LARGE-SCALE ANALYSIS OF MOUSE AND HUMAN DENDRITIC SPINES REVEALS NO EVIDENCE OF LOCAL SPATIAL CLUSTERING AND WEAK MORPHOLOGICAL SIMILARITY BETWEEN NEIGHBOURING SYNAPTIC CONTACTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Dendritic spines receive most excitatory synaptic inputs in the cortex, and their morphology and arrangement influence how neurons integrate signals. Whether spines form spatial clusters and whether neighbouring synaptic contacts resemble one another remain important questions for understanding synaptic organization and constructing computational representations of neurons. We investigated these questions using large-scale reconstructions of mouse (Minnie65) and human (H01) cortex, with morphological analyses covering more than 600,000 spines. Analyses of spine placement in the mouse dataset did not support local spatial clustering. Instead, spines were less closely spaced within dendritic branches than expected under independent placement; models incorporating short-range suppression improved, but did not fully reproduce, the observed spacing distributions. Across both datasets, neighbouring spines showed only weak excess morphological similarity after accounting for differences between branches. In the mouse dataset, repeated contacts from the same presynaptic axon showed correlated morphology, and their measurements improved spine-volume prediction beyond geometry and partner distances, although limited partner availability restricted the population-wide benefit. We compared our findings with previous studies and examined possible reasons for discrepant conclusions, including differences in tissue, morphological measures and statistical assumptions. An exploratory graph-based LeJEPA predictor collapsed to nearly constant outputs. In a separate neuron-disjoint experiment, hidden-volume recovery did not improve over geometry, and observed learned embeddings retained less volume information than randomly initialized embeddings. The procedure detected an imposed synthetic local signal, but these results establish limitations of the tested model rather than a causal link between weak morphological similarity and learning failure.

1 INTRODUCTION

Dendritic spines are small protrusions from the dendritic shafts of neurons, typically consisting of a bulbous head connected to the shaft by a narrow neck (Figure 1). They receive most excitatory synaptic inputs in the cortex. Their positions, shapes and presynaptic partners provide complementary descriptions of synaptic organization: spines may cluster along a dendrite, nearby spines may have similar morphology, and spines receiving input from the same axon may resemble one another even when they are spatially separated.

Previous studies have reported different patterns of spine organization. Yadav and colleagues found spatial clusters on apical terminal dendrites of monkey cortical neurons relative to uniform placement along each branch (Yadav et al., 2012), whereas Morales and colleagues found that spine attachment positions on most examined human cortical dendritic segments were compatible with spatial randomness (Morales et al., 2014). Morphological organization can occur independently of positional clustering: Jammalamadaka and colleagues reported associations between neighbouring

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

spine types in cultured rat hippocampal neurons even when their positions were compatible with spatial randomness (Jammalamadaka et al., 2013). Subsequent point-process analyses emphasized that changes in spine density and spine-type composition along a dendrite must be accounted for when testing for local organization (Baddeley et al., 2014; Christensen & Møller, 2020).

Shared presynaptic input offers another basis for spine similarity. Bartol and colleagues reported closely matched spine-head volumes in 17 pairs of spines receiving input from the same axon on the same dendrite in rat hippocampus (Bartol et al., 2015). Related connectomic studies found similarities in synaptic contact area and distinguished binary and continuous components of spine-head size variation (Motta et al., 2019; Dorkenwald et al., 2022). These anatomical observations complement reports of clustered spine formation, synaptic plasticity and functional co-tuning (Fu et al., 2012; Lu & Zuo, 2017; Scholl et al., 2017). Together, they motivate separating organization by spatial proximity from organization by shared input, while distinguishing anatomical similarity from correlations in neural activity.

In this work, we statistically analyse large-scale reconstructions of mouse (Minnie65) and human (H01) cortex, with morphological measurements from more than 600,000 spines. Our contribution is to quantify three forms of spine organization in these datasets (Figure 1). First, the mouse placement analysis provides no evidence of local spatial clustering; instead, nearby spine attachments are less frequent than expected under independent placement after accounting for branch geometry and density. Second, across both datasets, local morphological correlations are weak after accounting for branch composition, and neighbouring morphology provides no supported adjacency-specific predictive gain beyond branch and positional context. Third, repeated contacts from the same axon in Minnie65 show positive morphological correlations and improve volume prediction beyond geometry and partner distances.

The repeated-contact result supports the input-associated similarity reported by Bartol and colleagues, while yielding a more modest estimate in our cortical sample: the correlation of log whole-spine volume is $r = 0.1393$ (95% interval $[0.0775, 0.1989]$), based on 3,036 unique same-axon pairs. This extends the evidence with a larger sample and an explicit uncertainty estimate, although differences in tissue, morphological measurements and statistical procedures limit direct numerical comparison with Bartol et al. (Bartol et al., 2015). Our measurements characterize the anatomical similarity of contacts sharing an input; neural activity was not measured. As a further application, we test whether spatial graph representations improve recovery of hidden spine morphology, assessing the information available from neighbouring spines.

2 Datasets and spine measurements

We analysed three-dimensional reconstructions of mouse cortex from MICrONS Minnie65 and human cortex from H01 (MICrONS Consortium, 2021; Shapson-Coe et al., 2024). NEURD provided spine segmentation, dendritic skeletons and morphological descriptors (Celii et al., 2025). After selection, the local morphology analyses included 572,434 spines from 622 neurons in Minnie65 and 41,931 spines from 64 neurons in H01. The two datasets were analysed separately.

Each spine was associated with its parent neuron and dendritic branch. Its attachment position on the branch skeleton located it along the dendrite, allowing distances between spines to be measured along the branch. The surrounding anatomical context included branch length and calibre, distance from the soma and dendritic compartment. For the Minnie65 placement analysis, angular positions around the branch additionally represented the distribution of attachments over the dendritic surface.

Spine morphology was described by 17 measurements spanning overall size, shape and neck geometry (Table 4). These included whole-spine volume, surface area, skeletal length, bounding-box dimensions and neck measurements. Volume was estimated from the segmented three-dimensional mesh and refers to the whole spine; separate head and neck volumes were not included. This distinction matters when comparing our measurements with studies of spine-head size.

In Minnie65, presynaptic assignments also identified the axon contacting a spine. Contacts sharing an axon could therefore be grouped independently of their spatial proximity. These assignments were not available for every retained spine, and comparisons of repeated contacts required another observed contact from the same axon. Consequently, the analyses use overlapping subsets with different inclusion requirements: branch counts include branches without retained spines, local morphology

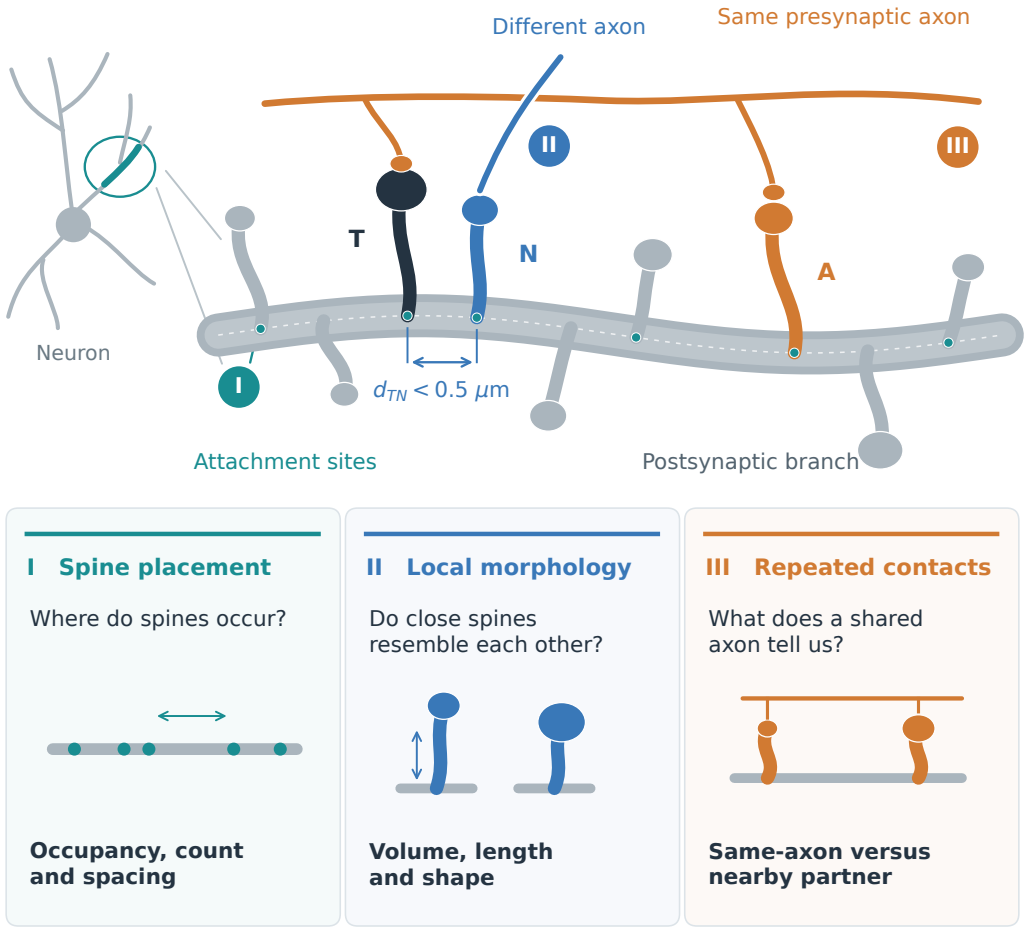

Figure 1: Three questions about organization on a dendritic branch. The schematic separates (I) spine placement, described by branch occupancy, spine number and attachment coordinates; (II) local morphology, assessed by comparing spines at nearby attachment positions; and (III) repeated contacts, assessed by comparing a target spine (T) with a nearby different-axon contact (N) and another contact of the same presynaptic axon (A). Distances are measured along the branch. The graph-recovery experiment extends question II by testing whether spatial aggregation improves hidden-morphology prediction. Axon colours indicate connectivity rather than activity; geometry and spine shapes are schematic and not to scale.

requires valid measurements and attachment coordinates, and partner comparisons require the relevant contacts to be observed. Table 3 in Appendix A gives the sample sizes for each analysis, alongside measurement definitions and selection criteria.

3 Part I: Distribution of spines along dendrites

We analysed spine distribution at two complementary levels. The first level asked whether a dendritic branch contained any retained spines and, if it did, how many it contained. The analysis unit was a maximal unambiguous, noncyclic chain of skeleton segments between branch points. The resulting risk set comprised 55,721 Minnie65 branch units, including branches with no retained spines; 29,391 units were occupied. Separating empty from occupied branches prevents variation in the number of available branch units from being conflated with variation in spine number along an occupied branch (Appendix A.2).

We used a two-stage model to predict whether a branch contained spines and, if so, how many. We compared prediction from branch length alone with prediction using additional anatomical context, evaluating both on held-out neurons (Appendix A.2).

Anatomical context improved prediction beyond branch length alone: held-out occupancy AUC increased from 0.9129 to 0.9699, and count error on occupied branches decreased (Table 10).

We then compared attachment positions within occupied branches, preserving each branch's observed spine count and broad density variation. Positions were represented on a cylindrical dendritic surface. The independent model allowed unrestricted placement; a hard-core model excluded very close pairs; and a soft-core model penalized close pairs without forbidding them. Parameters were selected without using the evaluation neurons. Model fit was assessed by how many of 207 observed spatial summaries fell within simulation intervals (Appendix A.2).

Short-range suppression improved model fit, with soft-core giving the highest coverage of spatial summaries and the lowest mean absolute error for nearest-neighbour summaries. Nevertheless, all three models remained below the operational 80% coverage reference (Table 11; Appendix E.2). Coverage is a descriptive adequacy measure: the summaries share branches and distance thresholds and are not independent hypothesis tests.

