Why do we have so many excitatory neurons?

Qingyang Wang*[1], Albert Cardona[2,4,5], Marta Zlatic[3,4,5], Joshua T. Vogelstein[6,1], and Carey E. Priebe[7]

[1]Department of Neuroscience, Johns Hopkins University
[2]Department of Physiology, Development, and Neuroscience, University of Cambridge
[3]Department of Zoology, University of Cambridge
[4]MRC Laboratory of Molecular Biology, Neurobiology Division
[5]Janelia Research Campus, Howard Hughes Medical Institute
[6]Department of Biomedical Engineering, Johns Hopkins University
[7]Department of Applied Mathematics and Statistics, Johns Hopkins University

Approximately four in five neurons are excitatory. This is true across functional regions and species. Why do we have so many excitatory neurons? Little is known. Here we provide a normative answer to this question. We designed a task-agnostic, learning-independent and experiment-testable measurement of functional complexity, which quantifies the network's ability to solve complex problems. Using the larval *Drosophila* whole-brain electron microscopy connectome, we discovered the optimal Excitatory-Inhibitory (E-I) ratio that maximizes the functional complexity: 75-81% percentage of neurons are excitatory. This number is consistent with the true distribution observed via scRNA-seq. We found that the abundance of excitatory neurons confers an advantage in functional complexity, but only when inhibitory neurons are highly connected. In contrast, when the E-I identities are sampled uniformly (not dependent on connectivity), the optimal E-I ratio falls around equal population size, and its overall achieved functional complexity is sub-optimal. Our functional complexity measurement offers a normative explanation for the over-abundance of excitatory neurons in the brain. We anticipate that this approach will further uncover the functional significance of various neural network structures.

1 Introduction

How do brains support complex behaviors? Despite the significant advancements in neuroscience, it remains challenging to identify the fundamental wiring rules or structural properties that are necessary for complex functionalities.

Part of the difficulty lies in expanding the repertoire of structural properties to be studied. Although graph theory and other theoretical fields offer many valuable insights into which network properties might be critical [1, 2], many of these properties are hard to estimate without large-scale, synapse-level connectivity data. Even basic graph metrics, such as in degree (the number of pre-synaptic neurons), cannot be accurately determined without a complete synapse-level connectivity map. The recent emergence of nanoscale electron microscopy (EM) connectomes solve this problem, enabling analysis at an unprecedented level [3–9].

Another challenge arises from the generality of the concept of functional complexity, rendering it difficult to define or quantify without relying too much on specific behaviors or tasks. Statistical theories, such as Vapnik–Chervonenkis dimension and Rademacher Complexity [10] offer analogous means of defining function complexity; yet, they depend on exhaustive metrics that are difficult to translate into experimentally testable measurements. To address this, we propose a functional complexity measurement that is task-agnostic, learning-independent and experimentally testable.

The brain structure property we focus on is the Excitatory-Inhibitory (E-I) neuron ratio. Although basic, the E-I ratio is highly conserved across species. Yet, its functional significance remain poorly understood. Across a wide range of species - from *Drosophila* [11], rodents [12, 13], cats [14, 15], to non-human primates [16], and humans [17] - the excitatory neuron population consistently takes up the majority ($\sim$70-80%, Figure 4A). This Excitatory-Inhibitory (E-I) ratio is also conserved across different functional regions of the neocortex [13, 14, 16, 17], with the hippocampus being an extreme case where $\sim$85-90% neurons are excitatory [12]. What is the necessity of having so many excitatory neurons? Our results offer a clue.

By sampling 8180 different recurrent neural networks (RNNs) constrained by the larval *Drosophila* EM connec-

*qwang88@jhu.edu

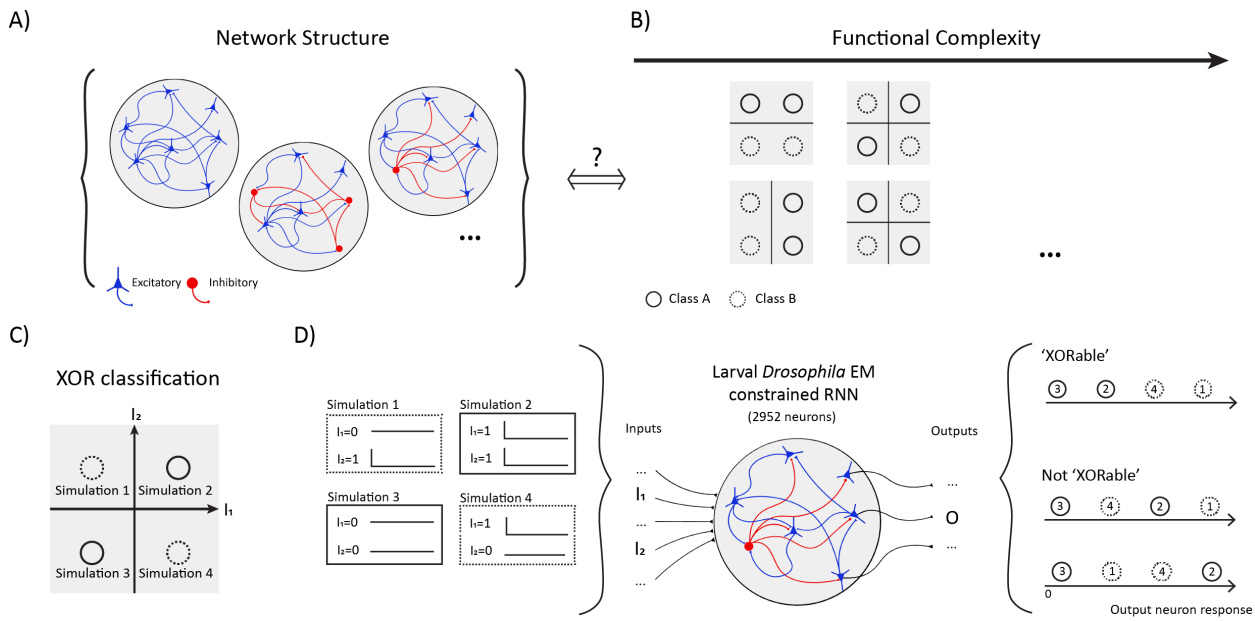


Figure 1: **Proposed methodology to measure functional complexity of a given network.** A) Illustration on relating network structure, specifically E-I composition, to the network's functional complexity. B) XOR is a binary classification that takes two inputs, with the desired output being the two scenarios where only one of the two inputs is activated are classified into the same class (second and forth quadrant, dashed circles). Whether a subnetwork is 'XORable' serves as the building block of the functional complexity measurement and it can be tested by four independent input stimulation pairs. C) Given an EM-constrained larval *Drosophila* whole-brain RNN (2952 neurons in total, E-I composition sampled as detailed in Section 4), we iterate through the subnetworks $\{I_1, I_2, O\}$ where each subnetwork is defined by $\{I_1, I_2\}$ as two input neurons and O as the output. Exhaustively, we measure through all olfactory and gustatory neurons as inputs and all 400 output neurons (RGN and descending neurons). Whether a subnetwork $\{I_1, I_2, O\}$ is XORable is tested by the following procedure: four independent simulations are run (illustrated on the left and in panel C). The four simulation inputs correspond to $\{(0,1),(1,1),(0,0),(1,0)\}$ respectively, where 1 means the input neuron is activated. XORable corresponds to the situation where simulations 1 and 4 elicit output response that is linearly separable from simulations 2 and 3. In another word, the relative order of 1 and 4 should be larger than 2 and 3. This is illustrated on the right along with counter-examples.

tome [3] and quantifying their functional complexity, we demonstrate that the optimal E-I ratio that maximizes functional complexity indeed favors a high proportion of excitatory neurons - the top 1% networks have an average of 73% excitatory neurons. The E-I ratio distribution of the top simulations closely mirrors that of real brains data [18]. This provides a normative explanation for the over-abundance of excitatory neurons across species: to readily perform more complex functions. Furthermore, along the process, we will also show it is important to consider the interplay between different structural properties, which is now possible thanks to the availability of EM connectomes.

2 Results

2.1 Measuring functional complexity based on nonlinearity

We give the full detail of the functional complexity measurement in Section 4; here we focus on building the intuition and reasoning behind the measurement. Consider an illustrative example where an animal has to classify toxic leaves apart from edible leaves (binary classification problem). In *world alpha*, toxic leaves are always red regardless of shape (Figure 1B left two illustrations). In *world beta*, toxic leaves (class A solid) could be either green straight or red curved; edible leaves (class B dashed) could be either green curved or red straight. Which world requires solving more complicated function? *World beta*, because it cannot be solved with a linear function.

World beta is actually a description of XOR, a binary classification problem that involves two dimensional input and a binary output. In fact, any binary classification task that is dependent on both features is an XOR task [19].

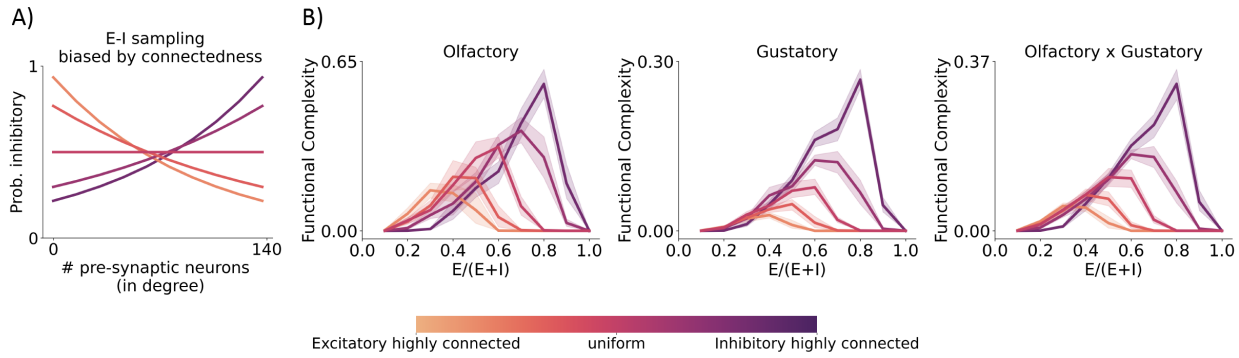


Figure 2: **The optimal E-I ratio that maximizes functional complexity is dependent on the connection degree of inhibitory neurons.** A) E-I probability as a function of its in degree (number of pre-synaptic neurons). For each neuron, the probability for it being sampled as inhibitory depends on its in degree. All colored curves result in a target E-I ratio of 0.5 but with different degrees of biasing the inhibitory population to be highly connected (noted by the colors given in the bottom colorbar). Darker colors = inhibitory neurons are more highly connected; Lighter colors = excitatory neurons are more highly connected. B) The resultant functional complexity with E-I identities sampled at various E-I ratio levels (x-axis) and various connectedness bias levels (colors). Shaded area mark the 95% confidence interval (n=10). When the E-I identity is uniformly sampled across all neurons (horizontal line in A and the curves of middle color in B), the optimal E-I ratio that maximizes functional complexity is 0.6 across sensory modalities. However, the maximum level of functional complexity reached by uniform sampling is sub-optimal; more optimal E-I compositions require the sampling procedure to bias the inhibitory neurons to be highly connected (darker colors). In contrast, when the E-I sampling procedure shifts the excitatory neurons to be highly connected (lighter colors), it generally results in lower functional complexity.

