

Extended Abstract Track

Topographic Soft-Matching Distance (TSMD): A Spatially-Constrained Extension of Soft-Matching

Editors: List of editors' names

Abstract

Neural populations are spatially organized according to their functional tuning, but it remains unclear whether this organization constitutes a conserved principle that recurs across individuals. Existing approaches require a predefined tuning variable (*e.g.*, orientation preference) and compare maps qualitatively, which becomes a limiting factor for responses to complex natural stimuli where tuning is high-dimensional, and its relevant axes may not be known *a priori*. Here, we introduce the Topographic Soft-Matching Distance (TSMD), which extends soft-matching via fused Gromov–Wasserstein optimal transport as a method to identify correspondences that preserve both tuning similarity and relative spatial geometry, without requiring predefined tuning features. We test these correspondences against a permutation null that disrupts the tuning–spatial pairing, yielding a statistical test for topographic organization. We validate TSMD on synthetic orientation maps, demonstrating its invariance to rigid transformations and sensitivity to disrupted organization. When applied to single-neuron recordings, TSMD finds little shared organization in mouse V1, but strong cross-animal correspondence in macaque V1 and V4. However, in inferotemporal cortex, topographic structure does not seem to be reliably shared across animals. TSMD thus offers a general approach for quantifying whether spatial organization of neural representations is conserved across individuals and systems.

Keywords: Topographical Organization, Neural Representations, Optimal Transport

1. Introduction

For over half a century, neuroscientists have documented systematic relationships between neural tuning and cortical space. Classical topographic organization refers to maps in which neural tuning across the cortex continuously tracks an interpretable latent variable, as in the orientation pinwheels of primary visual cortex (Hubel and Wiesel, 1962; Bonhoeffer and Grinvald, 1991). More broadly, topographic organization refers to the tendency for similarly tuned neurons to occur near one another, producing structures such as category-selective patches for faces, bodies, and scenes (Kanwisher et al., 1997; Downing et al., 2001; Epstein et al., 1999). This is often described by a decay in response similarity with cortical distance (Lee and DiCarlo, 2019; Margalit et al., 2024). Existing approaches to quantify this, however, are typically formulated with respect to interpretable response properties specified in advance, and do not extend to the high-dimensional response patterns elicited by natural stimuli, nor offer a general measure of whether spatial organization corresponds across populations.

A separate literature compares neural systems at the level of individual neurons and populations. Representational similarity metrics such as Centered Kernel Alignment (CKA) (Kornblith et al., 2019), Representational Similarity Analysis (RSA) (Kriegeskorte et al., 2008) and the Procrustes shape distance (Williams et al., 2021) quantify geometric similarity of

Extended Abstract Track

these systems, but are rotation-invariant, discarding how individual neurons respond to inputs. Recently proposed optimal transport methods (Peyré and Cuturi, 2019) match units by their response profiles. The soft-matching distance (Khosla and Williams, 2024) finds rotation-sensitive correspondences between tuning curves while remaining invariant to arbitrary unit ordering. While this compares tuning curve distributions across systems of unequal size, it only captures functional properties, not whether the spatial *arrangement* of neurons is preserved—a key property of topographic organization.

Here, we ask whether functional correspondence across populations is accompanied by correspondence in spatial organization. We define *topographic correspondence* as the extent to which populations with corresponding functional tuning exhibit corresponding relative spatial arrangements, irrespective of absolute position or orientation in cortical space. To quantify this, we introduce the topographic soft-matching distance (TSMD), which extends soft-matching into the fused Gromov–Wasserstein family (Vayer et al., 2020) by jointly matching neurons on tuning and spatial geometry. A hyperparameter α interpolates between functional and spatial matching. Comparison against a permutation null that disrupts the association between tuning and location provides a statistical test of whether the observed tuning–spatial correspondence exceeds that expected by chance, without requiring predefined or interpretable tuning dimensions. We validate TSMD on simulated orientation maps, where it is invariant to rigid transforms while distinguishing maps in which this organization has been disrupted (Sec. 3), and on neural recordings, where it recovers the canonical absence of topography in mouse V1 (Appendix C). In macaque V1, by contrast, TSMD reveals strong topographic correspondence across individuals (Sec. 3), which persists after accounting for retinotopy (Appendix F, B).

2. Methods

Consider two neural populations with tuning matrices (mean-centered, unit-normed) $\mathbf{X} \in \mathbb{R}^{M \times N_x}$ and $\mathbf{Y} \in \mathbb{R}^{M \times N_y}$, where M is the number of stimulus conditions and N_x and N_y neurons respectively. Each neuron also has a spatial location, $\mathbf{u}_i$ for neuron i in $\mathbf{X}$ and $\mathbf{v}_j \in \mathbb{R}^2$ for neuron j in $\mathbf{Y}$.