Figure 2 illustrates the remaining spacing discrepancy on held-out branches. The independent model predicts too many close neighbours at the evaluated thresholds. Soft-core reduces this excess, particularly at submicrometre distances, but does not eliminate it. Thus accurate prediction of branch occupancy does not imply an adequate model of within-branch spacing.

These comparisons describe spacing without identifying its biological mechanism. Their relation to earlier clustering studies depends on branch definitions, tissues and reference models (Appendix F). The following analysis asks whether spines at nearby positions also resemble one another in morphology.

4 Part II: Morphological correlation between nearby spines is consistently weak

We quantified morphological resemblance between spines on the same raw dendritic branch whose attachment points were separated by less than 0.5 μm along the branch. Every eligible unordered pair contributed, rather than only nearest neighbours. The cutoff was fixed before examining these correlations and also defines the immediate spatial neighbour in Part III. Restriction to the same branch permits a control for branch-level differences in morphological composition (Appendix A.3).

Figure S1 outlines the local morphology controls used here and the repeated-contact comparisons in Part III.

For descriptor m, we first applied its declared transformation f and centred it within branch b as $z_{bi} = f(m_{bi}) - \overline{f(m)}_b$. Volume, surface area and skeletal length were analysed on the logarithmic scale; the whole-spine volume-to-area ratio retained its linear scale. Both directions of each close pair, (z_i, z_j) and (z_j, z_i), entered the pooled Pearson correlation r_{observed}, giving symmetric pair weights. To remove correlation attributable merely to the set of morphologies present on a branch, measurements were randomly reassigned among the fixed spine positions within that branch. With 999 such permutations, the reported effect was $\Delta r = r_{\mathrm{observed}} - 999^{-1} \sum_{k=1}^{999} r_{\mathrm{shuffled}}^{(k)}$. Thus Δr measures excess resemblance at short separation beyond branch-level composition; it is not the raw correlation between all spines. Uncertainty was estimated with 2,000 neuron-bootstrap draws. The resulting 95% intervals are descriptive and are not adjusted for testing multiple descriptors.

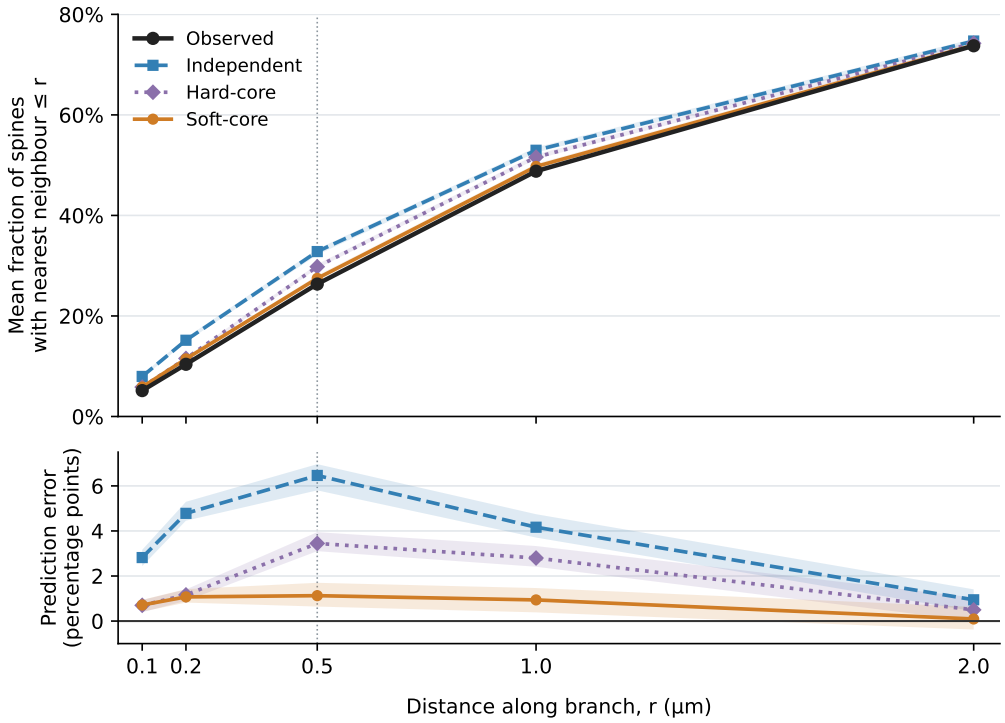

Figure 2: Nearest-neighbour distances along dendritic branches in Minnie65. Top: observed fractions of spines with a nearest neighbour within distance r, and mean predictions from independent, hard-core and soft-core placement models, on the same 4,000 held-out branches in the primary cohort. Fractions are computed within branches and then averaged with equal branch weights; simulations preserve each branch's observed spine count. Bottom: prediction error, defined as the model mean minus the observed fraction, in percentage points. Positive values indicate that the model predicts too many close neighbours; zero indicates agreement. Points show the five evaluated distance thresholds; connecting lines do not represent additional evaluations or smoothing. Shading shows pointwise 2.5–97.5 percentile intervals across 30 simulations per model, not uncertainty intervals for the observations. In the lower panel, these intervals are shifted by the observed fraction. The vertical dotted line marks 0.5 μm.

Table 1: Local morphological correlations for four whole-spine descriptors. Values are excess correlations Δr relative to whole-branch permutations for pairs separated by less than 0.5 μm; brackets give neuron-bootstrap 95% intervals without multiplicity correction. Minnie65 contributes 106,875 unordered pairs and H01 12,860. Volume, surface area and skeletal length use log-transformed measurements; volume/area retains its linear scale. Results for the full descriptor panel are reported in Appendix C.

Descriptor	Minnie65: Δr [95% CI]	H01: Δr [95% CI]
Volume	0.0038 $[-0.0019, 0.0095]$	0.0082 $[-0.0031, 0.0236]$
Surface area	0.0070 $[0.0006, 0.0136]$	0.0278 $[0.0142, 0.0469]$
Skeletal length	0.0112 $[0.0048, 0.0176]$	0.0213 $[0.0065, 0.0402]$
Volume/area	-0.0006 $[-0.0070, 0.0056]$	0.0049 $[-0.0081, 0.0176]$

Local excess correlations were small in both corpora (Table 1). Volume intervals included zero; surface area and skeletal length had positive intervals but small point estimates. Across all 17 descriptors, Δr ranged from -0.0006 to 0.0144 in Minnie65 and from 0.0022 to 0.0312 in H01 (Appendix C). These results support weak morphological ordering at the specified submicrometre scale; they do not establish general morphological independence.

As a secondary check, we asked whether neighbour morphology improved prediction of a spine's measurements beyond its anatomical context. We compared a geometry-only model with a model that additionally used the nearest neighbour's morphology and the mean morphology of neighbouring spines, evaluating both on held-out neurons. For volume, adding these measurements changed R^2 by only 0.0008 in Minnie65 and by -0.0034 in H01, indicating almost no improvement in the mouse dataset and worse prediction in the human dataset. Gains were larger for surface area: $\Delta R^2 = 0.0051$ in Minnie65 and 0.0310 in H01.

To determine whether these gains depended on the actual neighbours, we reassigned morphology within each branch and eight relative-position bins and refitted the models. This control preserves local anatomical context while disrupting the original neighbour assignments. For surface area, the mean gains under reassignment were 0.0050 in Minnie65 and 0.0266 in H01, reproducing much of the improvement obtained with actual neighbours. Across all 17 descriptors, no observed gain exceeded this control at a Benjamini–Hochberg-adjusted threshold of $q < 0.05$ in either dataset (Figure S3).

Thus, neighbour measurements can improve prediction, but the observed benefit is largely consistent with morphology shared by spines in the same branch and positional context. We found no supported additional predictive information specific to the original neighbour assignments. Together with the small excess correlations above, this result supports weak local morphological organization at the tested scale. It does not establish morphological independence, particularly given the lower sensitivity of the H01 test (Appendix E.1). The following experiment examines whether a spatial graph model can nevertheless use neighbouring spines to recover hidden morphology.

4.1 Graph-based recovery of hidden morphology

We tested whether a graph representation of neighbouring spines helps recover hidden morphology in Minnie65. In the graph-based LeJEPA model, nodes represent individual spines and edges connect spines on the same branch separated by less than 0.5 μm. Measurements from some spines were hidden, and the model used the observed spines on their branch to predict the missing information. Evaluation used neurons excluded from training and model selection, with log whole-spine volume as the primary outcome (Appendix H.2).

The graph model showed no supported improvement over prediction from geometry alone and performed worse than a model combining geometry with mean observed branch morphology: the paired difference in R^2 was -0.0069 (95% interval $[-0.0091, -0.0050]$). Removing encoder edges or reassigning morphology within branch and position bins also produced no supported difference in recovery (Figure 3A). The model without encoder edges still had access to observed spines on the same branch, so this comparison tests the added value of the spatial edges. Detailed scores and representation checks are reported in Appendix I.1.

These results provide no evidence that graph connectivity improves hidden morphology recovery in the tested setting. On synthetic data, the model detected an imposed local dependence, but recovered it less effectively than a simple supervised predictor using nearby observations (Figure 3B; Appendix I.2). This limited recovery means that the model's performance on real data cannot independently establish the absence of local biological dependencies. The result concerns the tested model and prediction task; it does not rule out useful graph representations for other purposes.

5 Part III: Morphological similarity and predictive information from repeated contacts

We asked whether spines receiving input from the same axon have similar morphology and whether one contact helps predict the size of another. The Minnie65 comparison included 3,605 target spines with both another contact from the same axon and a different-axon neighbour within 0.5 μm. Both partners were on the target's own dendritic branch, so repeated contacts also shared the postsynaptic neuron. Correlation and prediction analyses used this same population; neural activity was not measured.

Repeated contacts showed positive correlations in volume, surface area and skeletal length, while the volume-to-area interval included zero (Table 2).

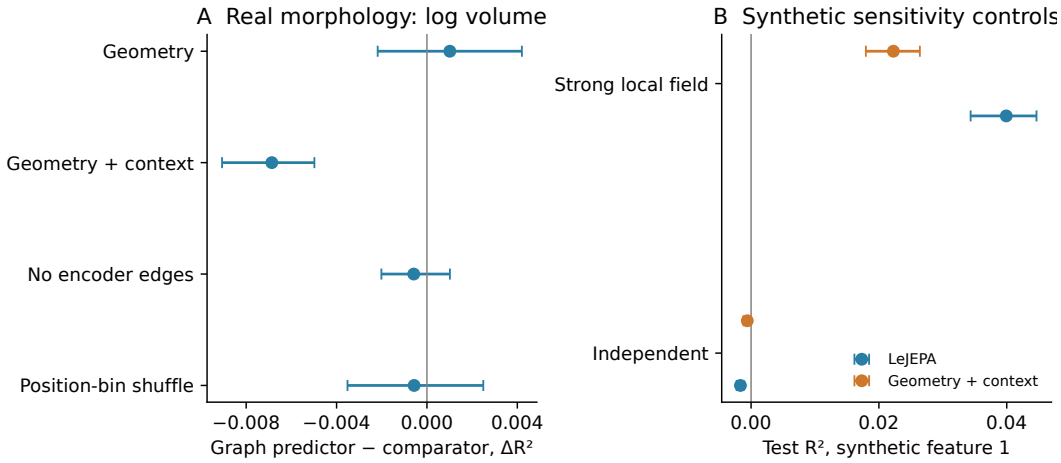

Figure 3: Independent-neuron recovery and sensitivity controls. (A) Paired differences in test log-volume R^2 between decoded graph predictions and each comparator. The no-edge model retains branch-context attention. The shuffle control is retrained and evaluated on reassigned 17-mark vectors within branch and eight relative-position bins. (B) Recovery of synthetic channel 1 under independent and strong locally correlated marks on the same positions. The transparent context comparator uses geometry and mean observed branch morphology. Points average scores from three training seeds; bars are paired 95% intervals from 2,000 resamples of complete test neurons. They do not represent retraining uncertainty or variability across biological replicates.

