2018) could allow for rich accounts of individual differences in judgment processes or stimulus representations, and allow for models that extend beyond the judgment of similarity to other cognitive capabilities like categorization and inference. As one example, Ennis (1992) considers extended assumptions about MDS representations that allow for the noisy representation of perceptual stimuli, which could be incorporated by adding hierarchical structure to the coordinate locations. As another example, there are extensions of the basic MDS model we considered that allow for structured

individual differences, such as INDSCAL (Carroll and Chang 1970; Carroll 1972). These would be easy to implement within our generative modeling framework. A model like INDSCAL, which assumes individuals weight the latent stimulus dimensions differently, relies on the appropriate number of dimensions being inferred, and evidence that the stimulus domain is separable. In this way, the potential of our method to make these inferences is especially important. As a final example, the rectangle and line segment stimuli are used by Kruschke (1993) to study category learning, but the similarity data and category learning data are analyzed independently. In effect, the similarity data are used to generate the MDS representation, and that representation is then assumed to provide the fixed basis for category learning. An alternative approach would be to infer the MDS representation jointly from both the similarity judgments and the category learning choices. This sort of flexibility raises the possibility of tackling more complicated cognitive phenomena, such as the ability to adapt representations in response to changes in the external environment, or the current context or goals.

### Conclusion

We adopted a Bayesian model-selection approach to the problem of determining the dimensionality and metric structure of MDS representations, while considering psychologically interpretable Euclidean and non-Euclidean metrics. Our methods for inferring the representations and choosing their dimensionality and metric structure show the promise of the approach, but computational challenges remain a barrier in terms of an easy-to-use general capability. Our methods and applications also show the promise of placing MDS representations in a generative cognitive modeling framework, offering the possibility of new models of how people represent stimuli, and how those representations help guide behavior.

**Acknowledgments** All code is available on the Open Science Framework: <https://osf.io/82g3r/>. We thank Rob Nosofsky and Mike D’Zmura for helpful discussions. QFG acknowledges the support by a Netherlands Organization for Scientific Research (NWO) grant (406.16.528). Correspondence should be sent to Quentin F. Gronau, University of Amsterdam, Nieuwe Achtergracht 129 B, 1018 WT Amsterdam, The Netherlands. E-mail may be sent to [Quentin.F.Gronau@gmail.com](mailto:Quentin.F.Gronau@gmail.com).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted

use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix 1. The ordering heuristic

Figure 13 provides a concrete example to motivate the need for the ordering heuristic. It is clear this is an inferior representation to the one presented in Fig. 8. In Fig. 13, the first- and second-line stimuli, which are the two shortest, are located at almost the same point, rather than being appropriately spaced to reflect their psychological dissimilarity. Consistent with this intuition, the posterior density is worse for the representation in Fig. 13 than the representation in Fig. 8.

This suboptimality is caused by the naive application of the constraints identified in Fig. 2 for a one-dimensional representation. The first stimulus is fixed at the origin, and the second stimulus is constrained to be positive. It is clear from Fig. 13 that the second stimulus is indeed inferred to be positive, but is extremely close to zero, with the remaining longer line stimuli “flipping” to negative values in the MDS space. This configuration still satisfies the proximity data reasonably well, because the required distance between the first two stimuli is small, and the distances from the first and second stimuli to all of the others is approximately conserved. Thus, it is the choice of the two similar stimuli as those that are constrained that leads to this potential for a local maximum and suboptimal representation.

Accordingly, we developed an ordering heuristic to try and assign the constraints for the various dimensionalities and metrics to stimuli that are sufficiently dissimilar. Because higher dimensionalities place constraints on more than two stimuli, the general approach is to order all of the

stimuli. Our heuristic for doing this is based on the across participants averaged pairwise dissimilarity ratings. The first two stimuli are chosen to be the ones with the largest averaged pairwise dissimilarity. The remaining stimuli are chosen, one at a time, by considering the minimum averaged pairwise dissimilarity to the already selected stimuli. Specifically, the next stimulus is always chosen to be the one with the maximum value for the minimum averaged pairwise dissimilarity to the already selected stimuli.

We used this ordering heuristic for the colors and colored shapes applications. For the line-length application, we used the heuristic as described but then, in an additional step, switched the first stimulus with the second stimulus. This switch helped prevent the posterior for the ninth stimulus, corresponding to the longest line, push against the upper bound of 1. For the rectangles with interior line segments and Shepard circles applications, we used the heuristic as a starting point, but we then reordered some of the stimuli manually since it seemed to help with convergence.

