Neuron Geometry Underlies a Universal Local Architecture in Neuronal Networks

Eyal Gal[1], Rodrigo Perin[2], Henry Markram[2], Michael London[1,3] and Idan Segev[1,3]

[1]The Edmond and Lily Safra Center for Brain Sciences, and [3]Department of Neurobiology, The Hebrew University of Jerusalem, Jerusalem, Israel.

[2]Blue Brain Project, École Polytechnique Fédérale de Lausanne, Geneva, Switzerland

Communication: Eyal Gal – eyal.gal@mail.huji.ac.il

Keywords: neocortex, microcircuit, connectome, morphology, topology, motifs, network science, generative models, spatially-embedded networks, random networks

ABSTRACT

Cortical microcircuits in a variety of brain regions express similar, highly nonrandom, synaptically-connected cell triplets. The origin of these universal network building blocks is unclear and was hypothesized to result from plasticity and learning processes. Here we combined *in-silico* modeling of dense cortical microcircuits with electrophysiological/anatomical studies to demonstrate that this recurring connectivity emerges primarily from the anisotropic morphology of cortical neurons and their embedding in the cortical volume. Using graph-theoretical and machine-learning tools, we developed a series of progressively more complex generative models for circuit connectivity that considers the geometry of cortical neurons. This framework provided predictions for the spatial alignment of cells composing particular cell-triplets which were directly validated via *in-vitro* whole-cell 12-patches recordings (7,309 triplets) in the rat somatosensory cortex. We concluded that the local geometry of cortical neurons imposes an innate, highly structured, global network structure, a skeleton upon which fine-grained structural and functional plasticity processes take place.

INTRODUCTION

Structural connectivity in neuronal circuits plays a key role in governing the network dynamics and the emerging brain functions[1,2]. Connectomes at both the microscale, the mesoscale and the macroscale exhibit structured, highly nonrandom, connectivity patterns[3–7]. A glimpse into the microscale connectivity diagram in several neocortical microcircuits was provided recently using *in vitro* whole-cell multi-patch recordings[3–5]. The circuits display specific reoccurring local connectivity patterns such as bidirectional connections and specific cell triplet configuration (motifs) that are much more prevalent than expected in reference random models (**Fig. 1**).

In the absence of a concrete theory about the principles underlying these structured connectivity patterns in neuronal microcircuits, the emergence of such patterns was considered as being shaped by specific genetic markers[8,9] or by plasticity and learning processes[5,10–12]. But another possibility comes to mind. Most types of cortical neurons have highly asymmetric geometry with dendritic and axonal trees typically extending in different directions. It could be the case that the anisotropic geometry of cortical neurons, and their relative embedding in the cortical volume *per se* "enforces" the complex profile of cortical motifs. If this was the case, then specificity and learning processes might operate on top of an innate, already structured, cortical skeleton rather than on a *tabula rasa* network structure. This will strongly constrain the degree by which these processes could further shape network connectivity.

Here we explored this challenging possibility by first using a reference circuit of a dense *in silico* reconstruction of a 0.3 mm^3 from rat somatosensory cortex (the neocortical circuit, NMC), consisting of ~30,000 3D-reconstructed neurons interconnected by ~40 million synapses[13]. The connectivity within the NMC is based on a set of pairwise geometrical rules. The profile of over/under represented triplet motifs emerging in the NMC was similar to those found experimentally[14]. Through a systematic construction of a series of progressively more complex generative geometrical circuit models we showed that, indeed, the asymmetrical geometry of neuron's morphology, and their embedding in cortical 3D space, is sufficient to explain the profile of triplet motifs in the NMC. We next studied experimentally the appearance of triplet motifs using 12-patch whole-cell recordings from layer 5 thick tufted pyramidal cells (L5-TTPC) in the rat somatosensory cortex. We found a close match between the physical position of L5-TTPCs forming a specific triplet motif in the NMC and their position in the biological tissue. This implies that the local geometry of L5-TTPCs and their distribution in the 3D cortical volume impose an innate global and highly structured skeleton in the L5-TTPCs circuit. We argued that plastic processes and/or genetic processes that shape specific cortical connectivity add second-order refinements to the already existing innate connectivity pattern of cortical microcircuits.