Table 2: Morphological correlations between contacts of the same axon in Minnie65. Values are connection-weighted Pearson correlations between each target and its nearest same-axon partner on the same branch, with crossed neuron/axon bootstrap 95% intervals. The 3,605 target–partner rows represent 3,036 unique unordered pairs; each connection has equal total weight. Volume, surface area and skeletal length use log-transformed measurements; volume/area retains its linear scale. These descriptive intervals have no multiplicity correction.

Descriptor	Same-axon r [95% CI]
Volume	$0.1393\ [0.0775, 0.1989]$
Surface area	$0.2230\ [0.1639, 0.2804]$
Skeletal length	$0.2022\ [0.1354, 0.2651]$
Volume/area	$0.0430\ [-0.0281, 0.1239]$

These correlations indicate modest morphological resemblance: contacts sharing an axon still vary substantially in size and shape. Their raw correlations cannot be directly compared with the local excess correlations in Part II, which subtract a branch-shuffle reference and use different populations and weights. We therefore tested predictive information on common targets, comparing models with the same geometry and partner-distance inputs. Each model additionally received either no partner morphology, morphology from the different-axon neighbour, or morphology from the same-axon contact. Evaluation held out entire neurons while retaining observed partner measurements as inputs (Appendix A.4).

Adding same-axon morphology improved log-volume prediction over the geometry-and-distance baseline ($\Delta R^2 = 0.0247$, 95% interval $[0.0052, 0.0445]$), although the resulting score remained low ($R^2 = 0.0285$). However, an advantage over the different-axon neighbour was not established: the paired interval for that comparison included zero. The subset of 1,680 targets with an additional distance-matched control likewise showed no supported advantage over either different-axon partner (Figure 4A). Thus, repeated contacts provided information beyond measured geometry, but these comparisons did not establish that they were more informative than spatial neighbours. Exact contrasts and weighting sensitivities are reported in Appendix D.

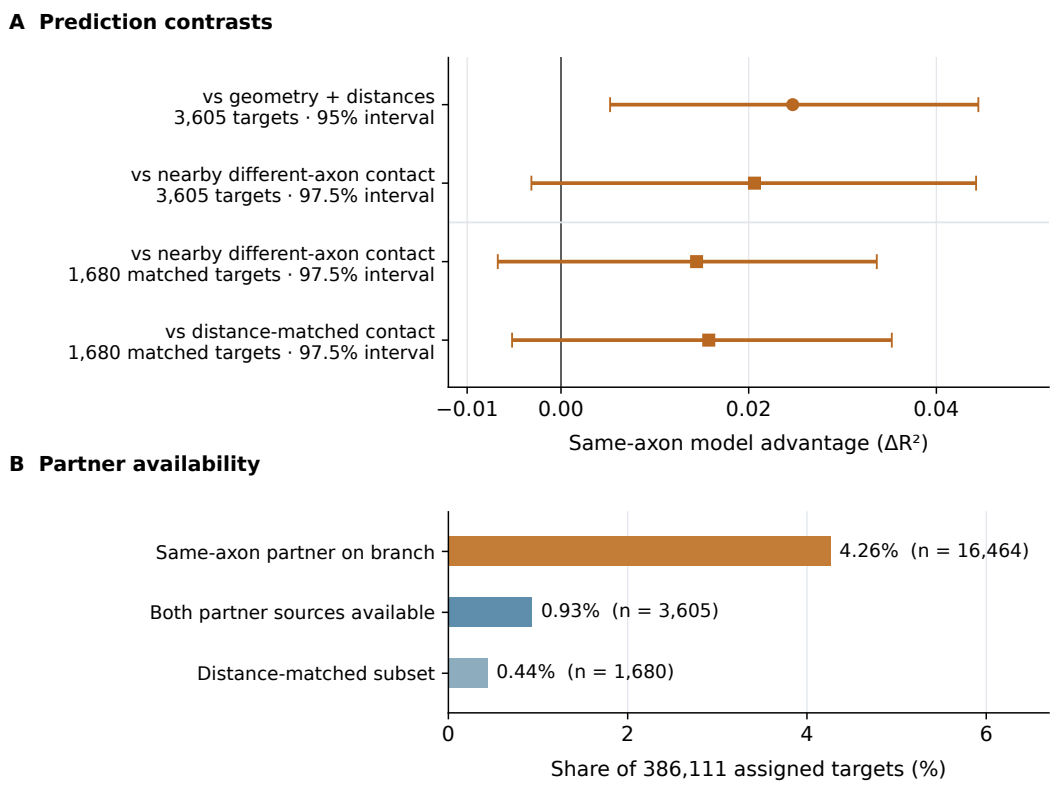

Figure 4: Predictive information from repeated same-axon contacts in Minnie65. (A) Differences in held-out log-volume R^2 between the model using same-axon morphology and each comparator. All models share geometry and partner-distance covariates; each contrast uses identical targets and connection weights. Positive values favour the same-axon model. Error bars use crossed neuron/axon bootstrap intervals: 95% for the increment above the baseline and 97.5% for comparisons between partner sources. The horizontal divider separates the 3,605 eligible targets from the 1,680-target distance-matched subset. The baseline increment is positive, whereas all three partner-source intervals include zero. (B) Availability of partner information as a percentage of all 386,111 strictly assigned targets. Bars represent nested subsets, not disjoint categories. These fractions describe the filtered reconstruction, not the biological prevalence of repeated connections.

The benefit across the full assigned population was limited by partner availability. Only 4.26% of assigned targets had another observed same-axon contact on their branch (Figure 4B; Appendix B). Using same-axon information where available and the geometric baseline elsewhere produced only a small whole-cohort improvement ($\Delta R^2 \approx 0.001$; Appendix D). This improvement includes information from partner availability and distance as well as morphology. The coverage fraction describes the filtered reconstruction and cannot be interpreted as the biological prevalence of repeated connections.

6 Discussion

Our large-scale analysis found no evidence of local spatial clustering in the Minnie65 spine-placement analysis and only weak local morphological resemblance in both mouse and human cortex. Repeated contacts from the same axon nevertheless showed positive morphological correlations and provided predictive information beyond geometry. These findings indicate that proximity alone offers limited information about spine morphology under the tested controls, while shared presynaptic input identifies a further source of anatomical similarity.

The placement results depend on distinguishing broad changes in density from local clustering. After conditioning on branch counts and geometry, models incorporating short-range suppression fitted the

data better than independent placement, although they did not fully reproduce the observed spacing. Differences in tissue, sampling and the density variation preserved by the reference model may therefore help explain discrepant findings from uniform-placement, surface-based and inhomogeneous analyses (Yadav et al., 2012; Morales et al., 2014; Christensen & Møller, 2020). Our comparisons do not identify the biological mechanism of suppression (Appendix F).

Weak local morphological resemblance can coexist with broader gradients in spine size. Our supplementary analysis found that soma-distance patterns depend on dendritic compartment and distance range (Appendix G). Shared branch and positional context can therefore contribute to neighbour similarity, consistent with the importance of accounting for spine-type composition in earlier analyses (Jammalamadaka et al., 2013; Baddeley et al., 2014). Static morphological measurements also leave open the possibility of clustered spine formation, coordinated plasticity and functional co-tuning (Fu et al., 2012; Lu & Zuo, 2017; Scholl et al., 2017).

The repeated-contact findings support the anatomical similarity associated with shared input reported by Bartol and colleagues and subsequent connectomic studies (Bartol et al., 2015; Motta et al., 2019; Dorkenwald et al., 2022). Bartol et al. reported closely matched spine-head volumes in rat hippocampus; our larger cortical sample shows modest correlations in whole-spine morphology. Differences in tissue, measurement definitions, pair selection and statistical weighting limit direct numerical comparison. Our estimates extend the evidence to this cortical population and quantify its uncertainty. The predictive analysis further shows that same-axon morphology adds information beyond geometry, although its advantage over spatial-neighbour morphology remains unconfirmed. Limited availability of observed repeated contacts restricts the benefit across the full sample. These associations concern anatomy; they do not directly measure functional correlations or establish that shared activity causes the resemblance.

The graph experiment found no added benefit of spatial encoder edges for recovering hidden morphology. This is consistent with the weak local associations, but does not establish them as the cause of the model's performance. Although the model detected an imposed synthetic dependence, a simple supervised predictor recovered more of that signal. The result therefore identifies a limitation of the tested learning procedure and supports evaluating spatial graph models against geometry and observed branch-context baselines. Other architectures, training choices and tasks may use the available information differently.

The main limitations concern measurement and sampling. Mesh-based volumes are geometric approximations whose anatomical accuracy has not been independently established. The local predictive tests have limited sensitivity, particularly in H01, so a lack of statistical support does not establish independence. Resampling neurons and axons measures uncertainty within the reconstructed volumes, without providing independent replication across animals or people. The conclusions therefore apply to the sampled tissue, measured properties and tested spatial scales.

AI USE STATEMENT

Generative AI tools assisted with writing code for working with the datasets, English-language editing, manuscript organization, presentation of statistical results, and LaTeX formatting.

REPRODUCIBILITY STATEMENT

The source datasets and morphological measurements are described in Section 2. Appendix A specifies selection criteria, statistical models and uncertainty estimation; the following appendices report analysis populations, sensitivity checks and additional results. Appendices H and I describe the graph-model training, evaluation and controls.

REFERENCES

Adrian Baddeley, Aruna Jammalamadaka, and Gopalan Nair. Multitype point process analysis of spines on the dendrite network of a neuron. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 63(5):673–694, 2014. doi: 10.1111/rssc.12054. URL https://doi.org/10.1111/rssc.12054.

Randall Balestriero and Yann LeCun. Lejepa: Provable and scalable self-supervised learning without the heuristics, 2025. URL https://arxiv.org/abs/2511.08544.

Thomas M. Bartol, Cailey Bromer, Justin Kinney, Michael A. Chirillo, Jennifer N. Bourne, Kristen M. Harris, and Terrence J. Sejnowski. Nanoconnectomic upper bound on the variability of synaptic plasticity. *eLife*, 4:e10778, 2015. doi: 10.7554/eLife.10778. URL https://elifesciences.org/articles/10778.

Brendan Celii, Stelios Papadopoulos, Zhuokun Ding, Paul G. Fahey, Eric Wang, Christos Papadopoulos, Alexander B. Kunin, Saumil Patel, J. Alexander Bae, Agnes L. Bodor, et al. Neurd offers automated proofreading and feature extraction for connectomics. *Nature*, 640(8058):487–496, 2025. doi: 10.1038/s41586-025-08660-5.

Heidi Søgaard Christensen and Jesper Møller. Modelling spine locations on dendrite trees using inhomogeneous Cox point processes. *Spatial Statistics*, 39:100478, 2020. doi: 10.1016/j.spasta.2020.100478. URL https://doi.org/10.1016/j.spasta.2020.100478.

Sven Dorkenwald, Nicholas L. Turner, Thomas Macrina, Kisuk Lee, Ran Lu, Jingpeng Wu, et al. Binary and analog variation of synapses between cortical pyramidal neurons. *eLife*, 11:e76120, 2022. doi: 10.7554/eLife.76120.

Min Fu, Xinzhu Yu, Ju Lu, and Yi Zuo. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. *Nature*, 483:92–95, 2012. doi: 10.1038/nature10844. URL https://doi.org/10.1038/nature10844.

Aruna Jammalamadaka, Sourav Banerjee, Bangalore S. Manjunath, and Kenneth S. Kosik. Statistical analysis of dendritic spine distributions in rat hippocampal cultures. *BMC Bioinformatics*, 14:287, 2013. doi: 10.1186/1471-2105-14-287. URL https://doi.org/10.1186/1471-2105-14-287.

Ju Lu and Yi Zuo. Clustered structural and functional plasticity of dendritic spines. *Brain Research Bulletin*, 129:18–22, 2017. doi: 10.1016/j.brainresbull.2016.09.008.

MICrONS Consortium. Iarpa MICrONS minnie. https://doi.org/10.60533/BOSS-2021-T0SY, 2021. BossDB Archive.