Soft-Matching Distance The soft-matching distance (Khosla and Williams, 2024) treats each population as a uniform empirical measure and computes the optimal transport cost between their tuning curves: $d_T(\mathbf{X}, \mathbf{Y}) = \min_{\mathbf{T} \in \mathcal{T}(N_x, N_y)} \langle \mathbf{C}, \mathbf{T} \rangle_F$, where $\mathbf{C}_{ij} = 1 - \mathbf{x}_i^\top \mathbf{y}_j$ is the cosine distance, $\langle \cdot, \cdot \rangle_F$ the Frobenius inner product, and $\mathcal{T}(N_x, N_y)$ the transportation polytope (De Loera and Kim, 2013): $\mathcal{T}(N_x, N_y) = \{\mathbf{T} \in \mathbb{R}_{\geq 0}^{N_x \times N_y} : \mathbf{T} \mathbf{1}_{N_y} = \frac{1}{N_x} \mathbf{1}_{N_x}, \mathbf{T}^\top \mathbf{1}_{N_x} = \frac{1}{N_y} \mathbf{1}_{N_y}\}$. The minimizing coupling is invariant to permutations of neuron indices yet sensitive to rotations, unlike CKA or Procrustes shape distance. Its cost, however, is a function of tuning alone: two populations with identical tuning curves are assigned an identical coupling regardless of spatial position.

Topographic Soft-Matching Distance (TSMD). TSMD augments the tuning cost $\mathbf{C}_{ij}$ with a geometric cost that compares intra-population distances $d_{\mathbf{X}}(i, i') = \|\mathbf{u}_i - \mathbf{u}_{i'}\|^2$ and $d_{\mathbf{Y}}(j, j') = \|\mathbf{v}_j - \mathbf{v}_{j'}\|^2$, thus penalizing couplings that disturb relative spatial arrangement. A hyperparameter $\alpha \in [0, 1)$ balances the two costs:

Extended Abstract Track

$$\mathcal{L}_\alpha(\mathbf{T}) = \underbrace{\alpha \sum_{i,i',j,j'} |d_{\mathbf{X}}(i,i') - d_{\mathbf{Y}}(j,j')|^2 \mathbf{T}_{ij} \mathbf{T}_{i'j'}}_{\text{Geometric term}} + \underbrace{(1 - \alpha) \sum_{i,j} \mathbf{C}_{ij} \mathbf{T}_{ij}}_{\text{Tuning term}},$$

minimized over $\mathcal{T}(N_x, N_y)$. We write $\text{TSMD}_\alpha(\mathbf{X}, \mathbf{Y}) = \min_{\mathbf{T} \in \mathcal{T}(N_x, N_y)} \mathcal{L}_\alpha(\mathbf{T})$. A pair is coupled strongly only when both neurons are functionally similar *and* lie in matching topographic neighborhoods. This is exactly the fused Gromov–Wasserstein distance (Vayer et al., 2018), and defines a semi-metric¹ on the space of tuning-and-location populations. At $\alpha = 0$, it reduces exactly to soft-matching (squared), while at $\alpha = 1$, it matches by spatial arrangement alone. Between these, TSMD balances tuning against topographic fidelity.

Null Test for Topography. Rather than fix α , we sweep $\alpha \in [0, 1)$, and summarize the alignment statistic $A(\alpha)$, the mean tuning correlation under the optimal coupling by its area $\text{AUC} = \int_0^1 A(\alpha) d\alpha$. We compare each AUC against a permutation null, constructed by randomly reassigning tuning curves to spatial locations. This conserves tuning distribution, while destroying spatial–tuning pairing, yielding a null area AUC_π . We assess significance over $K = 15$ resamples. For within-population comparisons, each split divides electrodes into two disjoint halves, and for cross-population comparisons, each split subsamples half of each population’s electrodes. In both cases, the null permutes tuning across electrode positions in one group. For each split k , we compute the paired difference $\Delta_k = \text{AUC}_k - \text{AUC}_{\pi,k}$ and test $\{\Delta_k\}_{k=1}^K$ against zero with a one-sided Wilcoxon signed-rank test.

3. Results

Synthetic Tuning Maps. To test whether TSMD jointly captures tuning and topographic correspondence, we validate it on synthetic orientation maps with known structure (construction in Appendix A). Let $\mathbf{X}_{\text{base}} \in \mathbb{R}^{M \times N}$ denote the base cortical sheet with M orientation directions and N neurons. We compute TSMD (over the α -continuum) between $\mathbf{X}_{\text{base}}$ and two families of transformed copies: **rigid transformations** (translation, rotation; see Appendix G) which preserve pairwise tuning and relative geometry, and should give near-constant alignment across all α , and **permutations**, which reassigns each location’s tuning via a random permutation $\pi : [N] \rightarrow [N]$, preserving the tuning distribution while destroying spatial correspondence, and should therefore monotonically reduce alignment from $\alpha = 0$ toward $\alpha = 1$. We demonstrate this in Fig. 1—rotation and translation leave alignment near-constant, while permutation degrades monotonically as α increases.

THINGS Ventral Stream Spiking Data. Macaque V1 exhibits well-established columnar organization (Ohki et al., 2006; Nauhaus et al., 2012), providing a positive control for whether TSMD can recover topographic structure from responses to natural images. We analyze multi-unit activity from the THINGS Ventral Stream Spiking Dataset (Papale et al., 2025), comprising recordings from two monkeys (N and F) across V1, V4, and IT in response to natural images from THINGS (Hebart et al., 2019), together with electrode coordinates inferred from array placement images (Appendix F). Within each monkey and region, we randomly partition the electrodes into two complementary halves and compare them with