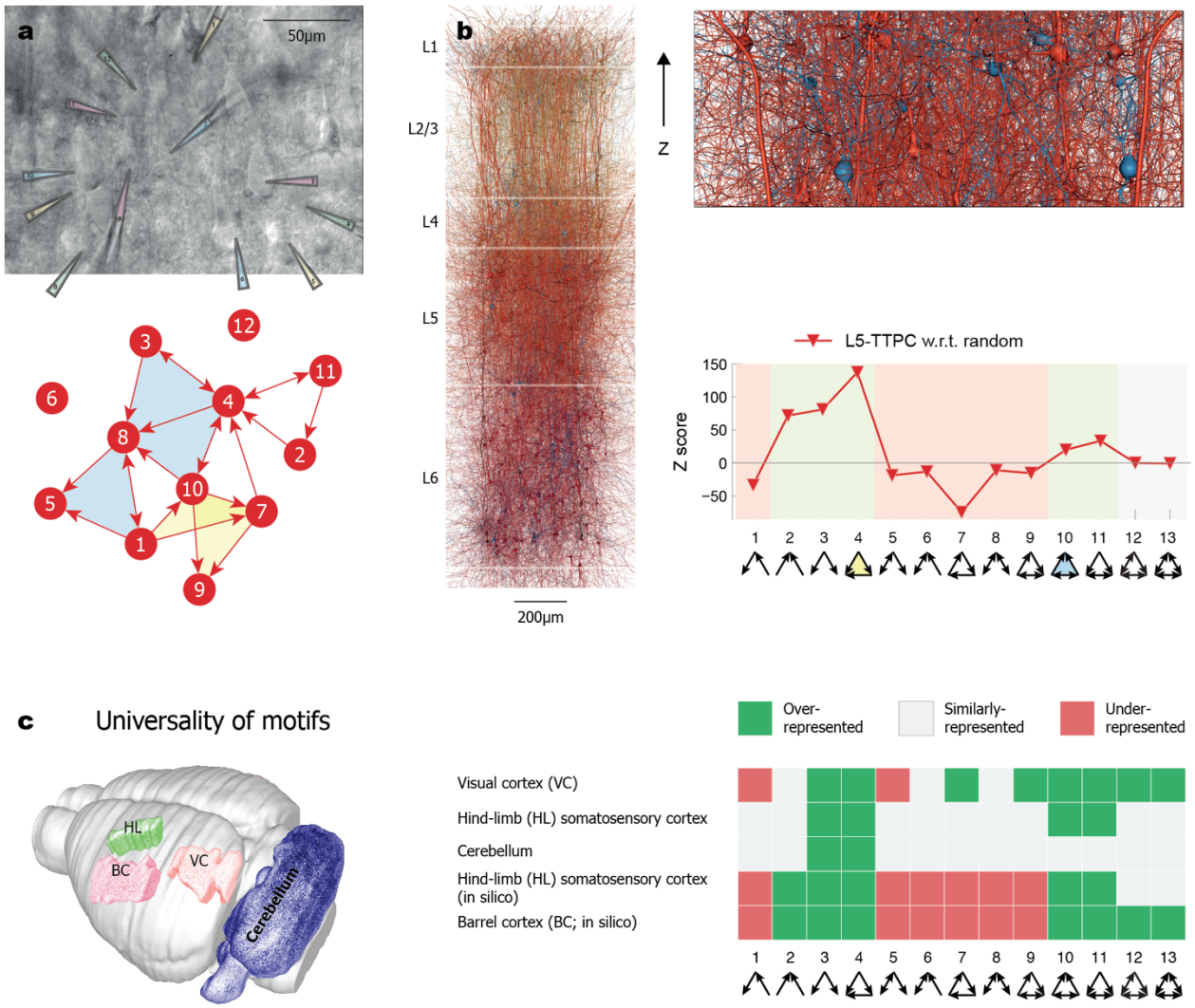


Figure 1. Universality of triplet motifs appearance in a variety of mammalian neuronal microcircuits. a, Differential contrast video image of *in vitro* 12-patches recording from L5 thick tufted cortical pyramidal cells in rat somatosensory cortex (top). Triggering a spike in a presynaptic cell and recording excitatory post-synaptic synaptic potentials in the post-synaptic cells enabled to establish circuit connectivity diagram (schematically depicted at the bottom) and to count the frequencies of specific triplet motifs (two types of motifs are highlighted by light green and yellow). **b**, Left - dense *in silico* reconstruction of neocortical microcircuit (NMC) from the rat somatosensory cortex[13]; its volume is of ~0.3 mm^3 with ~40 million synapses. Top right - a slice in L5 from the NMC circuit shown at left, with L5 thick tufted pyramidal cells (L5-TTPC) colored in red and L5 Martinotti cells (L5-MC) in blue. Lower frame - graph theoretical analysis of the connectivity diagrams for the two cell types shown above, highlighting the overrepresented triplet motifs (e.g., motifs #3 and #4), the underrepresented (e.g., motif # 7 for L5-TTPCs) ones, and those that are similarly represented, as compared to the respective distance-dependent random network (**Methods**). Triplet #4 is marked in yellow and triplet #10 in light blue, as in a, bottom. **c**, Universality of overrepresented triplet motifs in a variety of neuronal microcircuits. Note that triplets #3 and #4 are overrepresented in all circuits. The brain regions from which these microcircuits reside in is depicted schematically at left; the expression of the various triplet motifs in these circuits is shown at right. Triplet motifs expression data is displayed for primary visual cortex[3]; somatosensory cortex[4]; interneurons in the cerebellum[5]; dense *in silico* somatosensory cortex[14], and for dense *in silico* barrel cortex[15].

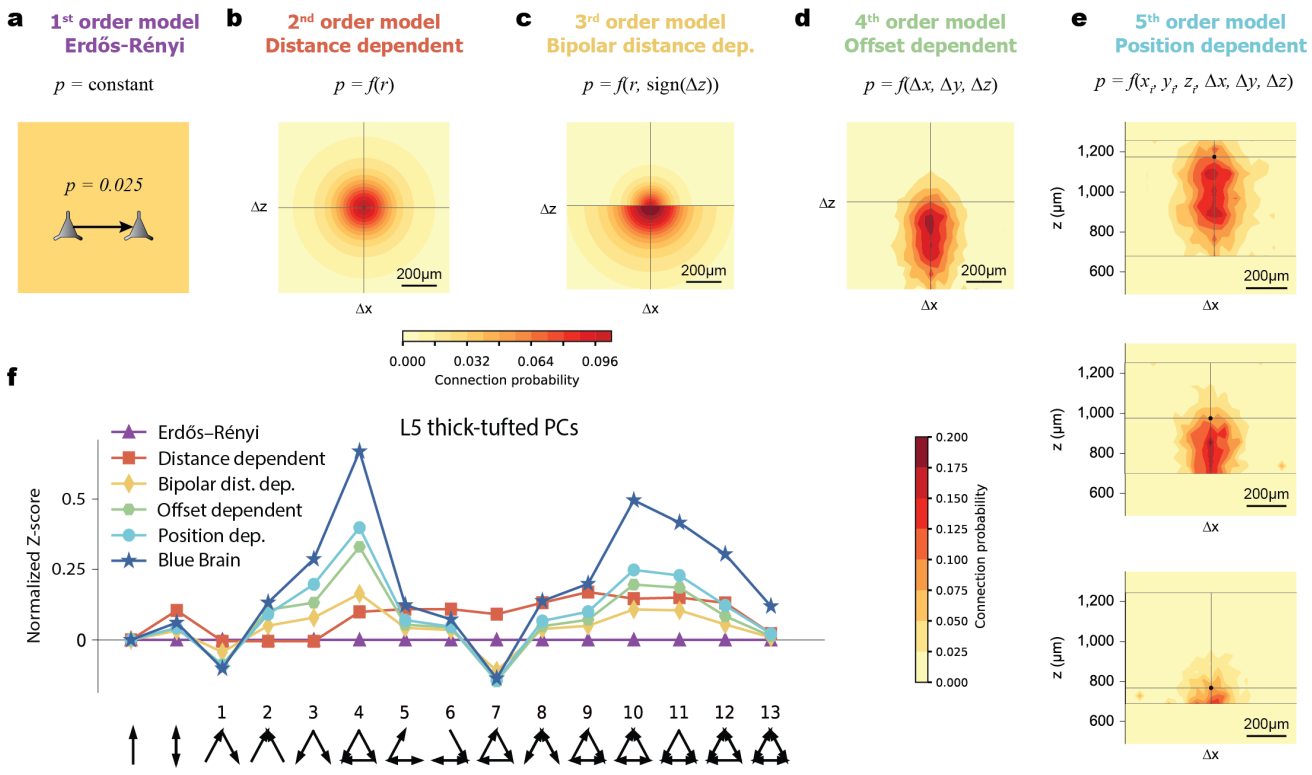


Figure 2. The anisotropic connectivity among cortical neurons and their spatial embedding in cortical volume underlie the emergence of triplet motifs in the NMC. a-e, a series of progressively more complex generative models for the circuit connectivity of L5 thick-tufted pyramidal cells (L5-TTPC). **a**, 1st order Erdős-Rényi model, with uniform connection probability, p = 0.025, among the cells as is the average connection probability in the NMC. **b**, 2nd order model, in which p (color coded) is distance dependent, as derived from the NMC. p is depicted with respect to the distance from the postsynaptic cell located at the origin (Δx, Δz intersection point). Note the decrease in p with distance. **c**, 3rd order "bipolar" model, in which p also depends on whether the postsynaptic neuron is above or below the presynaptic cell (along the z-axis). Note that, in the NMC, p is larger for downwards connection. **d**, 4th order "offset dependent" model, in which p depends also on the relative direction, in 3D, to the postsynaptic cell as it does in the NMC. Note that p is larger for postsynaptic neurons that are directly below (along the z-axis), as compared to cells that are obliquely below the presynaptic cell. **e**, 5th order model, in which p depends on the absolute position in 3D of both the presynaptic and the postsynaptic cells. E.g., L5-TTPCs whose somata are located close to the border of layer 4 (top horizontal line) have a larger span, in the z-direction, for having a postsynaptic partner as compared to L5-TTPCs whose somata reside near the border of layer 6 (bottom). **f**, Distribution of the various triplet motifs for each of the models shown in a-e, with respect to the reference Erdős-Rényi model (x-axis). Motifs distribution in the NMC is shown by the blue line. Note that, as the geometry captured by model becomes more realistic, their motifs appearance becomes closer to that of the NMC (blue line with stars). See Methods for detailed explanation regarding the construction of the different statistical models used in this figure.

