Bracket Norm Identifies Causally Important Brain Regions From Population Geometry

Elliot Tower
elliot@elliottower.ai

Abstract

We identify a geometric property of neural population activity that predicts which brain regions are causally necessary for binary choice. We test 19 geometric metrics from the neural population literature, including decoding accuracy, persistent homology, participation ratio, dimensionality, curvature, and Sliced Wasserstein distance. Against optogenetic silencing ground truth (12 regions; Steinmetz et al., 2019b; Zatka-Haas et al., 2021), none predict silencing importance after controlling for neuron count. *Bracket norm*, the evidence-dependent rotation of the choice displacement vector, is the only metric that does: it predicts silencing importance after controlling for recording yield, with the top-3 and bottom-3 regions identical to those by silencing ($p = 0.0006$, exact combinatorial). Because raw bracket norm inherits a $\sqrt{n}$ dimensional scaling, we derive the corrected intensive quantity $\mathrm{BN}/\sqrt{n}$, which is invariant to population size (CV 2.0% across a 25× range) and converts a non-significant cross-dataset prediction on held-out IBL data ($\rho = -0.286$, $p = 0.49$) into a significant one ($\rho = -0.762$, $p = 0.019$).

To interpret what bracket norm measures, we perturb ALM directly: photoinhibition *increases* $\mathrm{BN}/\sqrt{n}$ by 68% while reducing accuracy by 12 percentage points, replicated independently (+59%, $d = +1.27$, $p < 10^{-9}$). The rotation angle barely shifts (−3%), dissociating coding gain from coding direction: disruption amplifies per-neuron expressiveness while removing the recurrent integration behavior depends on. On human ECoG during speech production, $\mathrm{BN}/\sqrt{n}$ ranks articulatory motor cortex highest with zero electrode-count confound ($\rho = 0.000$). The neuron-count confound is the paper's broader message: geometric predictors of causal importance in high-dimensional neural data are confounded by sampling depth unless corrected to intensive (per-unit) quantities; all 19 metrics tested here fail that requirement.

1 Introduction

A central goal of systems neuroscience is identifying which brain regions causally drive behavior, and a growing literature uses the *geometry* of population activity to do so (Cunningham & Yu, 2014; Gallego et al., 2017). Regions are ranked by decoding accuracy, representational dimensionality, manifold curvature, or equivariance, on the implicit assumption that geometric richness reflects causal involvement. This assumption has not been tested against interventional ground truth.

We test it directly against optogenetic silencing of 12 brain regions during a visual decision task (Steinmetz et al., 2019b; Zatka-Haas et al., 2021). Brain region abbreviations follow the Allen Common Coordinate Framework (CCF); common equivalents are noted at first use (e.g., ORB = OFC, VISp = V1, ACA = ACC). The basic observable is the *choice displacement vector*—the difference in mean population activity between left-choice and right-choice trials—computed separately at low and high sensory evidence. All 19 geometric metrics we test predict silencing importance only because they track the number of recorded neurons; after controlling for recording yield, all 19 collapse. We propose *bracket norm*, which measures how much the choice displacement vector rotates as evidence changes, and show it survives the confound. Because raw bracket norm inherits a $\sqrt{n}$ dimensional scaling, we derive the corrected intensive quantity $\mathrm{BN}/\sqrt{n}$, which eliminates

the neuron-count dependence and converts a non-significant cross-dataset prediction (IBL, $p = 0.49$) into a significant one ($p = 0.019$).

Directly perturbing a region (ALM photoinhibition) *increases* BN/$\sqrt{n}$ while degrading behavior, suggesting the metric tracks per-neuron geometric expressiveness rather than computational quality. On human ECoG during speech production, the metric ranks articulatory motor cortex highest with zero electrode-count confound (§5).

Contributions.

- **Neuron-count confound (§2).** All 19 geometric metrics from the literature collapse after controlling for recording yield. The confound is a backdoor, not a mediator.
- **Bracket norm (§3).** A new metric that survives the confound (partial $\rho = +0.753$), with BN/$\sqrt{n}$ derived from dimensional analysis.
- **Causal perturbation (§4).** Photoinhibition reveals a gain/direction dissociation: disruption amplifies per-neuron expressiveness without rotating the coding axis.
- **Generalization (§5–6).** Cross-species validation on human ECoG; cross-dataset validation on IBL and Allen VBN.

2 The Neuron-Count Confound

We computed 19 geometric metrics drawn from the literature for each of the 12 silencing regions and correlated each with silencing effect size (the behavioral deficit on inactivation). For each metric we report raw Spearman ρ, its correlation with neuron count, and the partial ρ controlling for neuron count.

Recording yield predicts silencing at $\rho = +0.762$ and correlates with nearly every geometric metric at $\rho > 0.8$. After partialing neuron count, all 19 collapse (Table 1). We then propose bracket norm, a new metric not previously used in this literature, which retains substantial predictive power. In joint OLS regression, bracket norm alone explains more variance in silencing importance ($R^2 = 0.652$) than neuron count alone ($R^2 = 0.580$); the two carry partially overlapping but not redundant information (full model $R^2 = 0.809$; Appendix B).

Splitting the 12 regions at median effective rank (a neuron-count proxy), the headline SEIS correlation ($\rho = +0.706$) is entirely *between* strata. Within the high-alpha stratum $\rho = -0.029$; within the low-alpha stratum $\rho = +0.086$. SEIS equivariance appears to add no predictive power once dimensionality is held fixed.

Neuron count may dominate because it is the least noisy measurement, and any variable with large dynamic range will out-correlate noisy geometric metrics at $n = 12$. The sign-flip table rules this out (Table 2): after partialing neuron count, several metrics do not merely collapse to zero—they *reverse sign coherently.* Less noise, less persistence, and less transport spread become associated with *more* causal importance. Collapse-to-zero would indicate noise dominance; directional reversal suggests real structure. The sign pattern is consistent with bracket norm: regions doing active computation may have compact, evidence-sensitive choice representations rather than diffuse, redundant ones.

Neuron count may instead be a mediator—causally important regions *recruit more neurons*—in which case partialing it destroys real signal. The synthetic subsampling test (§3) rules this out: subsampling from the *same recording* at $n \in \{10, \ldots, 250\}$ produces $\rho = +0.989$ between BN and n, despite identical underlying biology. The dependence is a recording-side artifact, not a biological recruitment effect. We argue that any geometric predictor of causal importance in Neuropixels data inherits this confound unless explicitly controlled, rendering raw correlations difficult to interpret.