1. See Prop. 5.3 of Vayer et al. (2020)

Extended Abstract Track

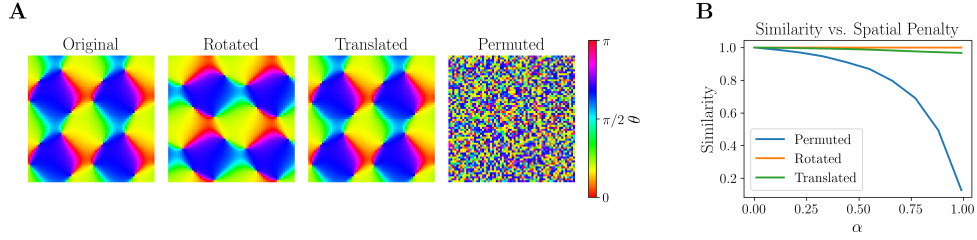


Figure 1: **Geometric Transforms of a Simulated Orientation Map.** (A) Synthetic pinwheel map under rotation, translation, and random permutation. Color indicates preferred orientation $\theta \in [0, \pi)$. (B) TSMD alignment between the original and transformed maps across α .

TSMD. This reveals reliable topographic organization within animals (Appendix E). TSMD reveals significant cross-monkey topographic correspondence in V1 and V4 ($p = 9.2e-05$ and $3.1e-05$, respectively), providing quantitative evidence that aspects of tuning-spatial organization are shared across individuals. In contrast, we find no reliable cross-monkey correspondence in IT ($p = 0.339$; Fig. 2). Although this may reflect either genuine divergence or limitations of the available recordings, the dissociation highlights how TSMD can formally test where topographic organization does and does not converge across individuals.

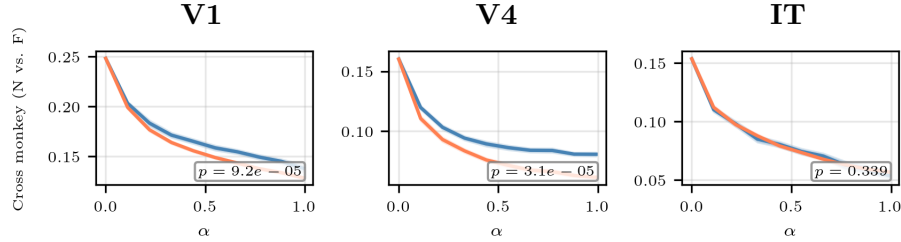


Figure 2: **Cross-monkey Topographic Correspondence.** TSMD scores comparing different monkeys are evaluated against a permutation null. Original scores (blue) exceed the null (orange) in V1 and V4, providing evidence for conserved topographic organization. No significant separation is observed in IT ($p = 0.339$).

4. Conclusion

We introduced the Topographic Soft-Matching Distance (TSMD), a framework for comparing neural populations jointly by their functional tuning and relative spatial organization. TSMD recovers known structure in synthetic maps and neural recordings, and reveals shared topographic organization across macaque visual areas even after controlling for retinotopy. By operating directly on high-dimensional response profiles without predefined tuning features, TSMD provides a general approach for testing whether the spatial organization of neural representations converges across individuals and systems.

Extended Abstract Track

References

- Tobias Bonhoeffer and Amiram Grinvald. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, 353(6343):429–431, 1991.
- Vincent Bonin, Mark H Histed, Sergey Yurgenson, and R Clay Reid. Local diversity and fine-scale organization of receptive fields in mouse visual cortex. *Journal of Neuroscience*, 31(50):18506–18521, 2011.
- Jesús A De Loera and Edward D Kim. Combinatorics and geometry of transportation polytopes: An update. *Discrete geometry and algebraic combinatorics*, 625:37–76, 2013.
- Saskia EJ de Vries, Jerome A Lecoq, Michael A Buice, Peter A Groblewski, Gabriel K Ocker, Michael Oliver, David Feng, Nicholas Cain, Peter Ledochowitsch, Daniel Millman, et al. A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nature neuroscience*, 23(1):138–151, 2020.
- Paul E Downing, Yuhong Jiang, Miles Shuman, and Nancy Kanwisher. A cortical area selective for visual processing of the human body. *Science*, 293(5539):2470–2473, 2001.
- Russell Epstein, Alison Harris, Damian Stanley, and Nancy Kanwisher. The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, 23(1):115–125, 1999.
- Martin N Hebart, Adam H Dickter, Alexis Kidder, Wan Y Kwok, Anna Corriveau, Caitlin Van Wicklin, and Chris I Baker. Things: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PloS one*, 14(10):e0223792, 2019.
- DH Hubel and TN Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.
- Nancy Kanwisher, Josh McDermott, and Marvin M Chun. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of neuroscience*, 17(11):4302–4311, 1997.
- Meenakshi Khosla and Alex H Williams. Soft matching distance: A metric on neural representations that captures single-neuron tuning. In *Proceedings of unireps: the first workshop on unifying representations in neural models*, pages 326–341. PMLR, 2024.
- Simon Kornblith, Mohammad Norouzi, Honglak Lee, and Geoffrey Hinton. Similarity of neural network representations revisited. In *International conference on machine learning*, pages 3519–3529. PMIR, 2019.
- Nikolaus Kriegeskorte, Marieke Mur, and Peter A Bandettini. Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2:249, 2008.
- Hyodong Lee and James J DiCarlo. Topographic deep artificial neural networks (tdanns) predict face selectivity topography in primate inferior temporal (it) cortex. *arXiv preprint arXiv:1909.09847*, 2019.