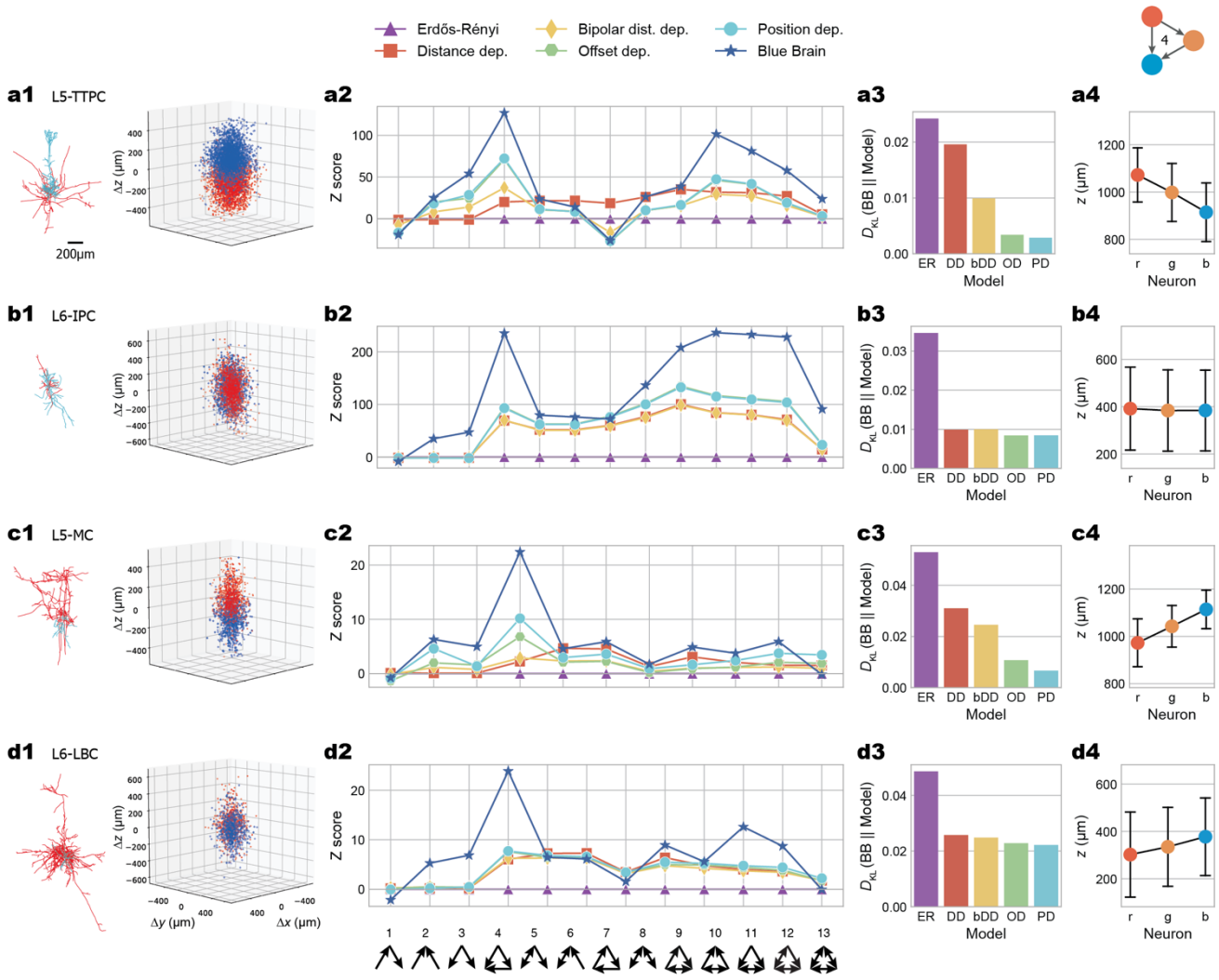


Figure 3. Impact of morphology of different cortical cell types on the emergence of triplet cell motifs in the NMC. a, Appearance of triplet motifs for L5-TTPC. **a1**, Left, an exemplar 3D reconstructed TTPC cell. Right, the relative soma position of all the presynaptic TTPC cells (blue) and all the postsynaptic TTPC cells (red) with respect to the cell somata centered at 0,0 (superposition of 50 TTPC cells). **a2**, Distribution of triplet motifs for TTPC cells (n=2,003). **a3**, Kullback–Leibler divergence between the motif distribution observed in the NMC and each of the 5 models shown in a2. Note the gradual improvement with model complexity. **A4**, spatial embedding of neurons composing motif #4 (mean ± s.d.). **b-d**, As in a, but for (**b**) L6 inverted pyramidal cells (n = 3,476), (**c**) L5 Martinotti cells (n=395) and (**d**) L6 large basket cells (n=463). Note that when the neuron is more isotropic as in **b** and **d**, models beyond the distance-dependent model do not provide further improvement.

