

Hypertrix: An indicatrix for high-dimensional visualizations

Shivam Raval*
Harvard University

Fernanda Viégas†
Harvard University
Google Research

Martin Wattenberg‡
Harvard University
Google Research

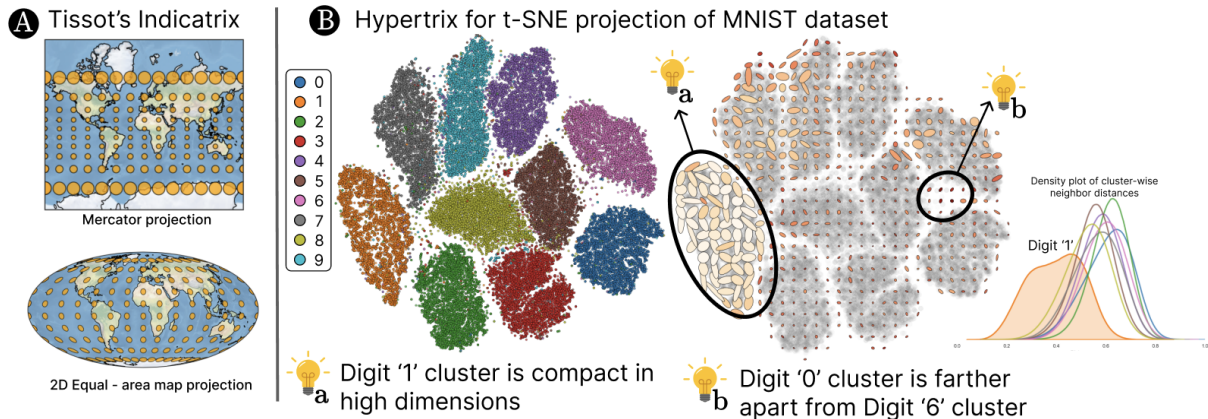


Figure 1: (A) Tissot’s indicatrix for two classic map projections, providing an intuitive view of cartographic distortion. (the images are produced using cartopy [14] in Python) (B) The “hypertrix” is a high-dimensional extension of Tissot’s Indicatrix. At left is a standard t-SNE projections of the MNIST digits dataset. At right is a hypertrix overlay. The hypertrix makes it clear that the cluster for the digit 1 is (non-conformally) expanded compared to other digits. There is also significant information compression at cluster boundaries.

ABSTRACT

Visualizing high dimensional data is challenging, since any dimensionality reduction technique will distort distances. A classic method in cartography—Tissot’s Indicatrix, specific to sphere-to-plane maps—visualizes distortion using ellipses. Inspired by this idea, we describe the *hypertrix*: a method for representing distortions that occur when data is projected from arbitrarily high dimensions onto a 2D plane. We demonstrate our technique through synthetic and real-world datasets, and describe how this indicatrix can guide interpretations of nonlinear dimensionality reduction.

Index Terms: Dimensionality Reduction, High-dimensional data—Distortion—Text Visualization, Clustering—

1 INTRODUCTION

Nonlinear dimensionality reduction techniques, such as t-SNE and UMAP, are notoriously difficult to interpret. Visualizing high dimensional data necessarily introduces a variety of distortions [26]. This leads to a natural question: is there a way to explicitly identify distortions, to help viewers read and understand these visualizations?

Of course, this is not a new question, and several techniques have been proposed by the visualization community [4, 9, 22] But the question predates computer science: for centuries, cartographers have known that map projections distort spherical distances on a globe. In 1859, French mathematician Nicolas Tissot invented a technique to make these distortions explicit [10]. The idea, today known as Tissot’s indicatrix, visualizes the local distortion with small ellipses overlaid on the map. Each ellipse represents the

derivative of the mapping—roughly speaking, how a standard-sized circle would be transformed by the projection at that point.

