Multi-sensory integration in biological and artificial systems through an hourglass network architecture

Kamal Shadi (kamalshadi@gatech.edu)
School of Computer Science, Georgia Institute of Technology, Atlanta GA, United States

Eva Dyer (evadyer@gatech.edu)
Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta GA, United States

Majid Mohajerani (mhmohajerani@gmail.com)
Canadian Centre for Behavioral Neuroscience, University of Lethbridge, Lethbridge AB, Canada

Constantine Dovrolis (constantine@gatech.edu)
School of Computer Science, Georgia Institute of Technology, Atlanta GA, United States

Abstract

Multi-sensory integration is a fundamental problem for any embodied cognitive system – both biological and artificial. We pursue a network diffusion approach to model the flow of evoked activity, initiated by stimuli at primary sensory regions. In particular, we apply the Asynchronous Linear Threshold (ALT) diffusion model on the mesoscale cortical connectome of the mouse. The ALT model captures how evoked activity at a given cortical region ripples through the rest of the cortex. Our results show that a small number of regions (the Claustrum being at the top of the list) integrate almost all sensory information paths, suggesting that the cortex relies on an "hourglass architecture" to integrate and compress sensory information before utilizing that lower-dimensionality representation in association areas and higher cognitive tasks.

Keywords: Multi-sensory integration; mouse connectome; network diffusion; hourglass architecture; Claustrum

Introduction

Consider a system that has several sensory inputs (e.g., visual, auditory, haptic) and that needs to encode, integrate and compress this multi-modal stream into a consistent perceptual state before it can respond. Further, the environment may be noisy and time-varying, and so any single sensory stream may not be sufficient for robust perception. Both biological and artificial intelligence systems must solve the problem of *Multi-Sensory Integration* (MSI), and thus this is a fundamental problem for any embodied cognitive agent.

In the brain, MSI takes place across widely distributed networks (Meredith, 2002; Stein, Stanford, Ramachandran, Perrault, & Rowland, 2009). The neural basis of how sensory streams are encoded, integrated and compressed is an increasingly active research area in neuroscience (Meijer, Mertens, Pennartz, Olcese, & Lansink, 2019).

In artificial agents, and in deep learning in particular, we presume that it would not be sensible to perform MSI with a monolithic, very deep network that operates directly on all sensory inputs and learns all higher-level tasks at the same time. Instead, a modular architecture in which different sub-networks operate either on different sensory modalities or on different tasks would be presumably easier to learn, adapt and integrate. But how should this modular architecture be structured? Which are the salient structural principles and constraints it should satisfy? Can we get some insight from neuroscience?

The brain, and especially the mammalian cortex, is a hierarchically modular system (Meunier, Lambiotte, & Bullmore, 2010). Different Regions of Interest (ROIs) in the cortex are associated with different functions (unisensory, multisensory, association, motor control, excutive control, etc). Additionally, these modules are organized in complex hierarchies in which together with the feedforward flow of information that starts from primary sensory ROIs, there are also many feedback connections from higher-level to lower-level modules, as well as lateral connections between modules of the same hierarchical level (Markov et al., 2013).

Our high-level objective is to examine the modular architecture through which the cortex performs MSI, identify the salient properties of this architecture, and to potentially apply these properties in the design of modular artificial neural networks. In this work we focus on the first two steps of this objective.

In particular, we rely on the mouse mesoscale connectome, which has been mapped by the Allen Institute for Brain Science (Oh et al., 2014). The connectome gives us the anatomical substrate upon which we apply a network diffusion model. We model the flow of evoked activity, initiated by stimuli at primary sensory ROIs, using the *Asynchronous Linear Threshold* (ALT) model. ALT captures how evoked activity that originates at a given brain ROI "ripples through" the rest of the brain. The weighted version of the ALT model assumes that a node becomes active when more than a weighted fraction of the neighboring nodes are active.