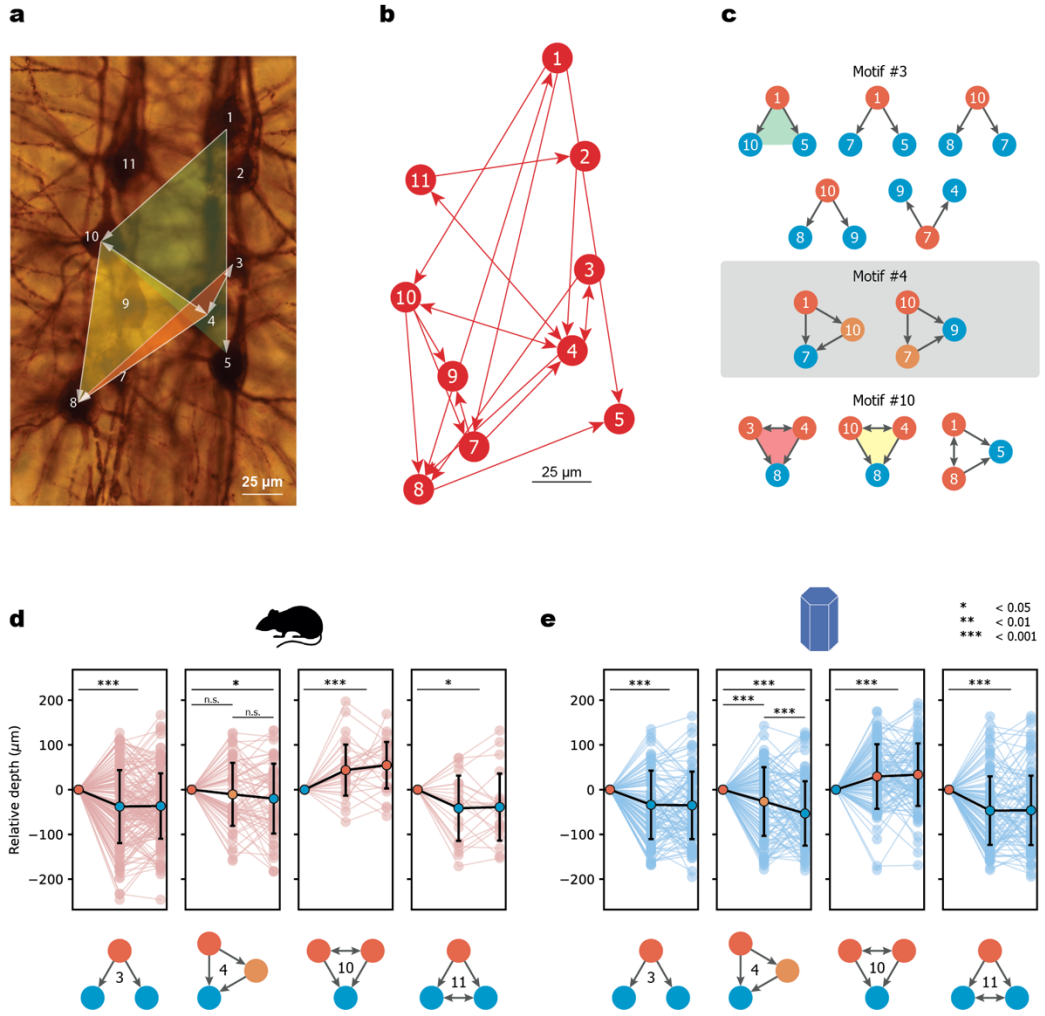


Figure 4. Experimental validation that neuron morphology and 3D position explain the appearance of triplet cortical motifs. **a**, Biocytin staining of multi-patch recordings from 12 L5-TTPCs. Three exemplar connected triplet cells are shown by the transparent (green, yellow and red) triangles. **b**, Connectivity map for the 12-neurons circuit shown in a. **c**, All instances for three triplet motifs (#3, #4 and #10) that were extracted from the circuit diagram shown in b. The relative depth (z axis) of the cells composing these triplets is preserved. E.g., note that, for triplet #3, in most (but not in all) cases the presynaptic cell (red circle) is located above the two postsynaptic cells (blue circles). Green, yellow and pink triangles correspond to the respective cell triplets shown in a. **d**, Spatial embedding of triplet motifs (#3, #4, #10 and #11) for L5-TTPCs in the experimental data set (n_3 = 163, n_4 = 63, n_{10} = 31, n_{11} = 21 triplets); the relative depth of the cells forming individual triplets (light brown lines) and mean ± s.d. (black line). For clarity, cell depths were aligned in each triplet according to one member (the left-most cell in the plot). *P < 0.05, **P < 0.01, ***P < 0.001. A paired-samples t-test. P values for R-vs-Y and Y-vs-B in motif #4 are 0.2470 and 0.3029, respectively. **e**, As in e but for a 400µm-thick slice from the NMC (n_3 = 219,377, n_4 = 22,137, n_{10} = 833 and n_{11} = 765 triplets of which only 100 instances are shown). Note the very close similarity between experiments (in e) and model (in d).

RESULTS

Universality of triplet motifs appearance in a variety of mammalian neuronal microcircuits

The motivation that ignited this study is summarized in **Figure 1**. It demonstrates the universality of the appearance of over/under represented of synaptically-connected cells triplets (motifs) in a variety of cortical networks and in the cerebellum. These same triplet motifs also appear in two different densely reconstructed *in silico* neocortical microcircuits (NMCs), whose building blocks are 3D reconstructed neurons of various types, positioned at their respective 3D cortical volume. Experimentally, identifying synaptically-connected triplet motifs in local microcircuits utilizes *in vitro* whole-cell patch-recordings (**Fig. 1a**, top). From such recordings it is possible to extract a connectivity graph for this 12-cell circuit (**Fig. 1a**, bottom). The frequency of these same triplet motifs could be measured in the NMC; an example is shown in **Figure 1b** for the NMC of the rat the somatosensory cortex[13]. This ~0.3mm^3 reconstructed cortical volume contains ~30,000 neurons interconnected by ~40 million intrinsic synapses and distributed in 6 cortical layers (**Fig. 1b**, left). Layer L5 of this NMC is shown in **Figure 1b**, right; layer 5 thick tufted pyramidal cells, L5-TTPC, are colored in red and L5 Martinotti cells (L5-MC) in blue. **Figure 1b** bottom shows the analysis of the frequency of appearance of the triplets for L5-TTPC, demonstrating the highly nonrandom occurrence of specific triplet motifs (**Methods**). The different microcircuits (three experimental systems and another *in silico* reconstruction) in which the statistics of such cell triplets where characterized is shown in **Figure 1c**; the universality of overrepresentation of the same triplet motifs in all these circuits, most notably motifs #3, #4, #10 and #11, is highlighted by the yellow squares in **Figure 1d**.

Origin of triplet motifs in dense *in silico* cortical microcircuits

What could be the origin for such universality in the prevalence of overrepresented triplet motifs? This question is examined theoretically using the NMC as a benchmark circuit. Connectivity in the NMC is based on the proximity of axons and dendrites ("touch detection" algorithm) together with a set of synaptic pruning rules, which are set to maintain an observed distribution of contacts per connection but are consistent with purely geometry driven connectivity pattern[13,16]. Therefore, the emergence of triplet motifs in this circuit must arise from certain geometrical rules (see Discussion). We uncovered these rules by building a series of network models which progressively take into account higher-order features of the geometry underlying neuronal connectivity in the NMC, exploring whether these triplet motifs will emerge in these generative network models as they do in the NMC target circuit (**Fig. 2, Supplementary Fig. 1**).

We started by focusing our attention on L5-TTPCs and constructed the simplest, 1st order, statistical network: an Erdős-Rényi random network with the same average connection probability of $p = 0.025$ as in the NMC (**Fig. 2a**, and the x-axis in **Figure 2f**, see also [3–5,14,15]). The over- and under- represented triplet motifs in the NMC, as compared to the Erdős-Rényi reference circuit, are depicted by the blue line in **Figure 2f**.

Next, we constructed a 2nd-order statistical network, which considers the dependence of connection probability on intersomatic distance as in the NMC (**Figure 2b** and red line in **Figure 2f**, see also[4,5,14]). This statistical circuit does not explain the profile of motifs that were found in the NMC. In particular, motif #7 which is the underrepresented in the NMC is overrepresented in the distance-dependent statistical circuit.

The next 3rd-order model considers also the implication of the asymmetrical geometry of L5-TTPCs whose dendrites tend to project upwards towards the pia (in the z-direction) and their axons tends to project in the opposite direction. Because the prerequisite for the formation of a synapse is that the axon of the presynaptic cell should come close to the dendrite of the post-synaptic cell, the probability of connection between pairs of L5-TTPCs in the NMC is larger in the downwards (z) direction. Indeed, taking into account this bipolar effect when quantifying the connection probability uncovers highly anisotropic connectivity function (**Fig. 2c**). This model starts to capture the representation trends of the observed motifs. Specifically, in this "bipolar" distance-dependent model, motifs #3, #4, #5, and #6 are overrepresented with respect to Erdős-Rényi reference model, whereas motif #7 is underrepresented, as is the case of the NMC (**Fig. 2f**, yellow line). It is important to note that, in both Erdős-Rényi and distance-dependent models, the probabilities for motif #4 and motif #7 are identical. Thus, by taking into consideration the anisotropy imposed by the bipolar nature of L5-TTPCs, the symmetry between motifs #4 and #7 is properly broken.