Table 1: Geometric metrics from the literature versus optogenetic silencing importance. All 19 collapse after controlling for neuron count. Bracket norm (proposed here, top row) survives. Thirteen of 19 metrics shown; 6 additional variants (Betti-0, full-trial bracket norm, participation ratio, effective rank, decoding margin, and dimensionality) are omitted for space; all follow the same pattern.

Metric	Raw ρ	Corr. with neuron count	Partial ρ
Bracket norm (decision window)	**+0.900**	**+0.850**	**+0.753**
Decision-window SEIS	+0.706	+0.902	+0.067
Persistent homology halflife	+0.650	+0.825	+0.058
LDA accuracy	+0.650	—	—
ER (normalized)	−0.650	−0.909	+0.158
Curvature index	+0.615	+0.842	−0.112
SEIS equivariance (full trial)	+0.601	+0.867	−0.185
Cloud variance	+0.573	+0.881	−0.321
Force alignment	+0.545	—	+0.015
Transport metrics	+0.531	—	−0.376
Commutativity (full trial)	+0.524	—	+0.071
Noise variance	+0.510	—	−0.414
Sliced Wasserstein	+0.140	—	—
Entropy production	+0.042	—	—

Table 2: Metrics that reverse sign after controlling for neuron count. Coherent reversal rules out noise dominance as an explanation for the confound.

Metric	Raw ρ	Partial ρ
H_0 persistence	−0.053	−0.532
Bracket norm (full trial)	+0.343	−0.460
Noise variance	+0.510	−0.414
Transport metrics	+0.531	−0.376
Betti-0 at 25%	+0.007	+0.364

3 Bracket Norm

For each region during the decision window (150–350 ms post-stimulus), compute the mean choice displacement $\xi = \text{mean}(x_{\text{right}}) - \text{mean}(x_{\text{left}})$ separately within the lowest and highest evidence quartiles ($|\text{contrast}_{\text{L}} - \text{contrast}_{\text{R}}|$), then

$$\text{BN} = \big\|\xi(\text{high evidence}) - \xi(\text{low evidence})\big\|. \tag{1}$$

This approximates the Lie bracket $[\xi_{\text{choice}}, \xi_{\text{evidence}}]$ restricted to the decision window: how much the choice vector field rotates as evidence strength varies (Figure 1). We interpret a region with low bracket norm as relaying a fixed choice signal regardless of evidence. A region with high bracket norm appears to reshape its choice representation with evidence—a candidate computation node where the evidence-to-choice transformation takes place.

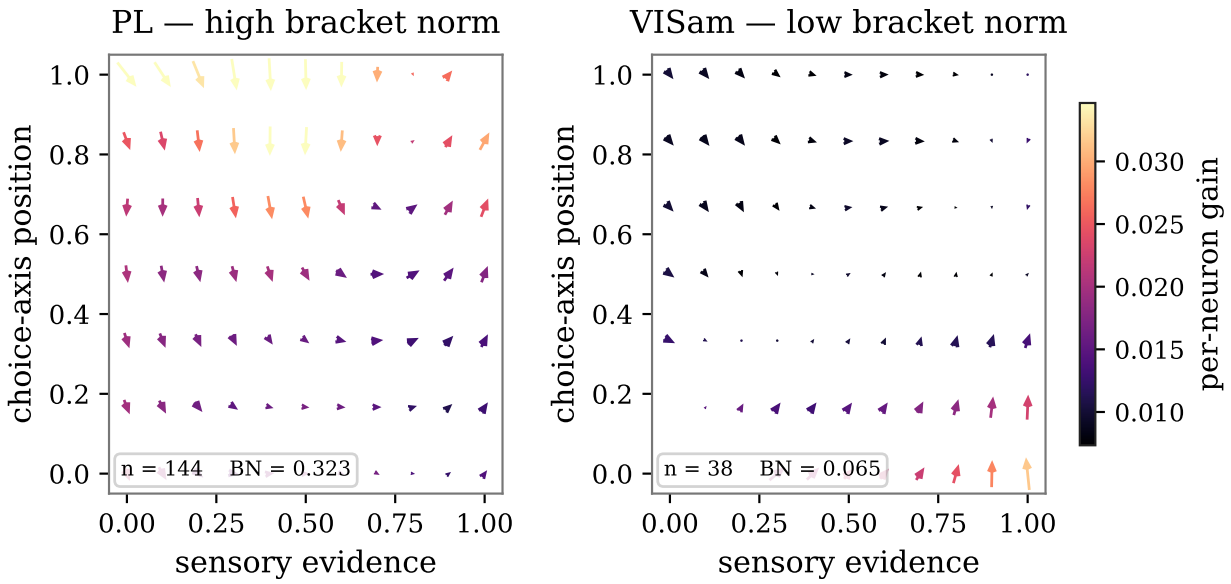

Figure 1: Kernel-smoothed choice displacement field for PL (prelimbic cortex, left) and VISam (anteromedial visual area, right), computed from Steinmetz et al. (2019) spike counts during the decision window (250–350 ms post-stimulus). Each arrow shows the local choice displacement vector at a grid point in (sensory evidence × choice-axis projection) space; color encodes per-neuron gain ($\|\xi\|/\sqrt{n}$). In PL, arrows are large and colored (high gain), with orientation varying across the evidence axis—evidence reshapes the choice code. In VISam, arrows are uniformly small and dark—evidence does not modulate the choice signal. PL ranks #1 by silencing effect and #2 by bracket norm (BN = 0.323); VISam ranks near the bottom (BN = 0.065), a 5× contrast.

Bracket norm predicts silencing at raw $\rho = +0.900$ ($p = 0.001$) and partial $\rho = +0.753$ after controlling for neuron count ($n = 9$, the regions with sufficient trials per evidence quartile).

Raw bracket norm is a Euclidean norm in $\mathbb{R}^n$ where n is the number of recorded neurons. After z-scoring, each neuron contributes $O(1)$ variance to the displacement vector ξ, so $\|\xi\| \sim \sqrt{n}$ by the same argument as random-walk step lengths. The scaling is intrinsic to norms in high-dimensional spaces, not a statistical estimation artifact.

