Distinct residual spatial organization of inputs and outputs in human cortical neurons

Xueqiang Xin[1*]

[1]School of Advanced Interdisciplinary Sciences, University of Chinese Academy of Sciences, Beijing, China.

Key Words: human cerebral cortex; connectomics; synaptic organization; neurite morphology; spatial statistics

*Corresponding authors: Xueqiang Xin, email: xqxin2025@163.com. Address: School of Advanced Interdisciplinary Sciences, University of Chinese Academy of Sciences, Beijing, P. R. China.

Abstract

The spatial arrangement of synapses shapes how neurons integrate incoming signals and distribute outputs. Yet observed synaptic maps are strongly constrained by neurite geometry and local cable availability. It remains unclear whether input and output maps retain distinct organization after these morphology-derived opportunities are controlled. We analysed 272,786 cable-mapped events across 104 proofread neurons from H01 human temporal cortex using hierarchical within-neuron null models that preserve geometry, compartment, proximodistal position and local cable length. Relative to N2.5, inputs were more dispersed at 2–5-μm observation scales, whereas outputs were more concentrated at 2–50 μm and showed excess same-branch pairs only below 1 μm. Fixing each branch's observed event count removed the negative input residual at 1–5 μm but retained a positive submicron output residual, separating branch allocation from within-branch organization. Thus, inputs and outputs exhibit distinct

residual spatial statistics beyond tested intrinsic cable opportunity, although these single-volume associations do not establish a causal organizing mechanism.

Introduction

Synaptic position is a central determinant of neuronal communication. The electrotonic impact of an input depends on dendritic location, branch context and coactivation with nearby synapses, whereas the distribution of presynaptic sites along an axon constrains how output is delivered to local targets. Functional imaging and anatomical reconstruction have revealed spatially structured inputs and outputs in several systems [1–6]. These observations show that synaptic position can carry circuit structure, but they do not by themselves separate biological placement from the geometry of the neurites on which synapses are observed.

Morphology creates heterogeneous opportunity. Branching, compartment identity, proximodistal position and local cable density change the number and arrangement of locations available to an observed event. A cluster in Euclidean space can therefore arise because cable is locally dense, and a gradient can follow compartment or root distance. The relevant counterfactual is not uniformity in tissue volume, but redistribution along the same neuron while preserving specified structural opportunity [7–9]. Such conditioning controls intrinsic cable-length opportunity; it does not reconstruct nearby partner membrane or complete contact opportunity.

The H01 human temporal-cortex volume provides synaptic-resolution reconstructions and a set of manually proofread neurons [10]. A polarity comparison in this bounded volume is nevertheless complicated by the large imbalance between input and output counts, repeated relationship rows at presynaptic sites, incomplete attachment of dendritic spines, axonal truncation and uncertainty in mapping synaptic coordinates to cable. These features require neuron-level inference and explicit sensitivity analyses rather than treating individual synapses as independent replicates.

Here we test whether input and output maps retain different residual statistics after conditioning on the geometry of the same proofread neuron. We analyse 104 H01 neurons with nested within-neuron null models. N2.5 preserves compartment, proximodistal bin and local own-neuron cable-length class; it does not model apposed partner membrane or complete contact opportunity. We separate two scales of description: voxel-entropy residuals quantify organization as observation resolution changes, whereas a zero-inclusive same-branch pair-separation residual quantifies excess or depletion at specified separations. We then audit event-count imbalance, projection threshold, opportunity scale, branch occupancy, output representation, missing spines, anisotropic compression and boundary censoring. The primary conclusion is deliberately statistical: inputs and outputs show distinct residual spatial organization after conditioning on intrinsic neurite opportunity, without implying a causal synaptic organizing principle.

Results

Morphology-constrained null models

The analysis population comprised 104 proofread neurons from the H01 human temporal-cortex volume. The source sampling selected neurons whose soma centres lay 23–30 μm from the start of the imaged block in z; this set is therefore conditional on a restricted sampling slab. Continuous cable-segment projection identified 268,536 inputs and 8,607 outputs before distance filtering. A 2-μm projection cutoff retained 264,737 inputs and 8,049 outputs (272,786 events). Primary direction-level inference included all 104 input-eligible neurons and 88 output-eligible neurons, comprising 264,737 input events and 7,837 output events. Neurons, not individual synapses, were the inference units.

Four nested within-neuron null models progressively conditioned the redistribution of each event set (Fig. 1). N0 preserved eligible cable geometry, N1 additionally preserved compartment, and N2 additionally preserved one of ten root-distance bins. N2.5 further preserved local intrinsic cable-length opportunity, defined by cable length in 5-μm voxels and discretized into four weighted quantile bins.

This hierarchy addresses several major morphology-derived opportunities but not membrane surface area, branch order, spine density, nearby partner membrane, axodendritic apposition or subcellular target preference. N2.5 is therefore an intrinsic cable-length null, not a complete contact-opportunity model.

The continuous projection and its generator were restored and rerun (Extended Data Fig. 2). Median projection distance was 0.192 μm for inputs and 0.281 μm for outputs; the 95th percentiles were 0.780 and 2.744 μm, respectively. The 2-μm cutoff retained 98.6% of inputs and 93.5% of outputs. We therefore report strict 0.25-, 0.5-, 1- and 2-μm cutoff sensitivities for both entropy and the zero-inclusive same-branch residual, and distinguish changes in effect estimates from the substantial loss of output-eligible neurons at the strictest cutoff.

Distinct residual spatial statistics

We first measured multiscale entropy residual, defined as Hnull − Hobs. Positive values indicate that observed events are more concentrated than N2.5; negative values indicate greater dispersion. Median input residuals were −0.268 bits at 2 μm (95% bootstrap confidence interval, −0.319 to −0.231), −0.312 bits at 3 μm (−0.359 to −0.271) and remained negative at 5 μm (Fig. 2a). Output residuals had the opposite sign from 2 to 50 μm: 0.151 bits at 2 μm, 0.162 at 3 μm, 0.190 at 5 μm, 0.229 at 10 μm, 0.285 at 25 μm and 0.273 at 50 μm. These values describe residual distributional organization, not attraction or repulsion.

Because inputs outnumbered outputs, we reran count matching under the final N2.5 model. Within each of 88 paired neurons, inputs were subsampled without replacement to that neuron's output count over 500 iterations; observed N2.5 strata and an independent N2.5 null realization were recomputed for every subsample. Each iteration therefore compared 7,837 inputs with 7,837 outputs. The median output-minus-input entropy residual remained positive from 2 to 50 μm (0.166–0.259 bits; all Benjamini–Hochberg-adjusted exact sign-test $q \leq 1.19 \times 10{-}9$). Count imbalance does not explain the

full-compartment entropy polarity. Polarity was assigned from the incident relationship, whereas cable compartment was annotated independently; the primary analysis therefore included mapped events from all compartments. This audit addresses entropy, not the pair-separation statistic. A canonical-compartment restriction to axonal outputs and dendritic inputs reduced the cohort to 36 neurons and attenuated the six-scale medians to 0.009–0.107 bits (all $q \geq 0.272$); independence from compartment composition is therefore not established.

We next quantified pair-separation scale with a zero-inclusive same-branch statistic (Fig. 2b). For each bin, the numerator counted unordered event pairs assigned to the same branch whose absolute root-distance difference fell in that bin; the denominator was all unordered event pairs in the neuron, including cross-branch pairs. The mean N2.5-null rate was subtracted from the observed rate and multiplied by 1,000, retaining neurons with zero observed pairs. Inputs were most negative at 0–1 μm (−0.0785 pairs per 1,000 candidate pairs; −0.0936 to −0.0550; FDR-adjusted $P = 2.23 \times 10{-}9$), remained negative but smaller at 1–2 μm (−0.0238; $q = 5.46 \times 10{-}7$), 2–3 μm (−0.00634; $q = 1.39 \times 10{-}6$) and 3–5 μm (−0.00223; $q = 0.00470$), and were not significant at 5–10 μm (−0.00136; $q = 0.493$). Outputs showed a positive residual only at 0–1 μm (1.978, 1.530 to 2.973; FDR-adjusted $P = 4.15 \times 10{-}14$) and negative residuals from 1 to 10 μm (all $q \leq 0.00228$). Zero inclusion keeps the same 88 output-eligible neurons in every bin. In the fixed 88-neuron paired cohort, the output-minus-input residual was positive at 0–1 μm (median, 2.038; $q = 2.30 \times 10{-}15$) and negative from 1 to 10 μm (all $q \leq 0.001824$), so the distance pattern is not caused by changing cell eligibility. An earlier log-ratio summary excluded zero-pair neurons and therefore overstated output enrichment at 1–3 μm; it is not used for inference here.