In our 4th-order model, we further added the dependency of connection probability on the 3D offset between the connected cells (**Figure 2d** demonstrated the offset in a x-z cross section); this further improved the similarity of motif appearances between the model and the NMC (green line in **Figure 2f**). To ensure that the improvement indeed results from extra information captured by the geometric offset features, and not merely as a consequence of a higher number of features, the probability functions of all geometrical models were fitted using a machine learning classifier having the same capacity (**Methods**). Finally, in our 5th order model, we considered the absolute positions of the L5-TTPCs within layer 5. It was found that L5-TTPCs that are located at the upper part of L5 have larger probability span, in the z-direction, for having a postsynaptic L5-TTPCs partner as compared to L5-TTPCs whose somata are located deeper in layer 5 (**Fig. 2e**). When this effect was taken into account in the 5th order model, the reproduced motif profile was even closer to that of the NMC (**Fig. 2f**, light blue line).

Geometrical considerations explain the emergence of cortical motifs

The very close similarity between triplet distribution in experiments and *in silico* (geometrically-based) NMC model, strongly suggests that the emergence and frequency of triplet motifs in the biological tissue can be explained, to the most part, based solely on geometrical considerations. The emergence of triplet motifs for highly polar neurons, such as L5-TTPCs, depends strongly on their exact bipolar structure and position in 3D (**Fig. 3a**). When the neuron is more isotropic, as is the case for the inverted L6 pyramidal neurons (in which dendrites and axons tend to overlap), then models beyond the 2nd-order distance-dependent model do not provide further improvement for explaining the triplet distribution of networks composed of such cells (**Fig. 3b**).

This implies that, in cortical microcircuits, processes such a "cortical specificity" and/or structural plasticity play a second order role in shaping motifs structure and frequency. We thus concluded that the skeleton of motifs is dictated, innately, by geometrical considerations; namely by the position of the neurons in 3D and the asymmetry of their morphology.

Interestingly, the series of the generative network models provided an additional prediction for the embedding of particular motifs in space. For instance, the models beyond the 2nd-order distance-dependent model predict that, for polarized cells, the embedding of the feedforward motif (#4) is such that the "source" neuron (the neuron that projects to the other two neurons) is located above the other two cells whereas the "sink" neuron (receiving synapses from the other two neurons) is the lowest (**Supplementary Fig. 2)**. Indeed, we found a tendency for such an alignment in the NMC circuitry (**Fig. 3a4**) as well as in the experimental results (see **Figure 4** and text below).

Experimental analysis of triplet motifs in L5 cortical pyramidal cells

To experimentally examine the predictions about the embedding of motifs in space we recorded from L5-TTPCs from the somatosensory cortex of the rat using *in vitro* 12-patch whole-cell recordings (**Fig. 4a**). A total of 7,309 triplets were mapped and characterized, as in the example shown in **Figure 4a-c**. The 3D positions of each of the recorded cells forming a given triplet type were measured. Having such large experimental dataset of L5-TTPCs triplet motifs, we then analyzed to what extent does their relative depth (z-axis) explain the appearance of particular types of motifs, as predicted by the geometrically-based NMC. **Figure 4d** depicts the relative positions of cells composing four types of triplet motifs (triplets #3, #4, #10 and #11) in the experimental dataset). The relative position of the cells forming individual triplets is shown by the light brown lines; the dark line depicts the average position of the cells for that particular triplet. The same analysis, for the same "thickness" of a "cortical slice", was performed in the NMC for L5-TTPCs (**Fig. 4e, Methods**). The analysis indeed shows that, as predicted by the model, that there is a close match between the specific triplet motifs and the relative position of their constituting neurons also in the biological tissue.

DISCUSSION

Neuronal circuits in a variety of mammalian brain areas share similar highly nonrandom expression pattern of network motifs (**Fig. 1**). This observed connectivity structure could be innate and/or the consequence of active plasticity and learning processes. Our study shows that the universal tendency of networks of pyramidal neurons for overrepresenting transitive triplet motifs and underrepresenting intransitive triplets emerges from the asymmetric geometry of neurons' morphology and their relative spatial position (**Fig. 2**). The impact of different neuron geometries on the emerging motifs was further evaluated using *is silico* circuits of several other cell types (**Fig. 3**). Theoretical predictions for the spatial alignment of cells were then directly validated via *in-vitro* simultaneous whole-cell recordings from up to 12 neurons in the rat somatosensory cortex (**Fig. 4**). We therefore

concluded that, with innate bipolarized neuron morphology[17], a structured network connectivity is essentially unavoidable and the observed complex profile of over/underrepresented motifs in neuronal networks is largely innate.

The present study focuses on the relation between geometrical features and connectivity pattern. Preliminary studies show a relationship between the fine structure of neuronal connectivity and neuron-level computations[18,19], network-level dynamics and information processing[20,21], and, ultimately, behavior[7,22]. It would therefore be essential for future studies to isolate the contribution of the expression level of specific network motifs to the network activity dynamics in order to advances in unravelling the long-standing structure-to-function problem in neuroscience.

The generalization of our prediction is also supported by a recent study in cerebellar microcircuits where transitive triplet motifs were overrepresented among asymmetrical interneurons, showing that the relative position in 3D space of the neurons within a triplet is correlated with the formation of a particular motif[5]. In addition, the preference for local transitive patterns has already been reported in different real-world networks from diverse disciplines, including the *C. elegans* connectome[23], social networks and the World Wide Web[24,25]. It would be interesting to study, using the present method, the extent to which this transitivity emerges simply from the underlying geometry in such spatially-embedded real-world networks too.

Our theoretical analysis is based on a detailed dense *in silico* model of a neocortical microcircuit. This provided us with the opportunity to dissect systematically which geometrical features (asymmetric neuronal morphology, neuron's position in the cortical tissue) explain the emergence of patterns in the triplet distribution (**Fig. 2**). By adding increasingly more elaborate aspects to the reference generative theoretical network, we have reproduced the general trends of neuronal motifs as found in the full circuit. Despite obtaining a significant reduction in the discrepancy between the real and expected distribution of motifs, we were unable to completely predict the absolute numbers of the various motifs as found in the full circuit (**Fig. 2f, Fig. 3 a2-d2**). This residual discrepancy likely stems from the fact that the statistics upon which our generative models is built considers mostly average values and neglecting, to a large extent, individual differences between cells' morphologies within classes (dendritic and axonal) as well as possible interaction of the morphology with position (e.g. cells towards the bottom of the layer have less probability of making contacts downwards).