Mite Mijalkov, Giovanni Volpe, Isabel Fernaud-Espinosa, Javier DeFelipe, Joana B. Pereira, Paula Merino-Serrais, et al. Dendritic spines are lost in clusters in Alzheimer's disease. *Scientific Reports*, 11:12350, 2021. doi: 10.1038/s41598-021-91726-x. URL https://www.nature.com/articles/s41598-021-91726-x.

Juan Morales, Ruth Benavides-Piccione, Mor Dar, Isabel Fernaud, Angel Rodríguez, Laura Anton-Sanchez, Concha Bielza, Pedro Larrañaga, Javier DeFelipe, and Rafael Yuste. Random positions of dendritic spines in human cerebral cortex. *Journal of Neuroscience*, 34(30):10078–10084, 2014. doi: 10.1523/JNEUROSCI.1085-14.2014. URL https://doi.org/10.1523/JNEUROSCI.1085-14.2014.

Alessandro Motta, Manuel Berning, Kevin M. Boergens, Benedikt Staffler, Marcel Beining, Sahil Loomba, Philipp Hennig, Heiko Wissler, and Moritz Helmstaedter. Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science*, 366(6469):eaay3134, 2019. doi: 10.1126/science.aay3134.

Benjamin Scholl, Daniel E. Wilson, and David Fitzpatrick. Local order within global disorder: Synaptic architecture of visual space. *Neuron*, 96(5):1127–1138.e4, 2017. doi: 10.1016/j.neuron.2017.10.017. URL https://doi.org/10.1016/j.neuron.2017.10.017.

Alexander Shapson-Coe, Michał Januszewski, Daniel R. Berger, Art Pope, Yuelong Wu, Tim Blakely, Richard L. Schalek, Peter H. Li, Shuohong Wang, Jeremy Maitin-Shepard, et al. A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science*, 384 (6696):eadk4858, 2024. doi: 10.1126/science.adk4858. URL https://doi.org/10.1126/science.adk4858.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations*, 2019.

Aniruddha Yadav, Yuan Z. Gao, Alfredo Rodriguez, Dara L. Dickstein, Susan L. Wearne, Jennifer I. Luebke, Patrick R. Hof, and Christina M. Weaver. Morphologic evidence for spatially clustered spines in apical dendrites of monkey neocortical pyramidal cells. *Journal of Comparative Neurology*, 520(13):2888–2902, 2012. doi: 10.1002/cne.23070. URL https://doi.org/10.1002/cne.23070.

A MEASUREMENT AND ANALYSIS METHODS

This appendix defines the measurements and methods for spine placement (Section 3), local morphology (Section 4) and repeated contacts (Section 5). The subsections follow that order after the common measurement and selection procedures. Appendix B documents partner availability; Appendix C gives the full local morphology results; and Appendix D examines the robustness of repeated-contact comparisons. Appendix E covers predictive-test sensitivity and positional model evaluation. Appendices F and G extend the Discussion with comparisons to previous studies and soma-distance gradients. The graph experiment in Section 4.1 has separate methods in Appendix H and additional results in Appendix I.

A.1 MEASUREMENTS AND SELECTION

The shared morphology threshold is $\log[V/(1\,\mathrm{nm}^3)] \geq 17.67$ (rounded for presentation). Corpus-specific detection floors and alternative selection masks are evaluated separately. Branch occurrence and position models use the Minnie65-specific floor. Analysis populations therefore have different inclusion requirements, as summarized in Table 3.

Table 3: Principal analysis populations. Rows refer to different observation units and overlapping subsets. The spatial position evaluation further requires valid cylindrical coordinates and conditions on branch counts.

Analysis	Corpus	Observations
Branch occupancy and counts	Minnie65	55,721 branches, including empty branches
Local morphology	Minnie65	572,434 spines from 622 neurons
Local morphology	H01	41,931 spines from 64 neurons
Local context prediction	Minnie65	161,530 targets with neighbour context
Local context prediction	H01	13,207 targets with neighbour context
Assigned prediction population	Minnie65	386,111 spines from 622 neurons
Both partner sources available	Minnie65	3,605 targets from 569 neurons
Distance-matched comparison	Minnie65	1,680 targets from 443 neurons
Graph recovery test	Minnie65	34,347 masked spines from 125 neurons

Raw branches are the NEURD prefix identifiers used for exact partner matching. Merged/maximal unambiguous branches combine successive unambiguous single-child segments and are used in the positional risk set, not as replacements for a failed raw-branch match. Mesh lengths, areas and volumes are in nm, nm^2, nm^3; arc distances are in micrometres. Neck SDF is the weighted upstream segmentation descriptor, not assumed to be a physical neck width. Whole-spine V/A is $V/(1000A)$ in micrometres.

The volume estimator centres coordinates, attempts the fixed fan closure, checks watertightness and winding, positive finite volume, bounding-box bounds and translation stability, then uses the convex hull if the attempt fails. Surface areas are measured on the original meshes. Missing descriptors, failed geometry and correspondence failures are separate recorded outcomes. Invalid volume estimates are excluded. Open boundaries at spine attachment are not by themselves segmentation errors. The 17 descriptors in Table 4 enter the morphology analyses; head and neck subvolumes are excluded.

Branch QC excludes length below 1 micrometre, raw density above 5 spines/micrometre or finite calibre above 5 micrometres. Artifact QC uses robust log-volume $z = (\log V - \mathrm{median}(\log V))/(1.4826\,\mathrm{MAD} + 10^{-9})$. A chunk combines centreline offset below 0.30 micrometres, $z > 2$ and elongation above its 90th percentile. Noise combines the same low offset with a face count below half the median in the offset> 0.8, $z < 0$ reference group. The upper-volume artifact rule is $z > 4$. Face counts are used only for this QC, not as morphology predictors. Native detection floors use the fixed KDE bandwidth 0.15 and the left-of-mode crossing at 5% of peak density on 500 grid points between the 0.0005 and 0.999 log-volume quantiles. These are operational selection rules.

A.2 BRANCH OCCURRENCE AND CONDITIONAL POSITION MODELS

The occupancy/count model includes all 55,721 eligible branch units, of which 29,391 are occupied.

Table 4: The 17 morphology descriptors used in the analysis.

Descriptor	Variable name
Whole-spine volume	`volume`
Whole-spine surface area	`area`
Whole-spine skeletal length	`skeletal_length`
Whole-spine volume/area	`spine_volume_to_spine_area`
Bounding-box side, minimum	`bbox_side_min`
Bounding-box side, middle	`bbox_side_mid`
Bounding-box side, maximum	`bbox_side_max`
Number of heads	`n_heads`
Neck surface area	`neck_area`
Neck skeletal length	`neck_skeletal_length`
Neck width	`neck_width`
Neck ray width	`neck_width_ray`
Neck ray width, 80th percentile	`neck_width_ray_80_perc`
Neck segmentation SDF score	`neck_sdf`
Neck bounding-box side, minimum	`neck_bbox_side_min`
Neck bounding-box side, middle	`neck_bbox_side_mid`
Neck bounding-box side, maximum	`neck_bbox_side_max`

We represented these two outcomes with a hurdle model. For branch b, let N_b denote its observed spine count and $Y_b = \mathbb{1}(N_b > 0)$ its occupancy indicator. The model was

$$\operatorname{logit} \Pr(Y_b = 1) = \beta_0 + \boldsymbol{\beta}^{\mathsf{T}} \mathbf{x}_b, \tag{1}$$

$$N_b - 1 \mid Y_b = 1 \sim \operatorname{Poisson}(\mu_b), \qquad \log \mu_b = \gamma_0 + \boldsymbol{\gamma}^{\mathsf{T}} \mathbf{x}_b. \tag{2}$$

Thus the predicted unconditional count was $\widehat{\mathbb{E}}[N_b] = \hat{p}_b(1 + \hat{\mu}_b)$. We compared a length-only specification, using transformed branch length, with a full specification that additionally included soma distance, branch depth, calibre, the number of merged segments, terminal status, dendritic compartment and geometry/QC indicators. Models were cross-fitted with entire neurons held out. Occupancy was evaluated by the area under the receiver operating characteristic curve (AUC), whereas positive counts were evaluated by mean absolute error (MAE) on occupied branches.

The conditional position analysis additionally requires at least two spines and usable spatial geometry; its prepared population contains 18,698 branch units and 388,672 events before cylindrical-geometry and evaluation-fold restrictions.

The second level asked how attachment positions were distributed within an occupied branch after its observed spine count had been fixed. Each attachment was projected onto the branch skeleton and represented on a cylindrical surface by axial position s_i and angle θ_i. For a branch of radius R, pairwise surface distance was

$$d_{ij} = \sqrt{(s_i - s_j)^2 + R^2 \Delta\theta_{ij}^2}, \tag{3}$$

where $\Delta\theta_{ij}$ is the shortest wrapped angular difference. Given the observed count n_b, configurations had conditional density

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_{n_b} \mid n_b) \propto \prod_{i=1}^{n_b} \lambda_b(\mathbf{x}_i) \exp\left[- \sum_{i<j} h(d_{ij}) \right], \tag{4}$$

where λ_b is the fitted first-order density describing broad positional variation. The independent model set $h(d) = 0$; the hard-core model excluded any configuration containing a pair within a fixed radius; and the soft-core model used $h(d) = \kappa \exp[-(d/\delta)^2]$ to penalize, rather than forbid, close pairs. All three models therefore preserved the observed count and the same first-order spatial structure, isolating their assumptions about pairwise spacing. Parameters were selected without using the held-out evaluation neurons, and model adequacy was measured by the fraction of 207 spatial summaries contained in simulation intervals.

A proposal replaces one point by a draw from the first-order density and accepts with probability $\min(1, \exp[E_{old} - E_{new}])$. Hard-core rejects configurations with a pair within its radius. This

is a retrospective evaluation of fixed candidates in the same reconstruction used to develop the model family. Neuron folds used for candidate selection are preserved. New neurons receive a deterministic hash-based fold. First-order density fitting excludes evaluation folds 3–4, including when constructing the nuisance-model training inputs. The fixed sampler and 15 evaluation sweeps generate 30 replicate endpoint summaries; reported intervals are their 2.5/97.5 percentiles. The 207 summaries share branches, distance thresholds and anatomical groups, so coverage is descriptive rather than an independent-test count. Central errors, interval widths and signed deviations are reported in Appendix E.2.

A.3 Local morphology and conditional prediction

Figure S1 summarizes the controls used in Parts II and III.

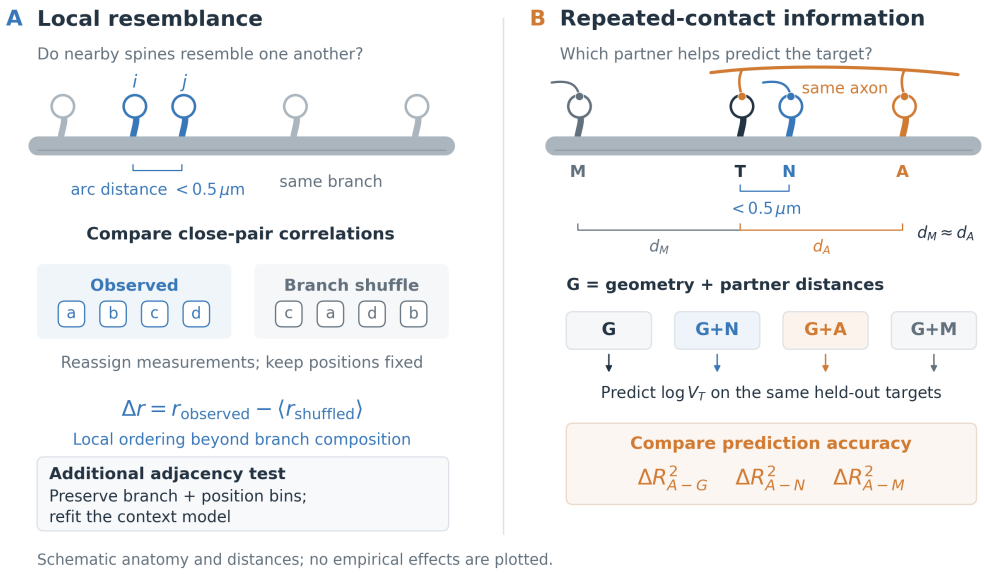

Figure S1: Overview of the morphological comparison workflow. (A) Close-pair correlations are compared with a whole-branch shuffle that reassigns measurements while preserving spine positions; a separate adjacency test preserves branch and relative-position bins and refits the context model. (B) For a target T, N denotes the nearest different-axon contact within 0.5 μm, A the nearest other contact of the same axon, and M a different-axon control matched to A's distance. G contains geometry and partner-distance covariates. The paired contrasts compare prediction of the target's log-volume on common held-out targets. The diagram specifies the comparisons; it does not display empirical effect sizes.