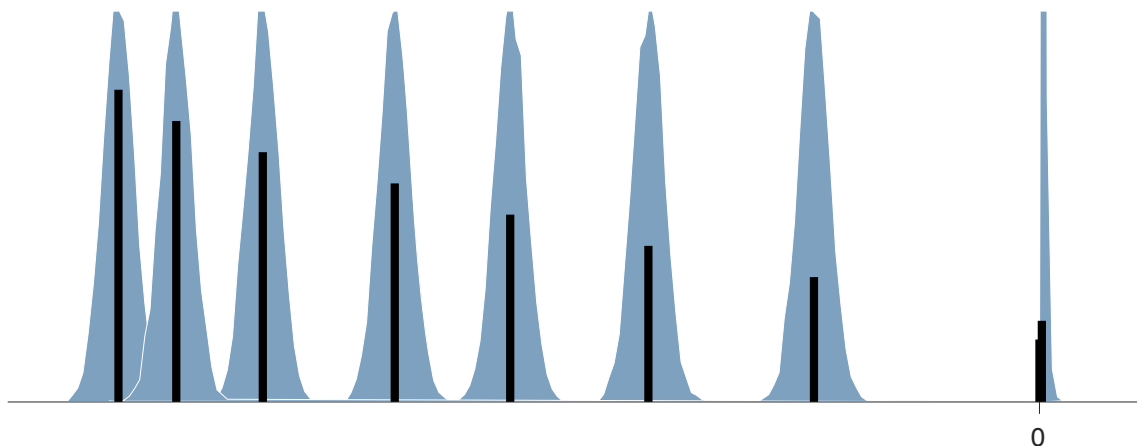
## Appendix 2. Transformation ordered vector (0–1 bounded)

The constrained vector  $\mathbf{x}$ ,  $0 \leq x_1 \leq x_2 \leq \dots \leq x_K \leq 1$ , can be transformed to an unconstrained vector  $\mathbf{y} \in \mathbb{R}^K$  as follows:

$$y_k = \begin{cases} \Phi^{-1}(x_k) & \text{if } k = 1, \\ \Phi^{-1}\left(\frac{x_k - x_{k-1}}{1 - x_{k-1}}\right) & \text{if } 1 < k \leq K, \end{cases}$$

where  $\Phi^{-1}(\cdot)$  denotes the inverse of the normal CDF. The inverse transformation is given by:

$$x_k = \begin{cases} \Phi(y_k) & \text{if } k = 1, \\ x_{k-1} + (1 - x_{k-1}) \Phi(y_k) & \text{if } 1 < k \leq K, \end{cases}$$



**Fig. 13** A suboptimal one-dimensional representation of the line-length similarity data from Cohen et al. (2001), motivating the need for the ordering heuristic. The black lines show the stimuli

at their inferred locations in the representation, and the blue histograms show the marginal posterior distributions for these locations

where  $\Phi(\cdot)$  denotes the normal CDF. Note that  $x_k$  is a function of  $y_1, y_2, \dots, y_k$  (the dependence on  $y_1, y_2, \dots, y_{k-1}$  is “hidden” in  $x_{k-1}$ ). Crucially,  $x_k$  does not depend on  $y_{k+1}, y_{k+2}, \dots, y_K$ . Consequently, the Jacobian matrix  $\mathcal{J}$  of the transformation is lower triangular so that its determinant  $|\mathcal{J}|$  is obtained by multiplying its diagonal entries. The diagonal entries are given by:

$$\mathcal{J}_{k,k} = \begin{cases} \phi(y_k) & \text{if } k = 1, \\ (1 - x_{k-1}) \phi(y_k) & \text{if } 1 < k \leq K, \end{cases}$$

where  $\phi(\cdot)$  denotes the normal PDF. Hence, the determinant of the Jacobian matrix is given by:

$$|\mathcal{J}| = \phi(y_1) \prod_{k=2}^K [(1 - x_{k-1}) \phi(y_k)].$$

## References

- Arnold, J.B. (1971). A multidimensional scaling study of semantic distance. *Journal of Experimental Psychology*, 90, 349.
- Attneave, F. (1950). Dimensions of similarity. *The American Journal of Psychology*, 63, 516–556.
- Bakker, R., & Poole, K.T. (2013). Bayesian metric multidimensional scaling. *Political Analysis*, 21, 125–140.
- Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling: theory and applications*. New York: Springer-Verlag.
- Bortz, J. (1974). Kritische Bemerkungen über den Einsatz nicht-euklidischer Metriken im Rahmen der multidimensionalen Skalierung. *Archiv für Psychologie*, 126, 196–212.
- Carroll, J.D. (1972). Individual differences and multidimensional scaling. In Shepard, R.N., Romney, A.K., Nerlove, S. (Eds.) *Multidimensional scaling: theory and application in the behavioral sciences* (pp. 105–155). New York: Seminar Press.
- Carroll, J.D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of ‘Eckart–Young’ decomposition. *Psychometrika*, 35, 283–319.
- Carroll, J.D., & Wish, M. (1974). Multidimensional perceptual models and measurement methods. *Handbook of Perception*, 2, 391–447.
- Cohen, A.L., Nosofsky, R.M., Zaki, S.R. (2001). Category variability, exemplar similarity, and perceptual classification. *Memory & Cognition*, 29, 1165–1175.
- Cortier, J.E. (1996). *Tree models of similarity and association*. Thousand Oaks, CA Sage.
- Cox, T.F., & Cox, M.A.A. (1991). Multidimensional scaling on a sphere. *Communications in Statistics: Theory and Methods*, 20(9), 2943–2953.
- Cox, T.F., & Cox, M.A.A. (1994). *Multidimensional scaling*. London: Chapman and Hall.
- Ekman, G. (1954). Dimensions of color vision. *The Journal of Psychology*, 38, 467–474.
- Ennis, D.M. (1992). Modeling similarity and identification when there are momentary fluctuations in psychological magnitudes. In Ashby, F.G. (Ed.) *Multidimensional models of perception and cognition* (pp. 279–298). Hillsdale: Erlbaum.
- Estes, W.K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, 53(2), 134–140.
- Fechner, G.T. (1966 [1860]). (H. E. Adler, Trans.) In Howes, D.H., & Boring, E.G. (Eds.) *Elements of psychophysics [Elemente der Psychophysik]* (Vol. 1). New York: Holt, Rinehart and Winston.
- Frank, L.E. (2006). Feature network models for proximity data: statistical inference, model selection, network representations and links with related models. Leiden University <http://hdl.handle.net/1887/4560>.
- Gamerman, D., & Lopes, H.F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Boca Raton, FL: Chapman & Hall/CRC.
- Garner, W.R. (1974). *The processing of information and structure*. Potomac, MD Erlbaum.
- Groenen, P.J., Heiser, W.J., Meulman, J.J. (1998). City-block scaling: smoothing strategies for avoiding local minima. In *Classification, data analysis, and data highways* (pp. 46–53). Springer.
- Gronau, Q.F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Steingrover, H. (2017). A tutorial on bridge sampling. *Journal of Mathematical Psychology*, 81, 80–97.
- Gronau, Q.F., Singmann, H., Wagenmakers, E.J. (in press). bridge-sampling: an R package for estimating normalizing constants. arXiv:1710.08162.
- Gronau, Q.F., Wagenmakers, E.J., Heck, D.W., Matzke, D. (2019). A simple method for comparing complex models: Bayesian model comparison for hierarchical multinomial processing tree models using Warp-III bridge sampling. *Psychometrika*, 84, 261–284.
- Heathcote, A., Lin, Y.S., Reynolds, A., Strickland, L., Gretton, M., Matzke, D. (in press). Dynamic models of choice.
- Hebart, M., Zheng, C.Y., Pereira, F., Baker, C. (2020). Revealing the multidimensional mental representations of natural objects underlying human similarity judgments. <https://doi.org/10.31234/osf.io/7wrgb>.
- Helm, C.E. (1964). Multidimensional ratio scaling analysis of perceived color relations. *Journal of the Optical Society of America*, 54, 256–262.
- Hubert, L., Arabie, P., Hesson-McInnis, M. (1992). Multidimensional scaling in the city-block metric: a combinatorial approach. *Journal of Classification*, 9, 211–236.
- Jäkel, F., Schölkopf, B., Wichmann, F.A. (2008). Similarity, kernels, and the triangle inequality. *Journal of Mathematical Psychology*, 52, 297–303.
- Kass, R.E., & Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795.
- Kruschke, J.K. (1993). Human category learning: implications for backpropagation models. *Connection Science*, 5, 3–36.
- Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1–27.
- Lee, M.D. (2001). Determining the dimensionality of multidimensional scaling representations for cognitive modeling. *Journal of Mathematical Psychology*, 45(1), 149–166.
- Lee, M.D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, 15(1), 1–15.
- Lee, M.D. (2014). Applications of Bayesian graphical modeling to psychophysics. In *Paper presented at 30th Annual Meeting of the International Society of Psychophysics*. <http://fechnerday.com/fd2014/pdfs/FD14.Proceedings.book.pdf>.
- Lee, M.D. (2018). Bayesian methods in cognitive modeling. In Wixted, J., & Wagenmakers, E.-J. (Eds.) *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience. Volume 5: Methodology* (Fourth ed., pp. 37–84). Wiley.
- Lee, M.D., & Navarro, D.J. (2002). Extending the ALCOVE model of category learning to featural stimulus domains. *Psychonomic Bulletin & Review*, 9, 43–58.
- Lee, M.D., & Pope, K.J. (2003). Avoiding the dangers of averaging across subjects when using multidimensional scaling. *Journal of Mathematical Psychology*, 47, 32–46.
- Lee, M.D., & Vanpaemel, W. (2018). Determining informative priors for cognitive models. *Psychonomic Bulletin & Review*, 25, 114–127.