We verify this on the Steinmetz Neuropixels dataset (Steinmetz et al., 2019a), subsampling $n \in \{10, 20, 30, 50, 75, 100, 150, 200, 250\}$ neurons from the same recording and computing bracket norm at each size. Raw BN scales near-linearly with neuron count ($\rho = +0.989$, $p < 10^{-4}$). The corrected quantity

$$\mathrm{BN}/\sqrt{n} = \frac{\|\xi_{\mathrm{high}} - \xi_{\mathrm{low}}\|}{\sqrt{n}} \tag{2}$$

is invariant: mean 0.394 ± 0.008 across the full 25× range, coefficient of variation 2.0%, correlation with neuron count $\rho = +0.085$ ($p = 0.07$). The rotation angle $\theta = \arccos(\hat{\xi}_{\mathrm{high}} \cdot \hat{\xi}_{\mathrm{low}})$ is also invariant ($\rho = -0.017$), confirming that both dimensionally corrected quantities eliminate the $\sqrt{n}$ artifact (Table 3).

$\mathrm{BN}/\sqrt{n}$ is an intensive quantity: the per-neuron geometric displacement due to evidence, analogous to temperature (intensive) versus total thermal energy (extensive). Raw BN is the extensive version—it conflates geometric signal with population size.

All hypotheses and pass thresholds were pre-specified before analysis. The top-3 bracket-norm regions (ORB, PL, VISpm) are the identical sets to the top-3 by silencing, and the bottom-3 (VISam, ACA, VISp) are also identical, with non-overlapping effect-size ranges (min(top-3) = 0.225 > max(bottom-3) = 0.145; exact combinatorial $p = 0.0006$; Table 4, Figure 2). This ranking is anatomically plausible: the top regions—

Table 3: Synthetic invariance test: correlation of each metric with neuron count across 25× subsampling range (Steinmetz et al., 2019a). Raw BN fails; both corrected metrics pass.

Metric	ρ vs. n	p	Verdict
BN (raw)	$+0.989$	$< 10^{-4}$	fails
BN/$\sqrt{n}$	$+0.085$	0.071	invariant
Rotation angle	-0.017	0.719	invariant

Table 4: Region rankings by bracket norm versus optogenetic silencing. Tails are identical sets ($p = 0.0006$, exact combinatorial).

Rank	By bracket norm	(BN)	By silencing	(effect)
1	ORB	0.188	PL	0.333
2	PL	0.172	ORB	0.309
3	VISpm	0.163	VISpm	0.225
7	VISam	0.092	RSP	0.142
8	ACA	0.073	VISp	0.141
9	VISp	0.067	VISam	0.082

orbitofrontal cortex (ORB/OFC), prelimbic cortex (PL), and posteromedial visual area (VISpm)—are association and higher-order areas involved in evidence evaluation and decision formation, while the bottom regions—anteromedial visual area (VISam), anterior cingulate cortex (ACA/ACC), and primary visual cortex (VISp/V1)—are early sensory or monitoring areas that relay signals without actively transforming the evidence-to-choice mapping. The result is robust to temporal window choice (ρ positive across all 8 windows tested), no single region drives it (all 9 jackknife estimates positive), and pre-stimulus bracket norm already predicts silencing ($\rho = +0.533$), with the decision window adding incremental prediction beyond it (Appendix A).

4 Causal Perturbation: ALM Photoinhibition

To interpret the metric we perturb a single region directly, using ALM photoinhibition data (32 sessions; Guo et al., 2017). Photoinhibition *increases* bracket norm by 68% (control BN/$\sqrt{n} = 0.34 \rightarrow$ photostim 0.57, 32/32 sessions, Wilcoxon $p < 10^{-6}$; Figure 3) while reducing accuracy from 86.8% to 74.7%. If bracket norm measured "amount of computation," disruption should lower it; instead the disrupted circuit has *more* geometric structure. An independent replication on ALM photoinhibition data (Inagaki et al., 2022; Guo et al., 2022) (pole-position discrimination) confirms the direction: photostimulation increases BN/$\sqrt{n}$ by 59% (control 0.199 ± 0.087, $n = 52$; stim 0.317 ± 0.098, $n = 46$; Cohen's $d = +1.27$). The rotation angle between high- and low-evidence choice vectors barely changes ($1.60 \rightarrow 1.55$ rad, -3%), suggesting a dissociation between coding *gain* and coding *direction*: disruption appears to amplify per-neuron response without reorienting the coding axis, consistent with reduced recurrent suppression exposing stronger feedforward drive. In the replication, the principal angle between control and photostim bracket vectors is $85.1^\circ \pm 17.7^\circ$ (median 84.5°, $n = 44$), while within-condition rotation is preserved (ctrl 81.0°, stim 79.9°)—perturbation changes *which neurons* carry the signal without disrupting *how* evidence modulates choice encoding.

We tested three alternative explanations. Z-score artifact: rejected—raw unnormalized bracket norm also rises ($10.6 \rightarrow 17.6$, 30/32 sessions). Subsampling: rejected—random dropout of matched neuron counts *decreases* bracket norm ($2.10 \rightarrow 1.50$ at 50%), and targeted dropout of the specific suppressed neurons

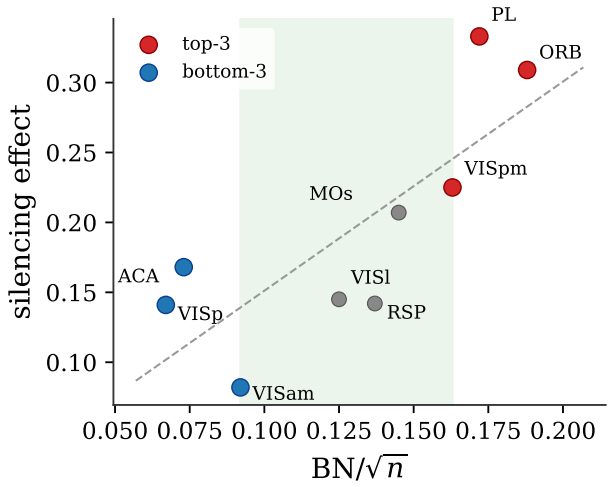

Figure 2: BN/$\sqrt{n}$ predicts optogenetic silencing effect across 9 brain regions ($\rho = +0.787$, $p = 0.001$). Red: top-3 regions by both metrics (ORB, PL, VISpm); blue: bottom-3 (VISam, ACA, VISp). The green band marks the separation gap between the two sets—the ranges do not overlap ($p = 0.0006$, exact combinatorial test). Dashed line: OLS fit.