Inspired by this approach, we propose a high-dimensional analog of the indicatrix, which we call a **hypertrix**. A hypertrix is a colored ellipse that indicates distortion in projections of high-dimensional data. The distortion is estimated from a grid of locally linear transformations of the high-dimensional data that maps to the pre-computed low-dimensional projection. The learned transformation matrix is used to quantify the magnitude and direction of distortions.

As in Tissot’s indicatrix, we draw ellipses that visualize the image of small spheres under this transformation, overlaid on the projection. This provides a view not just of the magnitude of expansion or contraction, but whether it is uniform in different directions. Because this 2D glyph cannot capture the full high-dimensional transformation, we also color the ellipses to show overall distortion. The Hypertrix plots intuitively highlight regions of high distortion, offering insights into underlying data distribution and complementing existing quantitative evaluation metrics.

We demonstrate the hypertrix technique by applying it to projections of synthetic and real-world examples. Applied to some of the misleading examples from [26], the hypertrix makes underlying distortions obvious. In real-world data sets, the hypertrix illuminates some subtle and non-obvious distortions, both within and between clusters. (See Figure 1 for one example.) Overall, the hypertrix presents a visual depiction of distortion that complements the previously proposed metrics, such as stress and trustworthiness, in qualitatively evaluating dimensionality reduction methods.

2 RELATED WORK

Tissot first analyzed the distortion in distances in cartographic projections [17, 24] and proposed a graphical representation of a field of ellipses on a map that describes the distortion. The idea is straightforward: visualize local distortion by showing the image of a standard-sized circle under the map projection. In effect, this is a visualization of the derivative of the map from sphere to plane. The power of the indicatrix stems from the fact that it visualizes multiple types of

*e-mail: sraval@g.harvard.edu

†e-mail: fernanda@g.harvard.edu

‡e-mail: wattenberg@g.harvard.edu

distortion in an intuitive fashion, clarifying directions of expansion and contraction independently.

Nonlinear dimensional projections bring new challenges, since they involve mappings that can be far more irregular and complicated than those seen in world maps. Indeed, a great deal of work has gone into measuring different types of distortion and uncertainty in these visualizations. Several scalar metrics have been proposed to assess the quality of visual projections [1, 12, 21]. Global measures of trustworthiness and preservation of local neighborhoods [7, 25] are informative while comparing projections computed using different methods. Visualizing quantities like uncertainty [3] or the distribution of data with plots can be informative, allowing viewers to quickly gain qualitative insights. Studies using these metrics [13, 19] concluded that Isomap and Locally Linear Embeddings (LLE) are better at extracting manifolds while t-SNE is better at preservation of original neighborhoods.

Closest to our work here are techniques that display these metrics by overlaying distortion data on a visualization. For example [2] proposes overlaying the projections with colored Voronoi cells that encode the degree of local stretching and compression, or [20], which visualizes precision measures directly on the visualizations themselves. ProxiViz and CheckViz [8, 11] integrate distortion metrics into interactive projections to make viewers aware of such artifacts while analyzing data. However, these visuals do not simultaneously show both the direction and magnitude of distortions, data that is important to understand the relative sizes of the clusters and visual artifacts in the projections. DynamicViz [23] proposes dynamic visualizations that give a sense of projection stability; however, it requires multiple embeddings.

3 HYPERTRIX FORMULATION

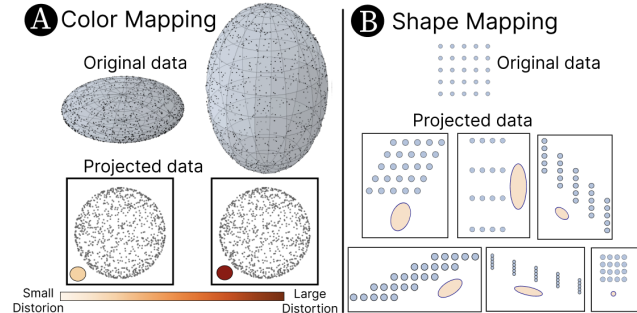


Figure 2: The Hypertrix reflects the local distortion due to scaling, rotation, and loss of information about distances, visualized as a visual glyph of an ellipse. (A) The color encodes the distortion in distances and (B) the shape encodes the stretching or compression and rotation relative to the original data.