Multiple animal behavior studies have shown that XOR is more challenging for humans and monkeys to learn compared to linearly solvable tasks [19–22]. Even in machine learning research, XOR has played a pivotal role in revealing the functional complexity of networks: the first AI winter started when Minsky and Papert pointed out perceptrons cannot solve XOR [23]. XOR has also revealed the functional significance of specific network structure - it has been mathematically proven that feedforward networks without any negative weights (equivalent to inhibitory connections) cannot solve XOR and thus are not universal approximators[1] [24].

Given a network with n neurons, out of which $|\mathcal{I}|$ are input neurons and $|\mathcal{O}|$ are output neurons, we measure its functional complexity by considering how many subnetworks can readily solve XOR, termed XORable subnetworks. The more XORable subnetworks, the more nonlinear input associations the network is able to form and the more complex decision boundaries the network can represent. A subnetwork is defined by the input-output combination $\{I_1, I_2, O\}$, which implicitly considers all the possible paths connecting the two input neurons to the output neuron within certain number of time steps. The procedure for deciding whether a subnetwork is XORable is illustrated in Figure 1D. Briefly, four independent simulations are run where the two input neurons are combinatorially activated at time 0 (left panel), corresponding to the four quadrants of the 2D input space (Figure 1C). The subnetwork $\{I_1, I_2, O\}$ is XORable when the output activity of the four simulations follow a particular order: the two simulations where both input neurons are simultaneously activated or not activated (solid borders in Figure 1C-D) should elicit output response that is linearly separable from the simulations where only one of the input neurons was activated. We iterate through all concerned subnetworks with the above procedure to obtain the functional complexity measurement.

We did not involve any parameter optimization in the functional complexity measurement - the connectome already encapsulates knowledge from evolution and life experiences, and our goal is to quantify the functional complexity on connectivity structures that is already naturally learnt.

2.2 Sampling E-I configurations for EM-constrained RNNs in search of maximum functional complexity

The recurrent neural network (RNN) (details in Section 4) is constrained by the larval *Drosophila* EM connectome [3]. There are in total 2952 neurons that form a connected graph (one single connected component), 400 of which are output neurons (RGNs and descending neurons) and 480 are input neurons. We will focus on olfactory (46) and gustatory (131) neurons as input throughout this paper. The neurons' firing rates $\mathbf{r}$ at time t are modeled as

$$\mathbf{r}(t) = ReLU(\mathbf{W}^T\mathbf{r}(t-1) + \mathbf{E}^{(t)}),\ t \in \{1 \dots T\}. \tag{1}$$

$\mathbf{E}$ is the external stimuli in Figure 1D. The connectivity from pre-synaptic neurons to post-synaptic neurons, $\mathbf{W}$, is the element-wise product of two components: 1) the strength of connections $\mathbf{M}$ is given by the synaptic count from the EM dataset (axon-dendrite synapses in main text; all-all connection results are in supplementary); 2) the excitatory/inhibitory identity $\boldsymbol{\alpha}$ of connections are randomly sampled from a distribution (details described below).

What network structure, specifically E-I composition, optimizes its functional complexity? We first varied the global fraction of excitatory neurons from 0 (all inhibitory) to 1 (all excitatory) by uniform sampling, i.e. all neurons have the same probability p of being excitatory (Figure 2, line with color indicated as uniform in the colorbar). As a sanity check of our method, consider the extreme case where all neurons are excitatory: previous theoretical results predicted no XORable subnetwork exists in this scenario [24]. This is exactly what we observed, where all simulations have zero functional complexity when E/(E+I)=1. This observation serves as a nice validation of our method (Figure 2). When E-I identities were uniformly sampled, the optimal E-I ratio that maximized functional complexity was observed at $\sim$50-60% excitatory neurons. This was the case across different sensory modalities (olfactory, gustatory and multimodal-olfactory$\times$gustatory). Such a trend greatly deviates from the expected $\sim$70-80% E-I ratio conserved across species. The discrepancy prompted us to wonder if more sophisticated E-I sampling procedure is required.

We resolved the discrepancy by making the E-I probability a degree dependent function (Figure 2A). For each neuron, the probability of it being sampled as inhibitory is dependent on its in degree, i.e. the number of pre-synaptic neurons (details of the sampling procedure in Section 4). As indicated by the bottom colorbar, such a connectedness-dependent sampling procedure allows one to systematically vary the network structure from having highly connected excitatory neurons (lighter colors) all the way to having highly connected inhibitory neurons (darker colors). When the inhibitory neurons are highly connected (darkest line), the optimal E-I ratio with maximum functional complexity peaked at 80% across sensory modalities. Furthermore, a nice trend was observed from the darkest colored curves to the lightest colored curves: the optimal E-I ratio peak gradually shifted from 80% to 30-40% as the network structure shifted from highly connected inhibitory neurons to highly connected excitatory neurons. The overall optimality is achieved when the inhibitory neurons are highly connected and at 80% E-I ratio. All these results still hold when the connection magnitudes are given by all synapse types, i.e. axo-dendritic, axo-axonic, dendro-dendritic, and dendro-axonic (Figure A.4).

Collectively, it is important to consider the connection degree of the inhibitory and excitatory populations when discussing the functional complexity of networks' E-I configurations. We will expand on this idea in the next section. It is also worth pointing out that the calculation of in degree for all neurons is *not* possible without the availability of the EM connectome data. This emphasizes on the importance of having the entire connectivity map, so that the unexplored but crucial structure properties and their interplay could be explored.

2.3 The top RNNs of highest functional complexity match the E-I ratio distribution of real brains

We further sampled E-I configurations with various forms of degree dependency, including ones where the trend is not monotonic, i.e. neurons of mid-level in degree having high probability of being inhibitory/excitatory (details in Section 4). In total, we performed 818 different types of sampling procedures according to different degree-dependent E-I probability functions, each type with 10 different samples, resulting in 8180 E-I configurations of various E-I ratios and E-I connectedness biasness (Figure 3A). Notably, networks of higher functional complexity (lighter color) not only have an over-abundance of excitatory neuron population (x-axis > 0.5, right to vertical gray

[1] Universal approximation theory is the foundational theoretical result that prompted deep neural networks' success of today.

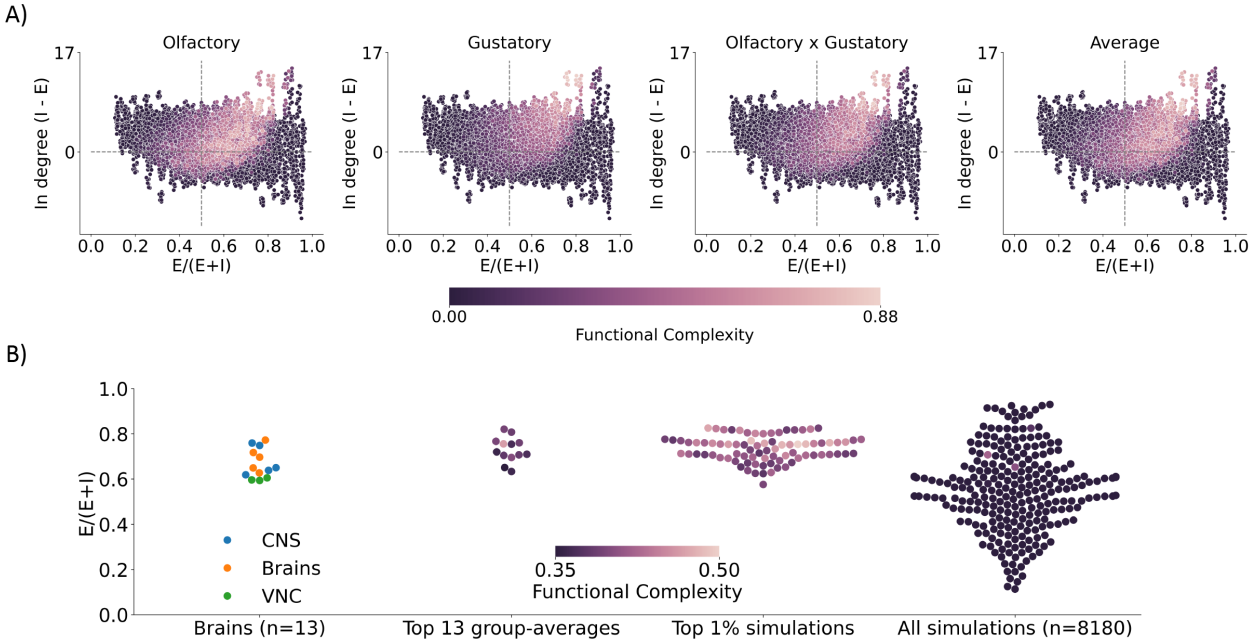


Figure 3: **The top simulations of highest functional complexity match the E-I ratio distribution observed in real laval *Drosophila* brains.** A) 8180 different E-I compositions are drawn from 818 different degree-dependent E-I probability functions (10 samples each). Networks of higher functional complexity (lighter color) lie on the upper right quadrant where the average in degree of inhibitory neurons are higher than excitatory neurons (y-axis > 0, above horizontal gray line) and there is an over-abundance of excitatory neurons (x-axis > 0.5, right to vertical gray line). B) Left: E-I ratios of 13 larval *Drosophila* obtained via scRNA-seq, colored by the source region; Middle left: the E-I ratios of top 13 simulations obtained via averaging across the 10 samples of each degree-dependent E-I probability function group (details see Supplementary Figure A.7); Middle right: the E-I ratios of top 1% simulations (same as red dots in A); Right: E-I ratios of all simulations. The mean E-I ratio of the top 13 group-averages and top 1% simulations are **not** statistical significantly different from the 13 brain E-I ratios (two-sided two-sample t-test, p-val: 0.62, 0.40 respectively). Multiple comparisons corrected by Dunnett's.

line), they also have more highly connected inhibitory population compared to the excitatory population (y-axis > 0, above horizontal gray line).

Across sensory modalities, maximum functional complexity was achieved when the excitatory neurons take majority - the top 1% networks have on average 73% of excitatory neurons. Furthermore, the top sampled networks' E-I ratio distribution closely aligns with real brain E-I ratios obtained via scRNA-seq [18]. This is true regardless of the exact criterion for selecting the *top* simulations (Figure 3A): the result holds when comparing brain data to top 1% simulations (p-val 0.62); it also holds when comparing to the top 13 probability function group averaged E-I ratios (p-val 0.40). When comparing the real brain E-I ratio distribution to all simulations, the means are statistical significantly different (p-val: 0.03). All multiple comparisons corrected by Dunnett's. Notably, we actually over sampled at 0.5 E-I ratios, making the observed optimality even stronger.