Restoration of the continuous projection enabled direct mapping- and opportunity-specification tests (Extended Data Fig. 2). With cutoff-specific eligibility, the output 0–1-μm residual remained positive at projection cutoffs of 0.25, 0.5, 1 and 2 μm (medians, 1.016, 1.758, 1.964 and 1.976 per

1,000; n = 64, 87, 88 and 88; all adjusted $q \leq 0.0112$). In the common 64-neuron cohort, corresponding medians were 1.016, 1.228, 1.198 and 1.454, with all 95% bootstrap intervals above zero. The strictest cutoff retained only 43.1% of output events and 64 eligible output neurons, so persistence is interpreted as cutoff robustness rather than evidence that projection uncertainty is negligible. Varying the opportunity voxel among 2, 3, 5 and 10 μm also preserved all three prespecified signs in the fixed cohort: 3-μm input entropy medians ranged from −0.377 to −0.205 bits, 25-μm output entropy medians were 0.265–0.288 bits and output 0–1-μm pair residuals were 1.438–1.473 per 1,000. The 5-μm model remains the prespecified primary analysis.

Branch-count conditioning decomposed the output result (Fig. 4b and Extended Data Fig. 4). On the same 88 neurons, the median output 0–1-μm residual decreased from 1.976 under N2.5 to 0.515 under N2.5-B (95% confidence interval, 0.341–0.694; adjusted $q = 6.44 \times 10{-}12$). The paired N2.5-B-minus-N2.5 change was −1.448 (−1.918 to −0.974; adjusted $q = 1.20 \times 10{-}12$). Thus, differential branch occupancy explains much of the primary residual, but a positive within-branch submicron component remains after every observed branch count is fixed. This structural component does not imply pairwise attraction or a causal placement mechanism. For inputs, branch conditioning reduced the 0–1-μm residual from −0.0681 to −0.00195 per 1,000. More decisively, the standard negative residuals at 1–2, 2–3 and 3–5 μm (−0.0238, −0.00487 and −0.00221) became +0.00102, +0.00262 and +0.000795 under N2.5-B; the 1–2-μm value was not significant after correction ($q = 0.0693$), and the latter two were small positive values. Thus, the primary 0–5-μm input depletion is principally a branch-allocation effect—lower same-branch co-occupancy—rather than evidence for negative within-branch spacing at 1–3 μm.

The entropy and pair-separation results address different scales of structure. Positive output entropy through 25–50 μm indicates broad distributional heterogeneity, whereas the positive pair residual confines strong local same-branch pair separation to below 1 μm. Conversely, input entropy is

most negative at 2–5 μm while pair depletion extends from 0 to 5 μm. The measures therefore distinguish organization scale from pair-separation scale rather than implying a single optimal spatial bin; neither statistic defines a biological interaction range.

Cell-level polarity and morphology

Eighty-eight neurons had at least 20 events in both directions. Seventy-three of 88 (83%) combined a negative input entropy residual at 3 μm with a positive output residual at 25 μm (Fig. 3a). This neuron-level prevalence shows that the population result is not produced by treating hundreds of thousands of synapses as independent replicates. Related residual formulations showed 96.4% sign concordance and a rank correlation of 0.989. The 83% fraction is descriptive at these selected organization scales and does not imply that every neuron has the same sign at every scale or pair-separation bin.

Morphology was associated with effect magnitude. Morphology PC2 correlated with the output-minus-input entropy contrast at 3 μm (Spearman $\rho = 0.569$, FDR-adjusted $q = 2.67 \times 10{-}7$; $n = 88$; Fig. 3b). A multivariate model using morphology PC1–PC3 explained 44.3% of the in-sample six-scale polarity variation. After median imputation, standardization and PCA were refitted independently within every training fold, predictive R^2 was 0.346 in fixed ten-fold cross-validation, 0.357 across 200 repeated ten-fold partitions (2.5–97.5% partition range, 0.329–0.377) and 0.359 in leave-one-out validation (permutation $P = 0.001$). Descriptive loadings indicate that PC2 (17.4% variance) contrasts axonal cable fraction (−0.580) with dendritic cable fraction (0.523) and also loads on maximum root distance (−0.403); loading signs are arbitrary. Dominant layer and morphology cluster showed exploratory multivariate associations, but sparse, imbalanced strata and unavailable molecular identities preclude a cell-type rule. These analyses support morphology-associated variation, not a causal classification.

Submicron output residual

Output relationship tables can contain multiple rows for one presynaptic site. We therefore repeated the zero-inclusive same-branch pair-separation analysis after collapsing events to unique presynaptic coordinates and after restricting to monadic unique sites (Fig. 4a). Submicron enrichment remained for relationship rows (median, 1.978 excess pairs per 1,000 candidate pairs), unique presynaptic sites (1.067) and monadic unique sites (1.060; all n = 88; FDR-adjusted $P \leq 9.61 \times 10{-}13$). Residuals from 1 to 5 μm were negative under all three definitions. Relationship multiplicity amplifies, but does not create, the submicron output signal.

We then partitioned monadic unique-site pairs by C3 agglomerate identifier (Fig. 4c). The submicron residual was strongly positive for same-C3 pairs (median, 249.875 pairs per 1,000 candidate pairs; 178.968–288.173) but slightly negative for different-C3 pairs (−0.106, −0.224 to −0.042). The paired same-minus-different contrast was 249.439 (178.704–289.821; FDR-adjusted sign-test $P = 1.13 \times 10{-}17$; n = 88). Thus, the output signal is strongly associated with repeated contacts assigned to the same agglomerate. C3 identifiers are automated fragments, not fully proofread biological partners; oversegmentation could split one biological cell across identifiers. This result is consistent with repeated contacts assigned to the same automated agglomerate, not clustering across verified distinct targets.

Input pair depletion under spine loss

Dendritic-spine attachment is incomplete in automated H01 segmentation. The source audit found that 33.7% of 1,550 spines were detached in one layer-2 pyramidal neuron; we used this cell-specific rate as a stress-test parameter, not a population estimate. A 33.7% missing fraction corresponds to adding approximately 50.8% of the observed count, calculated as nobs × 0.337/0.663. Random imputation moved pair ratios towards one but preserved depletion: median observed-to-N2.5 ratios were 0.736 at 0–1 μm, 0.820 at 1–2 μm and 0.879 at 2–3 μm (Fig. 5).

Adversarial gap-fill models tested spatially structured missingness. Preferentially inserting events into 1-µm cable gaps could reverse the 0–1-µm result, whereas the 1–3-µm bins remained below expectation; a 3-µm gap-fill model altered several bins. These are sensitivity bounds rather than biological generative models. The 1–3-µm input depletion is therefore the stronger inference. Submicron input depletion is supported under observed and random-missingness assumptions but remains conditional on the absence of strongly short-range-biased spine detachment.

Robustness to geometric artifacts

The H01 specimen underwent approximately 28% section-axis compression. Recomputing entropy residuals after scaling the compressed axis preserved the negative short-scale input values and positive output values across scales (Extended Data Fig. 1A). Censoring branches near reconstruction boundaries or likely artifact endpoints likewise preserved both signs (Extended Data Fig. 1B). Two unadjusted correlations between output summary measures and boundary proximity did not survive partial-correlation control for log output count and log axonal cable length (partial Spearman $\rho = -0.032$ and 0.059; FDR-adjusted $P = 0.765$ for both).

A representative neuron provides spatial context for the population statistics (Fig. 6). Its mapped events show dense dendritic inputs distributed across the arbor and sparser axonal outputs, including local groups of output sites. The reconstruction is illustrative rather than evidential; all hypothesis tests aggregate neuron-level statistics across the population.

Discussion

This study identifies distinct residual spatial organization of mapped inputs and outputs after conditioning on intrinsic neurite opportunity in 104 proofread neurons from one human temporal-cortex volume. Voxel entropy and the same-branch pair-separation residual provide complementary, not interchangeable, descriptions. The neuron is the inference unit in both analyses, and 73 of 88 paired neurons met the selected 3-µm-input/25-µm-output sign criterion. We therefore interpret the

findings as residual structural associations, not as evidence of synaptic repulsion, attraction or a causal organizing principle.

The positive output entropy residual from 2 to 50 μm and the positive output pair residual only at 0–1 μm are not contradictory. Entropy asks how concentrated a 3D event map appears as voxel resolution changes, whereas the primary pair residual combines same-branch co-occupancy with absolute root-distance separation. A map can be concentrated across spatial voxels while lacking excess same-branch pairs at intermediate separations. The supported local output statement is therefore a positive submicron pair residual; corrected residuals are negative beyond 1 μm. Branch conditioning reduced the submicron output residual by approximately three quarters but left a positive median in 84% of paired neurons. Differential branch occupancy and residual within-branch submicron concentration therefore both contribute. For inputs, N2.5-B nearly removed the negative 0–1-μm residual and converted the negative 1–5-μm spectrum to values near zero or slightly positive. The primary input depletion therefore reflects lower same-branch co-occupancy—a branch-allocation pattern—rather than negative within-branch spacing at 1–3 μm. This distinction also resolves the missing-spine result: 1–3-μm depletion is robust within the primary N2.5 estimand, but it is not a residual within-branch spacing effect once branch counts are fixed. These statistics define organization and pair-separation scales, not a biological interaction range.