Existing methods for showing high-dimensional distortions are confined to displaying just one distortion metric at a time, and provide no indication of systematic directionality related to compression or expansion. We would like to fill this gap, but there are obstacles to generalizing Tissot’s Indicatrix to higher dimensions.

To begin with, cartographic map projections are typically differentiable almost everywhere, so taking derivatives provides a natural way to define local distortion. By contrast, projection methods such as t-SNE or UMAP do not specify a mapping on the entire space—so taking derivatives is not an option. As a result, we need to use an approximation method, defining a local linear projection based on a sample of points in an appropriate neighborhood.

Using this local linear approximation, we can draw an elliptical glyph that represents the image of a small, standard-sized sphere. However, unlike the case of Tissot’s Indicatrix, this ellipse does not

capture all the information a reader might want. In particular, because this is a mapping from high-dimensional space, many different linear mappings might give rise to the same ellipse. Some of these may involve more distortion than others. To address this ambiguity, we encode a scalar metric of the overall distortion using color. This is similar to the technique used in quantifying distortion in cartographic distances [15] and is qualitatively similar to the projection-precision score proposed by [20]. Figure 2 is a visual view of these ideas.

3.1 General method

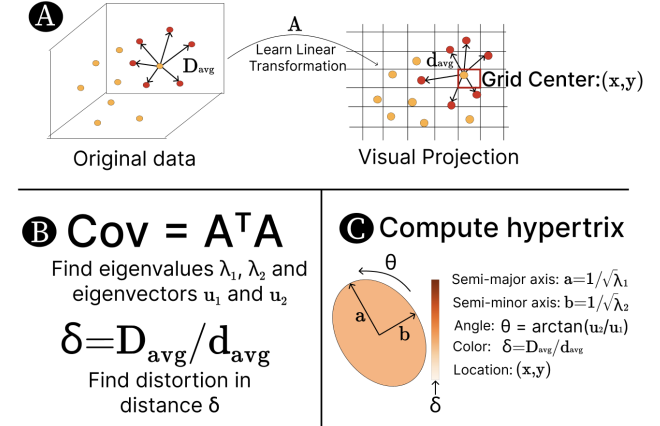


Figure 3: An illustration of the Hypertrix computation process. (A) We select a neighborhood of data points around a grid center in the projection space and a local linear transformation A . (B) The covariance of the transformation matrix A is a measure of distortions such as local scaling and rotation due to projection. We compute its eigenvalues, eigenvectors, and the distortion measure δ and (C) encode them visually as the indicatrix.

To define our method precisely, we first set up some notation. Consider a **data set** of points $X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^D$, where $D > 2$, along with a set of corresponding **projected points**, $Y = \{y_1, y_2, \dots, y_n\}$, where $y_i \in \mathbb{R}^2$. Estimating a single global transformation between them can be challenging since the transformation that results in the projected points might be highly complex and non-linear. Instead, we opt for computing a local linear mapping A that adequately captures the transformation within small localized regions. The **covariance matrix** $Cov = A^T A$ of A encodes the directional variations and local scaling due to the transformation, providing us with a quantifiable measure of distortion. In particular, from the singular value decomposition of A it follows that the image of a unit sphere under A is the same as the image of a unit disk under $\sqrt{A^T A}$.

Using this information about the distortion, we compute the parameters for the Indicatrix \mathcal{E} . This process is repeated uniformly across the entire projection space, allowing us to compare distortions in different regions. Figure 3 and Algorithm 1 outline our approach. We now describe the specifics of hypertrix algorithm.