Extended Abstract Track

- Eshed Margalit, Hyodong Lee, Dawn Finzi, James J DiCarlo, Kalanit Grill-Spector, and Daniel LK Yamins. A unifying framework for functional organization in early and higher ventral visual cortex. *Neuron*, 112(14):2435–2451, 2024.
- Ian Nauhaus, Kristina J Nielsen, Anita A Disney, and Edward M Callaway. Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex. *Nature neuroscience*, 15(12):1683–1690, 2012.
- Kenichi Ohki and R Clay Reid. Specificity and randomness in the visual cortex. *Current opinion in neurobiology*, 17(4):401–407, 2007.
- Kenichi Ohki, Sooyoung Chung, Yeang H Ch’ng, Prakash Kara, and R Clay Reid. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature*, 433(7026):597–603, 2005.
- Kenichi Ohki, Sooyoung Chung, Prakash Kara, Mark Hübener, Tobias Bonhoeffer, and R Clay Reid. Highly ordered arrangement of single neurons in orientation pinwheels. *Nature*, 442(7105):925–928, 2006.
- Paolo Papale, Feng Wang, Matthew W Self, and Pieter R Roelfsema. An extensive dataset of spiking activity to reveal the syntax of the ventral stream. *Neuron*, 113(4):539–553, 2025.
- Gabriel Peyré and Marco Cuturi. *Computational optimal transport: With applications to data science*. Now Foundations and Trends, 2019.
- Titouan Vayer, Laetita Chapel, Rémi Flamary, Romain Tavenard, and Nicolas Courty. Fused gromov-wasserstein distance for structured objects: theoretical foundations and mathematical properties. *arXiv preprint arXiv:1811.02834*, 2018.
- Titouan Vayer, Laetitia Chapel, Rémi Flamary, Romain Tavenard, and Nicolas Courty. Fused gromov-wasserstein distance for structured objects. *Algorithms*, 13(9):212, 2020.
- Alex H Williams, Erin Kunz, Simon Kornblith, and Scott Linderman. Generalized shape metrics on neural representations. *Advances in neural information processing systems*, 34:4738–4750, 2021.

Extended Abstract Track

Appendix A. Synthetic Orientation Map Construction

We generate synthetic orientation maps by bandpass-filtering complex white noise in the frequency domain by means of a Fourier transform. This simultaneously yields controllable spatial structure and known ground-truth orientation preference.

On a grid of $N_x \times N_y$ neurons, we draw complex white noise $\varepsilon(x, y) = \varepsilon_{\text{Re}}(x, y) + \iota \varepsilon_{\text{Im}}(x, y)$ with $\varepsilon_{\text{Re}}, \varepsilon_{\text{Im}} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. We filter it with a Gaussian bandpass centered at radial frequency ω_0 ,

$$H(\omega_x, \omega_y) = \exp\left(-\frac{(\omega - \omega_0)^2}{2\sigma_\omega^2}\right), \quad \omega = \sqrt{\omega_x^2 + \omega_y^2},$$

and obtain the filtered field $z(x, y) = \mathcal{F}^{-1}[H \cdot \mathcal{F}[\varepsilon]]$. The radius ω_0 and width σ_ω control the periodicity of the resultant map.

Orientation preference is obtained from the complex phase of z . Since $\arg[z] \in (-\pi, \pi]$, the half-angle $\frac{1}{2} \arg[z]$ lies in $(-\pi/2, \pi/2]$. We wrap this into the orientation domain $[0, \pi)$,

$$\theta_{\text{pref}}(x, y) = \frac{1}{2} \arg[z(x, y)] \bmod \pi.$$

Orientation selectivity is the normalized magnitude,

$$s(x, y) = \frac{|z(x, y)|}{\max_{x, y} |z(x, y)|} \in [0, 1].$$

From these maps, we proceed to build tuning curves. At each location, the response to a stimulus orientation θ follows a von-Mises profile,

$$\tilde{r}(\theta \mid \theta_{\text{pref}}, \kappa) = \frac{\exp(\kappa \cos 2\delta)}{2\pi I_0(\kappa)}, \quad \delta = \left[\left(\theta - \theta_{\text{pref}} + \frac{\pi}{2} \right) \bmod \pi \right] - \frac{\pi}{2},$$

with δ as the angular difference, κ the concentration and I_0 the modified Bessel function of the first kind. We introduce selectivity by scaling through amplitude $\kappa = \kappa_{\text{base}} s(x, y)$, giving the final response

$$r(\theta \mid x, y) = s(x, y) \tilde{r}(\theta \mid \theta_{\text{pref}}(x, y), \kappa).$$

Highly selective locations therefore have both sharper tuning and larger responses, while weakly selective locations respond broadly and weakly. Evaluating r over the stimulus orientations gives the tuning matrix $\mathbf{X}$ finally used by TSMD. We demonstrate the effect of choosing different k_0 and σ_k in Fig. 3.

Appendix B. Experimental Controls for Local Correlations

To account for local correlations introduced by spatial proximity of the electrodes, we run the following set of controls:

- **Random Split-Half:** We divide the electrodes into two equally sized halves and compute TSMD between them.