Volume, area, length, width, bounding-box dimensions and counts are transformed as $\log(1 + \max(x, 0))$ using the recorded numerical units. Whole-spine V/A and neck SDF retain their linear scale. The local volume outcome is thus $\log[1 + V/(1\,\mathrm{nm}^3)]$, whereas the connection outcome below is $\log[V/(1\,\mathrm{nm}^3)]$. Local estimates centre marks within raw branches and symmetrize every unordered pair below 0.5 micrometres. There are 999 branch permutations and 2000 neuron bootstrap draws. Rank and winsor thresholds are descriptive panel transforms. The geometry model uses six-knot splines, alpha=10 and five neuron folds; context contains nearest and mean-neighbour values for each of 17 marks. All learned transformations are fitted on the training folds. The predictive analysis uses 161,530 eligible targets in Minnie65 and 13,207 in H01, each with available neighbour context. Its 199 null permutations exchange complete morphology rows within raw branches, either over the whole branch or within eight equal intervals of relative arc position. They hold the target outcomes and coordinates fixed, refit the context model and prohibit permutation of a target's mark into its own neighbour input. These predictive nulls differ from the 999 whole-branch permutations used for descriptive excess correlations. BH families comprise 17 outcomes per corpus, variant and

null. Pooled and neuron-equal R^2 are distinct estimands; bootstrap intervals condition on the fitted out-of-fold (OOF) predictions.

A.4 Morphological correlations and prediction from repeated contacts

The descriptive correlations in Table 2 use the same 3,605 eligible target IDs and fixed connection weights as the predictive comparison. Each target is paired with its saved nearest same-axon partner on the same raw branch. Reciprocal rows are retained: the 3,605 directed target–partner rows cover 3,036 unique unordered pairs and 2,997 presynaptic–postsynaptic connections. Volume, area and skeletal length use $\log(1 + x)$ in the recorded numerical units; whole-spine V/A remains linear. Weighted Pearson correlations use no branch or neuron residualization and no permutation subtraction. Descriptive 95% intervals use 2,000 crossed neuron/axon bootstrap draws, multiplying the fixed weighted moments of each connection by its independently resampled neuron and axon multiplicities. These correlations are distinct from both the excess local correlations in Part II and the predictive scores below.

The main connection outcome is $\log[V/(1\,\mathrm{nm}^3)]$. All models in a common-target comparison contain the same geometric and partner-distance covariates. G is this shared baseline; G+N and G+A add the 17 morphology measurements of the different-axon neighbour and same-axon partner, respectively; G+N+A contains both. A matched different-axon partner is denoted M. Each same-axon pair also shares its postsynaptic neuron because partners are selected on the same raw dendritic branch.

For each target prediction, other observed partner measurements on the held-out neuron remain available. Connection models use five outer and three inner neuron folds, a fixed ridge grid $10^{-3}, \ldots, 10^{6}$, and transformations learned within the corresponding train fold. Every observed connection has total weight one on the comparison mask; neuron-equal weights are a separate refit. The same-axon and different-axon nearest partners are selected without morphology on the same raw branch. The neighbour radius is 0.5 micrometres. Matched partners minimize distance mismatch with a caliper $\max(0.25\ \mu\mathrm{m}, 0.1d_A)$ and no other-branch fallback. Distance standardized mean difference (SMD) must be below 0.1 in absolute value for a balanced-control interpretation. The observed connection-weighted SMD is -0.0001 on the common matched population; the neuron-weighted value is 0.0005. Two key volume contrasts use 97.5% intervals from 2000 paired draws, protecting their joint comparison; exploratory outcomes use descriptive 95% intervals. Crossed draws resample neuron and axon cluster frequencies independently. A positive hierarchy requires both key contrasts to remain positive across the declared weighting and clustering sensitivities. An upper 95% bound below $\Delta R^2 = 0.01$ indicates small gain only for that predictive score and weighting, not biological equivalence.

B Assigned population and partner support

This appendix documents the target selection and coverage underlying Part III (Section 5). The common-threshold Minnie65 morphology population contains 572,434 spines. Of these, 142 lie outside the geometric assignment inventory; other spines may lack a compatible presynaptic assignment. Missing assignments are not inferred. The resulting strict prediction population contains 386,111 targets. Table 5 separates the anatomical level at which a second observed contact from the same axon first appears. Categories are exclusive: same-branch support is counted first, followed by another branch of the same dendrite and then another dendrite of the same neuron.

Table 5: Availability of another observed spine contact from the same axon in the strict common-threshold population. Percentages use all 386,111 assigned targets as the denominator. "Branch" denotes the raw branch identifier used for partner matching.

Location of another same-axon contact	Targets	Percent
No other observed target contact	361,489	93.623
Same branch	16,464	4.264
Another branch, same dendrite	3,628	0.940
Another dendrite, same neuron	4,530	1.173

In total, 24,622 targets (6.38%) had another observed target contact from the same axon somewhere on the reconstructed neuron. Part III uses the stricter same-branch subset. "Eligible" additionally requires an immediate different-axon neighbour; "matched" requires a suitable distance control, and "unsupported" denotes eligible targets without that control. Table 6 gives the size and connectivity coverage of these comparison populations.

Table 6: Support for the common-threshold repeated-contact comparisons. Counts of neurons and axons describe reconstruction coverage, not independent biological replicates.

Population	Targets	Neurons	Axons	Connections
Eligible	3,605	569	2,935	2,997
Distance matched	1,680	443	1,302	1,322
Without matched control	1,925	537	1,820	1,842

The connection-weighted distance SMD on matched targets was -0.0001 (0.0005 with neuron-equal weights), satisfying the prespecified absolute SMD threshold of 0.1. Under the native-threshold sensitivity, 6,285 targets were eligible and 3,307 were distance matched.

C SUPPLEMENTARY LOCAL MORPHOLOGY

The complete prespecified descriptor panel confirms that local excess correlations were small in both corpora (Table 7). The secondary predictive test was also refitted after morphology was permuted within branch-position strata. After correction across 17 outcomes, none of the adjacency-specific gains was supported: the smallest BH-adjusted q was 0.1700 in Minnie65 and 0.2762 in H01. Figure S3 shows the observed gains and permutation reference ranges for volume, surface area and skeletal length.

Table 7: Excess close-pair correlations Δr for the complete 17-descriptor panel. The four principal descriptors and their neuron-bootstrap intervals are shown in Table 1.

Descriptor	Minnie65	H01
Volume	0.0038	0.0082
Surface area	0.0070	0.0278
Skeletal length	0.0112	0.0213
Whole-spine V/A	-0.0006	0.0049
Bounding-box side, minimum	0.0049	0.0117
Bounding-box side, middle	0.0009	0.0166
Bounding-box side, maximum	0.0101	0.0312
Number of heads	0.0005	0.0070
Neck area	0.0130	0.0093
Neck skeletal length	0.0093	0.0072
Neck width	0.0046	0.0063
Neck ray width	0.0050	0.0120
Neck ray width, 80th percentile	0.0023	0.0128
Neck SDF	0.0103	0.0064
Neck bounding-box side, minimum	0.0077	0.0022
Neck bounding-box side, middle	0.0063	0.0064
Neck bounding-box side, maximum	0.0144	0.0137

C.1 EXPLORATORY DEPENDENCE ON PAIR DISTANCE

Figure S2 extends the three whole-spine size measurements to non-overlapping distance bins on the same filtered cohorts. Measurements use the same transformations, branch centring and symmetric pair weighting as the primary analysis. Each bin includes all unordered pairs on the same raw branch with distances in $[0, 0.25)$, $[0.25, 0.5)$, $[0.5, 1)$ or $[1, 2)$ μm. Excess correlations subtract the mean of 999 whole-branch permutations. Descriptive 95% intervals use 2,000 neuron bootstrap draws, with the same mean-moment and offset construction as the primary analysis. Combining the first two bins

reproduces the primary < 0.5 μm estimates and intervals. Pair counts across the four bins are 54,421, 52,454, 105,868 and 233,450 in Minnie65, and 7,265, 5,595, 8,179 and 17,042 in H01.

Point estimates remain small across these bins. H01 surface-area estimates decrease with distance, but a common monotonic decline is not apparent across all three descriptors. No formal between-bin contrast was performed; changing pair support and branch composition also limits interpretation of a spatial decay scale. This exploratory extension does not replace the prespecified < 0.5 μm comparison across 17 descriptors.

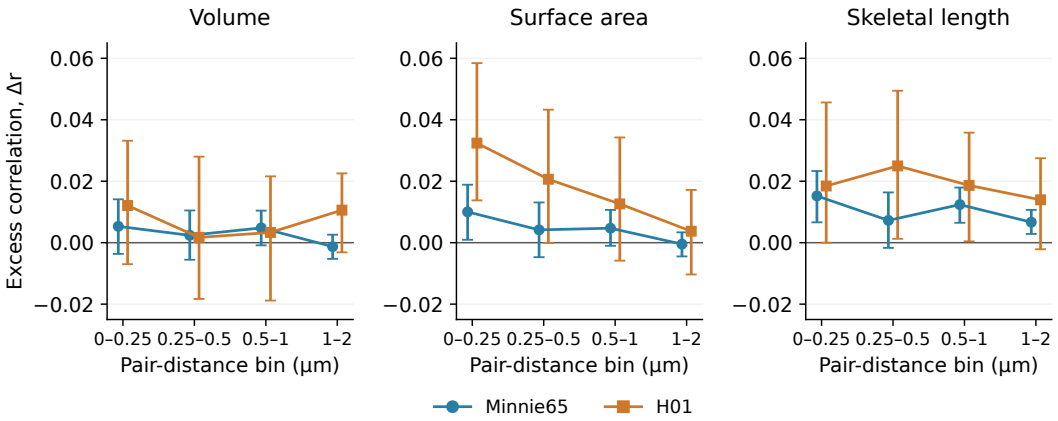

Figure S2: Exploratory distance dependence of local morphological resemblance. Excess correlations Δr are shown for log-transformed whole-spine volume, surface area and skeletal length in Minnie65 (blue circles) and H01 (orange squares). Error bars are descriptive 95% neuron-bootstrap intervals without multiplicity correction; zero denotes no excess correlation relative to the whole-branch permutation control. Distance bins include their lower boundary and exclude their upper boundary. They are displayed as equally spaced categories despite unequal physical widths; connecting lines guide the eye. Pair support differs between bins, and the plot does not establish a characteristic decay scale.

C.2 Predictive spatial-context control

Figure S3 compares actual-neighbour prediction gains with the branch-position permutation reference for three whole-spine descriptors. These results use the original 17-outcome correction family and complement the descriptive distance analysis above.

D Connection and coverage sensitivities

This appendix expands the repeated-contact comparisons in Section 5. Table 8 separates the effects of target selection, weighting and uncertainty estimation from the choice of partner information. "Connection" weights give each observed presynaptic–postsynaptic pair equal total weight; "neuron" weights give each neuron equal total weight. The resampling column specifies neuron-only or crossed neuron/axon resampling. Here A denotes the same-axon partner, N the nearest different-axon neighbour and M the distance-matched different-axon control. The contrasts A–N and A–M compare G+A with G+N and G+M, respectively. G includes the same partner-distance covariates in each common-target comparison. In the population hybrid, G uses geometry alone; the partner model additionally uses same-axon distance and morphology and is applied only where a peer is available.