3.2 Step-by-step algorithm

1. **Sampling points from a neighborhood:** We partition the projection space Y into a uniform grid of $M \times M$ cells. The grid size M determines the granularity of the hypertrix glyphs: we will display one ellipse for each cell that contains a data point, while skipping empty cells. For a grid cell containing data points, we sample from a local neighborhood around the grid center (g^x, g^y) to learn the linear mapping. The number of neighbors N defines the size of this neighborhood and the data points used to approximate the linear

transformation. If a cell has fewer points than N , we select the N closest points to the grid’s center, resulting in a local sample of projections \tilde{Y} corresponding to data points \tilde{X} . Note that in this case we end up sampling beyond the cell boundary—in early tests, we learned this was essential for a good linear approximation and avoiding overfitting.

2. **Computing the hypertrix parameters:** Using the sampled data, we solve the linear equation $\tilde{\mathbf{X}} \cdot \mathbf{A} = \tilde{\mathbf{Y}}$. The number of points influences the utility of the approximation \mathbf{A} . If fewer points than the dimensionality of the data are used, the equation becomes underparameterized, with the resulting collinearity leading to an ill-defined learned transformation. Thus we must use samples with more points than the data dimensionality. In cases of very high dimensionality, we reduce dimensions to 50 using PCA to maintain most of the data variance. (Note that for methods like t-SNE this is the first step of the computation as the data is initialized using the first 50 PCA dimensions before computing the projections, so no information is lost.) To quantify distortion, we need to know how the projection process rotated or scaled the high-dimensional relationships between data points. The covariance matrix of the linear transformation: $Cov = A^T A$ provides a measure of this. Its eigenvalues λ_1, λ_2 and eigenvectors $\mathbf{u}_1, \mathbf{u}_2$ encode the magnitudes of the direction of the most change due to transformation. We use them to obtain the parameters of the hypertrix ellipse a and b , as shown in Figure 3C. To quantify the distortion in distances, we calculate $\delta = D_{avg}/d_{avg}$, the ratio between average distances in projected d_{avg} and high-dimensional spaces D_{avg} .

This process is repeated for all grid cells that contain data to obtain a set of Indicatrix parameters.

3. **Hypertrix visualization:** We render ellipses with the parameters derived above onto the projection space. Each ellipse is centered at the grid cell’s midpoint (g^x, g^y) , with lengths of the semi-major and semi-minor axis corresponding to a and b respectively, and rotated by angle θ . The color intensity of each ellipse, representing δ , indicates the level of distance distortion, with deeper colors denoting higher distortions.

4 RESULTS

We illustrate the hypertrix as applied to a series of synthetic and real-world datasets and projections, each designed to probe different aspects of distortion inherent to dimensionality reduction.

4.1 Synthetic Datasets

1. *Two Gaussian clusters in 100D:* This synthetic dataset consists of 1,000 points sampled from two Gaussian distributions in 100-dimensional space, with one distribution having a higher variance than the other (Figure 4A). This scenario tests the methods’ ability to reflect differences in cluster density and scale. The projections by t-SNE and UMAP depict these clusters—inaccurately—as nearly identical in size. The Indicatrix provides a clear visualization of this discrepancy. This example showcases how the Indicatrix captures the true density and scale of clusters in data, a common occurrence in large-scale datasets.

2. *Linked circles in 3D:* This dataset consists of 1,000 points distributed along two circles of different radii in 3D space, linked and at an angle to each other (Figure 4B). Both t-SNE and UMAP projections show these circles as disjointed and similar in size, which misrepresents their original relationship and scale. The hypertrix in these cases displays inhomogeneously stretched ellipses for each circle. Particularly in the UMAP projection, the variation in ellipses’ shape and orientation in different sections of the circles indicate that there is non-uniform distortion in the projections. Overall, the distortion indicates a significant difference in circle sizes in the original data which is less apparent in the projections.