Previous works have hypothesized that connectivity is probably dependent on intersomatic orientation of cells[5,26], but the present study is the first to systematically quantify it and its impact on the emerging network topology and its embedding in cortical space. Additional studies highlight wiring specificities beyond what expected from simple geometry. Synaptic connections might have stronger preference to specific cell-type populations[27], layers[28] or perhaps even to specific cell identity[29]. Another possible source of nonrandom expression pattern on motifs is structural synaptic plasticity due to learning. While structural plasticity provides a huge number of connectivity patterns, we show here that the spatial organization of neurons and their asymmetric geometry, put strong constraints on the connectivity configuration that are likely and place a strong

bias for the formation of certain types of motifs. Hence, while plasticity and learning may change the specific members of a triplet motif, they are likely to keep the global pattern of motifs expression.

Our theoretical and experimental work provides the foundation for exploring the yet open question regarding the origin of the connectivity structure of cortical microcircuits: what is innate (simply due to asymmetrical neuronal geometry embedded in cortical volume) and what is specific (due to genetically-guided and/or learning processes)? The present work, combined with the soon-coming dense reconstructions of ~1 mm^3 of rodents and human cortical microcircuits, will provide the answer for this intriguing question. Systematic consideration of the fine structure of neuronal circuits, combined with functional modeling of its neurons and synapses, will significantly advance the long-standing structure-to-function problem in neuroscience.

ACKNOWLEDGEMENTS

We thank Oren Amsalem, Hanan Shteingart and all lab members of the Segev and London Labs for many fruitful discussions and valuable feedback regarding this work. This work was supported by the Drahi family foundation to IS, a grant from the EU Horizon 2020 program (720270, Human Brain Project) and a grant from the Gatsby Charitable Foundation. ML is a Sachs Family Lecturer in Brain Science and is supported by the ISF (1024/17) and the Einstein Foundation.

REFERENCES

1. Bonifazi, P. *et al.* GABAergic hub neurons orchestrate synchrony in developing hippocampal networks. *Science (80-.).* **326**, 1419–1424 (2009).

2. Avena-Koenigsberger, A., Misic, B. & Sporns, O. Communication dynamics in complex brain networks. *Nat. Rev. Neurosci.* **19**, 17–33 (2017).

3. Song, S., Sjöström, P. J., Reigl, M., Nelson, S. & Chklovskii, D. B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, 0507–0519 (2005).

4. Perin, R., Berger, T. K. & Markram, H. A synaptic organizing principle for cortical neuronal groups. *PNAS* **108**, 5419–5424 (2011).

5. Rieubland, S., Roth, A. & Häusser, M. Structured connectivity in cerebellar inhibitory networks. *Neuron* **81**, 913–929 (2014).

6. Fulcher, B. D. & Fornito, A. A transcriptional signature of hub connectivity in the mouse connectome. *Proc. Natl. Acad. Sci.* **113**, 1435–1440 (2016).

7. Bassett, D. S. & Sporns, O. Network neuroscience. *Nat Neurosci* **20**, 353–364 (2017).

8. Yamagata, M. & Sanes, J. R. Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. *Nature* **451**, 465 (2008).

9. Yu, Y.-C., Bultje, R. S., Wang, X. & Shi, S.-H. Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature* **458**, 501 (2009).

10. Song, S. & Abbott, L. F. Cortical Development and Remapping through Spike Timing-Dependent Plasticity. *Neuron* **32**, 339–350 (2001).

11. Ocker, G. K., Litwin-Kumar, A. & Doiron, B. Self-Organization of Microcircuits in Networks of Spiking Neurons with Plastic Synapses. *PLOS Comput. Biol.* **11**, e1004458 (2015).

12. Ravid Tannenbaum, N. & Burak, Y. Shaping Neural Circuits by High Order Synaptic Interactions. *PLOS Comput. Biol.* **12**, e1005056 (2016).

13. Markram, H. *et al.* Reconstruction and Simulation of Neocortical Microcircuitry. *Cell* **163**, 456–492 (2015).

14. Gal, E. *et al.* Rich cell-type-specific network topology in neocortical microcircuitry. *Nat Neurosci* **20**, 1004–1013 (2017).

15. Egger, R., Dercksen, V. J., Udvary, D., Hege, H.-C. & Oberlaender, M. Generation of dense statistical connectomes from sparse morphological data. *Front. Neuroanat.* **8**, 129 (2014).

16. Reimann, M. W., King, J. G., Muller, E. B., Ramaswamy, S. & Markram, H. An algorithm to predict the connectome of neural microcircuits. *Frontiers in Computational Neuroscience* **9**, 28 (2015).

17. Barnes, A. P. & Polleux, F. Establishment of Axon-Dendrite Polarity in Developing Neurons. *Annu. Rev. Neurosci.* **32**, 347–381 (2009).

18. Amsalem, O., Van Geit, W., Muller, E., Markram, H. & Segev, I. From Neuron Biophysics to Orientation Selectivity in Electrically Coupled Networks of Neocortical L2/3 Large Basket Cells. *Cereb. Cortex* **26**, 3655–3668 (2016).

19. Cossell, L. *et al.* Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518**, 399–403 (2015).

20. Zhao, L., Beverlin, B., Netoff, T. & Nykamp, D. Q. Synchronization from second order network connectivity statistics. *Front. Comput. Neurosci.* **5**, 28 (2011).

21. Denk, W., Briggman, K. L. & Helmstaedter, M. Structural neurobiology: missing link to a mechanistic understanding of neural computation. *Nat. Rev. Neurosci.* **13**, 351 (2012).

22. Seung, H. S. Reading the Book of Memory: Sparse Sampling versus Dense Mapping of Connectomes.

Neuron **62**, 17–29 (2009).

23. Azulay, A., Itskovits, E. & Zaslaver, A. The C. elegans Connectome Consists of Homogenous Circuits with Defined Functional Roles. *PLoS Comput. Biol.* **12**, (2016).

24. Milo, R. *et al.* Network motifs: Simple building blocks of complex networks. *Science (80-.).* **298**, 824–827 (2002).

25. Milo, R. *et al.* Superfamilies of Evolved and Designed Networks. *Science (80-.).* **303**, 1538 LP – 1542 (2004).

26. Schroter, M., Paulsen, O. & Bullmore, E. T. Micro-connectomics: probing the organization of neuronal networks at the cellular scale. *Nat. Rev. Neurosci.* **18**, 131–146 (2017).

27. Pfeffer, C. K., Xue, M., He, M., Huang, Z. J. & Scanziani, M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat Neurosci* **16**, 1068–76 (2013).

28. Harris, K. D. & Mrsic-Flogel, T. D. Cortical connectivity and sensory coding. *Nature* **503**, 51–58 (2013).

29. Kasthuri, N. *et al.* Saturated Reconstruction of a Volume of Neocortex. *Cell* **162**, 648–661 (2015).

30. Erdös, P. & Rényi, A. On random graphs, I. *Publ. Math.* **6**, 290–297 (1959).

31. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

MATERIALS AND METHODS

Slice preparation.

Experiments were carried out according to the Swiss national and institutional guidelines. Fourteen to eighteen day-old non-anesthetized Wistar rats were quickly decapitated, and their brains were carefully removed and placed in iced artificial cerebrospinal fluid (ACSF). Slices (300 µm) were cut on an HR2 vibratome (Sigmann Elektronik). Parasagittal slices, ~1.7–2.2 mm lateral to the midline, were cut to access primary somatosensory cortex (SSC; above the anterior extremity of the hippocampus ± 1 mm). Slices were incubated at 37 °C for 30–60 min and then left at room temperature until recording. Cells were visualized by infrared differential interference contrast video microscopy using a VX55 camera (Till Photonics) mounted on an upright BX51WI microscope (Olympus). Thick-tufted layer 5 Pyramidal Cells (PCs) were selected according to their large soma size (15–25 µm) and their apparent large trunk of the apical dendrite. Care was taken to use only parallel slices (i.e. slices that had a cutting plane parallel to the course of the apical dendrites and the primary axonal trunk). This ensured sufficient preservation of both the PCs' axonal and dendritic arborizations.