Extended Abstract Track

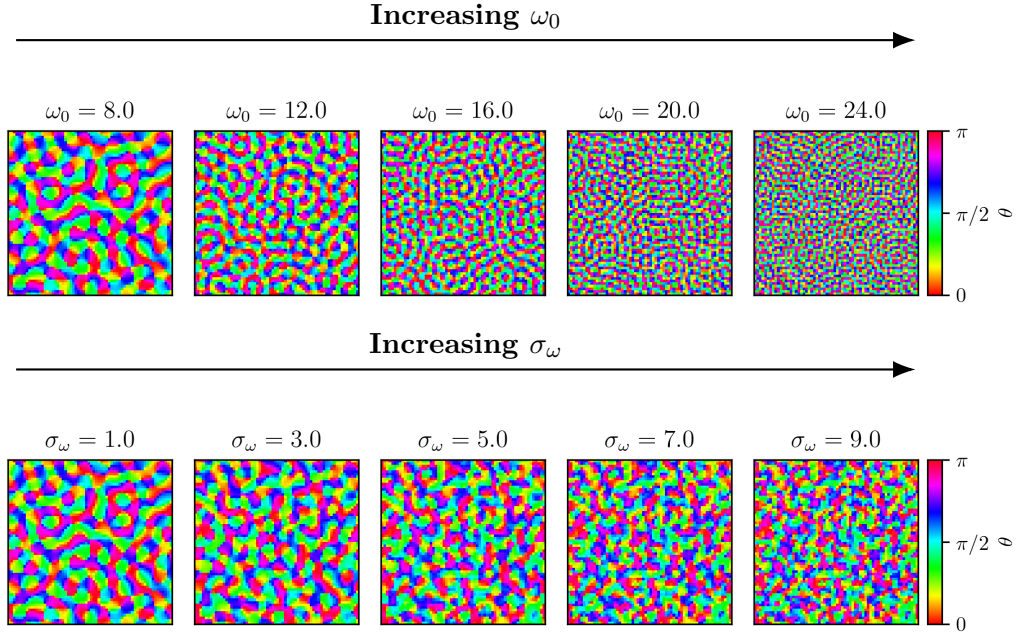


Figure 3: **Effect of Bandpass Filter Parameters on Synthetic Orientation Maps.** (**Top**) Increasing the center frequency ω_0 (with $\sigma_\omega = 1$ fixed) raises the spatial frequency of the map, producing finer orientation structure. (**Bottom**) Increasing the bandwidth σ_ω (with $\omega_0 = 8$ fixed) widens the passband in the Fourier domain, admitting a broader range of spatial frequencies and increasing ringing artifacts. Color indicates preferred orientation $\theta \in [0, \pi)$.

Extended Abstract Track

- **Alternate Electrodes:** We assign spatially adjacent electrodes to opposite halves in a checkerboard pattern, such that no two neighboring electrodes belong to the same subset. Through this, we remove the contribution of *purely* local correlations. Intuitively, if the detected topographic structure were driven by short-range proximity effects, separating adjacent electrodes would account for correlated signals.
- **Random Subsampling:** We randomly subsample a fixed number of electrodes ($N = 100$) from the full population to construct a tuning and spatial coordinate matrix, hence greatly reducing the spatial density of the recording. By controlling for the number of electrodes, we test whether topographic signatures exist solely due to the numerosity of electrodes.

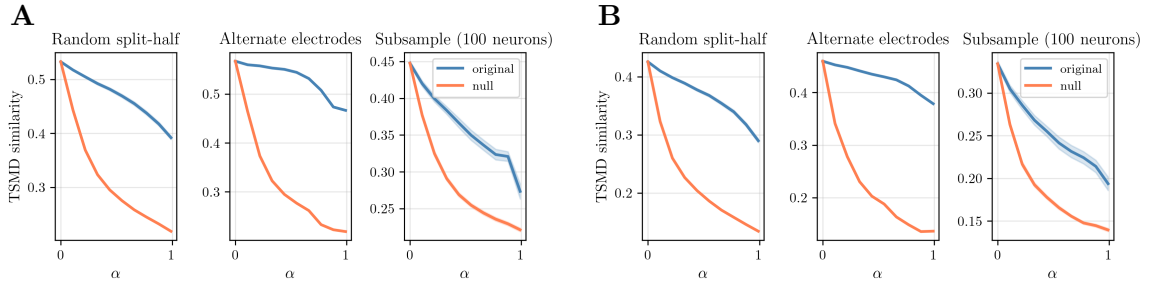


Figure 4: **Experimental Controls.** We run controls for different confounders of topographic signature as described earlier, for monkey N (**A**) and monkey F (**B**). In both cases, we observe a statistically significant original vs. null separation, indicating that the observed topographic signature through TSMD is not an artifact of inter-electrode spacing or the large number of electrodes used for recording neural activity.

Appendix C. Allen Brain Observatory – Mouse V1

To test for topography in the mouse V1, we use two-photon calcium imaging data from the Allen Brain Observatory (ABO) Visual Coding dataset (de Vries et al., 2020). This dataset contains somatic calcium fluorescence traces from genetically identified excitatory neurons across multiple visual areas during presentation of 8 drifting gratings at a fixed temporal frequency of 2 Hz, each shown approximately 15 times. We selected $N = 10$ mice with imaged populations in primary visual cortex (VISp), restricting to Layers 2/3 Cre-line-targeted excitatory neurons (Cux2-CreERT2 and Slc17a7-Cre). For each neuron, we computed the mean $\Delta F/F$ response across trials for each of the 8 orientations, yielding a tuning matrix $\mathbf{X} \in \mathbb{R}^{8 \times n}$ per mouse. Spatial coordinates were taken from the two-photon field of view, which provides a two-dimensional tangential plane through the cortical surface ($\approx 500 \mu\text{m}$ extent). Last, we preprocess the tuning matrices to remove all dead units, mean-center and L2 normalize tuning curves, and min-max normalize coordinates to the unit square.