- Lee, M.D., & Wetzels, R. (2010). Individual differences in attention during category learning. In Catrambone, R., & Ohlsson, S. (Eds.) *Proceedings of the 32nd annual conference of the cognitive science society* (pp. 387–392). Austin: Cognitive Science Society.
- Lindman, H., & Caelli, T. (1978). Constant curvature Riemannian scaling. *Journal of Mathematical Psychology*, 17, 89–109.
- Mair, P., & Leeuw, J.D. (2014). Unidimensional scaling. *Statistics Reference Online* 1–3.
- Meng, X.L., & Schilling, S. (2002). Warp bridge sampling. *Journal of Computational and Graphical Statistics*, 11, 552–586.
- Meng, X.L., & Wong, W.H. (1996). Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Statistica Sinica*, 6, 831–860.
- Navarro, D.J., & Lee, M.D. (2003). Combining dimensions and features in similarity-based representations. In Becker, S., Thrun, S., Obermayer, K. (Eds.) *Advances in neural information processing systems 15* (pp. 59–66). Cambridge: MIT Press.
- Nosofsky, R.M. (1985). Overall similarity and the identification of separable-dimension stimuli: a choice model analysis. *Perception & Psychophysics*, 38, 415–432.
- Nosofsky, R.M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, 43, 25–53.
- Nosofsky, R.M., Sanders, C.A., Meagher, B.J., Douglas, B.J. (2018). Toward the development of a feature-space representation for a complex natural category domain. *Behavior Research Methods*, 50, 530–556.
- Oh, M.S. (2012). A simple and efficient Bayesian procedure for selecting dimensionality in multidimensional scaling. *Journal of Multivariate Analysis*, 107, 200–209.
- Oh, M.S., & Raftery, A.E. (2001). Bayesian multidimensional scaling and choice of dimension. *Journal of the American Statistical Association*, 96, 1031–1044.
- Okada, K. (2012). A Bayesian approach to asymmetric multidimensional scaling. *Behaviormetrika*, 39, 49–62.
- Okada, K., & Mayekawa, S.i. (2018). Post-processing of Markov chain Monte Carlo output in Bayesian latent variable models with application to multidimensional scaling. *Computational Statistics*, 33, 1457–1473.
- Okada, K., & Shigemasa, K. (2010). Bayesian multidimensional scaling for the estimation of a Minkowski exponent. *Behavior Research Methods*, 42, 899–905.
- Pitt, M.A., Kim, W., Navarro, D.J., Myung, J.I. (2006). Global model analysis by parameter space partitioning. *Psychological Review*, 113, 57–83.
- Schiffman, S.S., Reynolds, M.L., Young, F.W. (1981). New York, NY Academic Press.
- Shepard, R.N. (1957). Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. *Psychometrika*, 22, 325–345.
- Shepard, R.N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27, 125–140.
- Shepard, R.N. (1974). Representation of structure in similarity data: problems and prospects. *Psychometrika*, 39, 373–422.
- Shepard, R.N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, 214, 390–398.
- Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237, 1317–1323.
- Shepard, R.N. (1991). Integrality versus separability of stimulus dimensions: From an early convergence of evidence to a proposed theoretical basis. In Pomerantz, J.R., & Lockhead, G.L. (Eds.) *The Perception of Structure: Essays in Honor of Wendell R. Garner* (pp. 53–71). Washington: American Psychological Association.
- Shepard, R.N., & Arabie, P. (1979). Additive clustering representations of similarities as combinations of discrete overlapping properties. *Psychological Review*, 86, 87–123.
- Steyvers, M. (2006). Multidimensional scaling. In Nadel, L. (Ed.) *Encyclopedia of Cognitive Science Wiley Online Library*.
- Treat, T.A., McFall, R.M., Viken, R.J., Kruschke, J.K. (2001). Using cognitive science methods to assess the role of social information processing in sexually coercive behavior. *Psychological Assessment*, 13, 549.
- Turner, B.M., Sederberg, P.B., Brown, S.D., Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods*, 18, 368–384.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.
- Tversky, A., & Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological Review*, 89, 123–154.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.