Chemicals and solutions.

Slices were continuously superfused with ACSF containing 125 mM NaCl, 25 mM NaHCO3, 2.5 mM KCl, 1.25 mM NaH2PO4, 2 mM CaCl2, 1 mM MgCl2, and 25 mM D-glucose bubbled with 95% O2 and 5% CO2. The intracellular pipette solution contained 110 mM potassium gluconate, 10 mM KCl, 4 mM ATP-Mg, 10 mM phosphocreatine, 0.3 mM GTP, 10 Hepes, and 13mMbiocytin adjusted to pH 7.3–7.4 with 5MKOH. Osmolarity was adjusted to 290–300 mOsm L−1 with D-mannitol (25–35 mM). The membrane potential values given were not corrected for the liquid junction potential, which is approximately −14 mV. Chemicals were purchased from Sigma Aldrich or Merck.

Electrophysiological recordings.

Multiple somatic whole-cell recordings (4–12 cells simultaneously) were performed with Multiclamp 700B amplifiers (Molecular Devices) in the current clamp mode at 34 ± 1 °C bath temperature. Data acquisition was performed through an ITC-1600 board (Instrutech) connected to a PC running a custom-written routine (PulseQ) under IGOR Pro (Wavemetrics). Sampling rates were 10 kHz, and the voltage signal was filtered with a 2-kHz Bessel filter. Patch pipettes were pulled with a Flaming/Brown micropipette puller P-97 (Sutter Instruments) or a DMZ puller (Zeitz Instruments) and had an initial resistance of 3–8 MΩ.

Stimulation protocols.

Monosynaptic, direct excitatory connections were identified by stimulation of a presynaptic cell with a 20–70 Hz train of 5–15 strong and brief current pulses (1–2 nA, 2–4 ms) followed by a recovery test response 0.5 s after the end of the train (not shown in the traces), all precisely and reliably eliciting action potentials (APs).

Final somatic positions.

The soma positions were recorded relative to an arbitrary reference point, and the z axis was oriented perpendicular to the surface of the slice. After morphological stainings were ready, the y axis was rotated around the z axis to match the orientation of the apical dendrites. The x axis was rotated by the same amount and remained orthogonal to the other two axes.

Constructing the series of geometrical network models

The observed (neuronal) network of n nodes (neurons) is described by $n \times n$ adjacency matrix A, whereas the A_{ij} entry is a binary variable denoting the existence of connection from the neuron i to the neuron i, and by a 3D vector $\mathbf{x}_i$ for each neuron i, denoting its 3D position is space, $\mathbf{x}_i = (x_i, y_i, z_i)$. Using this data, the geometrical profile of pairwise connection probability function $p = \Pr(A_{ij} = 1 \mid \mathbf{x}_i, \mathbf{x}_j)$ was computed at different orders of approximation, gradually taking into account additional features of pair's spatial embedding $(\mathbf{x}_i, \mathbf{x}_j)$. Then, to study the statistical dependence of network motifs on the geometrical profile a series of generative models for spatially-embedded networks was developed, in which each model gradually fits a higher order of approximation of the probability function. The distributions of motif counts emerging from each of the models were compared to the distribution observed in the original network using Kullback–Leibler divergence.

The 1st-order model that was used for the generation of reference random networks was the Erdős-Rényi model[30]. This model assumes that the probability of forming a connection between two neurons, p, is uniform among all pairs of nodes and independent from every other connection. Naïvely, this 1st order model does not take into account spatial features of the nodes whatsoever. Thus, the model requires matching only two parameters: the number of neurons as in the original network, $n = n_{L5-TTPC}$, and the mean connection probability over the whole network, $p(\mathbf{x}_i, \mathbf{x}_j) = \text{const} = \sum_{i \neq j} A_{ij} \,/(n(n-1))$.

The 2nd-order model generated random networks which preserved the non-uniform dependency of connection probability on intersomatic distance (r) as observed in the original network (**Fig. 2b**, **Supplementary Fig. 1a**). First, the original distance-dependent connectivity profile, $p^{L5-TTPC}(r)$ (**Supplementary Fig. 1a**, blue line), was fitted (using an ensemble of 500 decision trees each with a maximal depth of 5 levels[31]; **Supplementary Fig. 1a**, orange line). Next, using the fitted probability function, $p(r)$, we generated random networks with a matching distance-dependent connectivity. The original coordinates of all neurons were preserved, and the connections were defined as independent binomial random variables A_{ij}, with a success probability $p(\mathbf{x}_i, \mathbf{x}_j) = p(r_{ij})$ that corresponded to their inter-somatic distance r_{ij}.

The 3rd-order model captures the fact that the original probability function is anisotropic and depends not only on the intersomatic distance but also on the direction. Specifically, the probability may be different depending on whether the postsynaptic neuron is above or below (along the z-axis) the presynaptic cell. For example, in the L5-TTPC network, the probability is higher when the post-synaptic cell is located below (**Fig. 2c**). The 3rd-order model generated networks with a matching "bipolar" distance dependence (**Supplementary Fig. 1b**).

The 4th-order "offset dependent" model captures $p(\mathbf{x}_i, \mathbf{x}_j)$ dependence on the full 3D relative direction between the presynaptic and postsynaptic cell. Specifically, in the L5-TTPC network, $p(\mathbf{x}_i, \mathbf{x}_j)$ is larger for postsynaptic neurons that are directly below (along the z-axis), as compared to cells that are obliquely below the presynaptic cell (**Fig. 2d**).