N2.5 is stringent with respect to the proofread neuron's own cable: it preserves compartment, proximodistal position and local cable-length class. It is not a complete model of contact opportunity. It omits membrane surface area, branch order and identity, spine density, apposed partner membrane, subcellular target identity, extracellular accessibility and molecular compatibility. Residuals may therefore reflect omitted contact geometry as well as biological placement processes. Nevertheless, the three prespecified signs persisted across 2-, 3-, 5- and 10-μm opportunity voxels, and the output submicron residual persisted across 0.25–2-μm projection cutoffs in a fixed cohort. A 2,070-event

candidate-set audit found identical projection distance, root position and cutoff membership for k = 32 and k = 128, with 99.9% branch-label agreement; k = 128 is a sensitivity check, not proof of a global nearest segment. True N2.5 count matching shows that abundance imbalance does not explain the full-compartment entropy polarity, but canonical axonal-output/dendritic-input restriction attenuated it and reduced the cohort to 36 neurons. Independence from compartment composition is therefore not established. Leakage-controlled morphology prediction supports reproducible effect-magnitude variation, not a molecular cell-type rule or an explanation of all residual structure.

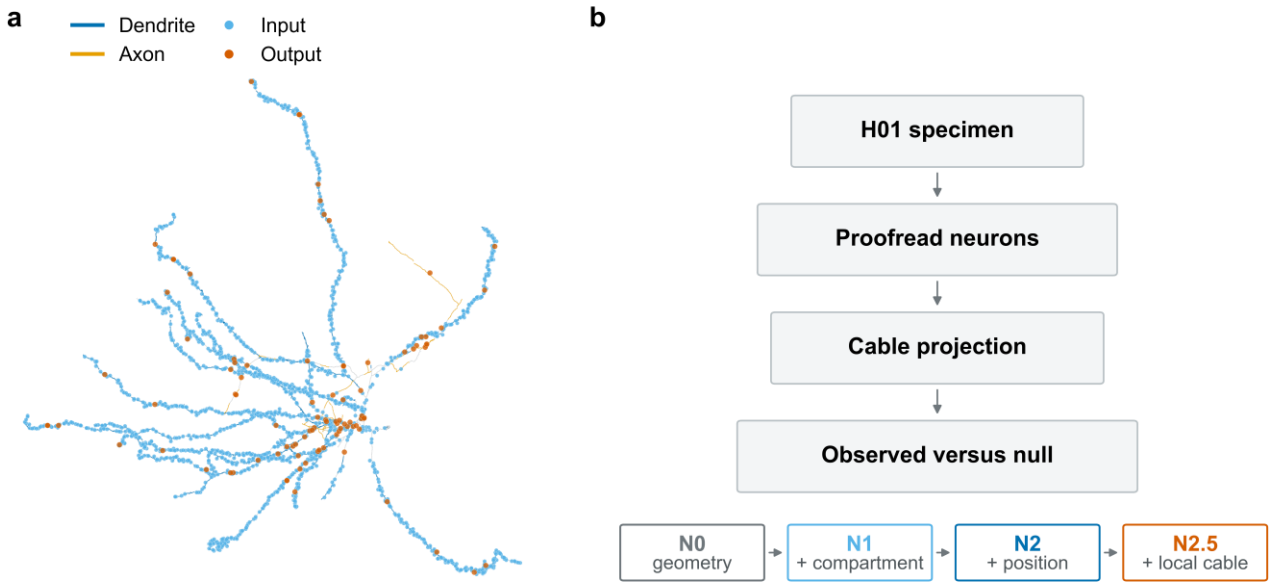
For outputs, site deduplication shows that relationship-table multiplicity is insufficient to generate the positive submicron residual, although it increases its magnitude. The same-C3 audit associates much of this signal with repeated contacts assigned to the same automated agglomerate. The null preserves overall C3-label abundance and marginal N2.5 stratum counts while breaking label–location association; it does not preserve label composition within each stratum. Because C3 labels may be fragmented or merged and are not proofread biological partners, the analysis cannot establish a partner-specific wiring rule or clustering across verified distinct targets.

The scope is further limited by one surgical specimen from anterior middle temporal gyrus in a donor with drug-resistant epilepsy [10], a restricted soma sampling slab, incomplete molecular identities and axons truncated by the imaged volume. The data describe local observed structural events, not complete neuronal output, synaptic efficacy, dynamics or causality. A next-generation N3 null should incorporate apposed partner membrane, subcellular target surfaces, branch-level covariates and proofread partner identities. Replication across cortical regions, donors and species, together with linked physiology, will be required to establish generality and function. Within the present scope, the data support a focused conclusion: input and output maps differ in how events are allocated across and within neurite branches after conditioning on tested intrinsic neurite opportunity. Inputs show broader branch-level dispersion, whereas outputs preferentially occupy particular branches and retain a

submicron within-branch concentration. The restored frozen projection source regenerated all prespecified cutoff, opportunity-scale and branch-conditioned sensitivity analyses.

References

1. DeBello, W. M. et al. Input clustering and the microscale structure of local circuits. *Front. Neural Circuits* **8**, 112 (2014). https://doi.org/10.3389/fncir.2014.00112
2. Gökçe, O., Bonhoeffer, T. & Scheuss, V. Clusters of synaptic inputs on dendrites of layer 5 pyramidal cells in mouse visual cortex. *eLife* **5**, e09222 (2016). https://doi.org/10.7554/eLife.09222
3. Iacaruso, M. F., Gasler, I. T. & Hofer, S. B. Synaptic organization of visual space in primary visual cortex. *Nature* **547**, 449–452 (2017). https://doi.org/10.1038/nature23019
4. Rah, J.-C. et al. Thalamocortical input onto layer 5 pyramidal neurons measured using quantitative large-scale array tomography. *Front. Neural Circuits* **7**, 177 (2013). https://doi.org/10.3389/fncir.2013.00177
5. Hellwig, B., Schüz, A. & Aertsen, A. Synapses on axon collaterals of pyramidal cells are spaced at random intervals: a Golgi study in the mouse cerebral cortex. *Biol. Cybern.* **71**, 1–12 (1994). https://doi.org/10.1007/BF00198906
6. Schmidt, H. et al. Axonal synapse sorting in medial entorhinal cortex. *Nature* **549**, 469–475 (2017). https://doi.org/10.1038/nature24005
7. Stepanyants, A., Hof, P. R. & Chklovskii, D. B. Geometry and structural plasticity of synaptic connectivity. *Neuron* **34**, 275–288 (2002). https://doi.org/10.1016/S0896-6273(02)00652-9
8. Chklovskii, D. B. Synaptic connectivity and neuronal morphology: two sides of the same coin. *Neuron* **43**, 609–617 (2004). https://doi.org/10.1016/j.neuron.2004.08.012
9. Reimann, M. W. et al. An algorithm to predict the connectome of neural microcircuits. *Front. Comput. Neurosci.* **9**, 120 (2015). https://doi.org/10.3389/fncom.2015.00120
10. Shapson-Coe, A. et al. A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science* **384**, eadk4858 (2024). https://doi.org/10.1126/science.adk4858
11. Beaulieu-Laroche, L. et al. Enhanced dendritic compartmentalization in human cortical neurons. *Cell* **175**, 643–651.e14 (2018). https://doi.org/10.1016/j.cell.2018.08.045
12. Benavides-Piccione, R. et al. Key morphological features of human pyramidal neurons. *Cereb. Cortex* **34**, bhae180 (2024). https://doi.org/10.1093/cercor/bhae180
13. Kasthuri, N. et al. Saturated reconstruction of a volume of neocortex. *Cell* **162**, 648–661 (2015). https://doi.org/10.1016/j.cell.2015.06.054
14. Schneider-Mizell, C. M. et al. Inhibitory specificity from a connectomic census of mouse visual cortex. *Nature* **640**, 448–458 (2025). https://doi.org/10.1038/s41586-024-07780-8
15. Kirchner, J. H. & Gjorgjieva, J. Emergence of synaptic organization and computation in dendrites. *Neuroforum* **28**, 21–30 (2022). https://doi.org/10.1515/nf-2021-0031



296
297 **Fig. 1 | Morphology-conditioned analysis of residual synaptic statistics.** a, Orthogonal projection of a representative
298 proofread neuron with dendritic cable, axonal cable, mapped inputs and mapped outputs. b, Analysis workflow. Incident
299 synaptic events from 104 proofread H01 neurons were projected to reconstructed cable and compared with nested within-
300 neuron null models. N0 preserves cable geometry; N1 additionally preserves compartment; N2 additionally preserves root-
301 distance bin; N2.5 additionally preserves local own-neuron cable-length class in 5-µm voxels. Arrows indicate the
302 sequential addition of constraints. N2.5 represents intrinsic cable-length opportunity, not complete contact opportunity. The
303 2-µm mapping cutoff retained 272,786 events.
304