3. *A square grid with a hole in 2D:* We generate points (x, y) on a uniform 2D grid within the range $[0, 1]$, except for a square-shaped

Algorithm 1: Hypertrix computation algorithm

Data: dataset $X = \{x_1, x_2, \dots, x_n\}$, projections $Y = \{y_1, y_2, \dots, y_n\}$, number of nearest neighbors N , grid size M

Result: Distortion ellipse parameters \mathcal{E}

```

begin
  if (dimensionality of X > 50) then
    |  $X \leftarrow \text{PCA}(n\_components = 50)(X)$ ;
  end
   $g_i^x, g_i^y, grid\_spacing \leftarrow \text{GenerateUniformGrid}(Y, M)$ ;
   $nm \leftarrow \text{InitializeNearestNeighbors}(Y, R = grid\_spacing, n = N)$ 
  for  $i=0$  to  $M \times M$  do
    |  $idx \leftarrow \text{FindNeighbors}([g^x, g^y], nm, N)$ ;
    | if  $idx$  is not empty then
      | |  $\tilde{X} \leftarrow X[idx]$ ;
      | |  $\tilde{Y} \leftarrow Y[idx]$ ;
      | |  $\text{ComputeEllipse}(\tilde{X}, \tilde{Y}, g_i^x, g_i^y)$ ;
    | end
  end
Function  $\text{ComputeEllipse}(\tilde{X}, \tilde{Y}, g^x, g^y)$ :
  Solve:  $\tilde{X} @ \mathbf{A} = \tilde{Y}$ ;
   $Cov \leftarrow A^T @ A$ ;
   $\lambda \leftarrow \text{eigenvalue}(Cov)$ ;
   $\bar{u} \leftarrow \text{eigenvector}(Cov)$ ;
   $\theta \leftarrow \arctan(\bar{u}_1 / \bar{u}_0)$ ;
   $\delta \leftarrow \text{AvgNeighborDistance}(\tilde{X}) / \text{AvgNeighborDistance}(\tilde{Y})$ ;
   $width, height \leftarrow 1 / \sqrt{\lambda_0}, 1 / \sqrt{\lambda_1}$ ;
  Add  $(g^x, g^y, width, height, \theta, \delta)$  to  $\mathcal{E}$ ;
  return  $\mathcal{E}$ 

```

hole offset from the center (Figure 4B). This example shows the value of elliptical glyphs. Even if the viewer does not know what the original data looks like, the ellipses clearly show that the UMAP projection has stronger stretching in one direction than another overall. Similar local non-conformal distortion is evident in the t-SNE view. In addition, in both the t-SNE and UMAP projections, the hole appears disproportionately enlarged. The hypertrix reveals this through ellipses that are elongated around both the perimeter of the hole and the edges of the grid.

4.2 Handwritten digits

The MNIST dataset [5], consisting of 60,000 images of handwritten digits, is a classic benchmark. The hypertrix for t-SNE projections of the data (Figure 1B) yields several qualitative insights:

1. *The projections inhomogeneously distort the data.* Notably, the cluster for the digit ‘1’ is disproportionately expanded compared to those for other digits. Indeed, after examining the pairwise intra-cluster distances for all digits, we find that distances within the ‘1’ cluster were consistently about half the magnitude of those in other clusters. This suggests that the cluster is tighter than the others, and should appear smaller in visual representation. This observation intuitively makes sense, since there is less variation in the ways people write the digit 1, as compared to other digits.

2. *The projection shows pronounced distortions in the region between clusters of the digits ‘0’, ‘5’, and ‘6’* The hypertrix for this region shows compressed, elongated ellipses with a darker color, indicating a high degree of distortion between these clusters. This distortion could suggest a greater separation in the original dataset than is represented in the scatterplot visualization. Comparing the

