TOPOLOGICAL MODEL OF NEURAL INFORMATION NETWORKS

MATILDE MARCOLLI

Abstract. This is a brief overview of an ongoing research project, involving topological models of neural information networks and the development of new versions of associated information measures that can be seen as possible alternatives to integrated information. Among the goals are a geometric modeling of a "space of qualia" and an associated mechanism that constructs and transforms representations from neural codes topologically. The more mathematical aspects of this project stem from the recent joint work of the author and Yuri Manin, [16], while the neuroscience modeling aspects are part of an ongoing collaboration of the author with Doris Tsao.

1. Motivation

Before describing the main aspects of this project, it is useful to present some of the main motivations behind it.

1.1. **Homology as functional to processing stimuli.** In recent experiments on rhesus monkey, Tevin Rouse at Marlene Cohen's lab at Carnegie Mellon showed that visual attention increases firing rates while decreasing correlation of fluctuations of pairs of neurons, and also showed that topological analysis of the activated networks of V4 neurons reveals a peak of non-trivial homology generators in response to stimulus processing during visual attention, [22].

A similar phenomenon, with the formation of a high number of non-trivial homology generators in response to stimuli, was produced in the analysis directed by the topologist Katrin Hess [21] of simulations of the neocortical microcircuitry. The analysis of [21] proposes the interpretation that these topological structures are necessary for the processing of stimuli in the brain cortex.

These findings are very intriguing for two reasons: (1) they link topological structures in the activated neural circuitry to phenomena like attention; (2) they suggest that a sufficient amount of topological complexity serves a functional computational purpose.

The first point will be discussed more at length below, in §3. The second point is relevant because it suggests a better mathematical setting for modeling neural information networks architecture in the brain. Indeed, although the work of [21] does not offer a theoretical explanation of why topology is needed for stimulus processing, there is a well known context in the theory of computation where a similar situation occurs, which may provide the key for the correct interpretation, namely the theory of concurrent and distributed computing [10], [12].

In the mathematical theory of concurrent and distributed computing, one considers a collection of sequential computing entities (processes) that cooperate to solve a problem (task). The processes communicate by applying operations to objects in a shared memory, and they are asynchronous, in the sense that they run at arbitrary varying speeds.

Concurrent and distributed algorithms and protocols decide how and when each process communicates and shares with others. The main questions are how to design such algorithms that are efficient in the presence of noise, failures of communication, and delays, and how to understand when a an algorithm exists to solve a particular task.

Protocols for concurrent and distributed computing are modeled using (directed) simplicial sets. For example, in distributed algorithms an initial or final state of a process is a vertex, any $d+1$ mutually compatible initial or final states are a d-dimensional simplex, and each vertex is labelled by a different process. The complete set of all possible initial and final states is then a simplicial set. A decision task consists of two simplicial sets of initial and final states and a simplicial map (or more generally correspondence) between them. The typical structure consists of an input complex, a protocol complex, and an output complex, with a certain number of topology changes along the execution of the protocol, [12].

There are very interesting topological obstruction results in the theory of distributed computing, [12], which show that a sufficient amount of non-trivial homology in the protocol complex is necessary for a decision task problem to be solvable. Thus, the theory of distributed computing shows explicitly a setting where a sufficient amount of topological complexity (measured by non-trivial homology) is necessary for computation.

The working hypothesis we would like to consider here is that the brain requires a sufficient amount of non-trivial homology *because* it is carrying out a concurrent/distributed computing task that can only run successfully in the presence of enough nontrivial homology. This suggests that the mathematical modeling of architectures of neural information networks should be formulated in such a way as to incorporate additional structure keeping track of associated concurrent/distributed computational systems.

An important aspect of the topological analysis of brain networks is the presence of directed structures on the simplicial sets involved. A directed structure is not an orientation, but rather a kind of "flow of information" structure, as in concurrent and distributed computing, [10]. A directed n-cube models the concurrent execution of n simultaneous transitions between states in a concurrent machine. In directed algebraic topology one considers topological (simplicial) spaces where paths and homotopies have a preferred direction and cannot be reversed, and directed simplicial sets can be regarded as a higher dimensional generalization of the usual notion of directed graphs. The categorical viewpoint is natural when one thinks of paths in a directed graph as forming a free category generated by the edges. A similar notion for higher dimensional directed complexes exists in the form of ω-categories, [13], [25].