Extended Abstract Track

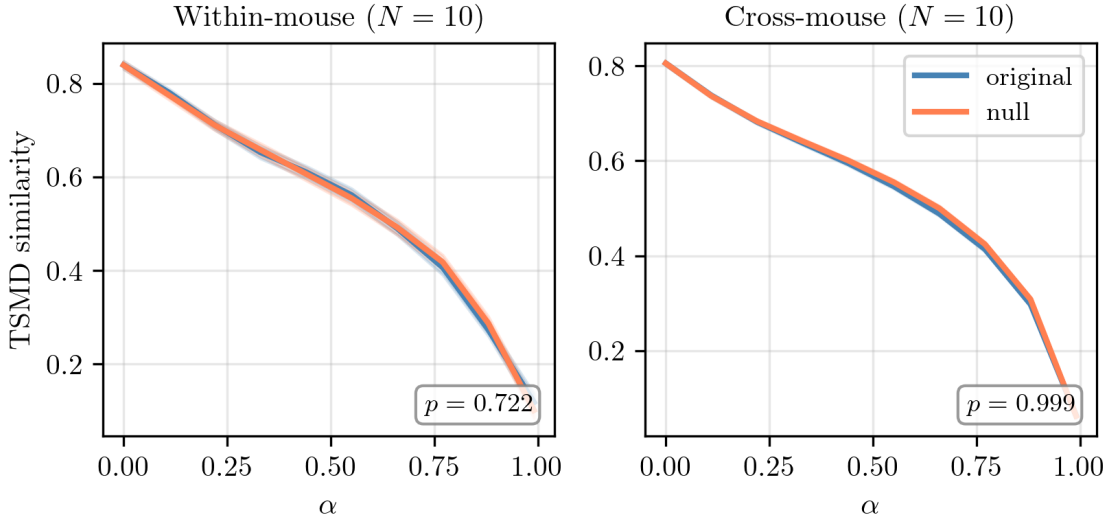


Figure 5: **Similarity vs. Spatial Penalty Within and Across Mice Pairs.** Original (blue) and permutation-null (orange) TSMD similarity are statistically indistinguishable across the entire α continuum, both within and across mice pairs. This indicates no significant structure in mouse V1.

We observe that similarities both within- and across-mice remain indistinguishable from the permutation null across α (Fig. 5). We use our permutation null defined in Section 2 to obtain significance values $p = 0.722$ within and $p > 0.999$ across $N = 10$ mice respectively. In other words, TSMD finds no significant topographic structure in mouse V1, which is consistent with the widely documented “salt-and-pepper” organization (Ohki et al., 2005; Ohki and Reid, 2007; Bonin et al., 2011). This motivates extending our analysis to a species with known columnar organization, where TSMD should detect topographic structure if the metric is sufficiently sensitive.

Appendix D. Retinotopic Confound with Natural Scenes Stimuli in Mouse V1

In Appendix C, we report TSMD results for mouse V1 using drifting grating stimuli. Here we show *why* natural scenes, despite having significantly higher-dimensional tuning vectors (118 scenes vs. 8 orientations) are unsuitable for testing topographic organization with TSMD.

We start by observing general patterns in natural scenes: they are spatially structured images, i.e.: luminance, edges, and texture vary across the visual field. Even in a small field of view of a two-photon imaging session, neighboring neurons have slightly offset receptive fields which sample overlapping image patches. As a result, their natural scene response vectors are correlated by virtue of their shared visual input, agnostic to their feature tuning (e.g., orientation) being spatially organized. When TSMD compares a neural population

Extended Abstract Track

to its null model in which tuning vectors are randomly reassigned to spatial locations, the receptive fields overlap and produce consistent differences, which manifest themselves as retinotopic correlations that permutation destroys.

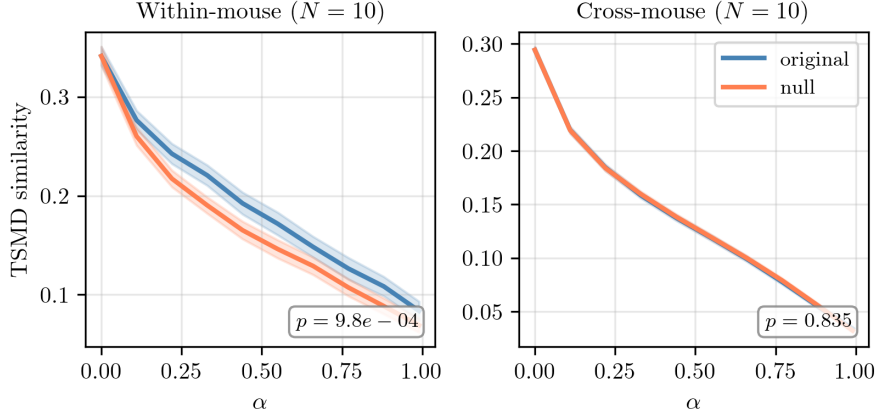


Figure 6: **Natural Scene Stimuli Introduce a Retinotopic Confound in Mouse V1.**

Left: Within-mouse split-half comparisons show significant separation between original (blue) and null (orange) conditions ($p < 0.001$), driven by receptive field overlap rather than feature topography. **Right:** Cross-mouse comparisons show no separation ($p = 0.835$), confirming the within-mouse signal does not reflect conserved topographic organization. Shaded regions indicate ± 1 standard error.