For a given activated source node, we calculate the activation time of all other nodes that participate in the cascade. We then construct a Directed Acyclic Graph (DAG), dubbed Activation-DAG (A-DAG), based on the nodes' activation times for each type of sensory stimulation. We analyze the result-

ing A-DAGs to identify salient architectural properties of the mouse cortex that enable MSI.

Our results show that a small number of cortical ROIs (the Claustrum being at the top of the list) integrate almost all sensory information streams. This suggests that the cortex relies on an *hourglass architecture* (see Fig. 1). An hourglass architecture has many input modules at one end of the architecture (one or more for each sensory modality), many output modules at the other end (one or more for each cognitive or motor task) and only a small number of core modules at the waist of the architecture (Friedlander, Mayo, Tlusty, & Alon, 2015; Sabrin & Dovrolis, 2017; Sabrin, Wei, van den Heuvel, & Dovrolis, 2019). The objective of those core modules is to compute representations of the input space that are both efficient (i.e., of low dimensionality compared to the inputs) and accurate (i.e., they can capture almost all of the input variance). Further, these core modules tend to remain invariant, even when there are changes in the input and/or output modules (Siyari, Dilkina, & Dovrolis, 2018).

To validate the ALT model, we use Voltage Sensitive Dye (VSD) imaging data (Mohajerani et al., 2013). VSD data was collected from mice while different sensory stimuli were introduced to evoke responses in the visual, somatosensory (upper limb, lower limb and whisker), and auditory cortices. This analysis corroborated the predictive power of ALT in modeling diffusion of activity from specific source ROIs.
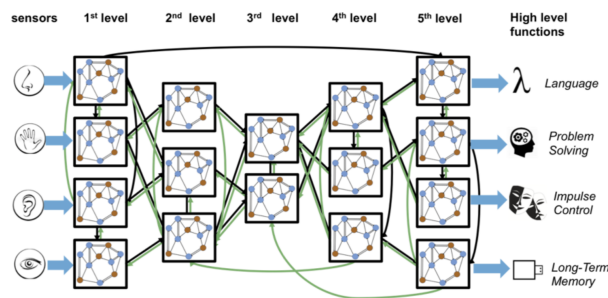

Figure 1: A hypothetical hierarchy with feedforward, feedback and lateral connections between modules. Input information is provided at sensory-specific modules, while the high-level cognitive tasks are performed by task-specific modules at the other end.

Summary of methods and results

The Allen Mouse Brain Atlas has 213 ROIs (Oh et al., 2014). We consider a network of 67 ROIs that reside in the right hemisphere of the isocortex, olfactory areas, hippocampal formation and cortical subplate. The edge weights we consider are referred to as "connection density" (Oh et al., 2014). This metric is roughly proportional to the average number of axons from the source ROI that target neurons connect to. We filter the anatomical edges with p-value higher than 5%. We denote

the resulting 67-node weighted network by N_c; its density is 13%. Edge lengths are calculated by the Euclidean distance between ROI centroids.

We model the diffusion of evoked neural activity using an Asynchronous Linear Threshold model. The incoming neighbors of a node (or set of nodes) are denoted by $N_{in}(\cdot)$. Initially, the binary state of every node is set to 0, except a given primary sensory ROI whose state is set to 1. Subsequently, the state of each node n_i is updated asynchronously based on the state of all its incoming neighbors' $N_{in}(n_i)$, as follows:

$$s_i(t) = 1 \quad \text{if} \quad \sum_{j \in N_{in}(i)} w_{ji} s_j(t - t_{ji}) > \theta \tag{1}$$

where θ represents the activation threshold, t_{ji} the propagation delay from node j to i and w_{ji} is the connection density of the connection $j \rightarrow i$.

The sensory ROIs we consider as source nodes are: **AUDp** (primary auditory), **VISp** (primary visual), **GU** (gustatory areas), **SSp-n** (somatosensory – nose), **SSp-bfd** (somatosensory – barrel field), **SSp-ll** (somatosensory – lower limb), **SSp-ul** (somatosensory – upper limb), **SSp-m** (somatosensory – mouth), **SSp-tr** (somatosensory – trunk), and **MOB** (main olfactory bulb).