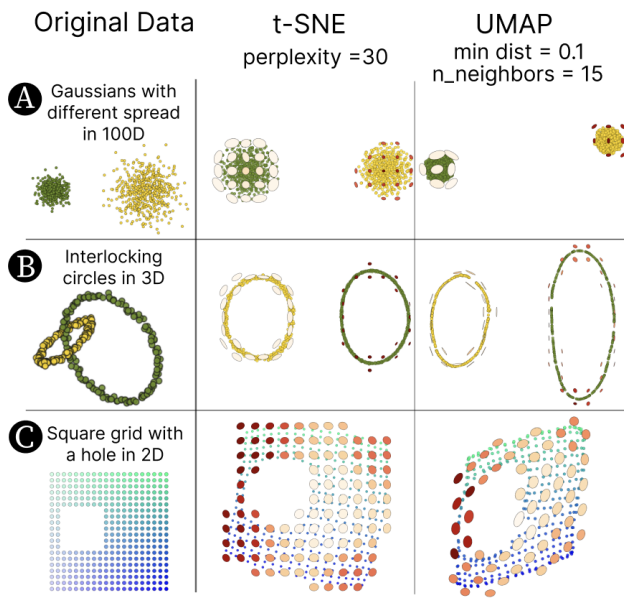


Figure 4: Hypertrix visuals overlaid over the t-SNE and UMAP projections of various synthetically generated datasets. The t-SNE and UMAP projections do not maintain certain crucial aspects of the data. In (A) and (B) the relative sizes of structures are distorted. In (C) the overall aspect ratio of the grid and shape of the “hole” are distorted. The hypertrix overlay aids in identifying these inaccuracies.

distance between the centroids of these clusters confirms our qualitative observation.

4.3 Embeddings of New York Times text

Transforming text into high-dimensional embeddings is a standard practice in natural language processing. Dimensionality reduction on these text embeddings generates spatial maps of text data useful for exploration and discovery. We use OpenAI’s text embedding API [16] to generate embeddings for a collection of metadata for 5,000 news articles [6]. Each data point consists of text with an article headline and keywords related to the article. The articles have been broadly classified into 8 main categories such as Sports, Crime, Politics, etc., but each category may contain several subtopics. Each embedding has 1,536 dimensions. The t-SNE projection of these embeddings along with annotated cluster labels generated using GPT4 API [18] are shown in Figure 5A.

1. In the “Wedding announcements” cluster, the hypertrix ellipses indicate that this cluster is relatively enlarged, however with low distortion in distances. We investigate this anomalous behavior by looking at some of the individual texts from the cluster and notice that they all contain the keywords “Weddings and Engagements” along with the title. Indeed, the embeddings in this cluster belong to a relatively small region in high-dimensional embedding space (Figure 5B, right). The hypertrix glyphs can alert a viewer that this cluster may be less diverse than others. The glyphs show other, more subtle effects as well. As with the MNIST projections, the color and size of glyphs suggest there is a higher distortion at the boundary of high-density clustered regions that separate different categories. Furthermore, several regions show eccentric ellipses, indicating non-conformal distortion of the data.

5 CONCLUSIONS AND FUTURE WORK

We have introduced the **hypertrix**, based on Tissot’s Indicatrix, to visualize the distortion introduced by nonlinear dimensionality-reduction methods. The technique computes distortions using a grid