Figure 6 confirms this empirically. Within-mouse comparisons yield a significant separation between original and null curves ($p < 0.001$), suggesting spatial-tuning structure. However, cross-mouse comparisons, where different animals have non-overlapping fields of view, and therefore independent retinotopic sampling shows no significant separation ($p = 0.835$). This indicates that the within-mouse signal reflects shared receptive field position rather than conserved topographic organization of feature preferences.

Drifting gratings avoid this confound entirely. Because they are full-field stimuli, every neuron “sees” identical stimulus content regardless of receptive field position. Any spatial-tuning structure detected by TSMD under drifting gratings must therefore reflect local clustering of feature tuning. Put differently, this is looking for columnar organization rather than retinotopy. Though grating responses yield lower-dimensional tuning vectors, this trade-off is necessary to study our desired question.

Appendix E. Within-Monkey Topographic Organization

Complementing the cross-monkey comparisons in Fig. 2, we also compute the within-monkey split-half comparisons for monkeys N and F across V1, V4, and IT. For each region, we split the electrodes into two disjoint halves, and compute TSMD between them. This is compared against a permutation null, which we had described earlier. We observe that original scores lie above the entire α continuum in all regions for both monkeys ($p = 3.1e-05$) throughout

Extended Abstract Track

(Fig. 7). This indicates that nearby electrodes share tuning to natural images within the recorded regions of the visual stream, establishing a notion of *local* topographic organization.

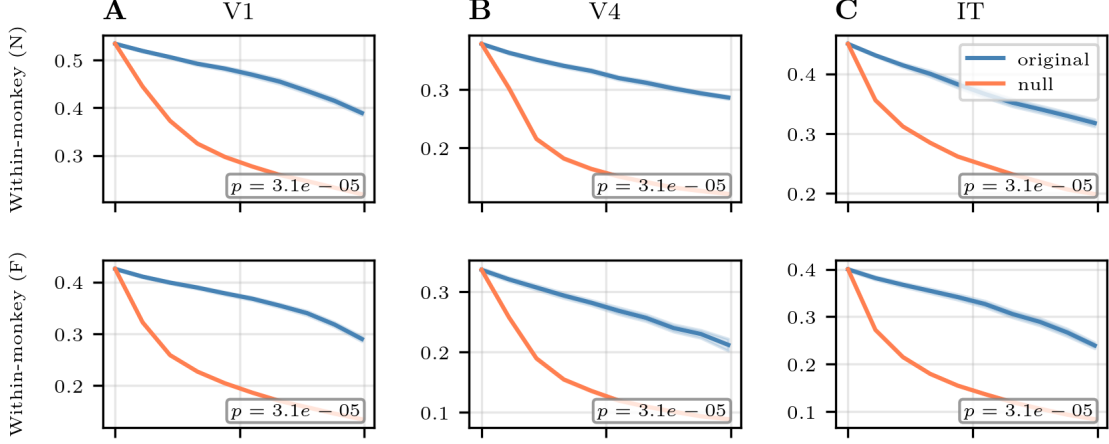


Figure 7: **Within-Monkey Topographic Structure in V1, V4, and IT.** Split-half TSMD comparisons for monkeys N (top) and F (bottom), for V1 (A), V4 (B), and IT (C). Original (blue) scores lie above the permutation null (orange) for all $\alpha > 0$, indicating topographic organization within every region for both monkeys.

Appendix F. Electrode Placement in Ventral Spiking Data

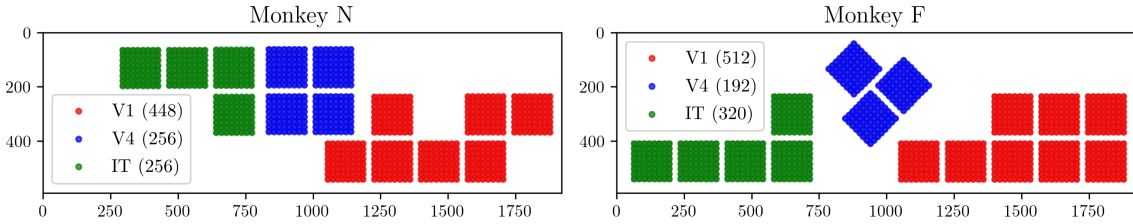


Figure 8: **Utah Array Placement in Monkeys.** We visualize the coordinates associated with every electrode in the Utah arrays implanted in the Macaques. Note that one of the arrays is discarded from the measurements in Monkey N (V1).

Appendix G. Invariance of the Spatial Loss

The spatial component of the TSMD objective is the Gromov-Wasserstein (GW) loss, given by:

$$\mathcal{L}_{\text{GW}}(T; D^R, D^S) = \sum_{i,j,k,l} |D_{ik}^R - D_{jl}^S|^2 T_{ij} T_{kl}, \quad (1)$$

Extended Abstract Track

where $\mathbf{D}_{ik}^R = \|\mathbf{r}_i - \mathbf{r}_k\|^2$ and $\mathbf{D}_{jl}^S = \|\mathbf{s}_j - \mathbf{s}_l\|^2$ are matrices of pairwise squared Euclidean distances among source coordinates $\{\mathbf{r}_i\}_{i=1}^n$ and target coordinates $\{\mathbf{s}_j\}_{j=1}^m$, and $\mathbf{T}$ ranges over the set of admissible couplings $\Pi = \{\mathbf{T} \geq 0 : \mathbf{T}\mathbf{1} = \mathbf{p}, \mathbf{T}^\top \mathbf{1} = \mathbf{q}\}$.