For each of these ten source nodes, we calculate the activation time of all nodes that participate in the ALT cascade. We construct a Directed Acyclic Graph (DAG), hereafter dubbed **Activation DAG** (A-DAG), based on the nodes' activation time. The A-DAG includes an edge $n_i \rightarrow n_j$ if and only if $n_j \in N_{out}(n_i)$ and $t_j^a > t_i^a + t_{ij}$, i.e., only those nodes that contributed to the activation of n_j should point to n_j in the activation DAG.

We have investigated the size of the cascade (number of activated nodes in the cascade) for different values of θ. Interestingly, either all nodes participate in the cascade or almost none of them does. Thus, we set θ to the highest value that causes a complete cascade ($\theta = 0.46$ for MOB, which is the only sensory ROI outside the isocortex, and $\theta = 0.98$ for all other sources). Complete cortical cascades of sensory activity have been experimentally observed (Mohajerani et al., 2013).

ALT is used to capture the activation propagation from a local perturbation to the whole cortical network. This propagation captures the first time the impact of a focal perturbation reaches each ROI. We do not expect ALT to model any subsequent feedback or sustained oscillations between ROIs.

We analyze the resulting A-DAGs based on the activation paths that originate from sensory nodes (the source of each A-DAG) and terminate at each sink node of that A-DAG. The *Path Centrality* (PC) of each node is defined as the fraction of source-to-target paths traversing the node across of all A-DAGs.

We then compute the $\tau - core\ of\ the\ network$ i.e., the minimum set of nodes that covers at least $\tau\%$ of all source-to-target paths across all activation DAGs (Sabrin & Dovrolis, 2017). This NP-Hard problem is approximated by a greedy
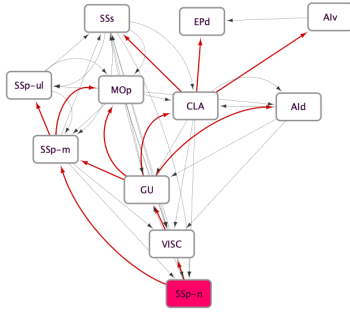
Figure 2: The first 10 activated nodes in the A-DAG of the somatosensory-nose ROI (SSp-n), according to the ALT model. Red edges are connectomic connections that participate in the activation cascade, while grey edges do not.

Table 1: Core nodes across the ten A-DAGs ($\tau = 90\%$).

Node	Description	Coverage	Rank
CLA	Claustrum	24	1
PTLp	Posterior parietal assoc.	19	3
AUDv	Ventral auditory area	19	4
SSs	Supp.somatosensory area	9	2
MOs	Secondary motor area	8	5
ACAd	Anterior cingulate - dorsal	6	16
VISl	Lateral visual area	5	9
AOB	Accessory olfactory bulb	4	7
ECT	Ectorhinal area	4	22

heuristic in which the node with maximum PC is removed from the network in each iteration, joining the $\tau - core$ set. The PC of the remaining nodes is updated after each iteration (Sabrin & Dovrolis, 2017).

Interestingly, when we compute the $\tau - core$ for τ=90% across all A-DAGs, we find only nine nodes. These core nodes are listed in table 1. With each core node, we also list the fraction of source-target paths it covers ("coverage") when it joins the core as well as its PC rank. The table identifies CLA, PTLp and AUDv as the three most important ROIs for MSI – these three ROIs cover more than 50% of the source-target paths in the ten A-DAGs.

The claustrum (CLA) is known for its anatomical uniqueness and its enigmatic function (Crick & Koch, 2005; Van Horn, 2019) – Francis Crick had hypothesized that the Claustrum plays a central role in the emergence of consciousness. In an intriguing experiment, Koubeissi et al. (Koubeissi, Bartolomei, Beltagy, & Picard, 2014) found that by delivering electric pulses to CLA in a human subject, she was immediately driven to a "frozen state" in which she could not continue a reading task – discontinuing the stimulation immediately resumed normal behavior.