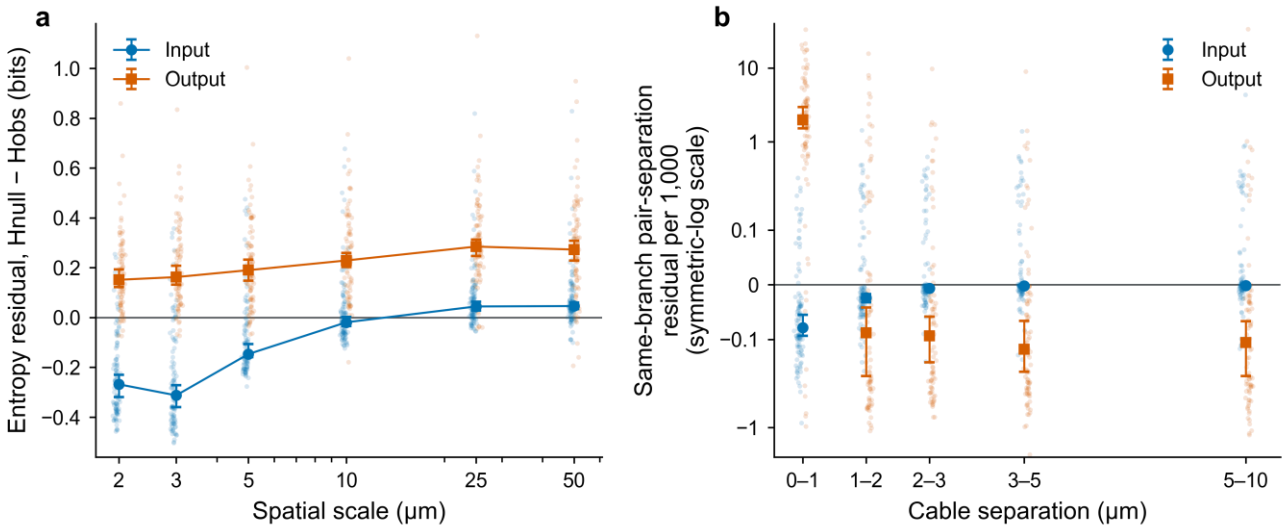


Fig. 2 | Residual synaptic statistics across organization and pair-separation scales. a, Entropy residuals under N2.5 across spatial scales. Small points show individual neurons (input, n = 104; output, n = 88); coloured symbols show population medians and error bars show 95% confidence intervals from 10,000 neuron-level bootstrap samples. Negative values indicate dispersion and positive values indicate concentration relative to N2.5. b, Zero-inclusive cable same-branch pair residual. The numerator counts same-branch unordered pairs within the stated absolute root-distance-difference bin; the denominator is all unordered event pairs in that neuron. The displayed statistic is 1,000 × (observed pair rate − mean N2.5-null pair rate). Small points show every eligible neuron (input, n = 104; output, n = 88), including neurons with no observed pair in a bin; coloured symbols and error bars show the population median and 95% neuron-bootstrap confidence interval. The symmetric-log axis is linear near zero and logarithmic at larger absolute values. P values are from two-sided exact sign tests with Benjamini–Hochberg correction across the ten direction-by-bin tests. The y axis is linear from −0.15 to 0.15 and logarithmic outside that interval. Neurons are the inference unit.

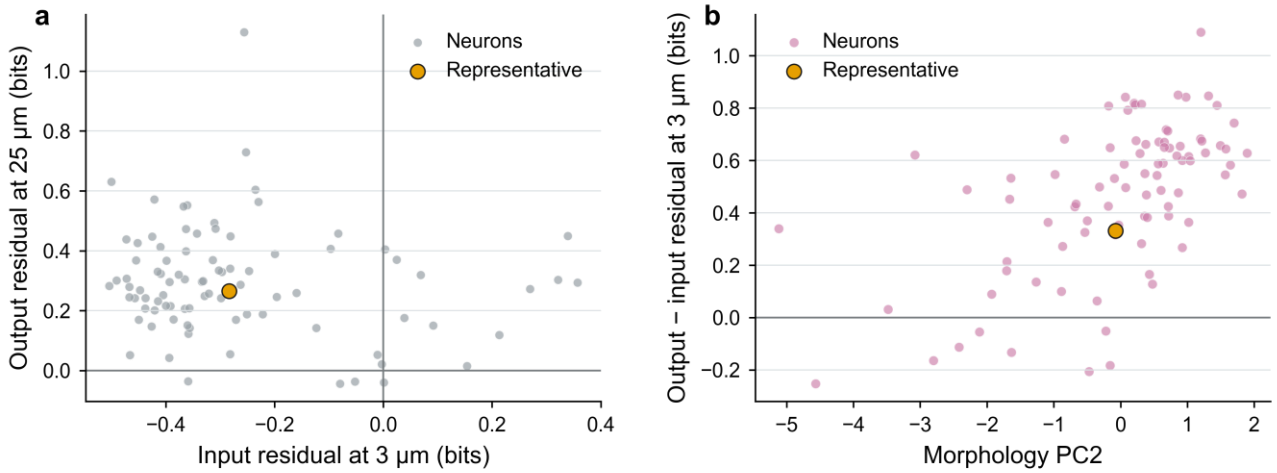


Fig. 3 | Cell-level prevalence and morphology-associated variation. a, Each point is one of 88 neurons with sufficient events in both directions. Input entropy residual at 3 µm is plotted against output entropy residual at 25 µm. Seventy-three of 88 neurons (83%) have a negative input residual and positive output residual under N2.5. b, The output-minus-input entropy contrast at 3 µm is associated with morphology PC2 (Spearman $\rho = 0.569$; FDR-adjusted $q = 2.67 \times 10^{-7}$; $n = 88$); 80 of 88 contrasts are positive. This association describes effect-magnitude variation and is not a causal model. The highlighted neuron is shown in Fig. 6.