In our setting the coordinate grid is a fixed property of the recording (*e.g.*, electrode positions, imaging field of view) and is identical for the two populations being compared. We therefore analyze invariance with the grid held fixed, perturbing only the assignment of tuning to locations. Concretely, both populations share the same coordinates $\{\mathbf{r}_i\}$, so $\mathbf{D}^S = \mathbf{D}^R =: \mathbf{D}$, and a transformation is a permutation σ that relabels which tuning vector sits at each location. We keep two statements distinct: **(i)** an exact invariance of the loss whenever the relabeling preserves the grid’s distance matrix, which follows directly from the form of (1); and **(ii)** a characterization of which transformations have this property.

Proposition 1 (Automorphism invariance) *Let the source and target share a fixed coordinate grid $\{\mathbf{r}_i\}_{i=1}^n$, with distance matrix $\mathbf{D}_{ik} = \|\mathbf{r}_i - \mathbf{r}_k\|^2$, and let the target tuning be obtained from the source by a permutation σ of locations (the cell at $\mathbf{r}_i$ receives the tuning previously at $\mathbf{r}_{\sigma^{-1}(i)}$). If σ is an automorphism of the grid, i.e.*

$$\mathbf{D}_{\sigma(i)\sigma(k)} = \mathbf{D}_{ik} \quad \text{for all } i, k, \quad (2)$$

then the permutation coupling $\mathbf{T}_{ij}^\sigma = \frac{1}{n} \delta_{j,\sigma(i)}$ achieves zero spatial cost, and the tuning of paired cells coincides. The GW loss is therefore unchanged from the unpermuted comparison, and TSMD similarity is unity.

Proof Take the coupling $\mathbf{T}_{ij}^\sigma = \frac{1}{n} \delta_{j,\sigma(i)}$, which is admissible for uniform marginals $\mathbf{p} = \mathbf{q} = \frac{1}{n}\mathbf{1}$. Substituting into (1) with $\mathbf{D}^R = \mathbf{D}^S = \mathbf{D}$,

$$\mathcal{L}_{\text{GW}}(\mathbf{T}^\sigma; \mathbf{D}, \mathbf{D}) = \frac{1}{n^2} \sum_{i,k} |\mathbf{D}_{ik} - \mathbf{D}_{\sigma(i)\sigma(k)}|^2 = 0,$$

where the final equality is the automorphism condition (2). Hence $\mathbf{T}^\sigma$ achieves minimal spatial cost. Moreover, by construction, the cell at target location $\mathbf{r}_{\sigma(i)}$ holds the tuning previously at $\mathbf{r}_{\sigma^{-1}(\sigma(i))} = \mathbf{r}_i$, so $\mathbf{T}^\sigma$ pairs each source cell with a target cell carrying identical tuning, and the tuning cost vanishes. Both terms of the fused objective are therefore minimized simultaneously, so the comparison is indistinguishable from the unpermuted one. $\blacksquare$

Grid Automorphisms. The permutations σ satisfying (2) are exactly those induced by Euclidean isometries that map the grid onto itself. For a regular grid under periodic (wrap) boundary conditions these comprise the lattice translations together with the point-symmetry group of the grid, i.e.: rotations by multiples of 90° and reflections about its axes and diagonals. Each such isometry $\phi(\mathbf{x}) = \mathbf{Q}\mathbf{x} + \mathbf{t}$ with $\mathbf{Q} \in O(2)$ preserves pairwise distances and, when it permutes the grid, induces a σ obeying (2). By Proposition 1, relabelings of the tuning by any of these transformations, realized in our synthetic experiments by rotating, translating, or reflecting the field on the fixed grid, hence leave TSMD similarity equal to unity across the entire α continuum. On the other hand, a random permutation of the tuning is not a grid automorphism: it does not satisfy (2), so no coupling zeroes the spatial cost, and TSMD faithfully distinguishes this transformation.

Extended Abstract Track

Remark. We make two points explicit here. First, the invariance is to grid automorphisms specifically, not to arbitrary continuous transforms: a rotation by an angle that is not a grid symmetry, or a sub-pixel translation, does not induce a permutation satisfying (2) and is not expected to be *exactly* invariant². Our synthetic perturbations use exact grid symmetries (90° rotations, axis reflections, integer translations under wrap boundaries) for this reason. Second, the reflection case is included because the relevant point group is generated by $Q \in O(2)$ rather than only $SO(2)$: since the loss depends on unsigned squared distances, it cannot distinguish orientation-reversing from orientation-preserving automorphisms, and is blind to the handedness of the map. A geometric mirror that flips the spatial layout of the tuning is thus an exact invariance, distinct from any operation that alters the tuning values themselves.

2. However, it is worth noting that this is a property of discretely sampled grids. In the continuous limit, a cortical sheet modeled as $\mathbb{R}^2$ with dense sampling, has *any* Euclidean isometry preserve pairwise distances, leaving the FGW objective unchanged.