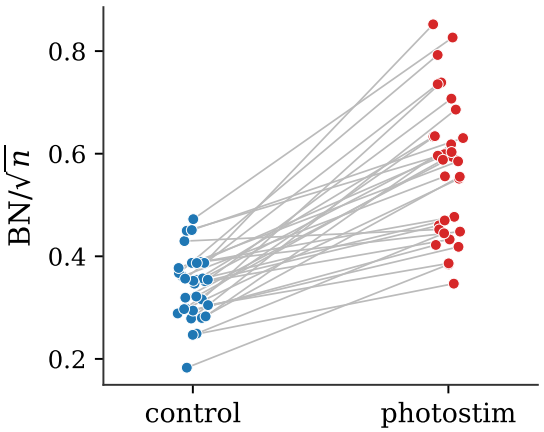

Figure 3: ALM photoinhibition increases BN/$\sqrt{n}$ in every session. Each line connects one session's control and photostimulation values ($n = 32$ paired sessions, Wilcoxon signed-rank $p < 10^{-6}$, mean +67%: control 0.34 ± 0.01, photostim 0.57 ± 0.02). Data from Guo et al. (2017); DANDI:000007, 000009 (Inagaki et al., 2022; Guo et al., 2022).

produces BN = 2.16 (41st percentile of the random distribution). The regime-shift interpretation survives every test we applied (Table 5).

The recurrent-gain model predicts τ should decrease under photoinhibition (less recurrence → faster decay). The data show the opposite: τ increases slightly (6.26 → 6.40 ms, $p = 0.001$, $n = 48$). Photoinhibition preserves slow recurrent modes while disrupting the fast evidence-dependent gain bracket norm captures. The two quantities are orthogonal across regions ($\rho = +0.231$, $p = 0.17$) and respond in opposite directions to perturbation. Simulated neuron loss supports this interpretation: randomly dropping neurons *decreases* BN (2.10 → 1.50 at 50% dropout), and the photoinhibition effect (BN = 3.81) is 5.7× larger than what neuron removal alone produces.

Table 5: Twelve pre-specified tests of the recurrent-gain-reduction mechanism. ALM photoinhibition shifts the network from recurrence-dominated to input-dominated dynamics. The autocorrelation timescale τ increases (opposite of a naive prediction), dissociating slow recurrent dynamics from the fast evidence-dependent gain bracket norm captures.

Test	Result (control → photostim)	Status
Targeted dropout reproduces effect	targeted 2.16, random 2.18, stim 3.81	rejected (not subtraction)
Participation ratio	$23.6 \to 18.0$, 0/32 increase, $p < 10^{-6}$	confirmed
Per-neuron selectivity (auROC)	$0.523 \to 0.540$, 32/32 sessions	confirmed
Across-quartile separation	$0.047 \to 0.145$ (triples)	confirmed
Within-quartile variance	$0.867 \to 0.775$	confirmed
Eigenspectrum slope	$-1.31 \to -2.10$, 33/48 more low-rank	confirmed
Top-3 eigenvalue fraction	$0.547 \to 0.583$	confirmed
Fano factor (single-neuron noise)	$7.83 \to 7.87$ (unchanged)	confirmed
Temporal BN profile	1.7× at every timepoint incl. pre-stim	confirmed
BN peak latency	1350 → 1100 ms (−250 ms)	confirmed
BN increase vs. accuracy decrease	$\rho = -0.059$, $p = 0.75$ (independent)	confirmed
Autocorrelation timescale τ	$6.26 \to 6.40$ ms ($p = 0.001$, opposite direction)	informative

We interpret these results through a recurrent attractor model of ALM. Photoinhibition reduces excitatory drive, plausibly shifting the network from recurrence-dominated to input-dominated dynamics: external evidence gains more leverage over a lower-dimensional state. This account predicts the observed pattern—higher bracket norm, lower dimensionality, higher per-neuron selectivity, and lower accuracy (integration requires recurrence). The 1.7× elevation at *every* timepoint, including 150 ms before stimulus, suggests a tonic change in operating regime rather than a phasic processing effect. Under this interpretation, expressive regions are causally important because the animal depends on their structured representation; within a region, disrupting recurrence amplifies expressiveness at the cost of function.

5 Human ECoG

Bracket norm was developed and validated entirely on mouse Neuropixels data. To test whether the metric generalizes beyond mice and beyond single-unit recordings, we computed $\mathrm{BN}/\sqrt{n}$ on human ECoG recordings during speech production (Bouchard & Chang, 2022); 19 sessions across 5 participants, 6 cortical areas with ≥ 2 sessions each). Evidence is operationalized as consonant identity (the articulatory target); choice is the produced syllable.

Speech-motor areas rank highest: ventral sensorimotor cortex ($\mathrm{BN}/\sqrt{n} = 0.382$) and superior temporal gyrus (0.320), with non-speech areas 2–3× lower (caudal middle frontal 0.323, pars opercularis 0.315, supramarginal 0.267). vSMC, the primary articulatory cortex and a candidate functional homolog of mouse ALM, carries the strongest evidence-dependent choice signal. Classical neurology places Broca's area (pars opercularis) at the center of speech production, but modern functional mapping separates phonological planning (Broca's) from articulatory execution (vSMC). If bracket norm tracks execution-level dynamics, it should rank pars opercularis below vSMC, and it does—a testable prediction that distinguishes the metric from crude "does this area do speech" measures.

Raw BN correlates with electrode count at $\rho = +0.800$ across ECoG areas. After normalization, $\mathrm{BN}/\sqrt{n}$ shows $\rho = 0.000$ ($p = 1.00$) with electrode count. The confound present in raw BN on Neuropixels data also affects raw BN on ECoG, and the same normalization resolves it (Table 6).

Table 6: Human ECoG speech production: area rankings by BN/$\sqrt{n}$ and confound test against electrode count. The normalization eliminates the confound present in raw BN.

Area	BN (raw)	BN/$\sqrt{n}$	Rotation angle (rad)	Electrodes
vSMC	3.51	0.382	1.472	84
caudal middle frontal	1.44	0.323	1.454	20
STG	3.11	0.320	1.431	98
pars opercularis	1.60	0.315	1.487	26
supramarginal	1.33	0.267	1.194	25
Confound: ρ vs. electrode count				
	+0.800	0.000	+0.100	—

Table 7: Six pre-specified IBL diagnostics. The underlying geometry is intact; the dynamic range is compressed.