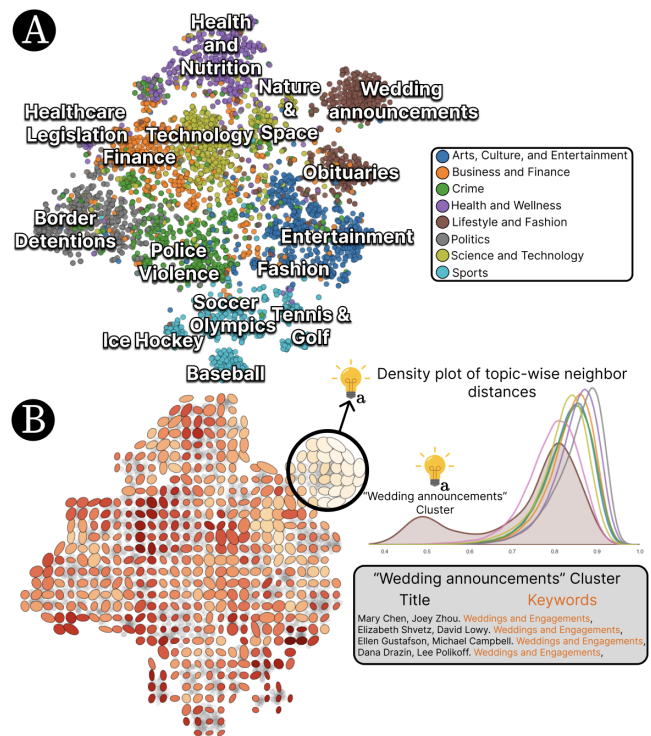


Figure 5: We compute hypertrix for (A) t-SNE projections of embeddings of metadata for a collection of New York Times articles. (B) The resulting indicatrices are overlaid on the projection plots. Notably, the color of the hypertrix in the “Wedding announcements” cluster suggests anomalous distortion; analysis of the underlying data indicates this cluster is more compact in high dimensions than other clusters, reflecting the limited diversity in keywords in the underlying text.

of local linear transformations that approximate the overall projection. It then represents this information as colored elliptical glyphs overlaid on the projection. We showed how applying this idea to both synthetic datasets and real-world examples brings insight into the true sizes and shapes of visual clusters in the underlying data. The hypertrix helps to visualize where these distortions occur but also quantifies these distortions in an interpretable way. Experience indicates that no single set of projection hyperparameters consistently minimizes distortion across various datasets. We believe the hypertrix overlays can assist users in making informed decisions about acceptable trade-offs when choosing dimensionality reduction methods and hyperparameters. We note that different grid sizes and sampling approaches may impact the hypertrix visuals. T-SNE and UMAP can produce highly nonlinear projections that may differ significantly between regions of the visual, thus we advise selecting as fine-grained grid size as possible without compromising on the interpretation of the hypertrix ellipse.

Several future directions look promising. A user study could shed light on how people read the hypertrix, and how it affects their decisions. One could also explore how to integrate the hypertrix with automated approaches, to create a human-in-the-loop system for hyperparameter optimization. Finally, there is no doubt room for visual improvements to the hypertrix glyphs, perhaps augmenting them with addition distortion information.

ACKNOWLEDGMENTS

The authors wish to thank Elena Glassman and Catherine Yeh for discussions and valuable suggestions during development of this work.