Collectively, these results provide a normative explanation to the E-I ratio observed in real larval *Drosophila*, from the perspective of functional complexity.

3 Discussion

Why do we have so many excitatory neurons? Our answer is: to readily solve more complex functions. To measure functional complexity, we proposed a novel normative measurement based on the nonlinearity of subnetwork functions. We leveraged the larval *Drosophila* EM connectome [3] and revealed a key structure property - in degree of neurons - that alters the functional complexity of the resultant networks. The top E-I configurations with the highest functional complexity nicely matched the E-I ratio distribution observed in real larval *Drosophila* brains.

The optimal E-I ratio has been previously explored in terms of the capacity on storing binary memory patterns[25]. The predicted optimal E-I ratio is a function of the degree of tuning (coefficient of variation) of the excitatory and inhibitory activities. Consequently, validating it requires knowing both the input statistics and the connectivity pattern. Capano et al. had trained RNNs with sparse connections (out degree following a bio-plausible distribution) to sequentially learn multiple binary classification tasks; consistent with the normative answers provided here, they also found the optimal performing RNNs to have 70% excitatory neurons [26]. These suggest the functional complexity measurement proposed here may have direct implications on networks' capabilities in multi-task learning.

We deliberately set all bias terms of the RNN, i.e. threshold of the neurons, to be zero. The main reason is we did not want the biases to obscure the effects of the connectivity patterns. It is known that any dynamics' trajectory can be approximated arbitrarily well by a wide-enough RNN with fixed weights and learnt/well-chosen biases [27]. This means if we choose to learn the bias terms or even just randomly set them, there is a good chance the subnetwork can produce XORable dynamics; but then such observed XORable pattern is not due to the connectivity pattern but because of the non-zero biases, resulting in high false positives when checking for XORable subnetworks. This logic is further backed by the fact that XOR can be solved by a network with only three hidden units and all zero biases [28]; although the 2952 neurons are not fully connected to each other in our case, there is still a high chance non-zero biases will greatly increase the number of XORable subnetworks and obscure the effects from the connectivity pattern. A dedicated separate study is required to carefully delineate the interplay between biases and connectivity pattern in terms of their effects on functional complexity.

The inspiring thread of work on manifold capacity touches upon the concept of representation complexity in classification settings [29–32]; however, the definition of complexity in manifold capacity fundamentally differ from the proposed functional complexity measurement, leading the two methods targeting completely different questions. Manifold capacity targets neuronal activity: essentially, a population of neurons have higher capacity when the desired output can be more readily linearly read-out for the pre-specified task output, leaving the discussion on the non-linearity part of the transformation implicit. In contrast, our functional complexity measurement directly measures the non-linearity of the network. Although the latest development incorporated non-linear classifications into the framework, it relies on context-dependent learning and the resultant capacity measurement is dependent on the definition of context [32]. In sum, manifold capacity focuses on neural representations and their linear readout; ours focuses on directly quantifying the non-linearity of networks.

In statistics and mathematics, some possible choices of quantifying the concept of functional complexity include Rademacher Complexity [33] and VC-dimension [34], which essentially measures the richness of a function class. However, both can not be directly applied since they require quantification across all possible data combinations in high-dimensional space. The proposed functional complexity measurement relates to VC-dimension in the sense that it considers the hardest problems out of the entire function set considered in VC-dimension (growth function to be precise) at 4 data points. In another word, our procedure takes the hardest slice of VC-dimension and makes an applicable measurement out of it.

The proposed functional complexity measurement is not just implementable *in silico*; it is also experimentally testable. With the high-throughput automated behavior technologies available for larva Drosophila [35, 36], one could potentially optogenetically stimulate pairs of input neurons according to patterns illustrated in Figure 1D while recording from the output neurons, such as RGNs and/or descending neurons. It will be exciting to see if the fraction of XORable subnetworks in real larvae match the results obtained via EM-constrained simulations presented here.

One key prediction awaiting to be tested in real brains is the highly connected inhibitory neurons (Figure 2-3). Once we know of the neurotransmitter predictions of the larval *Drosophila* EM in the future, we can calculate the in degree distribution of the excitatory and inhibitory neurons respectively, and compare them to the ones of our top networks with highest functional complexity and see if they match.

Another important future direction is to extend this analysis to other EM connectomes [4–9], exploring whether the E-I ratios of other species and adult *Drosophila* can also be explained by optimizing for high functional complexity. A comparative study across species would also be of interest. For example, it would be interesting to see if the proposed functional complexity measurement can explain for the expanded inhibitory subnetwork in humans compared to rodents [37]. Meanwhile, a macroscopic gradient of synaptic excitation and inhibition has been

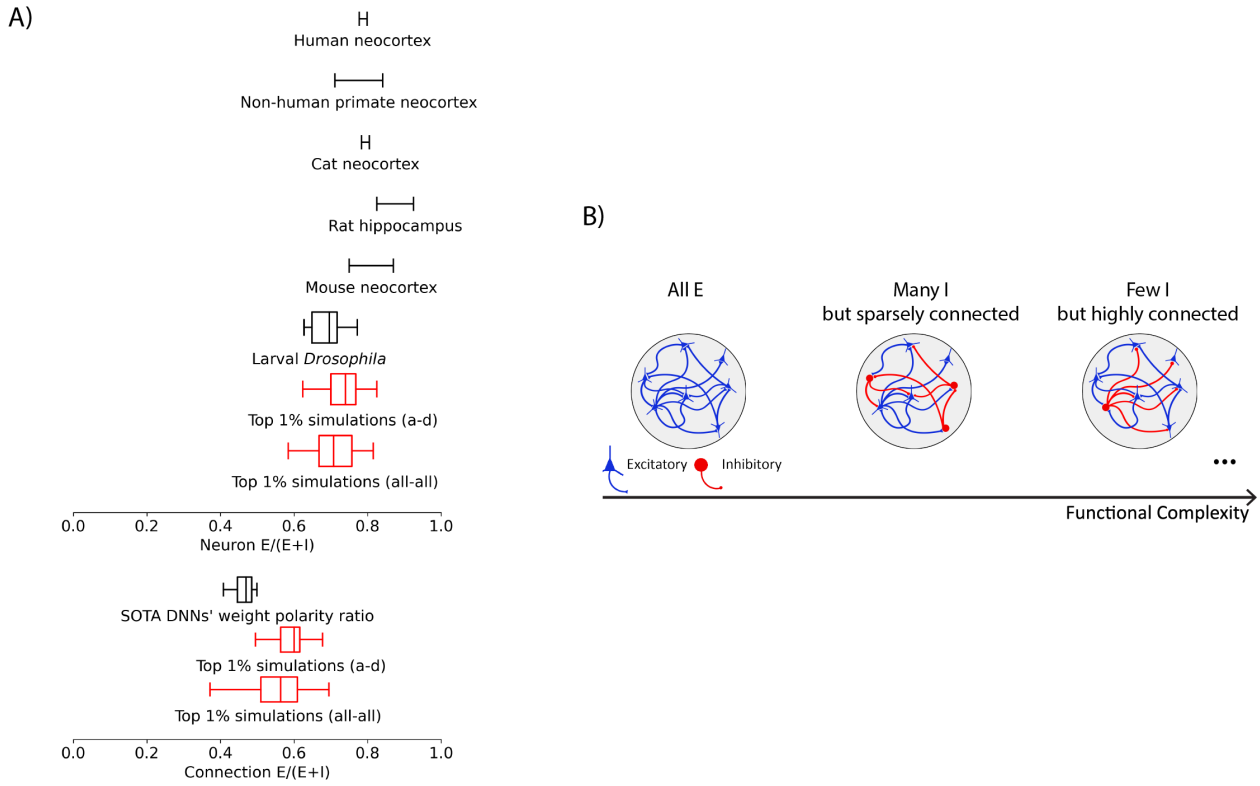


Figure 4: **Conceptual illustration on E-I composition and functional complexity.** A) Top plot: neuron-level E-I ratio. The over-abundance of excitatory neurons is observed across a wide range of species (from larva Drosophila [18] to mouse neocortex [13] to rat hippocampus [12], to cats [14, 15] to non-human primates [16] to humans [17]). Bottom plot: connection-level E-I ratio is closer to equal-abundance, a trend in sync with STA DNNs. Connection-level E-I ratio is quantified as the ratio of total number of E/I connections. One connection is not an individual synapse but the existence of a connection between two neurons (the synaptic count is considered as connection strength in our work). Bottom first box plot contain positive/negative weight ratio from 22 DNNs trained across a range of tasks: image classification (12), object detection (5), image generator (1), video classification (1), text embedding (1), audio embedding (2). The SOTA DNNs tend to have equal abundance of positive and negative weights. B) A conceptual illustration on the relationship between E-I composition and functional complexity.

observed along the hierarchy of brain regions [38], making it interesting to see if there is also a gradient of functional complexity along the hierarchy of processing.

Our work also prompts for an unexplored direction in building better artificial intelligence (AI). The state-of-the-art (SOTA) deep neural networks (DNNs) in AI generally have ∼40-50% positive weights (Figure 4A bottom first boxplot); this number is true across various tasks, from vision to auditory to text embedding. Since DNNs generally do not follow Dale's law, these numbers are not directly comparable to the neuron-level E-I ratios observed in the brains. Therefore we calculated connection-level E-I ratios of our top simulations, quantified as the total fraction of excitatory connections where each connection is not a single synapse but the existence of a connection between two neurons (Figure 4A bottom two boxplots). Indeed, the E-I ratio considered at connection-level is closer to 50%, thus more in sync with the positive/negative weight ratio observed in SOTA DNNs. Intriguingly, as the AI community progresses to find sparser network solutions [39], they have come to realize for sparse networks to properly learn, they have to have proper sign configurations [40]. Our results thus point out an important lesson for finding such proper sign configurations (Figure 3A & 4B) - having equal abundance of positive and negative weights may not be sufficient for networks to have high functional complexity, having highly-connected inhibitory neurons and an over-abundant population of excitatory neurons may be a more promising strategy for configuring the signs of sparse networks.

4 Methods

XORable subnetwork Consider a subnetwork defined by two input neurons I_1, I_2 and one output neuron O. It includes all possible paths connected between them. We are interested in checking if the binary classification problem XOR, depicted in Figure 1C, is readily solvable by the subnetwork $\{I_1, I_2, O\}$. To test this, four independent simulations are run (Figure 1D):

1. Simulation 1: only input neuron I_2 is activated at time zero;
2. Simulation 2: both input neurons are activated at time zero;
3. Simulation 3: both input neurons are *not* activated;
4. Simulation 4: only input neuron I_1 is activated at time zero.