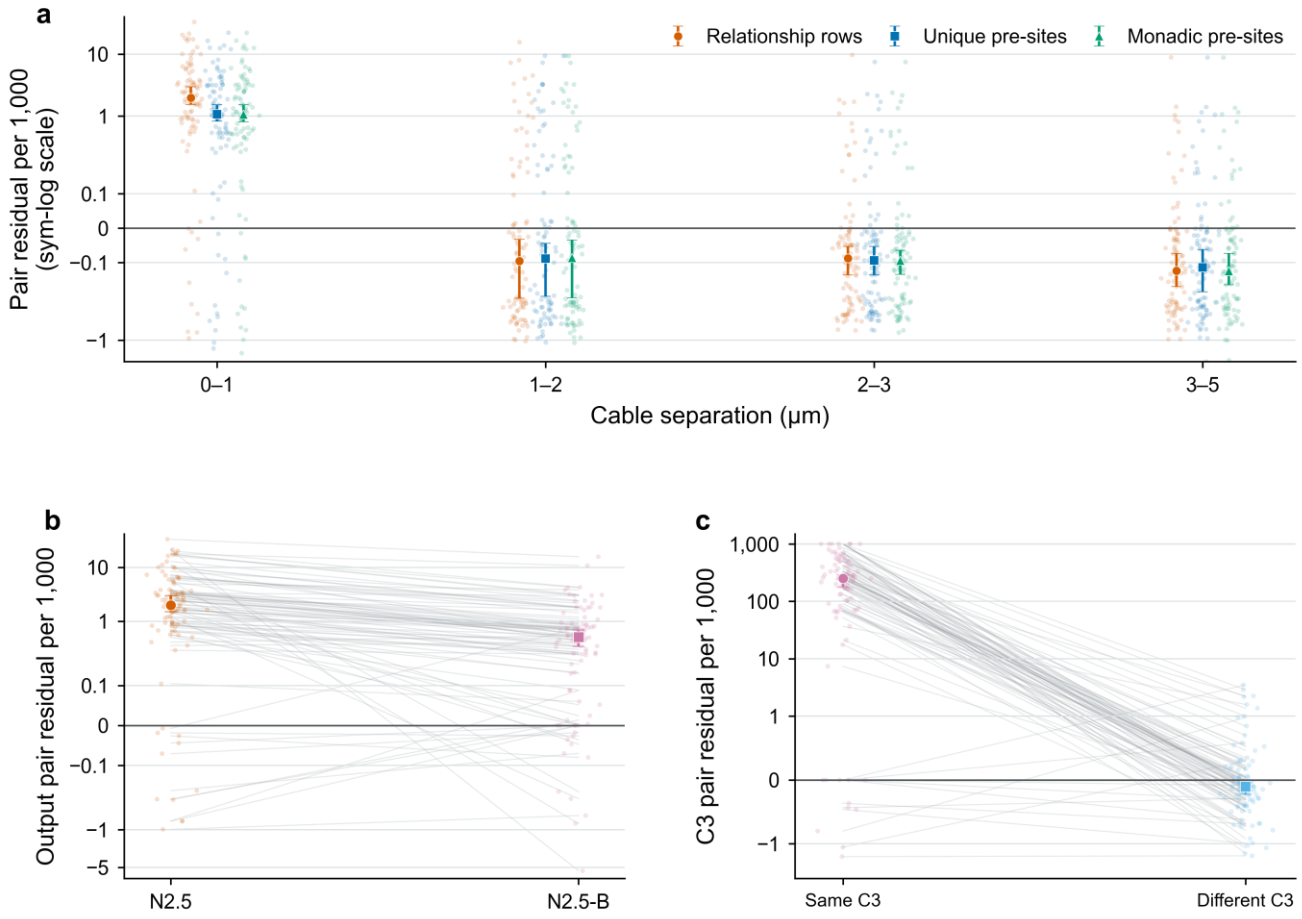


Fig. 4 | Output submicron concentration reflects both branch occupancy and within-branch organization. a, Zero-inclusive pair-rate residuals for relationship rows, unique presynaptic sites and monadic presynaptic sites. Small points show all 88 eligible neurons; coloured symbols show population medians and error bars show 95% neuron-bootstrap confidence intervals. Site deduplication does not remove the positive 0–1-μm residual. b, Paired output 0–1-μm residuals under standard N2.5 and the branch-count-preserving N2.5-B null in the same 88 neurons. The median falls from +1.976 to +0.515 per 1,000 pairs; the paired median change is −1.448 (95% bootstrap confidence interval, −1.918 to −0.974; two-sided sign-test q = 1.20 × 10−12). Under N2.5-B, 74 of 88 neurons remain positive (q = 6.44 × 10−12), indicating that branch occupancy explains much, but not all, of the output signal. c, Paired 0–1-μm residuals for event pairs assigned to the same or different C3 agglomerates, using monadic unique presynaptic sites in the same 88 neurons. The paired contrast uses a two-sided exact sign test with Benjamini–Hochberg correction (q = 1.13 × 10−17). C3 identifiers are automated agglomerates rather than fully proofread biological partners. Lines join measurements from the same neuron; summary symbols denote medians and 95% bootstrap confidence intervals.

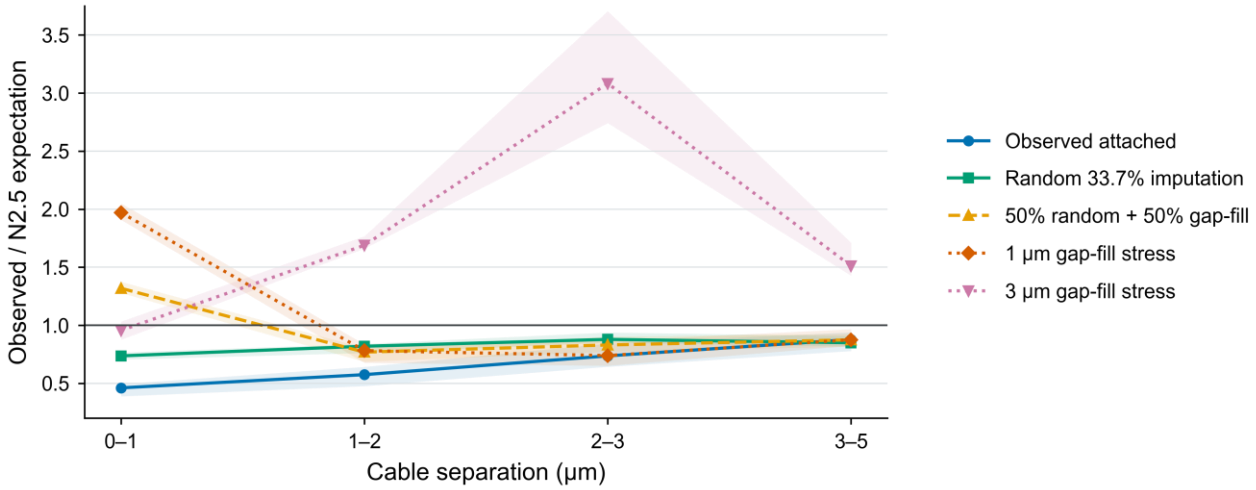


Fig. 5 | Input pair depletion under spine-loss stress tests. Observed attached inputs are compared with random imputation at the 33.7% rate measured in one audited layer-2 pyramidal neuron. The imputed count is nobs × 0.337/0.663, or approximately 50.8% of the observed count. A mixed random and gap-fill model and adversarial 1- and 3-µm gap-fill scenarios. Values below one indicate fewer nearby input pairs than expected under N2.5. Random imputation preserves depletion, whereas adversarial 1-µm gap filling can reverse the 0–1-µm bin; the 1–3-µm conclusion is therefore more robust than the submicron input result. Symbols show population medians and shading shows 95% neuron-bootstrap confidence intervals. Gap-fill models are sensitivity bounds rather than estimates of biological missingness.

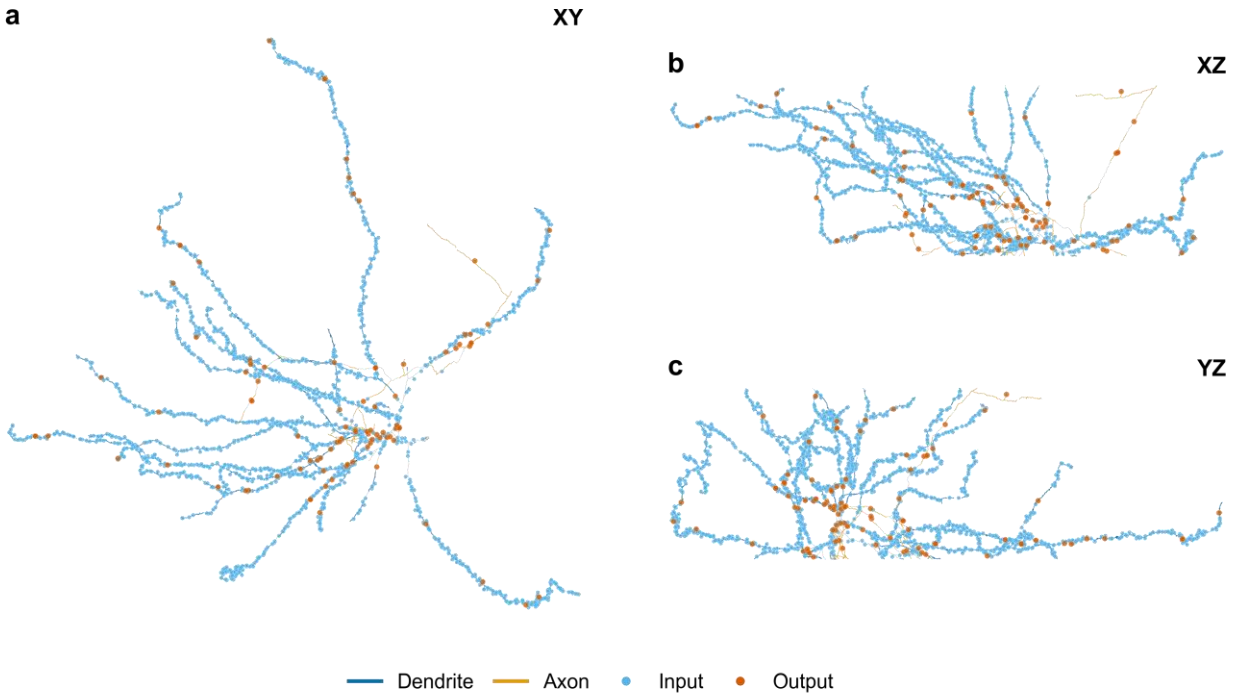


Fig. 6 | A representative proofread neuron places the population statistics in spatial context. a–c, XY, XZ and YZ projections, respectively, of proofread neuron 5,175,880,292 show dendritic and axonal cable with mapped input and output events. This reconstruction is illustrative rather than an independent replicate; all hypothesis tests use neuron-level statistics across the population.

Methods

H01 sample

We analyzed derived tables from the H01 human temporal-cortex resource [10]. Tissue acquisition, fixation, staining, sectioning, imaging, alignment, automated segmentation, and synapse detection were performed in the source study. Briefly, the resource contains a serial-section electron-microscopy volume from anterior middle temporal gyrus removed during surgery for drug-resistant epilepsy. The proofread-neuron set comprised 104 neurons randomly sampled by the H01 team from cells whose soma centers lay 23–30 μm in z from the beginning of the block. Four proofreaders corrected each reconstruction, followed by expert review. The original study reported approximately 810 person-hours of proofreading for this set.