The neuron-only intervals support some positive partner-source contrasts, but the crossed neuron/axon intervals include zero for all of them. Accounting for uncertainty across both neurons and shared axons therefore does not establish an advantage of same-axon information over either alternative partner source, consistent with the conclusion in Part III.

The whole-population hybrid gain was 0.00101 (crossed 95% interval $[0.0007, 0.0013]$).

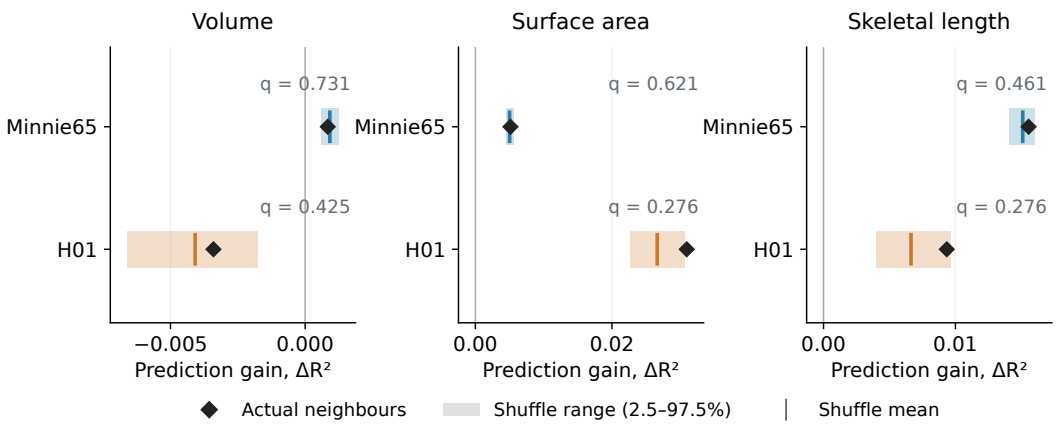

Figure S3: Neighbour information beyond shared spatial context. Black diamonds show the held-out prediction gain ΔR^2 from adding neighbour morphology to the geometric baseline for volume, surface area and skeletal length. Coloured bands show the 2.5–97.5 percentiles of gains from 199 refitted permutations of morphology within branches and eight relative-position bins; vertical ticks mark their means. Bands are permutation reference ranges, not confidence intervals for observed gains. Blue denotes Minnie65 and orange denotes H01. Horizontal scales differ between descriptors. The displayed q values use Benjamini–Hochberg correction across the original 17 outcomes in each corpus. The H01 area gain slightly exceeds its reference range but is not supported after this correction ($q = 0.276$). Lower sensitivity of the H01 test limits the interpretation of non-rejection.

Table 8: Key volume contrasts across common support, weights and clustering.

Weights	Targets	Contrast	Resampling	ΔR^2	97.5% interval	
					Lower	Upper
Connection	Eligible	A–N	Neuron	0.0206	0.0076	0.0339
Connection	Eligible	A–N	Neuron/axon	0.0206	-0.0032	0.0442
Connection	Matched	A–N	Neuron	0.0144	0.0025	0.0252
Connection	Matched	A–M	Neuron	0.0157	0.0038	0.0263
Connection	Matched	A–N	Neuron/axon	0.0144	-0.0067	0.0337
Connection	Matched	A–M	Neuron/axon	0.0157	-0.0052	0.0353
Neuron	Eligible	A–N	Neuron	0.0183	0.0064	0.0304
Neuron	Eligible	A–N	Neuron/axon	0.0183	-0.0030	0.0380
Neuron	Matched	A–N	Neuron	0.0058	-0.0088	0.0199
Neuron	Matched	A–M	Neuron	0.0147	0.0007	0.0290
Neuron	Matched	A–N	Neuron/axon	0.0058	-0.0199	0.0291
Neuron	Matched	A–M	Neuron/axon	0.0147	-0.0128	0.0406

E NULL SENSITIVITY AND POSITIONAL EVALUATION

E.1 SENSITIVITY OF THE LOCAL PREDICTIVE TEST

This check assesses how reliably the local prediction procedure in Section 4 detects an imposed spatial dependence. It is separate from the graph-model synthetic controls in Appendix I.2. The simulation evaluates prediction and permutation for one synthetic outcome, with detection defined by the unadjusted permutation $p \leq 0.05$. It does not calibrate the 17-outcome Benjamini–Hochberg decision. Synthetic marks combine a standardized positional mean, independent noise and a spatial field:

$$y_i = \mu_i + \sqrt{1-\alpha^2}\,\epsilon_i + \sqrt{\alpha^2}\,z_i.$$

At $\alpha^2 = 0.04$, the spatial field receives 4% of the stochastic variance weight, rather than 4% of total mark variance including the positional mean. Table 9 reports detection rates for this scenario across the tested field lengths. Detection was much lower in H01 than in Minnie65, supporting the more cautious interpretation of the H01 negative result in Part II. The eight-interval zero-field rejection frequency is 1/30 in Minnie65 (exact 95% interval $[0.001, 0.172]$) and 0/30 in H01 ($[0.000, 0.116]$). Detection frequencies for nonzero fields depend on their variance and spatial length. More local permutations preserve more of the neighbour context. Neither non-rejection nor input preservation alone establishes morphological independence.

Table 9: Detection frequency of the eight-position-bin predictive test for a simulated spatial field with $\alpha^2 = 0.04$. Each entry is based on 30 realisations; zero-field rejection is included as a calibration check.

Field length (μm)	Minnie65	H01
0.5	0.833	0.100
1.0	1.000	0.133
2.0	1.000	0.133
5.0	0.967	0.133
Zero field	0.033	0.000

E.2 POSITIONAL MODEL EVALUATION

The occurrence/count scores use held-out neurons. Positional coverage is the fraction of dependent spatial summaries captured by each model's simulation interval. The family-level table also reports central error, interval width and signed error so that coverage can be interpreted alongside prediction accuracy (Table 12).

Table 10: Branch occupancy discrimination and occupied-branch count error. Occupancy uses all 55,721 branches; count MAE uses the 29,391 occupied branches.

Model	Occupancy AUC	Count MAE (spines)
Length alone	0.9129	11.7056
Length and anatomical context	0.9699	7.4612

Table 11: Coverage of conditional positional summaries on held-out branches.

Interaction	Covered	Total	Coverage
Independent	83	207	0.4010
Hard-core	97	207	0.4686
Soft-core	128	207	0.6184

Table 12: Positional summary families in the primary cohort: coverage, mean absolute error (MAE), interval width and signed error.

Interaction	Summary family	Count	Coverage	MAE	Interval width	Signed error
Hard-core	Cable nearest neighbour	45	0.5333	0.0255	0.0480	0.0035
Hard-core	Cable pairwise	45	0.4444	0.0145	0.0374	-0.0023
Hard-core	First-order density	27	0.7778	0.0088	0.0332	-0.0000
Hard-core	Surface nearest neighbour	45	0.3778	0.0181	0.0289	0.0020
Hard-core	Surface pairwise	45	0.3333	0.0108	0.0245	-0.0011
Independent	Cable nearest neighbour	45	0.3556	0.0371	0.0503	0.0239
Independent	Cable pairwise	45	0.3556	0.0160	0.0408	0.0071
Independent	First-order density	27	0.8519	0.0079	0.0364	-0.0004
Independent	Surface nearest neighbour	45	0.3333	0.0293	0.0399	0.0227
Independent	Surface pairwise	45	0.2889	0.0135	0.0315	0.0076
Soft-core	Cable nearest neighbour	45	0.5778	0.0222	0.0496	-0.0063
Soft-core	Cable pairwise	45	0.5778	0.0145	0.0394	-0.0077
Soft-core	First-order density	27	0.7778	0.0084	0.0318	0.0008
Soft-core	Surface nearest neighbour	45	0.5778	0.0134	0.0363	-0.0038
Soft-core	Surface pairwise	45	0.6444	0.0095	0.0280	-0.0044

F Relation to previous clustering studies

The studies below are not treated as interchangeable replications. They use different tissues, imaging modalities, spine definitions and null models. This section therefore identifies which conclusions are in tension with ours and which address a different biological quantity.

F.1 Positional clustering and the reference model

Yadav and colleagues reported an excess of spatial clusters on terminal apical dendrites relative to uniform placement along each branch (Yadav et al., 2012). A Yadav-style C-score sensitivity did not reproduce terminal enrichment in the available Minnie65 geometric proxies: 8 of 93 terminal branches (8.6%) and 244 of 2,467 oblique branches (9.9%) exceeded $C > 0.90$, close to the nominal 10% expectation. In H01, a uniform positional null gave a submicrometre component-occupancy ratio of 1.165; retaining the observed four-bin intensity profile reduced it to 1.001, although a small occupancy excess remained near 2 μm.

These checks show why the apparent contradiction depends on the comparison. A uniform null can interpret a broad intensity gradient as clustering, whereas a position-conditioned null asks whether pairwise structure remains after that gradient is preserved. Our branch definitions are reconstructed geometric proxies, and the tissue, species and cortical region differ from those of Yadav et al.; the sensitivity therefore does not formally refute their result. The spatial-randomness result of Morales et al. (Morales et al., 2014) is likewise conditional on its dendritic sample and surface-based null, rather than evidence for a universal placement process.

F.2 Static morphology versus plasticity and functional tuning

Reports of nearby spine formation during motor learning (Fu et al., 2012), clustered synaptic plasticity (Lu & Zuo, 2017) and local receptive-field co-tuning (Scholl et al., 2017) concern temporal change or activity-defined groups. Such transient, task-dependent or input-specific organization need not produce a strong correlation among all mature spine shapes in a static EM volume. Conversely, our weak average excess correlation does not establish the absence of functional clusters. It narrows the claim to static morphological ordering at the declared submicrometre scale and under the branch- and position-conditioned controls.

Jammalamadaka et al. reported associations between discrete neighbouring spine types (Jammalamadaka et al., 2013). Our continuous and count descriptors address a related but nonidentical estimand. The small effects we observe after controlling branch composition are therefore best interpreted as a limit on average morphological resemblance, not as a direct rejection of all neighbour-type associations.

F.3 Methodological stress test of localized spine loss

We also reanalysed the dendrite-level summaries published by Mijalkov et al. (Mijalkov et al., 2021). Because their original coordinates and analysis code were unavailable, this is a stress test of identifiability rather than a reproduction of the original analysis or part of the main inferential argument. Conditional Poisson patterns fixing each published dendrite's spine count and length still produced approximately 10–12 contiguous Louvain communities with modularity 0.67–0.71; published community size lay within the null 95% interval in five of six case-by-τ groups.

Independent deletion left the grouping coefficient near zero, whereas local depletion windows increased it. Near-complete physical windows fit the published attack trajectories better than literal fixed-rank blocks in 92.3% and 99.7% of bootstrap comparisons for the two block targets. These summary-level results are compatible with spatially localized loss, but they do not uniquely identify either a clustered healthy baseline or deletion of fixed adjacent blocks. They consequently qualify the mechanistic interpretation of that work without overturning its observation of localized depletion.

G Soma-distance gradients of spine size

Spine size varied with distance from the soma, but the direction and magnitude depended on dendritic compartment, distance range and the comparison being made. This supplementary analysis addresses a broad gradient in mean morphology, separately from the resemblance of nearby spines in Part II.

G.1 Population and comparisons

We used the same filtered Minnie65 morphology population as Part II and selected 143,632 apical and 194,761 basal spines from the same 384 neurons. Compartment labels follow an operational dendritic-tree length rule: among neurons with at least two non-axon trees, a longest-to-second-longest length ratio of at least two identifies the longest tree as apical and the others as basal. Uncertain, axon-like and unlabelled trees are excluded. The labels are geometric proxies for dendritic compartment.