The activity of the output neuron O is recorded for all time steps. At each times step, we compare the output activity of the four simulations (Figure 1D right panel) and their relative order determines whether the subnetwork is XORable. Specifically, the desired XORable order request simulations 2 & 3 (both input activated or not) to be linearly separable from simulations 1 & 4 (only one of the input activated). Since simulation 3 receives no input activation thus always having zero output, there are only two XORable ordering out of the six possible orderings in total (1 & 4 interchangeable in the illustration). As long as there is at least one time step that have XORable output order then we say the subnetwork is XORable.

Relative order of the four simulations' output neuron response	XORable?
3 < 1 < 2 < 4	n
3 < 1 < 4 < 2	n
3 < 2 < 1 < 4	y
3 < 2 < 4 < 1	y
3 < 4 < 1 < 2	n
3 < 4 < 2 < 1	n

Table 1: **All possible orders of output responses of the four simulations at each time step.** (related to Figure 1)

Functional Complexity The functional complexity measurement counts the fraction of input pair (I_1, I_2) that have a robust number ($\geqslant 20$) of XORable subnetworks associated with it, i.e. $\sum \mathbf{1}_{\{I_1, I_2, O\}\text{is XORable}} \geqslant 20$.

Whole-brain larval *Drosophila* EM connectome-constrained RNN The whole-brain neuronal activity of larval *Drosophila* is simulated with a biologically-constrained recurrent neural network (RNN). Concretely, the neurons' firing rates $\mathbf{r}$ at time t are given by

$$\mathbf{r}(t) = ReLU(\mathbf{W}^T\mathbf{r}(t-1) + \mathbf{E}^{(t)}),\ t \in \{1 \dots T\} \tag{2}$$

where $\mathbf{W}$ is the weight matrix describing the connections from pre-synaptic neurons to post-synaptic neurons, with the signs (E-I) sampled and the magnitudes given by the axon-dendrite synaptic count from the EM dataset [3] (all-all synaptic count results are given in supplementaries and are always consistent with the axon-dendrite results). $\mathbf{E}^{(t)}$ denotes the time-series external input delivered to the input neurons. All simulations were run for $T = 20$ time steps, a duration long enough for the network dynamics to stabilize when single input neurons were initially activated.

Uniform sampling on E-I identities For a target E-I ratio p, where $p \in [0, 1]$, the E-I identity of each neuron is independently sampled from a Bernoulli distribution of probability p, which gives the signs of the columns of the weight matrix $\mathbf{W}$.

Degree-dependent sampling on E-I identities For each neuron with an in degree of d, its E-I identity is sampled from a Bernoulli distribution of probability p_d. The degree-dependent probability p_d is set as following: all neurons are grouped into deciles ($q \in [10]$) based on their in degree, and each decile is assigned a probability p_d, forming a probability vector of length 10. In section 2.2, the probability vector is set following exponential function, i.e. $p_d \propto \beta^q$, and the base β controls the rate of biasness (Figure 2A). A linear relationship was also explored(Figure A.5), where $p_d \propto \beta \times q$ with β controlling the rate of biasness. In section 2.3, more degree-dependent probability vectors

are explored by sampling non-monotone relationships. The most exhaustive way would be search in the $[0,1]^10$ space for all possible probability vectors, which is computationally prohibitive. Instead, we explored a subset of all length-10 probability vectors featuring various peak positions. Specifically, we generated single-peak probability vectors where the first, fifth and last elements take values from [0, 0.25, 0.5, 0.75, 1], [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1], and [0, 0.25, 0.5, 0.75, 1] respectively. The fifth element is constrained to be the extrema, resulting in 184 distinct probability vectors. In addition, three other groups of single-peak vectors were created by shifting the peak two deciles higher, two deciles lower, or five deciles.

Real brain E-I ratios. The scRNA-seq data were obtained from [18] where 13 individuals were collected for their brains, ventral nerve cord (VNC) or the entire central nervous system (CNS = brain + VNC). Cholinergic cells and Kenyon cells were considered as excitatory. For CNS and VNC samples, only GABAergic cells were considered as inhibitory. For brain samples, glutamatergic cells were additionally considered to be inhibitory.

Acknowledgements

This work was supported by NSF NeuroNex Award (Grant no. 2014862) (to J.V.), NSF CAREER Award (Grant no. 1942963) (to J.V.), the JHU MINDS Fellowship (to Q.W.).

References and Notes

[1] Joshua T. Vogelstein, Eric W. Bridgeford, Benjamin D. Pedigo, Jaewon Chung, Keith Levin, Brett Mensh, and Carey E. Priebe. Connectal coding: discovering the structures linking cognitive phenotypes to individual histories. *Current Opinion in Neurobiology*, 55:199–212, 2019. ISSN 18736882. doi: 10.1016/j.conb.2019.04.005. URL https://doi.org/10.1016/j.conb.2019.04.005. 1

[2] Patrick Rubin-Delanchy, Joshua Cape, Minh Tang, and Carey E. Priebe. A statistical interpretation of spectral embedding: The generalised random dot product graph. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 84(4), 2022. ISSN 14679868. doi: 10.1111/rssb.12509. 1

[3] Michael Winding, Benjamin D. Pedigo, Christopher L. Barnes, Heather G. Patsolic, Youngser Park, Tom Kazimiers, Akira Fushiki, Ingrid V. Andrade, Avinash Khandelwal, Javier Valdes-Aleman, Feng Li, Nadine Randel, Elizabeth Barsotti, Ana Correia, Richard D. Fetter, Volker Hartenstein, Carey E. Priebe, Joshua T. Vogelstein, Albert Cardona, and Marta Zlatic. The connectome of an insect brain. *Science*, 379(6636), 2023. ISSN 10959203. doi: 10.1126/science.add9330. 1, 2, 4, 5, 8

[4] Sven Dorkenwald, Arie Matsliah, Amy R Sterling, Philipp Schlegel, Szi-chieh Yu, Claire E McKellar, Albert Lin, Marta Costa, Katharina Eichler, Yijie Yin, Will Silversmith, Casey Schneider-Mizell, Chris S Jordan, Derrick Brittain, Akhilesh Halageri, Kai Kuehner, Oluwaseun Ogedengbe, Ryan Morey, Jay Gager, Krzysztof Kruk, Eric Perlman, Runzhe Yang, David Deutsch, Doug Bland, Marissa Sorek, Ran Lu, Thomas Macrina, Kisuk Lee, J Alexander Bae, Shang Mu, Barak Nehoran, Eric Mitchell, Sergiy Popovych, Jingpeng Wu, Zhen Jia, Manuel Castro, Nico Kemnitz, Dodam Ih, Alexander Shakeel Bates, Nils Eckstein, Jan Funke, Forrest Collman, Davi D Bock, Gregory S X E Jefferis, H Sebastian Seung, Mala Murthy, and the FlyWire Consortium. Neuronal wiring diagram of an adult brain. *bioRxiv*, page 2023.06.27.546656, 1 2023. doi: 10.1101/2023.06.27.546656. URL http://biorxiv.org/content/early/2023/07/11/2023.06.27.546656.abstract. 6

[5] David Grant Colburn Hildebrand, Marcelo Cicconet, Russel Miguel Torres, Woohyuk Choi, Tran Minh Quan, Jungmin Moon, Arthur Willis Wetzel, Andrew Scott Champion, Brett Jesse Graham, Owen Randlett, George Scott Plummer, Ruben Portugues, Isaac Henry Bianco, Stephan Saalfeld, Alexander David Baden, Kunal Lillaney, Randal Burns, Joshua Tzvi Vogelstein, Alexander Franz Schier, Wei Chung Allen Lee, Won Ki Jeong, Jeff William Lichtman, and Florian Engert. Whole-brain serial-section electron microscopy in larval zebrafish. *Nature*, 545(7654), 2017. ISSN 14764687. doi: 10.1038/nature22356.

[6] Fabian Svara, Dominique Förster, Fumi Kubo, Michał Januszewski, Marco dal Maschio, Philipp J. Schubert, Jörgen Kornfeld, Adrian A. Wanner, Eva Laurell, Winfried Denk, and Herwig Baier. Automated synapse-level reconstruction of neural circuits in the larval zebrafish brain. *Nature Methods*, 19(11), 2022. ISSN 15487105. doi: 10.1038/s41592-022-01621-0.

[7] The MICrONS Consortium, J. Alexander Bae, Mahaly Baptiste, Agnes L. Bodor, Derrick Brittain, JoAnn Buchanan, Daniel J. Bumbarger, Manuel A. Castro, Brendan Celii, Erick Cobos, Forrest Collman, Nuno Maçarico da Costa, Sven Dorkenwald, Leila Elabbady, Paul G. Fahey, Tim Fliss, Emmanouil Froudarakis, Jay Gager, Clare Gamlin, Akhilesh Halageri, James Hebditch, Zhen Jia, Chris Jordan, Daniel Kapner, Nico Kemnitz, Sam Kinn, Selden Koolman, Kai Kuehner, Kisuk Lee, Kai Li, Ran Lu, Thomas Macrina, Gayathri Mahalingam, Sarah McReynolds, Elanine Miranda, Eric Mitchell, Shanka Subhra Mondal, Merlin Moore, Shang Mu, Taliah Muhammad, Barak Nehoran, Oluwaseun Ogedengbe, Christos Papadopoulos, Stelios Papadopoulos, Saumil Patel, Xaq Pitkow, Sergiy Popovych, Anthony Ramos, R. Clay Reid, Jacob Reimer, Casey M. Schneider-Mizell, H. Sebastian Seung, Ben Silverman, William Silversmith, Amy Sterling, Fabian H. Sinz, Cameron L. Smith, Shelby Suckow, Marc Takeno, Zheng H. Tan, Andreas S. Tolias, Russel Torres, Nicholas L. Turner, Edgar Y. Walker, Tianyu Wang, Grace Williams, Sarah Williams, Kyle Willie, Ryan Willie, William Wong, Jingpeng Wu, Chris Xu, Runzhe Yang, Dimitri Yatsenko, Fei Ye, Wenjing Yin, and Szi-chieh Yu. Functional connectomics spanning multiple areas of mouse visual cortex. *bioRxiv*, 2021.

[8] 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, Neha Karlupia, Sven Dorkenwald, Evelina Sjostedt, Laramie Leavitt, Dongil Lee, Jakob Troidl, Forrest Collman, Luke Bailey, Angerica Fitzmaurice, Rohin Kar, Benjamin Field, Hank Wu, Julian Wagner-Carena, David Aley, Joanna Lau, Zudi Lin, Donglai Wei, Hanspeter Pfister, Adi Peleg, Viren Jain, and Jeff W Lichtman. A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science*, 384(6696):eadk4858, 2024. doi: 10.1126/science.adk4858. URL https://www.science.org/doi/abs/10.1126/science.adk4858.