1.2. **Neural code and homotopical representations.** It is known from the work of Curto and collaborators (see [7] and references therein) that the geometry of the stimulus space can be reconstructed *up to homotopy* from the binary structure of the neural code. The key observation behind this reconstruction result is a simple topological property: under the reasonable assumption that the place field of a neuron (the preferred region of the stimulus space that causes the neuron to respond with a high firing rate) is a convex open set, the binary code words in the neural code represent the overlaps between these regions, which determine a simplicial complex (the simplicial nerve of the open covering of the stimulus space). Under the convexity hypothesis the homotopy type of this simplicial complex is the same as the homotopy type of the stimulus space. Thus, the fact that the binary neural code captures the complete information on the intersections between the place fields of the individual neurons is sufficient to reconstruct the stimulus space, but only up to homotopy.

This suggests another working hypothesis. One can read the result about the neural code and the stimulus space reconstruction as suggesting that the brain *represents* the stimulus space through a *homotopy type*. We intend to use this idea as the basic starting point and propose that the experience of *qualia* should be understood in terms of the processing of external stimuli via the construction of associated homotopy types.

The homotopy equivalence relation in topology is weaker but also more flexible than the notion of homeomorphism. The most significant topological invariants, such as homotopy and homology groups, are homotopy invariants. Heuristically, homotopy describes the possibility of deforming a topological space in a one parameter family. In particular, a *homotopy type* is an equivalence class of topological spaces up to (weak) homotopy equivalence, which roughly means that only the information about the space that is captured by its homotopy groups is retained. There is a direct connection between the formulation of topology at the level of homotopy types and "higher categorical structures". For our purposes here it is also worth mentioning that simplicial sets provide a good combinatorial models for the (weak) homotopy type of topological spaces.

Thus, these results suggest that a good mathematical modeling of network architectures in the brain should also include a mechanism that generates homotopy types, through the information carried by the network via neural codes.

1.3. **Consciousness and Informational Complexity.**

The idea of computational models of consciousness based on some measure of informational complexity is appealing in as it proposes quantifiable measures of consciousness tied up to an appropriate notion of complexity and interrelatedness of the underlying system.

Tononi's *integrated information* theory (or Φ function) is a proposal for such a quantitative correlate of consciousness, which roughly measures of the least amount of effective information in a whole system that is not accounted for by the effective information of its separate parts. More precisely, effective information is computed by a Kullback–Leibler divergence (relative entropy) between an observed output probability distribution and a hypothetical uniformly distributed input. The integrated information considers all possible partitions of a system into subsystems (two disjoint sets of nodes in a network for instance) and assigns to each partition a symmetrized and normalized measure of effective information with inputs and outputs on the two sides of the partition, and minimizes over all partitions. The main idea is therefore that integrated information is a measure of informational complexity and interconnectedness of a system, [26]. It is proposed as a measure of consciousness because it appears to capture quantitatively several aspects of conscious experience (such as its compositional and integrated qualities, see [26]).

An attractive feature of the integrated information model is an idea of consciousness that is not a binary state (an entity either does or does not have consciousness) but that varies by continuous degrees and is tied up to the intuitive understanding that sufficient complexity and interconnectedness is required for the acquisition of a greater amount of consciousness. The significant ethical implications of these assumptions are evident. On the other hand, one of the main objections to using integrated information as a measure of consciousness is the fact that, while it is easy to believe that sufficient informational complexity may be a *necessary* condition for consciousness, the assumption that it would also be a *sufficient* condition appears less justifiable.

Indeed, this approach to a mathematical modeling of consciousness has been criticized on the ground that it is easy to construct simple mathematical models exhibiting a very high value of the Φ function (Aaronson's Vandermonde matrix example). Generally, one can resort to the setting of coding theory to generate many examples of sufficiently good codes (for example the algebro-geometric Reed-Solomon error-correcting codes) that indeed exhibit precisely the typical from of high interconnectedness that leads to large values of integrated information.

Another valid objection to the use of integrated information lies in its computational complexity, given that it requires a computation and minimization over the whole set of partitions of a given set (which in a realistic example like a biological brain is already in itself huge). Other computational difficulties with integrated information and some variants of the same model are discussed in detail in [18].

Our goal is not to improve the mathematical theory of integrated information, but to construct a different kind of mathematical model based on topology and in particular homotopy theory. However, assuming that high values of (some version of) integrated information may indeed provide a valid necessary condition for consciousness, it is natural to ask how a topological model of the kind we will describe in the next section would constrain the values of integrated information.

In our setting, the best approach to infer the existence of lower bounds for some integrated information is to reformulate integrated information itself from a topological viewpoint as discussed in [16], along the lines of the cohomological formulations of information of [3] and [27], [28].