We retained the source study's ethical and data-use framework by analyzing public derived data without recruiting participants or obtaining new tissue. Any journal-specific ethics statement should cite the source H01 approval and consent language verbatim at submission.

Synapse mapping

Incident relationship rows were classified by synaptic polarity relative to each proofread neuron: an event was an input when the proofread neuron was postsynaptic and an output when it was presynaptic. Compartment identity was assigned independently from the annotation of the cable segment receiving the projection and could be axon, dendrite, cell body or unknown in the primary analysis. A separate canonical-compartment sensitivity restricted inputs to dendritic cable and outputs to axonal cable. Same-cell relationships were excluded from directional population inference. The restored projection generator constructed a midpoint k-d tree for each neuron's cable segments. For every event, the 32 nearest segment midpoints were retrieved, the orthogonal point-to-segment projection was computed with the segment parameter clipped to ^{0,1}, and the candidate with minimum Euclidean projection distance was selected. The resulting projected_events.npy stores projected coordinates, projection

distance, branch assignment and continuous root position. This is an exact projection within the 32-candidate set rather than a formal all-segment search; candidate-set sensitivity was audited separately. Events farther than 2 μm were excluded. Of 277,143 projected events, 264,737 inputs and 8,049 outputs passed this cutoff (272,786 total). All 104 neurons met the ≥20-event input criterion; 88 met the output criterion and contributed 7,837 output events to direction-level inference.

Cable representation

Each proofread reconstruction was represented as a cable graph with physical edge lengths derived from voxel coordinates. Dendritic and axonal compartments followed the curated reconstruction annotations. Root distance was computed as geodesic distance along the cable from the relevant compartment root. For null conditioning, root distance was divided into ten bins within neuron and direction.

Hierarchical nulls

Null models redistributed the observed number of mapped events separately within each neuron and polarity. N0 sampled eligible cable segments in proportion to segment length. N1 additionally preserved neurite compartment, and N2 additionally preserved one of ten root-distance bins. For N2.5, own-neuron cable length was summed in 5-μm 3D voxels and divided into four length-weighted classes within compartment-by-root-distance strata. Each event was resampled onto a segment in the same compartment, root-distance bin and cable-length class, with probability proportional to segment length. Core N2.5 analyses used 2,000 Monte Carlo replicates per neuron and polarity. Analyses required at least 20 mapped events for the relevant neuron and polarity. N2.5 represents intrinsic own-neuron cable-length opportunity and does not model complete contact opportunity. The prespecified primary opportunity voxel was 5 μm with four length-weighted classes. Sensitivity analyses recomputed N2.5 at 2-, 3- and 10-μm opportunity voxels while retaining four classes and ten root-distance bins. Primary N2.5 estimates used 2,000 null replicates; strict-cutoff, opportunity-scale and

branch-conditioned sensitivities used 250 replicates per cell and were compared with the archived 2,000-replicate primary tables. Across the full cohort, 250-versus-2,000-replicate neuron-level rank correlations were 0.998–1.000 for entropy and 0.992–1.000 for pair residuals; median absolute differences were 0.00064–0.00429 bits and 0.00029–0.0161 pairs per 1,000, respectively.

Entropy residuals

Projected event coordinates were discretized into 3D cubic voxels of side length 2, 3, 5, 10, 25 or 50 μm for the primary analysis. Shannon entropy was calculated from event-occupancy proportions across non-empty voxels. For each neuron, polarity and scale, the residual was mean(HN2.5 null) − Hobserved, in bits. Negative values indicate greater observed dispersion and positive values indicate greater observed concentration relative to N2.5. Voxel side length is termed the organization scale; it is an analysis resolution and not a physical synapse–synapse interaction range. The 1- and 100-μm estimates were retained only in audit outputs, outside the six-scale primary spectrum.

Same-branch pair-separation residual

For a separation bin [a,b), the numerator was the number of unordered event pairs assigned to the same branch with absolute root-distance difference in [a,b). The denominator was the total number of unordered event pairs in that neuron and polarity, n(n−1)/2, including cross-branch pairs. The observed rate was compared with corresponding rates from 2,000 N2.5 null replicates. The displayed zero-inclusive residual was 1,000 × (observed rate − mean null rate). This additive statistic remains defined when the observed numerator is zero and retains every neuron meeting the direction-level event criterion. Primary bins were 0–1, 1–2, 2–3, 3–5 and 5–10 μm. This statistic combines branch co-occupancy and within-branch root-distance spacing; it is not a full cable-geodesic pair-correlation function or direct evidence of biological interaction. A branch-count-preserving sensitivity (N2.5-B) held each observed branch's event count exactly fixed. Within each branch it sampled the observed compartment/root-bin/opportunity-bin strata in proportion to cable length; fallbacks dropped

opportunity class and then compartment/root bin if necessary, but never moved an event between branches. Across 68,143,500 branch-conditioned event draws, 99.922% retained the full branch × compartment × root-bin × opportunity-bin stratum, 0.0569% dropped opportunity class and 0.0209% used branch identity only. N2.5-B therefore conditions on branch occupancy and isolates the remaining within-branch spacing component; it does not replace the primary N2.5 estimand. The finite log2(observed/expected) analysis excludes zero-observation cells and is diagnostic only. Figure 2 uses a symmetric-log display with linthresh = 0.15: values from −0.15 to 0.15 are linear and larger absolute values are logarithmic.

Count matching

For the final N2.5 count-matching sensitivity, 88 neurons with at least 20 mapped events in each direction were retained. Within each neuron, inputs were subsampled without replacement to the corresponding output count over 500 iterations. For each input subsample and the observed output set, compartment, root-position decile and 5-µm local-cable-opportunity quartile were recomputed and a new N2.5 null realization was drawn. Entropy residuals and the paired output-minus-input contrast were calculated at 2, 3, 5, 10, 25 and 50 µm. Neuron-level estimates were medians across iterations; population confidence intervals bootstrap neurons 10,000 times. Two-sided exact sign tests were adjusted across six scales by Benjamini–Hochberg. Each iteration compared 7,837 events per direction. This sensitivity addresses entropy residuals and does not count-match the pair-separation statistic.

Output-site audits

Output analyses were repeated for relationship rows, unique presynaptic coordinates and monadic unique sites. Unique-site analysis collapsed relationship rows sharing a presynaptic-site identifier; monadic analysis retained sites represented by one relationship row. For the C3 audit, monadic unique-site pairs were divided into same-C3 and different-C3 categories. Each category rate used its category-specific possible-pair denominator and was compared with the corresponding N2.5 null rate. C3-label

abundances were held fixed within every Monte Carlo replicate. N2.5 generated a new set of positions with the observed marginal compartment, root-distance and opportunity-stratum counts, and the randomized sampled-position order was assigned to the fixed label vector. The null therefore preserves label abundance and N2.5 stratum counts while breaking the observed label–location association. It does not preserve the observed label composition within each N2.5 stratum. Same-versus-different inference used a within-neuron paired contrast. C3 labels were treated as automated segmentation identifiers rather than proofread biological partners.

Spine-loss sensitivity

A 33.7% detached-spine fraction measured in one layer-2 pyramidal neuron was used as a stress-test parameter, not a population estimate. For each analysed neuron, the imputed count was round[nobserved × 0.337/(1−0.337)], or approximately 50.8% of the observed count. Random events were sampled from N2.5-eligible cable. Mixed and adversarial conditions increasingly assigned events to 1- or 3-µm gaps. The default analysis used 300 N2.5 null replicates and 100 imputation replicates per cell. These scenarios bound sensitivity and do not estimate population missingness.

Morphology analyses

Ten raw morphological features were available for 104 neurons. For prediction in the 88-cell paired cohort, median imputation, standardization and principal-components analysis were fitted independently within each training fold, and the first three components were retained. The response was the paired N2.5 entropy spectrum at 2, 3, 5, 10, 25 and 50 µm; summary predictive R^2 is the mean of six scale-specific held-out R^2 values. Performance was evaluated using one fixed ten-fold partition, 200 repeated ten-fold partitions and leave-one-out prediction. Significance used 999 fixed-fold response-label permutations. A global 104-cell PCA was used only to describe PC1–PC3 loadings and the Fig. 3 correlation, not for predictive validation. Dominant-layer and morphology-cluster

associations used 9,999-permutation pseudo-F tests. These analyses were exploratory because strata were imbalanced and reliable molecular cell-type annotations were unavailable for most neurons.