[9] Narayanan Kasthuri, Kenneth Jeffrey Hayworth, Daniel Raimund Berger, Richard Lee Schalek, José Angel Conchello, Seymour Knowles-Barley, Dongil Lee, Amelio Vázquez-Reina, Verena Kaynig, Thouis Raymond Jones, Mike Roberts, Josh Lyskowski Morgan, Juan Carlos Tapia, H. Sebastian Seung, William Gray Roncal, Joshua Tzvi Vogelstein, Randal Burns, Daniel Lewis Sussman, Carey Eldin Priebe, Hanspeter Pfister, and Jeff William Lichtman. Saturated Reconstruction of a Volume of Neocortex. *Cell*, 162(3), 2015. ISSN 10974172. doi: 10.1016/j.cell.2015.06.054. 1, 6

[10] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. The MIT Press, 2012. ISBN 026201825X. 1

[11] Philipp Schlegel, Yijie Yin, Alexander Shakeel Bates, Sven Dorkenwald, Katharina Eichler, Paul Brooks, Daniel S Han, Marina Gkantia, Marcia Dos Santos, Eva J Munnelly, Griffin Badalamente, Laia Serratosa Capdevila, Varun Aniruddha Sane, Markus William Pleijzier, Imaan F M Tamimi, Christopher R Dunne, Irene Salgarella, Alexandre Javier, Siqi Fang, Eric Perlman, Tom Kazimiers, Sridhar R Jagannathan, Arie Matsliah, Amy R Sterling, Szi-Chieh Yu, Claire E McKellar, FlyWire Consortium, Marta Costa, H Sebastian Seung, Mala Murthy, Volker Hartenstein, Davi D Bock, and Gregory S X E Jefferis. Whole-brain annotation and multi-connectome cell typing quantifies circuit stereotypy in Drosophila. *bioRxiv : the preprint server for biology*, 2023. doi: 10.1101/2023.06.27.546055. 1

[12] Kenneth A. Pelkey, Ramesh Chittajallu, Michael T. Craig, Ludovic Tricoire, Jason C. Wester, and Chris J. McBain. Hippocampal gabaergic inhibitory interneurons. *Physiological Reviews*, 97(4), 2017. ISSN 15221210. doi: 10.1152/physrev.00007.2017. 1, 7

[13] Viktória K. Vereczki, Kinga Müller, Éva Krizsán, Zoltán Máté, Zsuzsanna Fekete, Laura Rovira-Esteban, Judit M. Veres, Ferenc Erdélyi, and Norbert Hájos. Total number and ratio of GABAergic neuron types in the mouse lateral and basal amygdala. *Journal of Neuroscience*, 41(21), 2021. ISSN 15292401. doi: 10.1523/JNEUROSCI.2700-20.2021. 1, 7

[14] P. L.A. Gabbott and P. Somogyi. Quantitative distribution of GABA-immunoreactive neurons in the visual cortex (area 17) of the cat. *Experimental Brain Research*, 61(2), 1986. ISSN 00144819. doi: 10.1007/BF00239522. 1, 7

[15] Douglas L. Oliver, Jeffery A. Winer, Gretchen E. Beckius, and Richard L.Saint Marie. Morphology of GABAergic neurons in the inferior colliculus of the cat. *Journal of Comparative Neurology*, 340(1), 1994. ISSN 10969861. doi: 10.1002/cne.903400104. 1, 7

[16] S. H. Hendry, H. D. Schwark, E. G. Jones, and J. Yan. Numbers and proportions of GABA-immunoreactive neurons in different areas of monkey cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 7(5), 1987. ISSN 02706474. doi: 10.1523/jneurosci.07-05-01503.1987. 1, 7

[17] Jean Pierre Hornung and N. De Tribolet. Distribution of GABA-containing neurons in human frontal cortex: a quantitative immunocytochemical study. *Anatomy and Embryology*, 189(2), 1994. ISSN 03402061. doi: 10.1007/BF00185772. 1, 7

[18] Marc Corrales, Benjamin T. Cocanougher, Andrea B. Kohn, Jason D. Wittenbach, Xi S. Long, Andrew Lemire, Albert Cardona, Robert H. Singer, Leonid L. Moroz, and Marta Zlatic. A single-cell transcriptomic atlas of complete insect nervous systems across multiple life stages. *Neural Development*, 17(1), 2022. ISSN 17498104. doi: 10.1186/s13064-022-00164-6. 2, 5, 7, 9

[19] Britt Anderson, Jessie J. Peissig, Jedediah Singer, and David L. Sheinberg. XOR style tasks for testing visual object processing in monkeys. *Vision Research*, 46(11), 2006. ISSN 00426989. doi: 10.1016/j.visres.2005.11.023. 2, 3

[20] J David Smith, John Paul Minda, and David A Washburn. Category learning in rhesus monkeys: a study of the Shepard, Hovland, and Jenkins (1961) tasks. *Journal of Experimental Psychology: General*, 133(3):398, 2004.

[21] Chris I Baker, Marlene Behrmann, and Carl R Olson. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature neuroscience*, 5(11):1210–1216, 2002.

[22] S J Thorpe, K O'Regan, and A Pouget. Humans fail on XOR pattern classification problems. *Neural networks from models to applications*, pages 12–25, 1989. 3

[23] Marvin Minsky and Seymour Papert. *Perceptrons.* M.I.T. Press, Oxford, England, 1969. 3

[24] Qingyang Wang, Mike A Powell, Ali Geisa, Eric Bridgeford, Carey E Priebe, and Joshua T Vogelstein. Why do networks have inhibitory/negative connections? In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 22551–22559, 10 2023. 3, 4

[25] Ran Rubin, L. F. Abbott, and Haim Sompolinsky. Balanced excitation and inhibition are required for high-capacity, noise-robust neuronal selectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 114(44), 2017. ISSN 10916490. doi: 10.1073/pnas.1705841114. 6

[26] Vittorio Capano, Hans J. Herrmann, and Lucilla De Arcangelis. Optimal percentage of inhibitory synapses in multi-task learning. *Scientific Reports*, 5, 2015. ISSN 20452322. doi: 10.1038/srep09895. 6

[27] Ezekiel Williams, Avery Hee-Woon Ryoo, Thomas Jiralerspong, Alexandre Payeur, Matthew G Perich, Luca Mazzucato, and Guillaume Lajoie. Expressivity of Neural Networks with Random Weights and Learned Biases, 2024. URL https://arxiv.org/abs/2407.00957. 6

[28] Qingyang Wang, Michael Alan Powell, Eric W Bridgeford, Ali Geisa, and Joshua T Vogelstein. Polarity Is All You Need to Learn and Transfer Faster. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 36264–36284. PMLR, 11 2023. URL https://proceedings.mlr.press/v202/wang23ae.html. 6

[29] Sueyeon Chung, Daniel D. Lee, and Haim Sompolinsky. Classification and Geometry of General Perceptual Manifolds. *Physical Review X*, 8(3):1–27, 2018. ISSN 21603308. doi: 10.1103/PhysRevX.8.031003. 6

[30] Uri Cohen, Sue Yeon Chung, Daniel D. Lee, and Haim Sompolinsky. Separability and geometry of object manifolds in deep neural networks. *Nature Communications*, 11(1):1–13, 2020. ISSN 20411723. doi: 10.1038/s41467-020-14578-5. URL http://dx.doi.org/10.1038/s41467-020-14578-5.

[31] Chi-Ning Chou, Luke Arend, Albert J Wakhloo, Royoung Kim, Will Slatton, and SueYeon Chung. Neural Manifold Capacity Captures Representation Geometry, Correlations, and Task-Efficiency Across Species and Behaviors. *bioRxiv*, page 2024.02.26.582157, 1 2024. doi: 10.1101/2024.02.26.582157. URL http://biorxiv.org/content/early/2024/02/28/2024.02.26.582157.abstract.

[32] Francesca Mignacco, Chi-Ning Chou, and SueYeon Chung. Nonlinear classification of neural manifolds with contextual information, 2024. URL https://arxiv.org/abs/2405.06851. 6

[33] V. Koltchinskii. Rademacher penalties and structural risk minimization. *IEEE Transactions on Information Theory*, 47(5), 2001. ISSN 00189448. doi: 10.1109/18.930926. 6

[34] V. N. Vapnik and A. Ya. Chervonenkis. On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. *Theory of Probability & Its Applications*, 16(2), 1971. ISSN 0040-585X. doi: 10.1137/1116025. 6

[35] Elise C. Croteau-Chonka, Michael S. Clayton, Lalanti Venkatasubramanian, Samuel N. Harris, Benjamin M.W. Jones, Lakshmi Narayan, Michael Winding, Jean Baptiste Masson, Marta Zlatic, and Kristina T. Klein. High-throughput automated methods for classical and operant conditioning of Drosophila larvae, 2022. ISSN 2050084X. 6

[36] Joshua T. Vogelstein, Youngser Park, Tomoko Ohyama, Rex A. Kerr, James W. Truman, Carey E. Priebe, and Marta Zlatic. Discovery of brainwide neural-behavioral maps via multiscale unsupervised structure learning. *Science*, 344(6182):386–392, 2014. ISSN 10959203. doi: 10.1126/science.1250298. 6

[37] Sahil Loomba, Jakob Straehle, Vijayan Gangadharan, Natalie Heike, Abdelrahman Khalifa, Alessandro Motta, Niansheng Ju, Meike Sievers, Jens Gempt, Hanno S. Meyer, and Moritz Helmstaedter. Connectomic comparison of mouse and human cortex. *Science*, 377(6602), 2022. ISSN 10959203. doi: 10.1126/science.abo0924. 6

[38] Xiao Jing Wang. Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nature Reviews Neuroscience*, 21(3), 2020. ISSN 14710048. doi: 10.1038/s41583-020-0262-x. 7

[39] Jonathan Frankle and Michael Carbin. The lottery ticket hypothesis: Finding sparse, trainable neural networks. In *7th International Conference on Learning Representations, ICLR 2019*, 2019. URL https://openreview.net/forum?id=rJl-b3RcF7. 7

[40] Hattie Zhou, Janice Lan, Rosanne Liu, and Jason Yosinski. Deconstructing lottery tickets: Zeros, signs, and the supermask. In *Advances in Neural Information Processing Systems*, volume 32, 2019. 7