The distance variable is path length along the dendritic skeleton from the soma to the spine attachment, in micrometres. Regressions used natural log volume or natural log surface area as the outcome, with distance divided by 100 μm as the predictor of interest. The adjusted model included neuron fixed effects, log branch calibre, log branch length and relative position along the branch. Individual spines received equal weight. We fitted the full observed distance range and restricted ranges of 0–200 and 0–300 μm; Table 13 reports the full and 0–300 μm adjusted comparisons. The effect is expressed as $100[\exp(\beta) - 1]$ percent change per additional 100 μm. Intervals use 2,000 bootstrap resamples of neurons and have no multiplicity correction.

G.2 Compartment and distance-range dependence

Across the full observed range, basal volume decreased by 6.80% per additional 100 μm (95% interval $[-8.30, -5.40]$%), whereas the apical volume estimate was close to zero, $+0.29$% ($[-0.25, 0.82]$%). Surface area decreased in basal dendrites and showed a small positive full-range apical slope. Restricting the analysis to 0–300 μm revealed decreases in both volume and surface area in both compartments: the volume estimates were -3.93% per 100 μm apically and -8.44% basally.

Table 13: Soma-distance gradients after accounting for neuron identity and branch geometry. Values are percentage changes per additional 100 μm, with neuron-bootstrap 95% intervals. Each row includes 384 neurons. The full-range and restricted fits describe different distance distributions.

Compartment	Spines	Volume, % [95% CI]	Area, % [95% CI]
Full observed soma-distance range			
Apical	143,632	0.29 $[-0.25, 0.82]$	0.40 $[0.04, 0.76]$
Basal	194,761	-6.80 $[-8.30, -5.40]$	-2.62 $[-3.49, -1.77]$
Soma distance below 300 μm			
Apical	128,992	-3.93 $[-4.70, -3.23]$	-0.68 $[-1.17, -0.17]$
Basal	193,989	-8.44 $[-9.62, -7.24]$	-3.20 $[-3.94, -2.42]$

The compartment comparison is descriptive; no direct apical-minus-basal interaction test was performed. Distance coverage also differed: the 5th–95th percentile range was 43–376 μm apically and 31–151 μm basally. A single full-range slope therefore summarizes different portions of the arbor in the two compartments.

G.3 Within-branch and paired-distance comparisons

Regressions with branch fixed effects address changes within a branch, without the geometric covariates of the adjusted neuron model. These full-range volume slopes were negative in both compartments: -5.21% per 100 μm apically (95% interval $[-6.49, -4.00]\%$) and -10.87% basally ($[-11.97, -9.70]\%$). Corresponding surface-area slopes were -3.44% ($[-4.19, -2.73]\%$) and -4.21% ($[-5.04, -3.38]\%$). The negative within-branch apical estimate and the near-zero full-range adjusted apical volume estimate thus concern different levels of anatomical variation.

We also compared mean log size at 150–200 μm and 50–100 μm within neurons represented in both distance bins. This comparison gives neurons equal weight and reports a finite-distance contrast, without geometric adjustment (Table 14). Volume was lower in the farther bin in both compartments. Surface area was lower basally, while the apical area interval included zero.

Table 14: Within-neuron comparison of farther and nearer soma-distance bins. Values are percentage differences in geometric mean size for 150–200 μm relative to 50–100 μm, with neuron-bootstrap 95% intervals. Only neurons represented in both bins contribute. These are finite-distance contrasts, not slopes per 100 μm.

Compartment	Neurons	Volume, % [95% CI]	Area, % [95% CI]
Apical	376	-3.24 $[-5.05, -1.46]$	-0.86 $[-2.33, 0.64]$
Basal	300	-6.61 $[-9.08, -4.11]$	-3.75 $[-5.79, -1.77]$

Together, these comparisons show compartment- and scale-dependent distal-size trends. They describe the selected reconstructed spines and do not identify a universal distance law or its biological mechanism. Variation in mean size over the arbor can coexist with the small excess correlations between immediate neighbours reported in Part II.

H Graph representation experiments and recovery controls

This appendix describes the graph experiment in Section 4.1. The architecture, neuron-disjoint recovery procedure and controls are presented first. The earlier 35-descriptor model is documented separately in Appendix H.5; its results are not pooled with the primary recovery experiment.

H.1 Encoder and prediction objective

A shared three-layer residual GIN encoder (Xu et al., 2019) mapped each spine to 32 coordinates, with a frozen random input projection. A two-head cross-attention predictor received target positions and observed context embeddings. Training minimized a weighted sum of latent mean squared prediction error and a characteristic-function regularizer (SIGReg), with regularizer weight 0.01. This is our graph adaptation of the LeJEPA regularization approach (Balestriero & LeCun, 2025). There was no exponential-moving-average teacher. The regularizer acted on encoded context and target measurements, not on predicted embeddings. Thus regularized, nonconstant encoder outputs can coexist with nearly constant predictor outputs.

H.2 Independent-neuron recovery on the 17-descriptor panel

The experiment in Section 4.1 starts from the 572,434 Minnie65 spines in Part II. Removing 6,803 spines on singleton branches leaves 565,631 spines on 29,815 branches from 622 neurons. A single fixed random split assigns 373 neurons to training, 124 to validation and 125 to test. The respective target counts are 101,892, 35,215 and 34,347. Masks hide the rounded 30% of each branch, clamped to leave at least one target and one context node. The same masks are used in training, validation and

testing across all compared models and seeds, without overlap between these neuron populations. This design conditions on one neuron split and one set of masks.

The primary outcome is natural log whole-spine volume. The remaining 16 descriptors use the declared morphology transformations, including log-one-plus for nonnegative extensive quantities and the unchanged linear ratios/SDF values. Descriptor means and scales, geometric imputations and positional scales are estimated using training neurons only. The predictor's three positional inputs are relative skeleton-branch position, log-one-plus soma distance and log-one-plus branch length. Neither these inputs nor the context contain the hidden target's morphology.

Encoder edges join distinct nodes on the same branch at arc separation strictly below 0.5 μm and have unit weights. The full graph contains 106,875 unordered pairs, matching the Part II close-pair population exactly. Context and target induced subgraphs are encoded separately. The predictor may attend to all unmasked nodes on its own branch, but never to another branch in the batch. Its target embeddings serve only as prediction labels. Consequently this experiment tests the full branch-context procedure, not an encoder whose entire receptive field is limited to 0.5 μm.

The three real-data variants use these physical edges, no encoder edges, or physical edges after jointly reassigning the 17-dimensional morphology vectors within branch and eight relative-position bins. Relative position uses skeleton branch length. The reassignment moves 70.5% of node vectors and preserves sparse strata, including singletons. Both training and evaluation in this control use the reassigned outcomes; it is a refitted conditional-null comparison rather than prediction of original labels after corrupting their context.

We use the three-layer, 32-coordinate encoder and two-head predictor, dropout 0.05, SIGReg weight 0.01 and 32 random projection slices. Three training seeds are evaluated per variant. AdamW uses learning rate 0.01, weight decay 10^{-5} and cosine decay over at most 200 epochs. Batches contain at most 256 branches, grouped by size. Training stops after 15 epochs without improvement in validation loss. Checkpoint selection minimizes validation JEPA loss with dropout disabled and fixed regularizer projections; it never uses test recovery scores.

H.3 Decoding, baselines and uncertainty

Separate standardized Ridge decoders, with penalty 1, are fitted on training hidden targets to map predicted latents or directly observed target latents to the 17 morphology outcomes. A third decoder uses the same-seed encoder at random initialization. Direct-target decoding measures information retained when the target morphology is supplied; it is not hidden-target recovery. The random-encoder comparison is necessary because the frozen random input projection and residual path can preserve morphological information before representation learning.

Transparent comparators predict from the training-target mean, from the structural covariates of Part II, or from those covariates plus the mean morphology of all available unmasked spines on the branch. The latter never uses hidden target measurements. Geometry models and decoders use training data only. We report test R^2 for each descriptor, and latent MSE skill relative to the mean target latent estimated from training neurons.

Paired 95% intervals use 2,000 resamples of complete test neurons, retaining spine weights within each resample and recomputing the R^2 denominator. Reported mean scores and contrasts average the three seed-specific scores, not their predictions. Identical neuron resamples are used for each contrast. Seed ranges are reported separately. The intervals condition on the fitted models and do not capture variability across neuron splits, mask realizations, retraining or independently sampled tissue. Secondary descriptor intervals are descriptive and do not support multiplicity-adjusted discovery claims.

H.4 Synthetic sensitivity controls

The same branch positions, masks, neuron split and graph-model training procedure are used for two 17-channel synthetic controls, each with three training seeds. The first assigns independent standard Gaussian values. The second generates independent stationary Ornstein–Uhlenbeck channels along each branch, with correlation length 2 μm, 95% field variance and 5% independent noise. Synthetic normalization uses training neurons only. Channel 1 is the primary synthetic endpoint; it is not

biological volume. Detectability requires a positive lower 95% bound for its graph recovery R^2 increment over the training-mean baseline. Failure of this control limits the negative biological interpretation of the neural result.

H.5 Historical models and the meaning of predictor collapse

The exploratory models used 35 classical descriptors, including historical head and neck measurements. These inputs differ from the validated 17-mark panel in Parts II and III.

Historical graphs used Euclidean distances between centroid coordinates standardized within each branch. The radius parameter $r = 1.5$ is therefore dimensionless. Setting $r = 0$ removes encoder edges but retains the predictor's context attention. Historical edge distances were min–max transformed within each encoded subgraph to obtain convolution weights. Saved training configurations specify 10% masking; the archived recovery experiment instead masked approximately 30% of nodes, with at least one context and target node. It examined the first 3,000 graph files and retained 1,806 graphs with at least four nodes, yielding 16,304 targets from 41 neurons. These are exploratory diagnostic targets, not the current Part II population or an independent test set.

The historical recovery scores used a target-sample mean for latent MSE skill and an ordinary five-fold Ridge decoding analysis without grouping spines by neuron. The classification probes used mean-pooled neuron representations, balanced logistic regression and ten repeats of five-fold stratified cross-validation. Their 95% intervals resample neurons. Holding out a neuron from the classifier does not establish that it was absent from encoder pretraining. No paired significance test between the archived classification estimates is claimed here.

We loaded the saved encoder and predictor weights strictly, replayed the historical masks and computed the variance of each latent coordinate across targets after centring that coordinate. Covariance spectra provide a separate assessment of encoder collapse. We also checked how adding an unrelated graph changed predictions. The historical predictor attended across the full batch during training, whereas the archived recovery evaluation processed individual graphs. Our replay preserves this difference when documenting the historical model; the neuron-disjoint experiment above explicitly restricts attention to each graph.

I Additional graph-model results and audit

These results support the graph-recovery analysis in Section 4.1. We first report recovery and representation controls on held-out neurons, followed by the synthetic benchmark. The audit of the earlier 35-descriptor model is retained in Appendix I.3 as exploratory background.

I.1 Recovery and representation controls on held-out neurons

On 34,347 hidden test spines from 125 neurons, the graph predictor's mean decoded log-volume R^2 was 0.0168 $[0.0122, 0.0215]$. Its paired increment above geometry was 0.0010 $[-0.0022, 0.0042]$, and above geometry plus observed branch-context morphology it was -0.0069 $[-0.0091, -0.0050]$. Figure 3 shows these contrasts alongside the no-edge and position-bin reassignment controls. The paired graph-minus-no-edge difference was -0.0006 $[-0.0020, 0.0010]$, and the graph-minus-reassignment difference was -0.0006 $[-0.0035, 0.0025]$. Neither comparison supported an added benefit of the original spatial encoder edges. These intervals quantify test-neuron uncertainty conditional on the fitted models; Table 15 reports the range across training seeds.

Directly observed learned embeddings yielded log-volume decoding R^2 0.8271 $[0.8234, 0.8306]$, compared with 0.9095 $[0.9074, 0.9115]$ from the corresponding randomly initialized encoders. The paired learned-minus-random increment was -0.0824 $[-0.0851, -0.0797]$. This comparison concerns encoder information retention when target measurements are supplied, and must not be read as recovery from neighbours.