The posterior parietal associative area (PTLp) has strong and direct connectivity to primary sensory ROIs and projections to motor areas, and it has been previously identified as a "hub" in the responses from a variety of sensory stimulation responses(Lim et al., 2012; Mohajerani et al., 2013; Nikbakht, Tafreshiha, Zoccolan, & Diamond, 2018).

To examine the robustness of this hourglass observation, we have examined whether the $\tau - core$ stays small, and whether it consists of the same nodes, if we randomize the underlying connectome. Randomizing the connection weights and/or lengths, but preserving their distribution, does *not* have a significant effect on the size or membership of the $\tau - core$ – the hourglass waist is not affected. If we randomize the

topology of the connectome however, by swapping edges in a degree-preserving manner, the $\tau - core$ doubles in size and it changes significantly in membership. We conclude that it is the architecture of the connectome, rather than the weights or lengths of the connections, that are the primary reason behind the hourglass structure of MSI.

We have used *Voltage Sensitive Dye* (VSD) imaging data (Mohajerani et al., 2013) to examine the accuracy of the ALT modeling predictions. VSD imaging enables neural population activity monitoring over large cortical areas and in temporal resolution of few milliseconds. Each VSD video corresponds to a single sensory stimulation experiment on a mouse. The resulting videos have 108 frames at a temporal resolution of 6.67ms. Each frame consists of 128×128 pixels at the spatial resolution of $50\mu m$/pixel. We compare ALT to VSD images for five animals and five sensory stimulation types: visual (VISp), auditory (AUDp), whisker (SSp-bfd), forelimb (SSp-ul) and hindlimb (SSp-ll).

To assign an activation frame for each ROI in the VSD data, we performed the following steps: 1) Find the peak signal for each VSD pixel (the frame at which the peak activity is observed). 2) Register the Allen Atlas ROIs to each animal's native space. 3) Find the activation frame for each ROI based on the activation frame of the majority of that ROI's pixels.

We then compare the temporal ordering of activated ROIs between VSD experiments and ALT modeling. There are three cases: 1) **Temporal agreement** comprises of all ROI pairs for which ALT predicts correctly the experimental activation order. 2) **Insufficient temporal resolution** comprises of all ROI pairs that appear to get activated in the same VSD frame. 3) **Temporal disagreement** comprises of all ROI pairs for which ALT does not predict correctly the experimental activation order. The results are summarized in Fig. 3 – note that the frequency of "temporal agreement" varies between 50-75% across five sensory stimulations and five animals per stimulation – while the frequency of "temporal disagreement" varies between 0-20%.

Brief discussion

The mechanism through which the brain integrates different sensory inputs and reaches a stable unified perception has been under investigation at different spatial and temporal scales, with classical works focusing on subcortical ROIs, such as the Superior Colliculus (Wallace & Stein, 1997; Stein
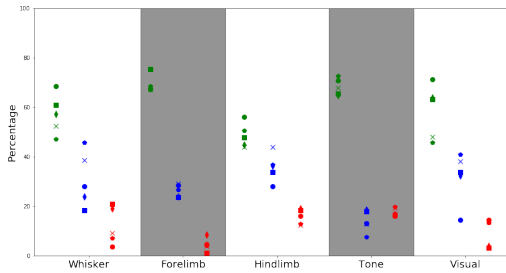
Figure 3: Percentage of ROI pairs that show temporal agreement (green), insufficient temporal resolution (blue), and temporal disagreement (red) for five different stimulation types (x-axis labels) and five different animals (different symbols).

et al., 2009) to more recent larger scale foci (Toker & Sommer, 2019; Worrell, Rumschlag, Betzel, Sporns, & Mišić, 2017). After compiling decades of MSI work, Meijer et al. concluded that MSI research has been overwhelmingly focused on behavioral-level integration of external stimuli and the associated micro-circuitry, and he proposed a shift of focus to "a systems neuroscience approach, with rodents as the prime model, to investigate how the neocortex combines sensory stimuli of different modalities" (Meijer et al., 2019). This is the line of inquiry we have followed here. We found that (a) information flow originating from primary sensory cortices is reasonably well predicted by a linear threshold-based network diffusion model, (b) a small number of cortical ROIs integrate and mediate almost all sensory pathways, and (c) these *"core nodes"* are anatomically close to the primary sensory ROIs in the cortex (not shown here due to space constraints).