A Supplementary Figures

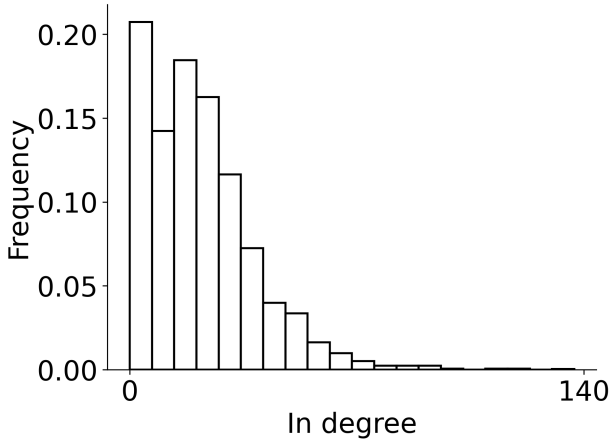


Figure A.1: **The in degree distribution of larval *Drosophila* EM connectome.**

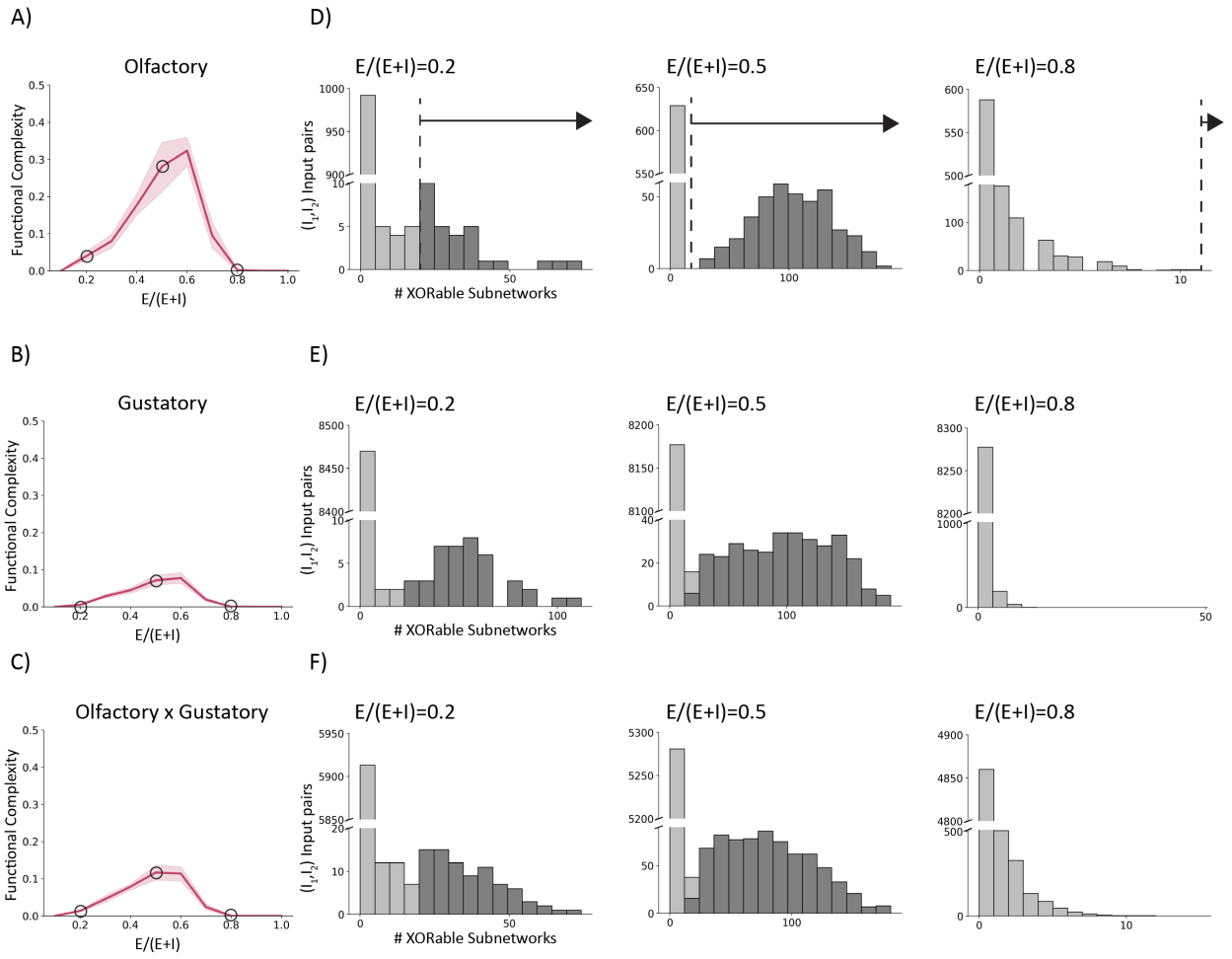


Figure A.2: **Distribution of XORable subnetworks when E-I identities are uniformly sampled across the entire network. (related to Figure 2)** A-C) Functional complexity of networks with different E-I ratios (E-I identities sampled uniformly for all neurons in the EM-constrained RNN). The functional complexities are measured with subnetworks receiving inputs from different sensory modalities, $\{I_1, I_2\} =$: A) $\{\text{Olfactory}_1, \text{Olfactory}_2\}$, B) $\{\text{Gustatory}_1, \text{Gustatory}_2\}$, C) $\{\text{Olfactory}, \text{Gustatory}\}$. Shaded area mark the 95% confidence interval (n=10). Circles mark the corresponding points shown in D-F on the right. D-F) Distributions of XORable subnetworks grouped by the $\{I_1, I_2\}$ input pairs. Each histogram is of a particular sensory modality (given on the left), and of a single RNN with E-I identities sampled at the ratio indicated on top. The functional complexity is then calculated by the fraction of XORable subnetworks, excluding the ones where a unique $\{I_1, I_2\}$ input pair have less than 20 output neurons exhibitng XORable pattern (out of 400 in total) (details in Sec 4). Subnetworks counted towards functional complexity are marked by dark gray.

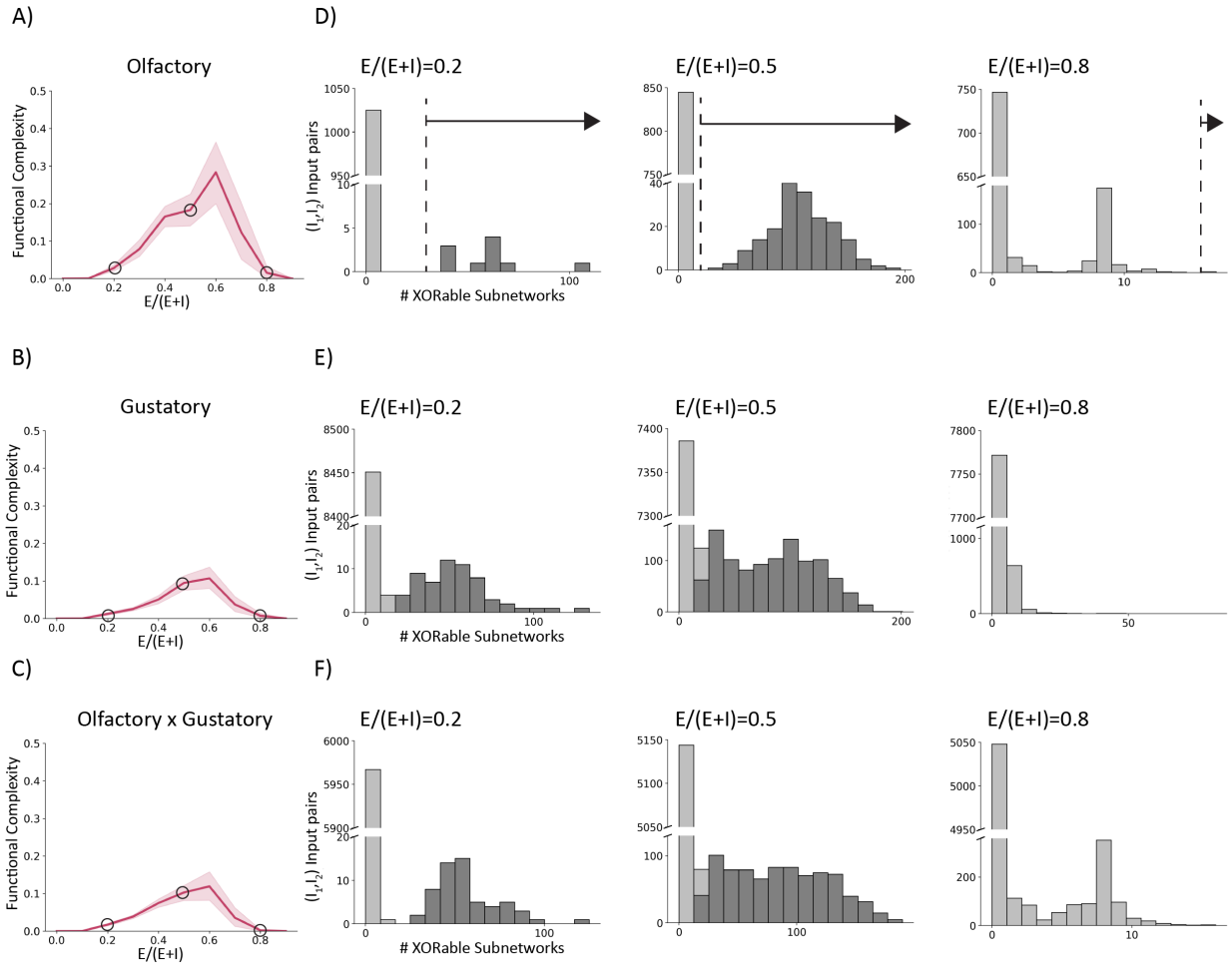


Figure A.3: **Same as Supplementary Figure A.2 except including all-to-all synaptic counts as the EM constraint. (related to Figure 2)**

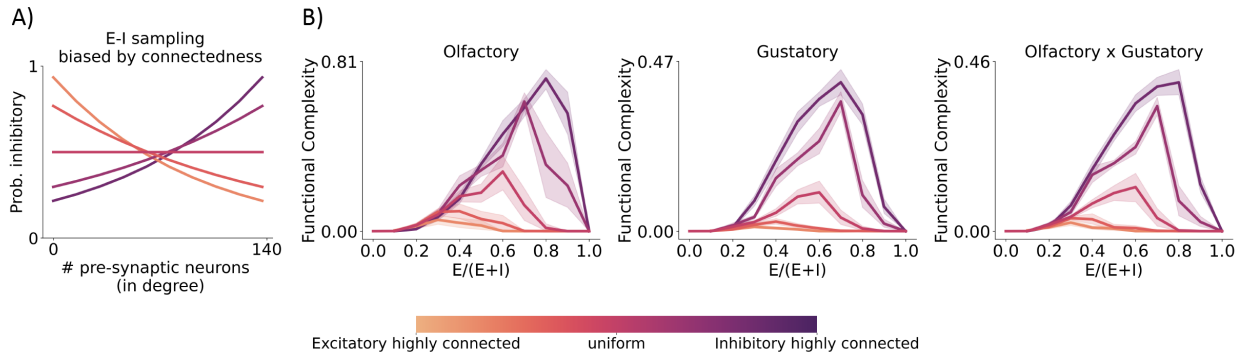


Figure A.4: **Same as Figure 2 except including all-to-all synaptic counts as the EM constraint.**