Treating integrated information as cohomological can have advantages in terms of computability: it is well known in the context of simplicial sets that computational complexity at the level of chains can be much higher than at the level of (co)homology.

2. Homotopical Models of Neural Information Networks

We summarize here the main mathematical aspects of this project, with special emphasis on the author's joint work with Manin in [16] on a mathematical framework for a homotopy theoretic modeling of neural information networks in the brain, subject to informational and metabolic constraints, performing some concurrent/distributed computing tasks, and generating associated homotopy types from neural codes structures.

2.1. **Gamma spaces and information networks.**

The starting point of our approach is the notion of Gamma spaces, introduced by Graeme Segal in the 1970s to model homotopy theoretic spectra. A Gamma space is a functor from finite (pointed) sets to simplicial sets. The Segal construction associated such a functor to a given category $\mathcal{C}$ with some basic properties (categorical sum and zero object), by assigning to a finite set X the nerve of a category of $\mathcal{C}$-valued summing functors that map subsets of a given finite set and inclusions between them to objects of the category $\mathcal{C}$ and morphisms in a way that is additive over disjoint subsets. One should think of summing functors as categorically valued measures that can be used to consistently assign a structure (computational, information-theoretic, etc.) to all subsystems of a given system. The enrichment of the notion of Gamma space with probabilistic data was developed in [17]. The nerve of the category of summing functors is a topological model that parameterizes all the possible ways of assigning structures described by the target category $\mathcal{C}$ to the system so that they are consistent over subsystems.

A general mathematical setting for a theory of resources and conversion of resources was developed in [5] and [11] in terms of symmetric monoidal categories. These can be taken as targets of the summing functors in the Gamma space construction. Thus, the language of Gamma spaces can be understood as describing, through the category of summing functors, all possible consistent assignments of resources to subsystems of a given system in such a way that resources of independent subsystems behave additively. The fact that Gamma spaces are built using data of subsystems and relations between subsystems is reminiscent of the ideas underlying the notion of integrated information, but it is also richer and more flexible in its applicability. One of the main advantages of the categorical framework is that the categorical structure ensures that many fundamental properties carry over naturally between different categories. As applications of category theory in other contexts (such as theoretical computer science and physics) have shown, this often renders the resulting constructions much more manageable.

The target monoidal category of resources can describe computational resources (concurrent/distributed computing structures that can be implementable on a given network) by using a category of transition systems in the sense of [29]. One can similarly describe other kinds of metabolic or informational resources. Different kinds of resources and constraints on resources assigned to the same underlying network can be analyzed through functorial relations between these categories. Gamma spaces, via spectra, have associated homotopy groups providing invariants detecting changes of homotopy types.

2.2. **Neural codes and coding theory.** In the theory of error-correcting codes, binary codes are evaluated in terms of their effectiveness at encoding and decoding. This results in a two-parameter space (transmission rate and relative minimum distance). There are bounds in the space of code parameters, such as the Gilbert-Varshamov bound (related to the Shannon entropy and expressing the typical behavior of random codes) or the asymptotic bound (related to Kolmogorov complexity and marking a boundary above which only isolated very good codes exist). Examples of good codes above both the Gilbert-Varshamov and the asymptotic bound typically come from algebro-geometric constructions. As mentioned above, these are precisely the type of codes that can generate high values of integrated information.

It is known, however, that neural codes are typically not good codes according to these parameters. This is essentially because these codes are related to the combinatorics of intersections of sets in open coverings and this can make the Hamming distance between code words very small. Manin suggests in [15] the possible existence of a secondary auxiliary level of neural codes involved in the formation of place maps consisting of very good codes that lie near or above the asymptotic bound. This idea that the bad neural codes are combined in the brain into some higher structure that determines good codes can now be made precise through the recent work of Rishidev Chaudhuri and Ila Fiete, [4], on bipartite expander Hopfield networks. Their model is based on Hopfield networks with stable states determined by sparse constraints with expander structure, which have good error correcting properties and are good models of memory, where the network dynamics extracts higher order correlations from the neural codes.

This suggests that a version of the model of Chaudhuri and Fiete, together with a categorical form of the Hopfield network dynamics discussed below, may be used to address the step from the (poor code) neural codes to the (good codes) bipartite expander Hopfield networks, and from these to a lower bound on the values of an appropriate integrated information measure.

2.3. **Dynamics of networks in a categorical setting.** The setting described in the previous subsection is just a kinematic scaffolding, in the sense that so far we have discussed the underlying geometry of the relevant configuration space, but not yet how to introduce dynamics. Hopfield network dynamics is a good model for how brain networks evolve over time with excitatory/inhibitory synaptic connections. Recent work of Curto and collaborators has revealed beautiful topological structures in the dynamics of Hopfield networks, [8], [9].