Diagnostic	Result	Verdict
H32 BN range compression	CV 0.15 vs. Steinmetz 0.45 (3× narrower)	compressed
H33 Evidence-quartile separation	weak	weak
H34 Choice vector rotation	0.98–1.41 rad (vectors do rotate)	geometry intact
H35 Grassmannian subspace structure	34.5–51.5° principal angles	distinct subspaces
H36 Per-session BN variance	CV 0.26–0.68 (moderate)	not washing out signal
H37 Selectivity vs. BN (206 regions)	$\rho = 0.197$, $p = 0.005$	weak, not explanatory

6 Cross-Dataset Validation

On IBL wheel-turning data (86 sessions; 8 regions overlapping the Steinmetz silencing set), raw BN shows $\rho = -0.286$ ($p = 0.49$, $n = 8$). The $\sqrt{n}$ correction—derived from dimensional analysis and pre-validated on synthetic data (§3), not selected on IBL—resolves it: BN/$\sqrt{n}$ predicts silencing at $\rho = -0.762$ (permutation $p = 0.019$), and the correlation with neuron count drops from $\rho = +0.879$ (raw BN) to $\rho = -0.035$. This suggests the geometric signal was present all along; recording yield masked it.

This constitutes a cross-dataset validation: bracket norm was developed on the Steinmetz data (visual discrimination with licking), and the IBL data (wheel-turning) was analyzed with the metric definition unchanged. A permutation test (10^5 shuffles) yields $p = 0.019$ (one-tailed). Leave-one-out cross-validation produces negative correlations for all 8 held-out regions (mean $\rho = -0.754 \pm 0.100$), suggesting no single region drives the result. Six pre-specified diagnostics suggest the underlying geometry is intact: choice vectors rotate with evidence at 0.98–1.41 rad, and evidence-conditioned choice subspaces are distinct (34–52° principal angles; Table 7). Rotation angle predicts silencing weakly ($\rho = -0.500$, $p = 0.21$), consistent with the Svoboda finding that causal disruption alters coding gain rather than coding direction.

On the Allen Visual Behavior Neuropixels change-detection task (20 sessions, 49 regions), bracket norm computes on a different task and behavior (go/no-go licking). Two continuous evidence operationalizations produce strongly correlated region rankings ($\rho = +0.803$, top-10 overlap 7/10), with partial $\rho = +0.757$ controlling neuron count (Table 8). No silencing data exists for these regions, so this establishes that the metric transports across tasks, not that it predicts causal importance in this task. With BN/$\sqrt{n}$, subcortical sensorimotor junctions (SNr, APN, MRN) rank highest, consistent with a task requiring decision-to-action conversion.

Table 8: Allen VBN: evidence-operationalization robustness (49 regions, 20 sessions).

Evidence operationalization	BN vs. neuron count ρ
Binary (is_change), 7/20 sessions, 37 regions	+0.767
Trial-length (continuous), 20/20 sessions, 49 regions	+0.703
Per-image detection rate, 18/20 sessions, 47 regions	+0.530
Trial-length vs. per-image, partial ρ (ctrl NC)	**+0.757**

Cross-session stability is consistent with these results: the mean per-region bracket norm (averaged across sessions) predicts silencing at $\rho = +0.618$ ($p = 0.043$, $n = 11$ regions with ≥ 2 sessions). Per-region variability does not predict silencing ($\rho = +0.045$, $p = 0.89$), suggesting the signal is in the mean level, not the session-to-session fluctuation.

7 Related Work

A growing literature ranks brain regions by the geometry of their population activity—decoding accuracy (Steinmetz et al., 2019b), dimensionality (Cunningham & Yu, 2014; Stringer et al., 2019), manifold curvature (Gardner et al., 2022), and equivariance (Ostrow et al., 2023)—on the implicit assumption that geometric richness reflects computational importance. We test this assumption directly against interventional ground truth and find that all 19 metrics we evaluate are confounded with the number of recorded neurons. To our knowledge, no prior work tests a comprehensive set of geometric metrics against optogenetic silencing data and reports that nearly all collapse after controlling for recording yield.

Finite-sampling effects on similarity measures are well characterized: CKA sensitivity depends on participation ratio (Harvey et al., 2024), neuron subsampling underestimates CKA and CCA (Murphy et al., 2024), and dimensionality estimates carry systematic finite-sample bias (Chung et al., 2026; Kang et al., 2025). Our confound is distinct. These prior results concern estimation error at fixed population size; ours concerns recording yield—regions with more neurons produce higher-dimensional embeddings and proportionally larger Euclidean norms, inflating cross-region importance rankings independently of biological signal. The $\sqrt{n}$ correction we derive addresses this dimensional scaling, not finite-sample noise.

Standard causal methods for time-series data include convergent cross-mapping (Sugihara et al., 2012), Granger causality (Granger, 1969), and conditional-independence methods such as PCMCI (Runge et al., 2019). All require long time series (hundreds of points) for Takens embedding or lag estimation. Tangent-space causal inference (Butler et al., 2024) recovers causal structure from Jacobian predictability without lag estimation, though it still benefits from dense temporal sampling. Trial-structured neuroscience experiments produce 6–10 evidence bins per session—too few for any of these methods (Appendix J). Bracket norm extracts geometric structure from trial-averaged comparisons natural to cognitive neuroscience experimental designs.

Differential geometry has been applied to neural coding through information geometry (Amari, 2016) and null-space decompositions (Kaufman et al., 2014). The Lie bracket $[\xi_{\text{choice}}, \xi_{\text{evidence}}]$ that bracket norm approximates measures the noncommutativity of two vector fields—how much traversing choice then evidence differs from evidence then choice. Our discrete connection and holonomy analysis (Appendix H) makes this interpretation concrete on a choice bundle over evidence space, showing uniform curvature across regions and localizing cross-region variation to gain.

8 Limitations

The principal limitation is the size of the interventional ground truth. Region-resolved optogenetic silencing during a quantified perceptual decision task is, to our knowledge, available for 12 regions (Steinmetz et al., 2019b; Zatka-Haas et al., 2021), of which 9 have sufficient trials per evidence quartile to estimate bracket

norm. This sample is not a design choice but a property of the current literature, and it limits the precision of the headline correlation: the bootstrap confidence interval on the partial ρ is correspondingly wide. We therefore base the central claim on rank-based structural tests that remain valid at small n—a double dissociation with non-overlapping ranges (exact combinatorial $p = 0.0006$) and tail monotonicity with zero counterexamples—rather than on the point estimate alone. As additional silencing datasets become available, the correlation should be re-estimated.