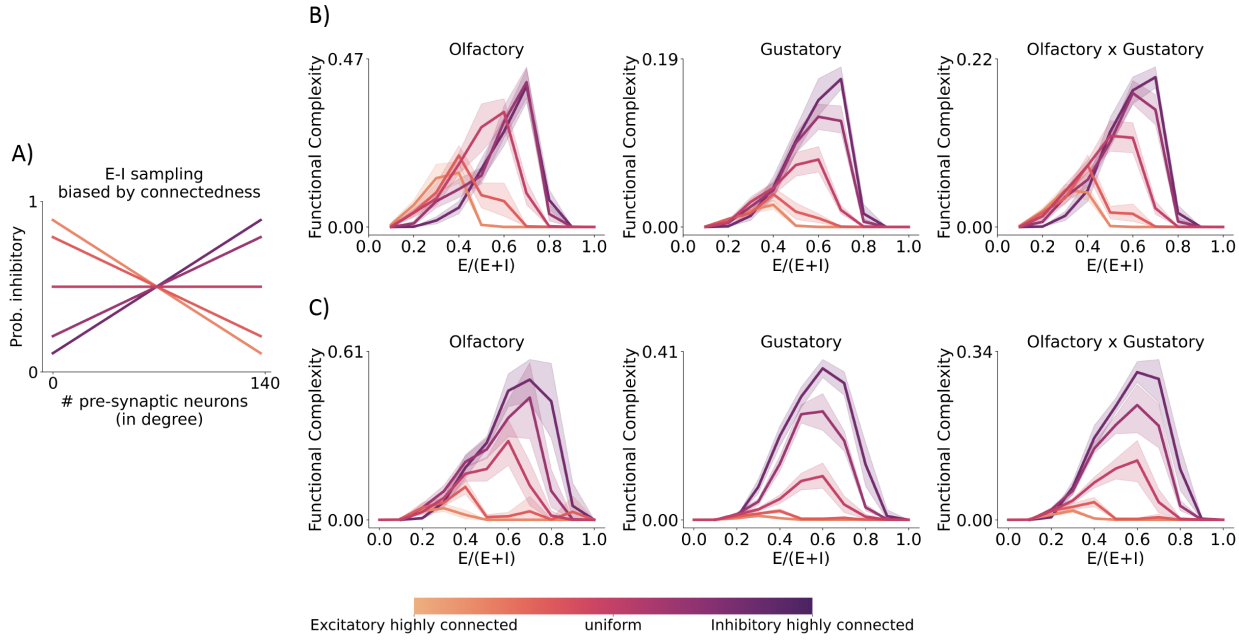


Figure A.5: **Same as Figure 2 except for linear degree-dependent probability vectors.** A) Linear degree-dependent E-I probability. B) Results for axon-dendrite synapses as the EM constraint. C) Results for all-to-all synapses as the EM constraint.

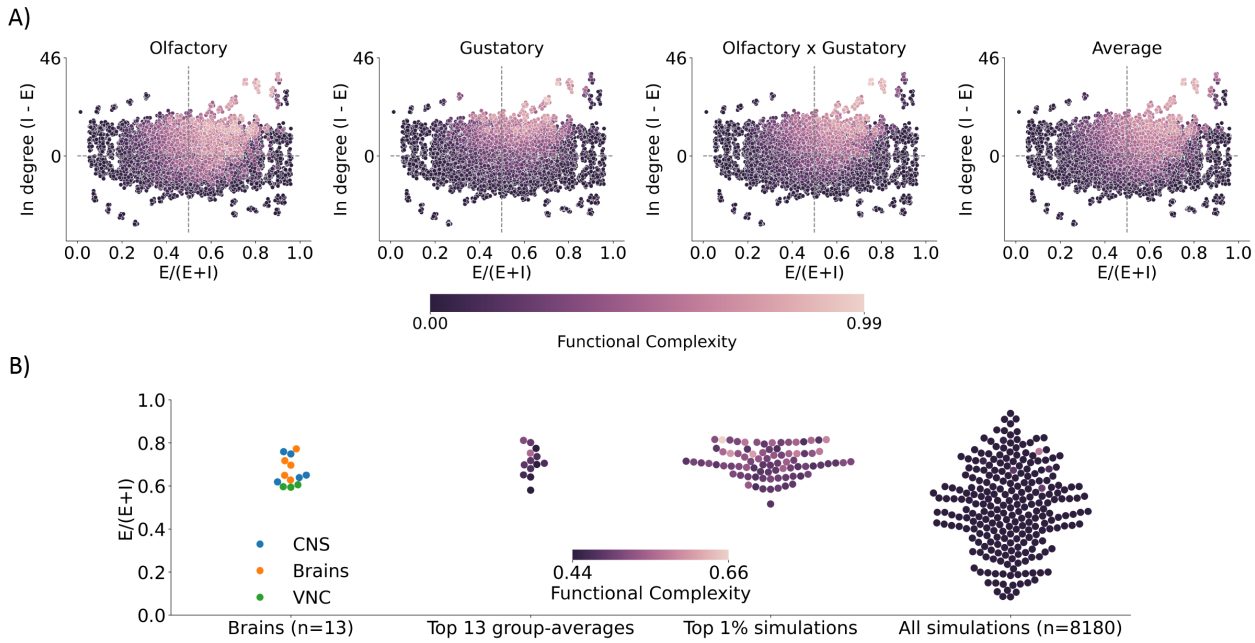
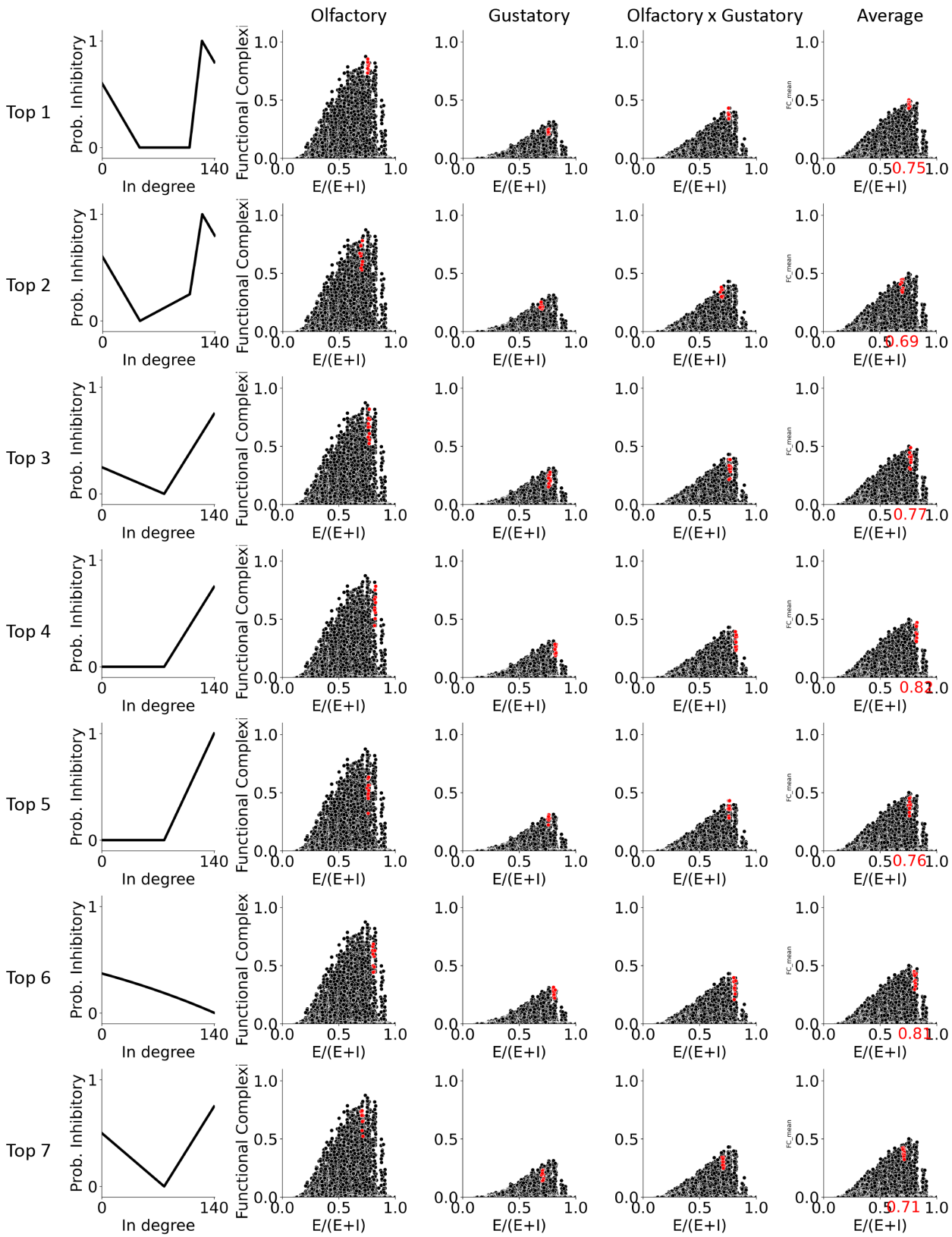


Figure A.6: **Same as Figure 3 except including all-to-all synaptic counts as the EM constraint.** The mean E-I ratio of the top 13 group-averages and top 1% simulations are **not** statistical significantly different from the 13 brain E-I ratios (two-sided two-sample t-test, p-val: 0.82, 0.70 respectively); while the mean of all simulations are statistical significantly different from the brain observations (p-val: 0.01). Multiple comparisons corrected by Dunnett's.



(cont'd on next page.)