Recovery scores are not directly comparable with Part III, whose partner-supported targets and available inputs differ. Table 16 extends the primary volume comparison to all 17 descriptors; Table 17 reports predictor variance and encoder diagnostics for real and synthetic data.

Table 15: Independent-neuron log-volume evaluation. Scores average the three seed-specific R^2 values. Bootstrap intervals reflect test-neuron variation; the last column shows the separate range over training seeds. Observed-target rows supply the hidden measurement to the encoder and are information-retention controls, not recovery results.

Method	Mean R^2 [95% CI]	Seed range
Training mean	-0.0004 $[-0.0023, -0.0000]$	$[-0.0004, -0.0004]$
Geometry	0.0158 $[0.0118, 0.0198]$	$[0.0158, 0.0158]$
Geometry + observed context	0.0237 $[0.0185, 0.0288]$	$[0.0237, 0.0237]$
Graph predictor	0.0168 $[0.0122, 0.0215]$	$[0.0141, 0.0201]$
No-edge predictor	0.0174 $[0.0126, 0.0219]$	$[0.0131, 0.0196]$
Position-bin shuffle predictor	0.0174 $[0.0123, 0.0225]$	$[0.0163, 0.0188]$
Observed learned target	0.8271 $[0.8234, 0.8306]$	$[0.7653, 0.8635]$
Observed random target	0.9095 $[0.9074, 0.9115]$	$[0.8979, 0.9178]$

Table 16: Recovery across the current 17 descriptors, using the transformations specified in the methods. Volume is the primary endpoint; intervals for the other descriptors are descriptive and are not adjusted for multiple comparisons. Each cell gives a three-seed mean and a 95% neuron-bootstrap interval.

Descriptor	Graph R^2	Geometry + context R^2	Paired difference
volume	0.0168 $[0.0122, 0.0215]$	0.0237 $[0.0185, 0.0288]$	-0.0069 $[-0.0091, -0.0050]$
area	0.0309 $[0.0211, 0.0410]$	0.0404 $[0.0304, 0.0507]$	-0.0095 $[-0.0124, -0.0064]$
skeletal_length	0.0801 $[0.0601, 0.0997]$	0.0880 $[0.0675, 0.1078]$	-0.0078 $[-0.0128, -0.0027]$
spine_volume_to_spine_area	0.0071 $[0.0043, 0.0099]$	0.0098 $[0.0069, 0.0128]$	-0.0026 $[-0.0042, -0.0013]$
bbox_side_min	0.0226 $[0.0157, 0.0301]$	0.0285 $[0.0215, 0.0362]$	-0.0059 $[-0.0083, -0.0035]$
bbox_side_mid	0.0438 $[0.0317, 0.0557]$	0.0513 $[0.0388, 0.0633]$	-0.0074 $[-0.0109, -0.0038]$
bbox_side_max	0.0726 $[0.0545, 0.0907]$	0.0801 $[0.0614, 0.0989]$	-0.0075 $[-0.0123, -0.0025]$
n_heads	0.0055 $[0.0030, 0.0079]$	0.0095 $[0.0066, 0.0125]$	-0.0040 $[-0.0057, -0.0023]$
neck_area	0.0223 $[0.0161, 0.0280]$	0.0271 $[0.0206, 0.0331]$	-0.0049 $[-0.0074, -0.0023]$
neck_skeletal_length	0.0391 $[0.0272, 0.0512]$	0.0448 $[0.0326, 0.0573]$	-0.0057 $[-0.0088, -0.0026]$
neck_width	0.0208 $[0.0157, 0.0255]$	0.0255 $[0.0201, 0.0307]$	-0.0047 $[-0.0068, -0.0027]$
neck_width_ray	0.0262 $[0.0192, 0.0331]$	0.0299 $[0.0226, 0.0369]$	-0.0037 $[-0.0060, -0.0016]$
neck_width_ray_80_perc	0.0218 $[0.0154, 0.0278]$	0.0247 $[0.0180, 0.0312]$	-0.0030 $[-0.0049, -0.0011]$
neck_sdf	0.0379 $[0.0284, 0.0466]$	0.0409 $[0.0315, 0.0494]$	-0.0030 $[-0.0054, -0.0008]$
neck_bbox_side_min	0.0202 $[0.0134, 0.0269]$	0.0252 $[0.0185, 0.0323]$	-0.0050 $[-0.0072, -0.0028]$
neck_bbox_side_mid	0.0385 $[0.0271, 0.0492]$	0.0449 $[0.0335, 0.0552]$	-0.0064 $[-0.0094, -0.0035]$
neck_bbox_side_max	0.0600 $[0.0450, 0.0743]$	0.0667 $[0.0515, 0.0812]$	-0.0067 $[-0.0106, -0.0026]$

Table 17: Graph-model diagnostics averaged over three seeds. Coordinate variances are centred across targets. Effective rank describes observed encoder outputs; relative parameter change refers to trainable encoder layers, excluding the frozen input projection. Parameter movement establishes optimization, not improvement in morphology decoding.

Data	Predicted variance	Encoded variance	Effective rank	Weight change
Real morphology	0.00522	0.8250	26.98	3.711
Independent synthetic	0.00110	0.8283	26.46	3.271
Strong synthetic field	0.02099	0.8403	26.41	3.233

I.2 Local supervised benchmark for synthetic sensitivity

The neural procedure recovered an imposed local synthetic field with channel-1 $R^2 = 0.0399$ $[0.0343, 0.0447]$, whereas independent synthetic marks yielded $R^2 = -0.0017$ $[-0.0023, -0.0012]$. These results establish sensitivity to the imposed signal but do not quantify how much of its predictable component the model recovered.

After inspecting the original synthetic results, we added a transparent supervised benchmark to assess how much of the imposed signal the neural procedure recovered. This is a post-hoc sensitivity analysis, not a change to the original neural model selection or sensitivity criterion. For each masked target, we averaged synthetic channel 1 over up to k nearest observed spines on its own branch, ranked by absolute along-branch distance. There is no radius cutoff; all hidden spines are excluded. A ridge predictor uses geometry and this local mean. Scaling and ridge fitting ($\alpha = 1$) use training masked targets only. We choose $k \in \{1, 2, 4, 8, 16\}$ by validation channel-1 R^2, separately for each synthetic condition, with ties resolved towards smaller k. The selected fit is evaluated without refitting on validation, using the existing test neurons and masks. Intervals and paired neural-minus-local differences use the same 2,000 test-neuron resamples as the original analysis; neural scores are averaged across the three seeds. This supervised channel-specific benchmark is not an architecture-matched ablation or a test on real morphology. Table 18 shows that it recovered more of the imposed local signal than the neural model, explaining the limited interpretation of the synthetic check in Section 4.1.

Table 18: Post-hoc synthetic local benchmark. The stronger recovery by a simple local predictor limits interpretation of the neural sensitivity check: detecting the imposed signal does not mean using it effectively.

Synthetic data	Selected k	Local R^2 [95% CI]	LeJEPA minus local
Independent marks	8	-0.0003 $[-0.0007, -0.0000]$	-0.0014 $[-0.0020, -0.0008]$
Strong local field	1	0.2276 $[0.2118, 0.2432]$	-0.1877 $[-0.2007, -0.1747]$

I.3 Historical model audit and classification controls

The following results concern the exploratory model described in Appendix H.5. They use different descriptors, graph construction and evaluation populations from the main recovery experiment in Appendix I.1.

In the exploratory 35-descriptor LeJEPA model, the predictor produced almost constant outputs. For 16,304 masked targets, latent prediction skill was approximately zero, and the archived decoded morphology score was $R^2 = -0.001$. In contrast, directly observed target embeddings retained morphology information (archived decoding $R^2 = 0.591$ with encoder edges and 0.287 without edges). These are exploratory, non-neuron-disjoint recovery scores, not generalization estimates.

Replaying the saved checkpoint confirmed that the graph predictor's mean coordinate variance across targets was 2.2×10^{-5}, compared with 0.737 for the encoded targets. Thus the predictor collapsed while the encoder remained nonconstant. This does not establish that encoder training learned a useful representation: morphology can also survive a random projection and residual path. The historical subtype probe likewise showed no demonstrated improvement over mean-pooled descriptors (Table 19). Excitatory/inhibitory (E/I) classification and count controls are reported in Table 21 and the accompanying text.

Table 19: Historical five-class excitatory-subtype prediction on 622 neurons. Scores use neuron-mean representations and repeated classifier cross-validation; intervals resample neurons. Encoder pretraining was not neuron-disjoint from this probe. The chance balanced accuracy is 0.20.

Representation	Balanced accuracy [95% CI]
Mean-pooled descriptors	0.617 $[0.580, 0.655]$
LeJEPA, encoder edges	0.584 $[0.545, 0.623]$
LeJEPA, no encoder edges	0.595 $[0.557, 0.633]$

The historical model is not a direct implementation of the submicrometre comparison in Part II. Its standardized-coordinate radius corresponds to a median along-branch edge distance of 9.63 μm in the audited sample, and its reconstructed training and validation sets share neurons.

The old encoder training split was reconstructed from the recorded seed and current sorted graph files. All 622 excitatory neurons occur in both the training and validation graph sets. Of the 41 neurons contributing to the replayed recovery sample, 38 occur in pretraining; 1,343 of its 1,806 graphs are training graphs. The historical file-list manifest is unavailable, so these are explicit reconstructions rather than a claim of exact historical membership. Recovery counts and near-zero latent skill reproduce the saved diagnostic output.

In the historical graph sample, 557,703 unique undirected edges have mapped arc coordinates; only 2.41% are shorter than 0.5 μm. The 5th–95th percentile arc-distance range is 1.00–29.40 μm. The centroid-distance median is 9.28 μm. These measurements refer to historical standardized- coordinate graphs, not to the new physical-radius graphs.

The old predictor's response changes when unrelated graphs enter its attention batch. The new graph-local implementation reproduces single-graph predictions to numerical tolerance when unrelated graphs are added. Historical checkpoint loading was strict; the nearly constant output is not due to missing predictor weights. Table 20 distinguishes the variance of predictions from the variance of encoded targets.

The recorded training seed also reproduces the frozen input projection exactly. Relative to this validated initialization, trainable encoder-layer weights changed substantially (relative Euclidean changes 3.25 and 3.70 for the edge and no-edge models). The encoder therefore underwent optimization. This parameter audit does not establish beneficial representation learning; the random-encoder decoding comparison addresses that separate question.

Table 20: Replayed historical recovery. Variances are first computed across targets for each coordinate and then averaged. MSE skill here uses the historical target-sample mean, unlike the training-mean baseline of the new experiment.

Model	Latent skill	Predicted variance	Encoded variance
Encoder edges	-0.00053	0.000022	0.7373
No encoder edges	-0.00019	0.000099	0.9087

Table 21: Historical E/I classification: 685 neurons (622 excitatory, 63 inhibitory), repeated classifier cross-validation and neuron-bootstrap intervals. Noise replaces normalized descriptors while preserving graph positions and structure. High E/I accuracy does not by itself demonstrate recovery of local morphology.

Representation	ROC-AUC [95% CI]
Spine count	0.963 $[0.929, 0.990]$
Mean-pooled descriptors	0.998 $[0.996, 1.000]$
All-cell graph encoder	0.999 $[0.998, 1.000]$
Same encoder, noisy descriptors	0.973 $[0.941, 0.996]$

The persistence of high graph-probe accuracy with noisy descriptors is consistent with structural and count-related information contributing to this endpoint. Morphology remains informative after count controls: using exactly 20 sampled spines per neuron gives AUC 0.985 $[0.972, 0.993]$, and fold-safe count-residualized morphology gives 0.910 $[0.859, 0.950]$. These additional controls use five repeats of five-fold CV and 1,000 neuron bootstraps; they are not numerically identical to the ten-repeat probe above.