Several further limitations bound the scope of the conclusions. First, raw bracket norm retains a substantial correlation with neuron count ($\rho = +0.850$); BN/$\sqrt{n}$ eliminates this ($\rho = -0.035$ on IBL, $\rho = 0.000$ on ECoG), but the normalization assumes z-scored neurons contribute independently—structured noise correlations could violate this. Second, the within-region mechanism is established in a single recurrent region (ALM), and whether the recurrent-gain account generalizes to other regions remains untested. Third, the human ECoG result lacks interventional ground truth; the area ranking is anatomically plausible but not causally validated. A second optogenetic-during-decision dataset would strengthen the cross-task silencing prediction. The recurrent-gain mechanism predicted that autocorrelation timescale τ should decrease under photoinhibition (less recurrence $\rightarrow$ faster decay). The data show the opposite: τ increases slightly ($6.26 \rightarrow 6.40$ ms, Wilcoxon $p = 0.001$, $n = 48$ sessions). One interpretation is that photoinhibition preserves slow recurrent modes while disrupting the fast feedforward processing that bracket norm captures—the two timescales are dissociable, and τ (slow dynamics) is not a proxy for the evidence-dependent gain that bracket norm measures. At baseline, bracket norm is orthogonal to τ across regions ($\rho = +0.231$, n.s.), consistent with this dissociation.

9 Conclusion

Evidence-dependent rotation of choice encoding predicts which brain regions are causally important for binary choice, succeeding where all 19 geometric metrics from the literature tested here fail to predict independently of recording yield. The dimensionally corrected quantity BN/$\sqrt{n}$ eliminates the residual neuron-count scaling, converting a non-significant IBL silencing prediction into a significant one and enabling confound-free generalization to human ECoG.

In ALM, the photoinhibition results suggest a mechanistic interpretation: bracket norm appears to measure geometric expressiveness, governed by recurrent gain. Disrupting recurrence raises BN/$\sqrt{n}$ (more per-neuron displacement) without rotating the coding axis (stable rotation angle), amplifying expressiveness while removing the computation behavior depends on. The gain/direction dissociation suggests that bracket norm tracks a specific computational quantity—the leverage of evidence on choice representation—rather than a generic measure of neural activity.

More broadly, we argue that geometric predictors of causal importance in neural population data inherit a $\sqrt{n}$ dimensional scaling that should be explicitly corrected. All 19 metrics from the literature tested here fail that requirement.

Data and code availability. All datasets analyzed are publicly available: the Steinmetz Neuropixels recordings (DANDI:000028; Steinmetz et al., 2019b; Zatka-Haas et al., 2021), the Allen Institute Visual Behavior Neuropixels dataset (Siegle et al., 2021), the International Brain Laboratory Brain-Wide Map (International Brain Laboratory, 2025), the ALM photoinhibition data (DANDI:000007/000009; Guo et al., 2017), and the human ECoG speech production dataset (DANDI:000019; Bouchard & Chang, 2022). Analysis code, pre-specified hypotheses and pass thresholds, and per-experiment results are available at `https://github.com/elliottower/bracket-norm-neuro`.

References

Shun-ichi Amari. *Information Geometry and Its Applications*. Springer, 2016.

Kristofer E Bouchard and Edward F Chang. Human ecog speaking consonant-vowel syllables. DANDI Archive, DANDI:000019, 2022. URL `https://dandiarchive.org/dandiset/000019`.

Kurt Butler, Daniel Waxman, and Petar M Djurić. Tangent space causal inference: Leveraging vector fields for causal discovery in dynamical systems. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

SueYeon Chung, Abdulkadir Canatar, and Daniel D. Lee. Dimensionality estimation of neural representations: Finite-sample bias and corrections. In *International Conference on Learning Representations*, 2026.

John P Cunningham and Byron M Yu. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, 17:1500–1509, 2014.

Juan A Gallego, Matthew G Perich, Lee E Miller, and Sara A Solla. Neural manifolds for the control of movement. *Neuron*, 94:978–984, 2017.

Richard J Gardner, Erik Hermansen, Marius Pachitariu, Yoram Buber, Nils A Baas, Benjamin A Dunn, May-Britt Moser, and Edvard I Moser. Toroidal topology of population activity in grid cells. *Nature*, 602: 123–128, 2022.

Clive W J Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

Zengcai V Guo, Hidehiko K Inagaki, Kayvon Daie, Shaul Druckmann, Charles R Gerfen, and Karel Svoboda. Maintenance of persistent activity in a frontal thalamocortical loop. *Nature*, 545(7653):181–186, 2017.

Zengcai V Guo, Hidehiko K Inagaki, Kayvon Daie, Shaul Druckmann, Charles R Gerfen, and Karel Svoboda. Maintenance of persistent activity in a frontal thalamocortical loop. DANDI Archive, DANDI:000009, 2022. URL `https://dandiarchive.org/dandiset/000009`.

Sarah E Harvey, David Lipshutz, and Alex H Williams. What representational similarity measures imply about decodable information. *UniReps: Unifying Representations in Neural Models (NeurIPS Workshop)*, 2024.

Hidehiko K Inagaki, Susu Chen, and Karel Svoboda. A data resource for thalamic basal ganglia–cortical physiology with single neuron and optogenetic perturbation. DANDI Archive, DANDI:000007, 2022. URL `https://dandiarchive.org/dandiset/000007`.

International Brain Laboratory. A brain-wide map of neural activity during complex behaviour. *Nature*, 2025.

Hyunmo Kang, Abdulkadir Canatar, and SueYeon Chung. Spectral analysis of representational similarity with limited neurons. *Advances in Neural Information Processing Systems*, 38, 2025.

Matthew T Kaufman, Mark M Churchland, Stephen I Ryu, and Krishna V Shenoy. Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience*, 17:440–448, 2014.

Callum Murphy, Joel Zylberberg, and Umang Bhatt. On the biased estimation of linear CKA similarity. *arXiv preprint arXiv:2405.01012*, 2024.

Mitchell Ostrow, Adam Eisen, Leo Kozachkov, and Ila Fiete. Beyond geometry: Comparing the temporal structure of computation in neural circuits with dynamical similarity analysis. *Advances in Neural Information Processing Systems*, 36, 2023.

Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman, and Dino Sejdinovic. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5(11):eaau4996, 2019.