In [16] we developed a version of the Hopfield network dynamics where the evolution equations, viewed as finite difference equations, take place in a category of summing functors, in the context of the Gamma space models described above. These induce the usual non-linear Hopfield dynamics when summing functors take values in a suitable category of weighted codes. This formulation has the advantage that all the associated structures, resources with constraints, concurrent/distributed computing architectures, informational measures and constraints, are all evolving dynamically according to a common dynamics and functorial relations between the respective categories. Phenomena such as excitatory-inhibitory balance can also be investigated in this context. One aspect of the dynamics that is interesting in this context is whether it involves changes in homotopy types, which can be especially relevant to the interpretation discussed in §3 below.

This leads to a related question, which is a mechanism in this categorical setting of dynamical information networks, for operations altering stored homotopy types, formulated in terms of a suitable "calculus" of homotopy types. This points to the important connection between homotopy types and logic (in the form of Church's type theory), which is provided by the developed in Voevodsky's "homotopy type theory", see for instance [19], a setting designed to replace set-theory based mathematical logic with a model of the logic theory of types based on simplicial sets.

3. Toward a model of qualia and experience

In this final section we describe some tentative steps in the direction of using the topological setting described above to model qualia and experience.

A homotopy type is a (simplicial) space viewed up to continuous deformations that preserve homotopy invariants. One can imagine a time series of successive representations of an external stimulus space that either remain within the same homotopy type or undergo sudden transitions that causes detectable jumps in some of the homotopy groups. This collection of discrete topological invariants given by the homotopy groups (or homology as a simpler invariant) as detectors triggered by sudden changes of the stimulus space that does not preserve the homotopy type of the representation. We propose that such "jumps" in the homotopy type are related to experience.

In "Theater of Consciousness" models like [1], it is observed how the mind carries out a large number of tasks that are only "intermittently conscious", where a number of processes continue to be carried out unconsciously with "limited access" to consciousness, with one process coming to the forefront of consciousness (like actors on the stage in the theater metaphor) in response to an unexpected change. In the setting we are proposing, changes that do not alter the stored homotopy type are like the processes running in the unconscious background in the theater of consciousness, while the jumps in the discrete homotopy invariants affect the foregrounding to the stage of consciousness. This suggests that one should not simply think of the stored homotopy type as the reconstruction of a stimulus space at a given time instant. It is more likely that a whole time sequence should be stored as a homotopy type, changes to which will trigger an associated conscious experience. Some filtering of the stored information is also involved, that can be modeled through a form of persistent topology. The Gamma spaces formalism can be adapted to this persistent topology setting, as shown in [16].

In this setting we correspondingly think of "qualia space" as a space of homotopy types. This can be compared, for example, with the proposal of a qualia space in the theory of integrated information [2], based on probabilities on the set of states of the system, with a conscious experience represented as a shape in this qualia space. In [2], a "quale" is described as the set of all probability measures on all subsystems and mappings between them corresponding to inclusions. Notice how this setting is very closely related to the notion of probabilistic Gamma spaces developed in [17].

One can fit within this framework the observed role of homology as a detector of experience. The neuroscience experiments of [20], for example, conducted measurement of persistent homology of activated brain networks, comparing resting-state functional brain activity after intravenous infusion of either a placebo or psilocybin. The results they obtained show a significant change in homology in the case of psilocybin, with many transient structures of low stability and a small number of more significant persistent structures that are not observed in the case of placebo. These persistent homology structure are a correlate of the psychedelic experience. This evidence supports the idea that changes in topological invariants are related to experience and different states of consciousness.

3.1. **Neural correlates of consciousness and the role of topology.** In recent work [24], Ann Sizemore, Danielle Bassett and collaborators conducted a detailed topological analysis of the human connectome and identified explicit generators for the most prominently persistent non-trivial elements of both the first and second persistent homology groups. The persistence is computed with respect to a filtering of the connectivity matrix. The resulting generators identify which networks of connections between different cortical and subcortical regions involve non-trivial homology. Moreover, their analysis shows that the subcortical regions act as cone vertices for many homology generators formed by cortical regions. Based on the fact that electrostimulation of the posterior, but not of the anterior cortex, elicit various forms of somatosensory experience, the posterior cortex was proposed as a possible region involved in a neural correlate of consciousness (see [6]). A question to investigate within the mathematical framework outlined in the previous section is whether nontrivial topological structures in the connectome can be linked to candidate substrates in the wiring of brain regions for possible neural correlate of consciousness.