REFERENCES

- [1] M. Aupetit. Visualizing the trustworthiness of a projection. In *14th European Symposium on Artificial Neural Networks, ESANN 2006, Bruges, Belgium, April 26-28, 2006, Proceedings*, pp. 271–276, 2006. 2
- [2] M. Aupetit. Visualizing distortions and recovering topology in continuous projection techniques. *Neurocomputing*, 70(7):1304–1330, 2007. Advances in Computational Intelligence and Learning. doi: 10.1016/j.neucom.2006.11.018 2
- [3] D. Brodbeck, M. Chalmers, A. Lunzer, and P. Cotture. Domesticating bead: adapting an information visualization system to a financial institution. In *Proceedings of VIZ '97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, pp. 73–80, 1997. doi: 10.1109/INFVIS.1997.636789 2
- [4] A. Chatzimparmpas, R. M. Martins, and A. Kerren. t-visne: Interactive assessment and interpretation of t-sne projections. *IEEE Transactions on Visualization and Computer Graphics*, 26(8):2696–2714, 2020. doi: 10.1109/TVCG.2020.2986996 1
- [5] L. Deng. The mnist database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine*, 29(6):141–142, 2012. 3
- [6] dsteffa. New york times topics, 2024. 4
- [7] A. Griparis, D. Faur, and M. Datcu. A dimensionality reduction approach for the visualization of the cluster space: A trustworthiness evaluation. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2917–2920, 2016. doi: 10.1109/IGARSS.2016.7729753 2
- [8] N. Heulot, M. Aupetit, and J.-D. Fekete. Proxilens: Interactive exploration of high-dimensional data using projections. In *VAMP@EuroVis*, 2013. 2
- [9] H. Jeon, A. Cho, J. Jang, S. Lee, J. Hyun, H. Ko, J. Jo, and J. Seo. Zadu: A python library for evaluating the reliability of dimensionality reduction embeddings. In *2023 IEEE Visualization and Visual Analytics (VIS)*, pp. 196–200. IEEE Computer Society, Los Alamitos, CA, USA, oct 2023. doi: 10.1109/VIS54172.2023.00048 1
- [10] P. H. Laskowski. The traditional and modern look at Tissot's indicatrix. *The American Cartographer*, 16(2):123–133, 1989. doi: 10.1559/152304089783875497 1
- [11] S. Lespinats and M. Aupetit. CheckViz: Sanity Check and Topological Clues for Linear and Non-Linear Mappings. *Computer Graphics Forum*, 2011. doi: 10.1111/j.1467-8659.2010.01835.x 2
- [12] R. Martins, D. Coimbra, R. Minghim, and A. Telea. Visual analysis of dimensionality reduction quality for parameterized projections. *Computers Graphics*, 41, 06 2014. doi: 10.1016/j.cag.2014.01.006 2
- [13] D. Meng, Y. Leung, and Z. Xu. A new quality assessment criterion for nonlinear dimensionality reduction. *Neurocomputing*, 74(6):941–948, 2011. doi: 10.1016/j.neucom.2010.10.011 2
- [14] Met Office. *Cartopy: a cartographic python library with a Matplotlib interface*. Exeter, Devon, 2010 - 2015. 1
- [15] J. Milnor. A problem in cartography. *The American Mathematical Monthly*, 76(10):1101–1112, 1969. 2
- [16] OpenAI. Text embedding api documentation. <https://platform.openai.com/docs/guides/embeddings>, 2023. Accessed: 04/2024. 4
- [17] A. Papadopoulos. Nicolas-auguste tissot: A link between cartography and quasiconformal theory, 2016. 1
- [18] S. Raval, C. Wang, F. Viégas, and M. Wattenberg. Explain-and-test: An interactive machine learning framework for exploring text embeddings. In *2023 IEEE Visualization and Visual Analytics (VIS)*, pp. 216–220, 2023. doi: 10.1109/VIS54172.2023.00052 4
- [19] B. Rieck and H. Leitte. Persistent homology for the evaluation of dimensionality reduction schemes. *Computer Graphics Forum*, 34:431–440, 07 2015. doi: 10.1111/cgf.12655 2
- [20] T. Schreck, T. von Landesberger, and S. Bremm. Techniques for precision-based visual analysis of projected data. *Information Visualization*, 9(3):181–193, 2010. doi: 10.1057/ivs.2010.2 2
- [21] C. Seifert, V. Sabol, and W. Kienreich. Stress Maps: Analysing Local Phenomena in Dimensionality Reduction Based Visualisations. In J. Kohlhammer and D. Keim, eds., *EuroVAST 2010: International Symposium on Visual Analytics Science and Technology*. The Eurographics Association, 2010. doi: 10.2312/PE/EuroVAST/EuroVAST10/013-018 2
- [22] J. Stahnke, M. Dörk, B. Müller, and A. Thom. Probing projections: Interaction techniques for interpreting arrangements and errors of dimensionality reductions. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):629–638, 2016. doi: 10.1109/TVCG.2015.2467717 1
- [23] E. D. Sun, R. Ma, and J. Zou. Dynamic visualization of high-dimensional data. *bioRxiv*, 2022. doi: 10.1101/2022.05.27.493785 2
- [24] A. Tissot. Mémoire sur la représentation des surfaces et les projections des cartes géographiques. *Nouvelles annales de mathématiques : journal des candidats aux écoles polytechnique et normale*, 18:337–356, 1879. 1
- [25] J. Venna and S. Kaski. Neighborhood preservation in nonlinear projection methods: An experimental study. In G. Dorffner, H. Bischof, and K. Hornik, eds., *Artificial Neural Networks — ICANN 2001*, pp. 485–491. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001. 2
- [26] M. Wattenberg, F. Viégas, and I. Johnson. How to use t-sne effectively. *Distill*, 2016. doi: 10.23915/distill.00002 1