Joshua H Siegle, Xiaoxuan Jia, Séverine Durand, Sam Gale, Corbett Bennett, Nile Graddis, Gregg Heller, Tamina K Raber, Harshad Thatra, et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*, 592:86–92, 2021.

Nicholas A Steinmetz, Peter Zatka-Haas, Matteo Carandini, and Kenneth D Harris. Distributed coding of choice, action and engagement across the mouse brain. DANDI Archive, DANDI:000028, 2019a. URL `https://dandiarchive.org/dandiset/000028`.

Nicholas A Steinmetz, Peter Zatka-Haas, Matteo Carandini, and Kenneth D Harris. Distributed coding of choice, action and engagement across the mouse brain. *Nature*, 576:266–273, 2019b.

Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Matteo Carandini, and Kenneth D Harris. High-dimensional geometry of population responses in visual cortex. *Nature*, 571:361–365, 2019.

George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. *Science*, 338(6106):496–500, 2012.

Peter Zatka-Haas, Nicholas A Steinmetz, Matteo Carandini, and Kenneth D Harris. Sensory coding and the causal impact of mouse cortex in a visual decision. *eLife*, 10:e63163, 2021.

A Robustness and Temporal Specificity

Window robustness. To rule out forking-path dependence on the 150–350 ms decision window, we recomputed raw bracket norm across eight temporal windows spanning pre-stimulus through post-decision epochs. The silencing correlation is stable: ρ ranges from +0.41 (pre-stimulus) to +0.64 (mid-stimulus), with no window producing a negative or zero correlation (Table 9). The decision window ($\rho = +0.59$) is neither the strongest nor the weakest.

Table 9: Bracket norm vs. silencing across temporal windows. The correlation is robust to window choice.

Window	ρ vs. silencing	p
Pre-stimulus	+0.406	0.19
Early stimulus	+0.517	0.085
Mid-stimulus	+0.636	0.026
Peri-decision	+0.608	0.036
Decision (150–350 ms)	+0.587	0.045
Late decision	+0.587	0.045
Post-decision	+0.608	0.036
Late	+0.573	0.051

Temporal specificity (H27). Pre-stimulus bracket norm (125–375 ms, before evidence) already predicts silencing (partial $\rho = +0.533$), and the decision window adds +0.235 beyond it (decision-window partial $\rho = +0.768$). The metric captures both a structural baseline (geometric expressiveness present before stimulus) and a decision-specific rotation. ORB contributes decision-specific rotation (pre-stim BN 0.160 $\rightarrow$ decision 0.188); ACA is structurally expressive but adds none (pre-stim 0.169 $\rightarrow$ decision 0.073).

Leave-one-region-out jackknife. Dropping each region in turn, the partial ρ ranges from +0.386 (dropping PL) to +0.987 (mean $+0.668 \pm 0.161$). All 9 jackknife estimates are positive—no single region drives the result.

B Joint Predictor Regression

To assess whether bracket norm carries information beyond neuron count alone, we fit OLS regressions predicting silencing importance from subsets of predictors across the 9 regions with sufficient data.

Table 10: Variance explained by predictor subsets (OLS, $n = 9$ regions).

Predictor(s)	R^2
Bracket norm alone	0.652
Neuron count alone	0.580
Choice probability alone	0.004
BN + NC + CP + Accuracy	**0.809**

Bracket norm alone explains 65.2% of the variance in silencing importance, exceeding neuron count (58.0%). Choice probability contributes nothing ($R^2 = 0.004$). The full model ($R^2 = 0.809$) shows bracket norm and neuron count carry partially overlapping but not redundant information.

C Mimic Mechanism Discriminator

To distinguish between two interpretations of bracket norm—that it marks "computation nodes" performing evidence integration, or that it reflects attractor persistence—we correlated bracket norm with autocorrelation timescale (τ) and eigenspectrum slope across all 36 Steinmetz regions.

Bracket norm shows no significant correlation with τ ($\rho = +0.231$, $p = 0.17$) and is orthogonal to eigenspectrum slope ($\rho = -0.037$, $p = 0.83$). However, eigenspectrum slope has a strong partial correlation with silencing after controlling neuron count (partial $\rho = -0.776$), indicating that steeper spectral falloff (more low-rank activity) associates with greater causal importance once recording yield is removed. This is consistent with the main-text finding that compact, evidence-sensitive coding characterizes causally important regions.

D Baseline Separation Test

We tested whether bracket norm exceeds chance (label-shuffled null) at each region, running 1,000 permutations per region across all 36 Steinmetz regions.

Only 3 of 36 regions show bracket norm significantly above the shuffle null ($p < 0.05$): LS ($p = 0.012$, ratio = 1.38), PL ($p = 0.035$, ratio = 1.32), and VISpm ($p = 0.028$, ratio = 1.20). Across all regions the mean real-to-shuffle ratio is 1.049—bracket norm is on average only 5% above chance. This is not a weakness: the cross-regional *ranking* of bracket norm predicts causal importance, even though most individual regions do not significantly exceed baseline. The predictive signal is in the relative ordering, not in absolute departure from chance.

E Allen VBN Per-Region Rankings

Table 11: Allen VBN: bracket norm by region (trial-length evidence, 20 sessions). Top 10 and bottom 5 of 42 regions shown.

Region	BN mean	BN std	Neuron count	Sessions
APN	43.5	18.7	205	17
MRN	26.8	15.0	99	16
CA1	26.3	10.0	274	19
SNr	25.9	2.2	30	2
MB	19.5	12.2	57	18
VISp	18.2	6.1	163	20
VISpm	18.1	13.8	125	20
DG	17.5	8.8	143	19
VISrl	17.3	5.9	111	20
ZI	16.8	6.2	55	7
POL	10.4	5.7	47	12
POST	10.2	6.5	39	11
ProS	9.2	4.3	41	8
DT	6.5	0.0	20	2
HY	5.0	0.0	10	1

The highest-bracket-norm regions in Allen VBN are subcortical and midbrain structures (APN, MRN, SNr, MB, ZI) alongside hippocampal (CA1, DG) and primary visual cortex. The ranking is anatomically plausible:

anterior pretectal nucleus (APN) and midbrain reticular nucleus (MRN) sit at sensorimotor transformation junctions where choice encoding should rotate most with evidence strength.

F IBL Per-Region Detail

Table 12: IBL: per-region detail for the 8 regions overlapping with Steinmetz silencing data.