Geometry sensitivity

To evaluate anisotropic tissue deformation, coordinates along the compressed section axis were corrected by the source-study factor (5.552/4), and entropy residuals were recomputed. A separate analysis censored branches near dataset boundaries or flagged artifact endpoints. Boundary associations were tested with partial Spearman correlations controlling log output count and log axonal cable length. Projection-threshold sensitivity filtered the restored continuous projection at 0.25, 0.5, 1 and 2 μm and recomputed entropy and same-branch pair-separation residuals. Results were summarized for all eligible profiles at each threshold and for the intersection of 64 neurons eligible in both directions at every threshold. A separate audit sampled 2,070 events across all 104 neurons and compared 32, 64 and 128 midpoint-candidate segments. The larger candidate sets are sensitivity checks rather than proof of exhaustive global nearest-segment projection.

Statistics

The proofread neuron was the biological inference unit. Population summaries are medians across eligible neurons with 95% percentile confidence intervals from 10,000 neuron-level bootstrap samples unless stated otherwise. Entropy residuals were tested against zero with two-sided Wilcoxon signed-rank tests. Zero-inclusive pair residuals and the paired C3 contrast were evaluated with two-sided exact sign tests after removing exact ties. Benjamini–Hochberg correction was applied within each prespecified family of scales or contrasts; q denotes an adjusted P value. Eligibility required at least 20 mapped events for the relevant neuron and polarity. The zero-inclusive pair residual retains eligible neurons with no observed pair in a bin. No synapse was treated as an independent biological replicate.

Software and reproducibility

Analyses and manuscript audits were implemented in Python with deterministic seed 20260803 for the new projection, cutoff, opportunity-scale, branch-conditioned, count-matching and morphology audits; bootstrap, cross-validation and permutation procedures used documented derived seeds. The accompanying frozen package contains projected_events.npy, the continuous-projection generator, cable-model hashes, per-cell and population analysis tables, analysis manifests, count-matching and nested-cross-validation outputs, figure-generation source, high-resolution raster figures and vector artwork. SHA-256 hashes and software versions are recorded in the package manifest. A public persistent data-and-code accession remains an author-team requirement and is not invented in this draft.

Data availability

The underlying H01 volume and source annotations are available through the access routes described by Shapson-Coe et al. [10]. This study uses derived proofread-neuron event and morphology tables. A public archival link for the frozen derived tables will be added when the author team deposits the submission package.

Code availability

Versioned analysis and figure-generation scripts are included in the project bundle used to generate this manuscript. A public repository and persistent release will be created before journal submission.

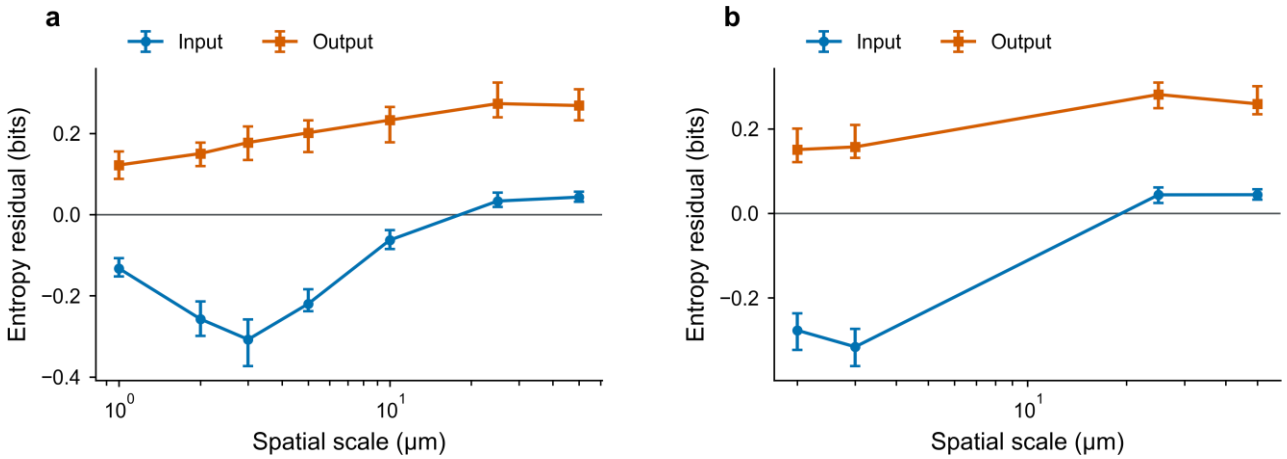
Acknowledgements

This work is developed during the 2026 PEBBLE BioFusion Workshop at Westlake University in Hangzhou, China, from July 24th to August 4th.