These findings support the existence of an "hourglass architecture" in the integration of multisensory information. The benefit of an hourglass architecture is that it first reduces the dimensionality of the inputs to a much lower dimensionality latent representation at the "hourglass waist", and second, it re-uses those compressed intermediate-level features in more than one higher-level tasks (Sabrin & Dovrolis, 2017).

References

Crick, F. C., & Koch, C. (2005). What is the function of the claustrum? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1271–1279.

Friedlander, T., Mayo, A. E., Tlusty, T., & Alon, U. (2015). Evolution of bow-tie architectures in biology. *PLoS computational biology*, *11*(3), e1004055.

Koubeissi, M. Z., Bartolomei, F., Beltagy, A., & Picard, F. (2014). Electrical stimulation of a small brain area reversibly disrupts consciousness. *Epilepsy & Behavior*, *37*, 32–35.

Lim, D. H., Mohajerani, M. H., LeDue, J., Boyd, J., Chen, S., & Murphy, T. H. (2012). In vivo large-scale cortical mapping using channelrhodopsin-2 stimulation in transgenic mice reveals asymmetric and reciprocal relationships between cortical areas. *Frontiers in neural circuits*, *6*, 11.

Markov, N. T., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., & Kennedy, H. (2013). Cortical high-density counterstream architectures. *Science*, *342*(6158), 1238406.

Meijer, G. T., Mertens, P. E., Pennartz, C. M., Olcese, U., & Lansink, C. S. (2019). The circuit architecture of cortical multisensory processing: Distinct functions jointly operating within a common anatomical network. *Progress in neurobiology*.

Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: a brief overview. *Cognitive brain research*, *14*(1), 31–40.

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, *4*, 200.

Mohajerani, M. H., Chan, A. W., Mohsenvand, M., LeDue, J., Liu, R., McVea, D. A., . . . Murphy, T. H. (2013). Spontaneous cortical activity alternates between motifs defined by regional axonal projections. *Nature neuroscience*, *16*(10), 1426.

Nikbakht, N., Tafreshiha, A., Zoccolan, D., & Diamond, M. E. (2018). Supralinear and supramodal integration of visual and tactile signals in rats: psychophysics and neuronal mechanisms. *Neuron*, *97*(3), 626–639.

Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., . . . Zeng, H. (2014). A mesoscale connectome of the mouse brain. *Nature*, *508*(7495), 207–214. Retrieved from http://dx.doi.org/10.1038/nature13186 doi: 10.1038/nature13186

Sabrin, K. M., & Dovrolis, C. (2017). The hourglass effect in hierarchical dependency networks. *Network Science*, *5*(4), 490–528.

Sabrin, K. M., Wei, Y., van den Heuvel, M., & Dovrolis, C. (2019). The hourglass organization of the c. elegans connectome. *bioRxiv*, 600999.

Siyari, P., Dilkina, B., & Dovrolis, C. (2018). Emergence and evolution of hierarchical structure in complex systems. *arXiv preprint arXiv:1805.04924*.

Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J., & Rowland, B. A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Experimental Brain Research*, *198*(2-3), 113.

Toker, D., & Sommer, F. T. (2019). Information integration in large brain networks. *PLoS computational biology*, *15*(2), e1006807.

Van Horn, J. D. (2019). *What is old is new again: Investigating and analyzing the mysteries of the claustrum.* Springer.

Wallace, M. T., & Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *Journal of Neuroscience*, *17*(7), 2429–2444.

Worrell, J. C., Rumschlag, J., Betzel, R. F., Sporns, O., & Mišić, B. (2017). Optimized connectome architecture for sensory-motor integration. *Network Neuroscience*, *1*(4), 415–430.