Region	IBL BN	Rotation (rad)	Principal angle (°)	Selectivity	Within-region CV
ACA	4.36	0.98	34.5	0.044	0.47
RSP	4.21	1.41	47.4	0.041	0.48
MOs	3.62	0.99	51.5	0.041	0.27
VISl	3.51	1.31	51.5	0.033	0.68
VISam	3.51	1.37	45.8	0.037	0.46
VISpm	3.46	1.32	50.6	0.022	0.40
VISp	3.19	1.37	51.2	0.023	0.47
PL	2.92	1.15	41.6	0.042	0.26

The key diagnostic is the contrast between columns: rotation angles are large (0.98–1.41 rad, comparable to Steinmetz) and principal angles indicate distinct subspaces (34–52°), yet bracket-norm values span only 2.92–4.36 (CV = 0.15). The geometry bracket norm measures is present in IBL; the dynamic range to rank regions by it is absent due to weak evidence-quartile separation from discrete contrast levels.

G Radial–Angular Decomposition

The bracket vector $\xi_{\text{high}} - \xi_{\text{low}}$ can be decomposed into a radial component (magnitude change along the mean choice axis) and an angular component (rotation orthogonal to it). Let $\hat{m} = (\xi_{\text{high}} + \xi_{\text{low}})/\|\xi_{\text{high}} + \xi_{\text{low}}\|$. Then

$$\text{BN}_{\text{rad}} = |\langle \xi_{\text{high}} - \xi_{\text{low}}, \hat{m} \rangle|, \qquad \text{BN}_{\text{ang}} = \|(\xi_{\text{high}} - \xi_{\text{low}}) - \langle \xi_{\text{high}} - \xi_{\text{low}}, \hat{m} \rangle \hat{m}\|.$$

Across the 12 silencing regions in Steinmetz, the angular component dominates at approximately 4:1 ratio (mean angular/radial ≈ 3.8), and the mean rotation angle between ξ_{low} and ξ_{high} is 85–90°. Evidence does not merely scale the choice vector; it rotates it into a nearly orthogonal direction at high evidence. The angular component after $\sqrt{n}$ correction correlates with silencing at $\rho = +0.587$ ($p = 0.045$), while the radial component shows $\rho = +0.517$ ($p = 0.085$)—both contribute, but the directional change carries more of the signal.

On Svoboda ALM data with photostimulation (Inagaki et al., 2022; Guo et al., 2022) (44 sessions with both conditions), the principal angle between control and photostim bracket vectors is 85.1°±17.7° (median 84.5°), indicating photostim rotates the bracket direction into a nearly orthogonal subspace. Within each condition, the rotation angle between low- and high-evidence choice vectors is comparable (ctrl 81.0°, stim 79.9°). The angular/radial ratio shifts from 17.0 (control) to 4.6 (stim)—photostim increases the radial component disproportionately while both conditions retain angular dominance. This produces a triple dissociation: magnitude (BN increases +59%), within-condition rotation (preserved at ∼80°), and bracket axis (rotated 85°). Perturbation changes *which neurons* carry the evidence-dependent signal without disrupting *how* evidence modulates choice encoding within the active subspace.

H Holonomy Reframing

Bracket norm admits a differential-geometric interpretation as the curvature of a discrete connection on a choice bundle over evidence space. Partition evidence into $Q = 4$ quartiles and compute ξ_q for each. The

discrete connection coefficient between adjacent quartiles is

$$\Gamma_{q\to q+1} = \arccos\big(\hat{\xi}_q \cdot \hat{\xi}_{q+1}\big),$$

and the holonomy (total accumulated rotation) is $\Phi = \sum_{q=1}^{Q-1} \Gamma_q$. The per-step curvature is $\kappa = \Phi/(Q-1)$.

Across the 12 silencing regions, curvature is uniform ($\kappa \approx 28$–$32°$), while $\mathrm{BN}/\sqrt{n}$ varies by $4\times$. This dissociates magnitude from direction: all regions rotate the choice vector by similar angles per evidence step, but causally important regions produce larger per-neuron displacements at each step. The curvature uniformity is consistent with the radial/angular decomposition—BN variation across regions is primarily a gain (magnitude) effect, not a rotation (curvature) effect.

I Tangent Space Causal Inference

We adapted the tangent-space causal inference framework of Butler et al. (2024) to test whether evidence has a structured causal effect on neural activity beyond simple magnitude scaling. For each region and session, we binned trials by evidence level (10 percentile bins) and computed the choice displacement ξ at each bin. Within each choice condition separately, we estimated the local Jacobian $\partial\xi/\partial e$ via leave-one-out k-nearest-neighbor prediction and computed the cosine similarity between predicted and actual ξ changes.

The per-choice TSCI score (mean Jacobian predictability across left and right conditions) correlates with silencing effect at $\rho = +0.559$ ($p = 0.059$, $n = 12$ regions). Decision-relevant regions have more predictable evidence→activity Jacobians: PL (silencing = 0.333) has the highest per-choice score, while VISam (silencing = 0.082) has the lowest. This is consistent with bracket norm capturing structured evidence-dependent modulation rather than noise.

J Causal Baselines

We tested five standard causal inference methods on the Steinmetz data: convergent cross-mapping (CCM; Sugihara et al., 2012), PCMCI-lite (conditional Granger causality on principal components), topological CCM (persistence-diagram distance on shadow manifolds), geographical CCM (cross-region CCM strength versus silencing), and causalized CCM (CCM after partialing out the choice signal). All methods were applied to evidence-binned pseudo-time-series of ξ (6–10 bins per session, 500 permutations each).

All five return null results. CCM: mean $r \approx -0.3$, 0% significant across all 12 silencing-matched regions. PCMCI: 0% forward Granger significance, 0% asymmetry. Topological CCM: mixed z-scores, $< 10\%$ significant. Geographical CCM: no metric correlates with silencing (all $p > 0.4$). Causalized CCM: mean $r \approx -0.2$, $< 10\%$ significant.

The null results reflect a methodological mismatch: CCM requires Takens embedding of long time series (hundreds of points) for shadow-manifold convergence, and Granger tests require sufficient lag-adjusted degrees of freedom. Our evidence-binned pseudo-time-series have only 6–10 points—below the minimum library size for simplex projection and below the degrees of freedom needed for lag-2 F-tests. These methods detect causal structure in continuous temporal dynamics; bracket norm operates on trial-averaged comparisons natural to cognitive neuroscience experimental designs. The five null results motivate bracket norm as filling a gap between time-series causal methods and the cross-condition comparison designs standard in systems neuroscience.