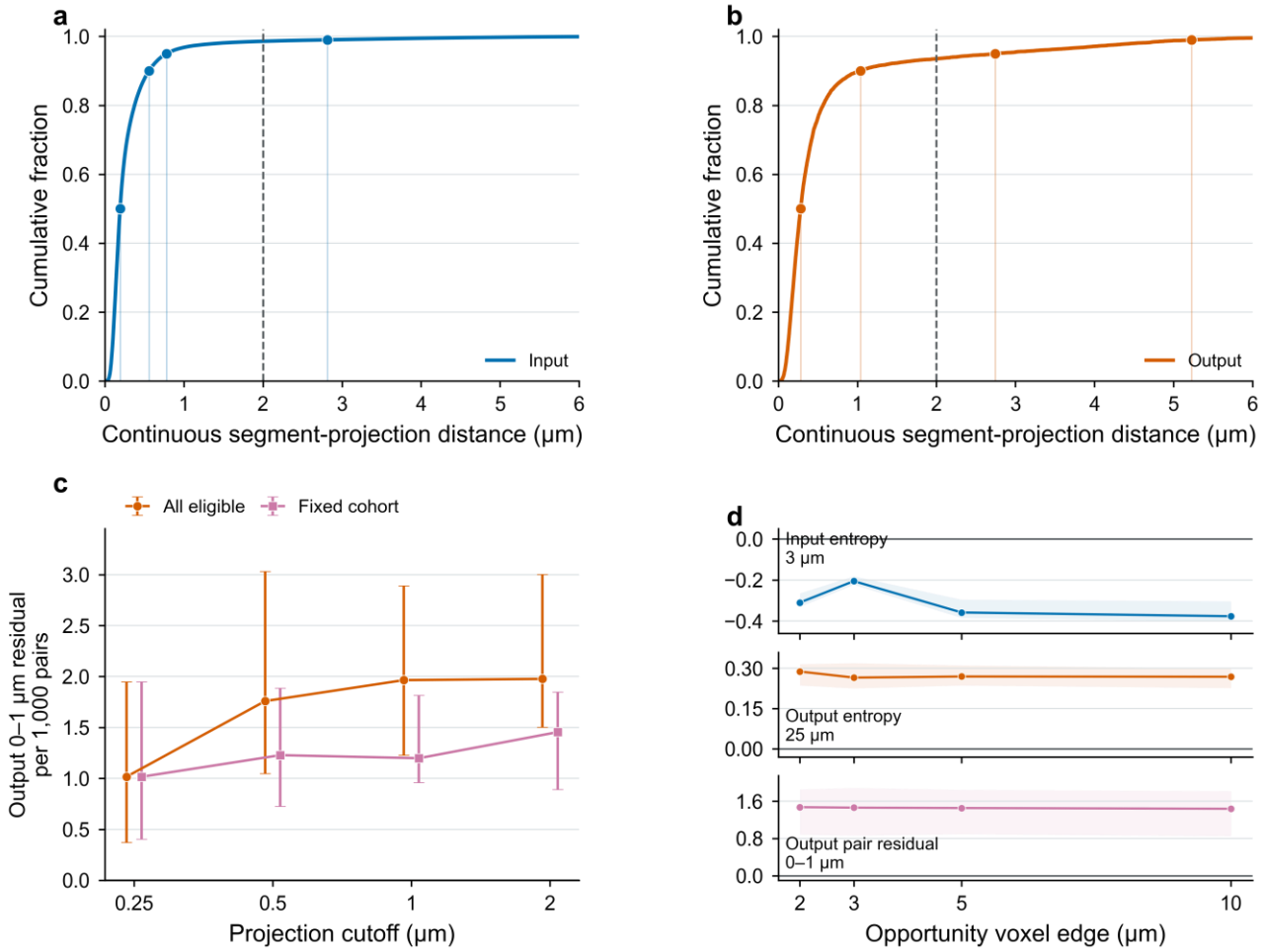
Funding

Author contributions

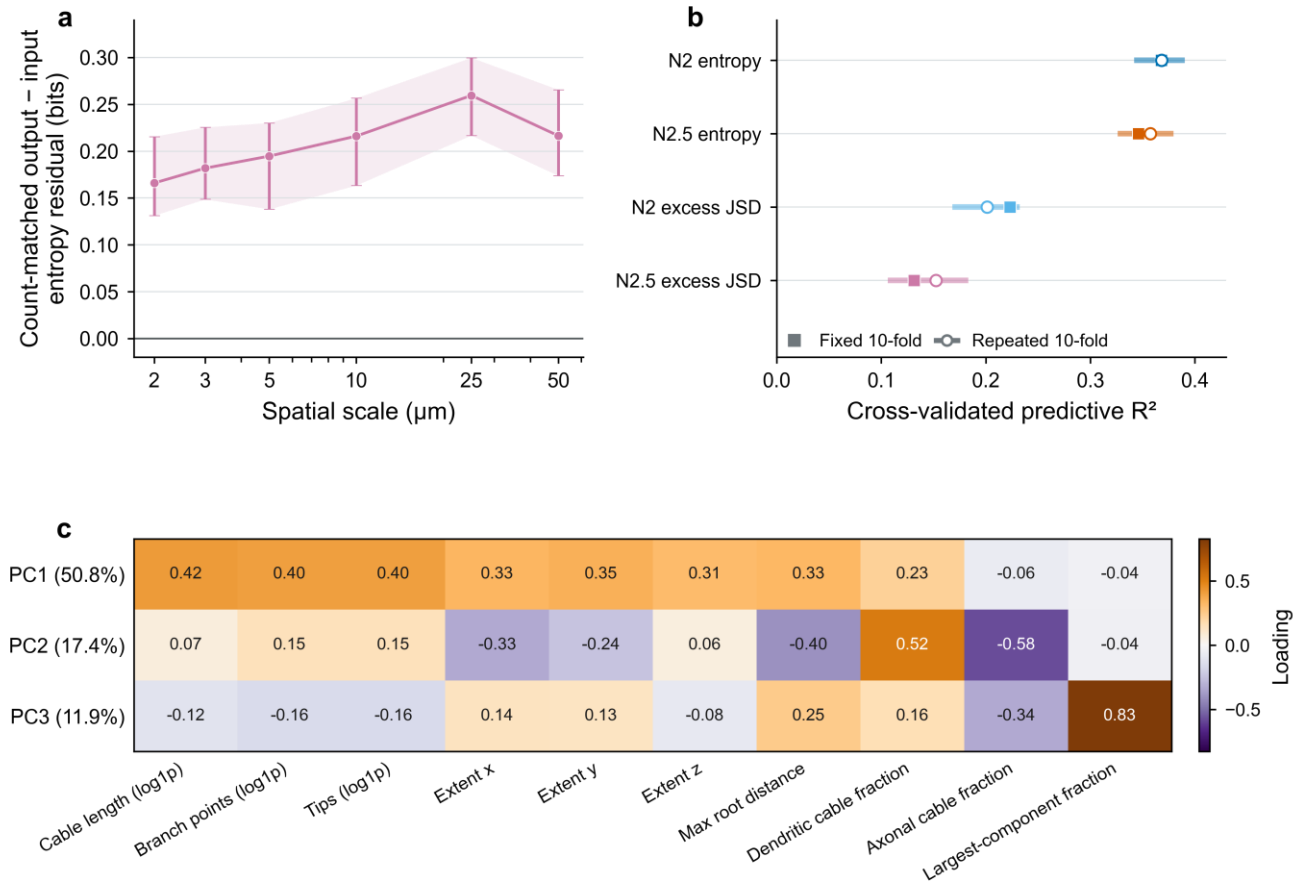
Competing interests



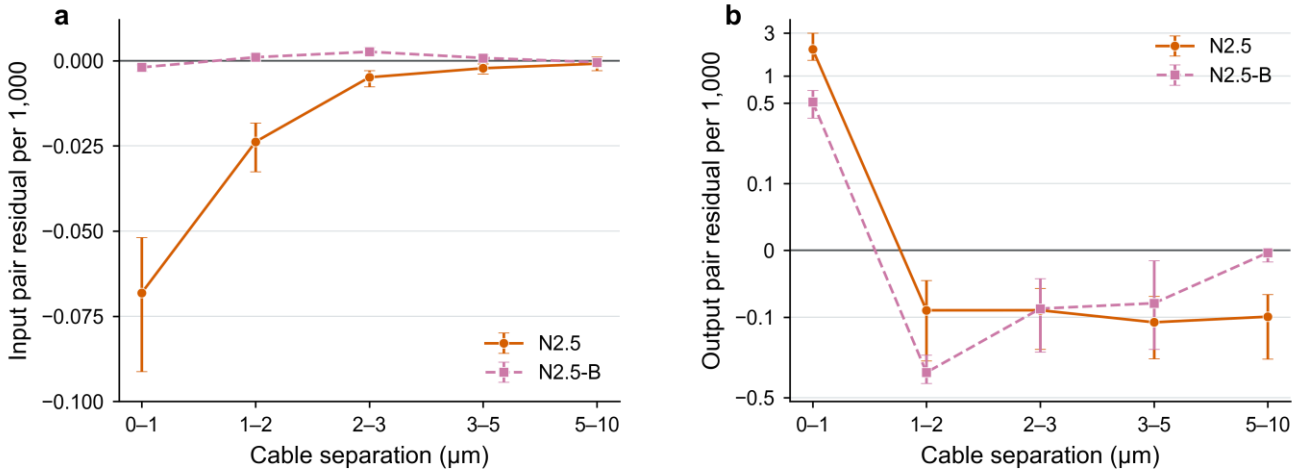
Extended Data Fig. 1 | Core polarity signs survive geometric sensitivity analyses. a, Entropy residuals after correction for section-axis compression. b, Residuals after censoring branches near volume boundaries or suspected reconstruction artifacts. Symbols show population medians and error bars show 95% neuron-bootstrap confidence intervals.



Extended Data Fig. 2 | Continuous-projection, cutoff and opportunity-scale sensitivity. a,b, Empirical cumulative distributions of continuous segment-projection distance for 268,536 input and 8,607 output events before the 2-μm validity cutoff. Input P50, P90, P95 and P99 values are 0.192, 0.558, 0.780 and 2.814 μm; output values are 0.281, 1.042, 2.744 and 5.228 μm. The 2-μm cutoff retains 98.6% of inputs and 93.5% of outputs. c, Output 0–1-μm same-branch pair-separation residuals at projection cutoffs of 0.25, 0.5, 1 and 2 μm. Cutoff-specific cohorts contain 64, 87, 88 and 88 neurons; the common fixed cohort contains 64 neurons. The medians remain positive in both analyses, although the strictest cutoff retains only 43.1% of output events. d, Prespecified readouts after varying the N2.5 opportunity-voxel edge among 2, 3, 5 and 10 μm in the fixed cohort: input entropy residual at a 3-μm observation scale, output entropy residual at 25 μm and output 0–1-μm pair-separation residual. All signs are preserved. Points denote neuron-level medians and error bars or bands denote 95% bootstrap confidence intervals. These sensitivity conditions use 250 null realizations per neuron; the primary estimates in Fig. 2 use 2,000.



Extended Data Fig. 3 | Count matching and leakage-free morphology analyses. a, Output-minus-input entropy residual after exact within-neuron input count matching under N2.5. For each of 88 paired neurons, inputs were subsampled to the observed output count in 500 iterations and the N2.5 strata and null were recomputed. Points show population medians and shading shows 95% neuron-bootstrap confidence intervals. The contrast remains positive at every observation scale. b, Cross-validated predictive R^2 for four prespecified targets from morphology features. Filled squares denote fixed ten-fold estimates; open circles denote medians over repeated ten-fold splits and horizontal intervals show the central 95% range over repeats. Median imputation, standardization and principal-component analysis were fitted within each training fold before transformation of held-out neurons. The N2.5 entropy contrast is predictively associated with morphology; the two excess-JSD targets are not. c, Loadings of the ten morphology features on the first three descriptive principal components. The components explain 50.8%, 17.4% and 11.9% of feature variance, respectively. PC2 loads negatively on axonal cable fraction (−0.58) and positively on dendritic cable fraction (+0.52).



Extended Data Fig. 4 | Branch-count conditioning resolves the complete pair-separation spectrum. a,b, Input and output zero-inclusive pair-separation residuals, respectively, under standard N2.5 and branch-count-preserving N2.5-B in the fixed 88-neuron cohort. Symbols show population medians and error bars show 95% neuron-bootstrap confidence intervals. For inputs, the standard negative residuals from 0 to 5 µm are reduced to −0.00195, +0.00102, +0.00262 and +0.000795 per 1,000 under N2.5-B; the 1–2-µm value is not significant after correction (q = 0.0693), whereas the 2–3- and 3–5-µm values are small and positive. Thus, the primary 1–3-µm input depletion does not persist after branch occupancy is fixed. For outputs, N2.5-B retains a positive 0–1-µm residual (+0.515) and negative residuals beyond 1 µm. N2.5-B used 250 null realizations per neuron; 99.922% of event draws retained the full branch × compartment × root-bin × opportunity-bin stratum, 0.0569% dropped opportunity class and 0.0209% used branch identity only.