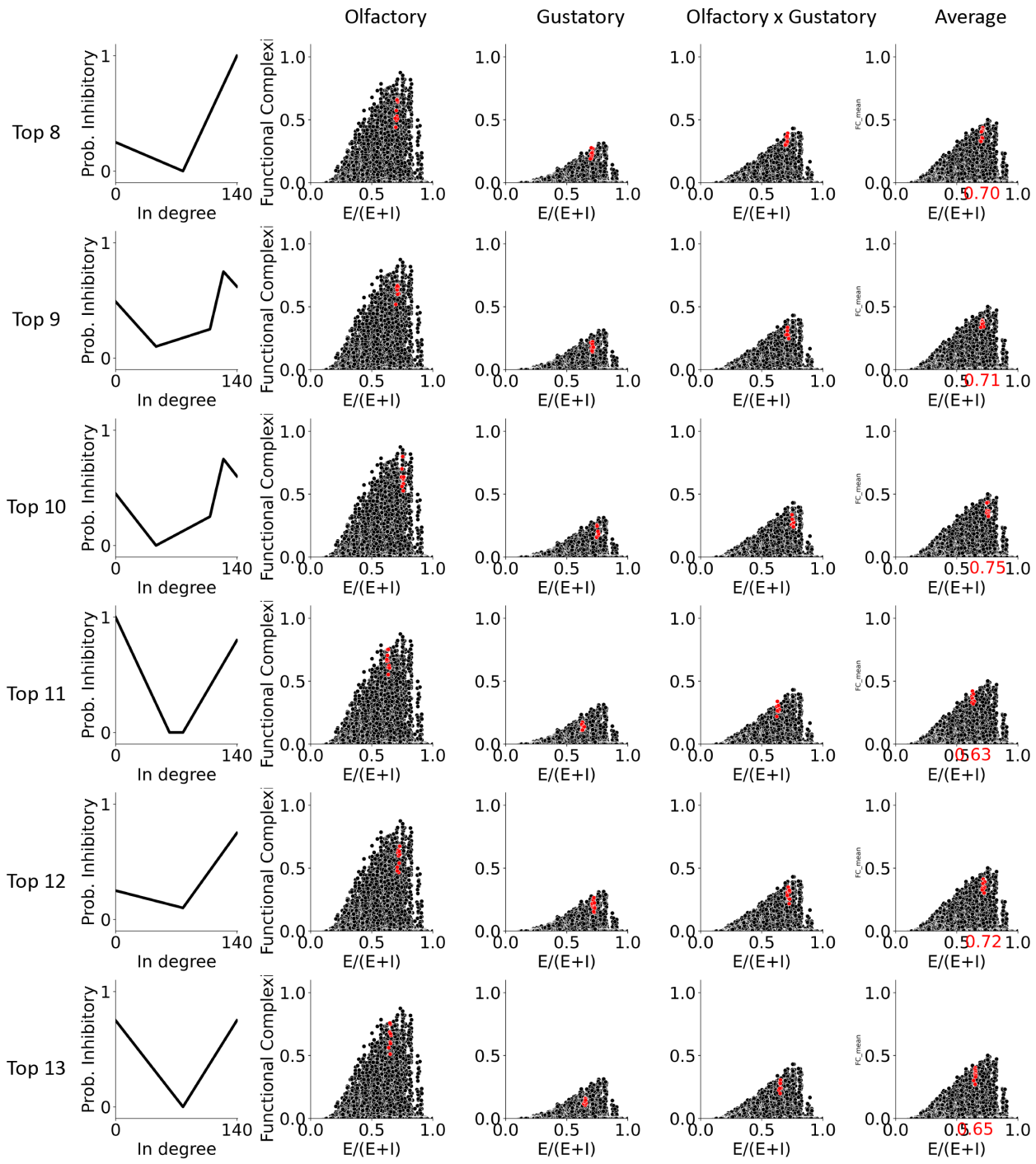
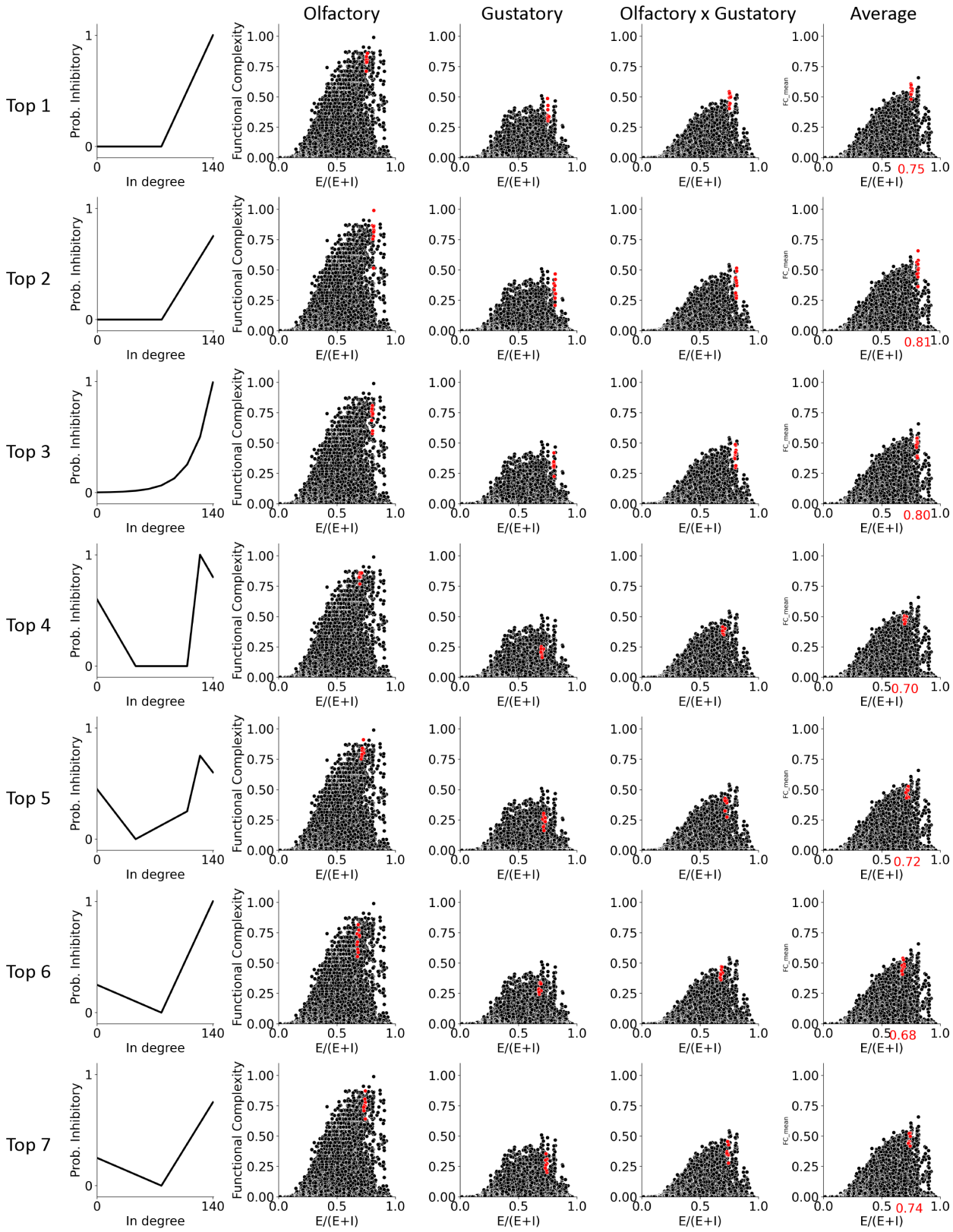


Figure A.7: **Top 13 degree-dependent E-I probability functions. (Related to Figure 3)** The top 13 E-I ratios plotted in Figure 3 middle left are delineated here. Top to bottom: the E-I probability function group that resulted in the highest functional complexity to lower.



(cont'd on next page.)

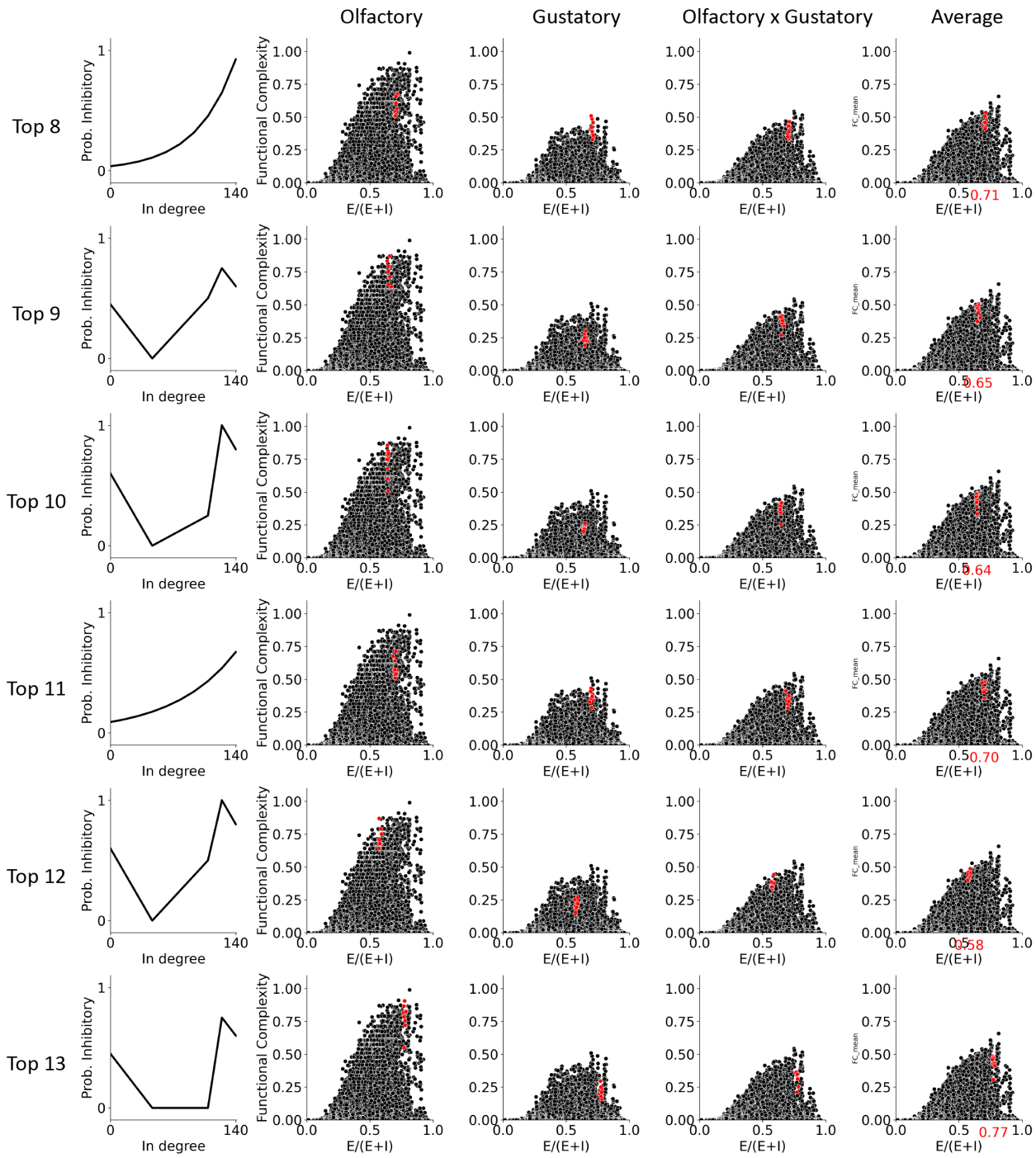


Figure A.8: **Same as Supplementary Figure A.7 except including all-to-all synaptic counts as the EM constraint.**