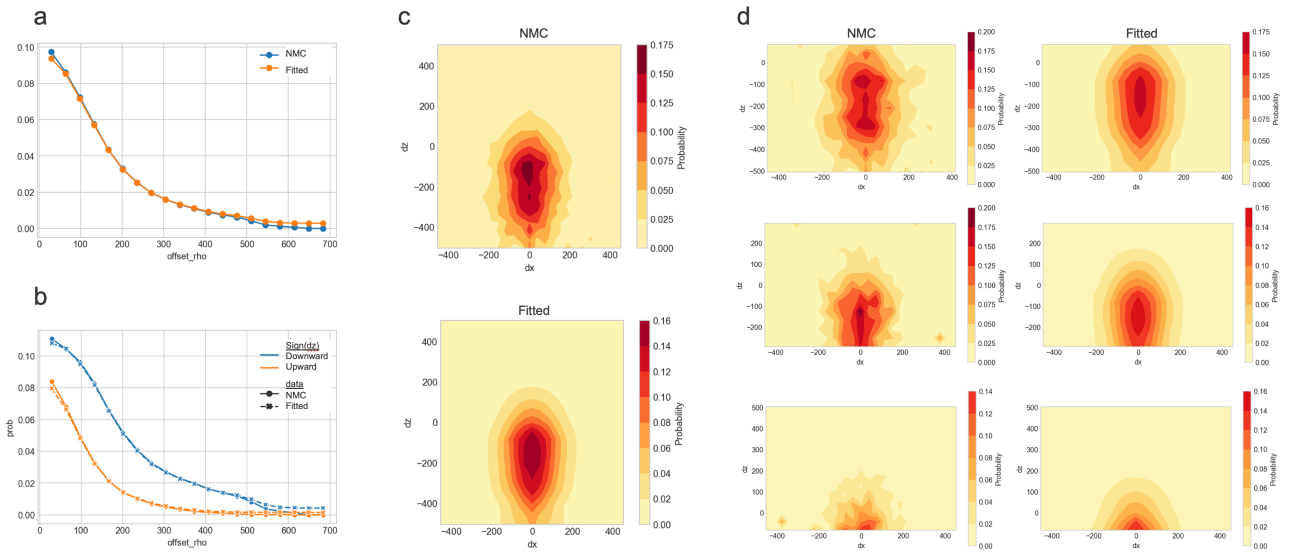
The 5th-order "position dependent" model captures $p(\mathbf{x}_i, \mathbf{x}_j)$ dependence on the absolute position in 3D of both the presynaptic and the postsynaptic cells. Specifically, L5-TTPCs whose somata are located close to the border of layer 4 (**Fig. 2e**, top horizontal line) have a larger vertical (z-axis) span for connecting a postsynaptic partner as compared to L5-TTPCs whose somata reside near the border of layer 6 (**Fig. 2e,** lower left "cloud").

In all four (2nd-5th) geometrical models we employed Machine-Learning approach to learn the connectivity observed in the NMC. First, given an $n \times n$ adjacency matrix A from the NMC, it was transformed into a labelled dataset consisting of $n(n-1)$ samples corresponding to all possible pairs of neurons. Each sample is composed of set of features describing the geometrical embedding of the respective pair (namely, their intersomatic distance, downward/upward alignment, 3D relative offset and absolute 3D positions) and a label indicating whether the pair is connected or not. Secondly, we chose the relevant subset of features for each geometrical model (e.g., for the 3rd order model we chose the intersomatic distance and the downward/upward

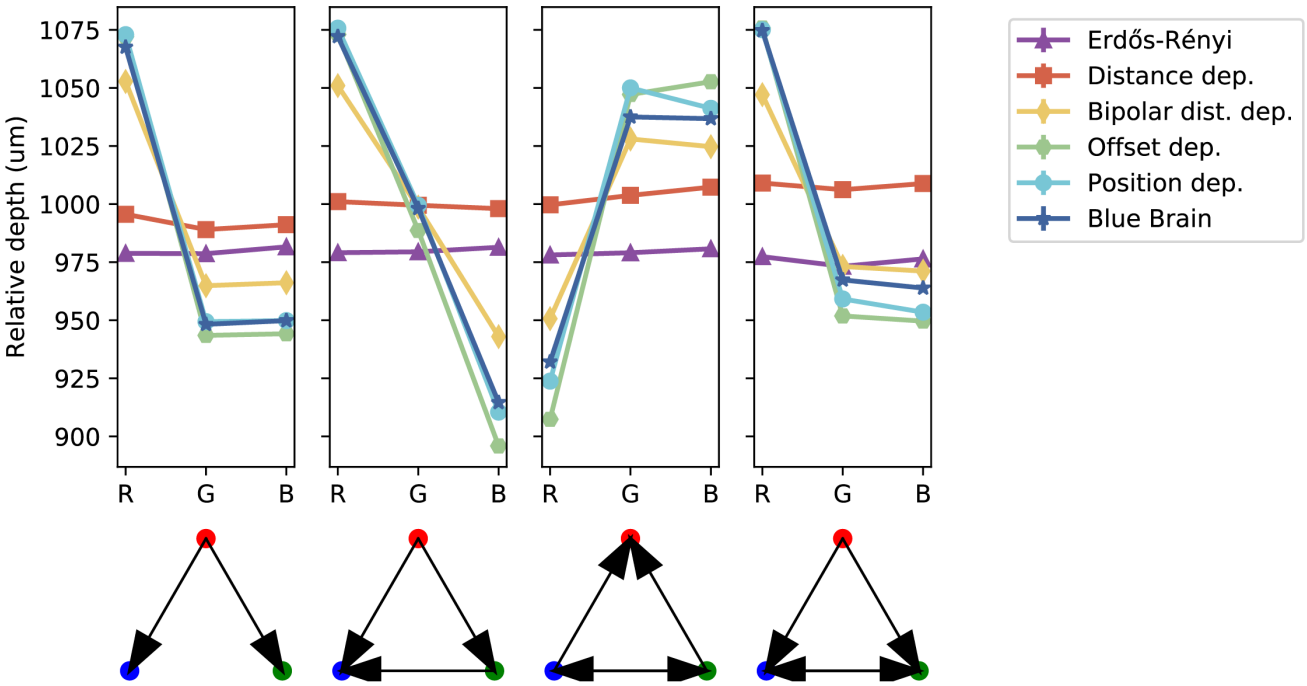
alignment as the two relevant features). Thirdly, we fitted the probability for a connection from the chosen labelled dataset. Importantly, all 2nd-5th order geometrical models were each fitted using an ensemble of 500 decision trees, each with a maximal depth of 5 levels (regression problem)[31]. This method ensures that the gradual improvement of the fit between the NMC and the progressively more complex geometrical model is indeed a consequence of the additional geometric features and not merely due to the higher degrees of freedom when considering a larger number of features. Finally, using the fitted probability functions, we generated the 2nd – 5th order random networks with matching connectivity as measured in the NMC.

Spatial Embedding Analysis.

For each motif pattern, a paired-samples t-test was conducted to compare the vertical position of its composing cells across all of its occurrences in the data. In a motif, the comparison was done between all different classes of cells (different colors in **Fig. 4d-e**). When a motif had two neurons belonging to the same class (e.g., two blue neurons in motif #3) the mean of their vertical position was taken for the purpose of the paired-samples t-test.



Supplementary Figure 1. Comparison of empirical and fitted connection probabilities for L5-TTPCs. a, The dependence of connection probability on intersomatic distance as observed in the NMC (blue) and as fitted by the 2nd-order distance-dependent model (orange). **b**, The distinct connection probability of the downward (blue) versus upward (orange) connections as observed in the NMC (circle) and as fitted by the 3rd-order bipolar distance-dependent model (cross). **c**, The offset dependence of connection probability as observed in the NMC (top) and as fitted by the 4th-order offset dependent model. **d**, Same as in **c**, but for position dependence of the connection probability as fitted by the 5th-order position dependent model.



Supplementary Figure 2. Spatial embedding of triplet motifs (#3, #4, #10 and #11) for L5-TTPCs as predicted by the series of the generative models. For each model, an instance of one network was generated, on which the spatial alignment of the cells involved in forming the specific triplet was measured. Note that models beyond the 2nd-order distance-dependent model predict the asymmetrical embedding found in the NMC model (blue) and in the slice (**Fig. 4**). For instance, the embedding of the feedforward motif (#4) is such that the "source" neuron (red) is located above the other two cells whereas the "sink" neuron (blue) is the lowest.