This work is supported by the Foundational Questions Institute and Fetzer Franklin Fund FFF grant number FQXi-RFP-1804, SVCF grant 2018-190467 and FQXi-RFP-CPW-2014, SVCF grant 2020-224047.

References

[1] B.J. Baars, *In the Theater of Consciousness*, Oxford University Press, 2001.

[2] D. Balduzzi, G. Tononi, *Qualia: the geometry of integrated information*, PLoS Comput. Biol. 5 (2009), no. 8, e1000462, 24 pp.

[3] P. Baudot, D. Bennequin, *The homological nature of entropy*, Entropy 17 (2015) no. 5, 3253–3318.

[4] R. Chaudhuri, I. Fiete, *Bipartite expander Hopfield networks as self-decoding high-capacity error correcting codes*, 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

[5] B. Coecke, T. Fritz, R.W. Spekkens, *A mathematical theory of resources*, Information and Computation 250 (2016), 59–86. [arXiv:1409.5531]

[6] F.C. Crick, C. Koch, *Towards a neurobiological theory of consciousness*, Seminars in Neuroscience, 2 (1990) 263–275.

[7] C. Curto, *What can topology tell us about the neural code?* Bull. Amer. Math. Soc. (N.S.) 54 (2017), no. 1, 63–78.

[8] C. Curto, J. Geneson, K. Morrison, *Stable fixed points of combinatorial threshold-linear networks*, arXiv:1909.02947

[9] C. Curto, C. Langdon, K. Morrison, *Robust motifs of threshold-linear networks*, arXiv:1902.10270

[10] L. Fajstrup, E. Goubault, E. Haucourt, *Directed Algebraic Topology and Concurrency*, Springer, 2016.

[11] T. Fritz, *Resource convertibility and ordered commutative monoids*, Math. Str. Comp. Sci. 27 (2017) N. 6, 850–938 [arXiv:1504.03661]

[12] M. Herlihy, D. Kozlov, S. Rajsbaum, *Distributed Computing Through Combinatorial Topology*, Elsevier, 2014.

[13] M. Kapranov, V. Voevodsky, *Combinatorial-geometric aspects of polycategory theory: pasting schemes and higher Bruhat orders (list of results)*, Cab. Top. Géom. Diff. Catégoriques Vol.32 (1991) 11–27.

[14] Yu. I. Manin, *Error-correcting codes and neural networks*, Selecta Math. (N.S.) 24 (2018) N.1, 521–530.

[15] Yu.I. Manin, M. Marcolli *Homotopy Theoretic and Categorical Models of Neural Information Networks*, arXiv:2006.15136

[16] M. Marcolli, *Gamma spaces and Information*, J. Geom. Phys. 140 (2019), 26–55.

[17] P.A.M. Mediano, A.K. Seth, A.B. Barrett, *Measuring integrated information: comparison of candidate measures in theory and simulation*, Entropy, 21 (2019), 17 [30p]

[18] A. Pelayo, M. A. Warren, *Homotopy type theory and Voevodsky's univalent foundations*, Bull. Amer. Math. Soc. (N.S.) 51 (2014) N.4, 597–648.

[19] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P.J. Hellyer, F. Vaccarino, *Homological scaffolds of brain functional networks*, J R Soc Interface. 2014 Dec 6;11(101):20140873

[20] M.W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dlotko, R. Levi, K. Hess, H. Markram, *Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function*, Frontiers in Computational Neuroscience, Vol.11 (2017) art.48 [16 pages]

[21] T. Rouse, *Topological Analysis of Attention*, talk at the Fields Institute Focus Program workshop New Mathematical Methods for Neuroscience, Toronto, March 2020
`http://gfs.fields.utoronto.ca/talks/Topological-Analysis-Attention`

[22] G. Segal, *Categories and cohomology theories*, Topology, Vol.13 (1974) 293–312.

[23] A.E. Sizemore, C. Giusti, A. Kahn, J.M. Vettel, R.F. Betzel, D.S. Bassett, *Cliques and cavities in the human connectome*, J Comput Neurosci, 44 (2018) 115–145.

[24] R. Steiner, *The Algebra of Directed Complexes*, Applied Categorical Structures, Vol.1 (1993) 247–284.

[25] G. Tononi, *An information integration theory of consciousness*, BMC Neurosci 5 G (2004) 42

[26] J.P. Vigneaux, *Generalized information structures and their cohomology*, arXiv:1709.07807.

[27] J.P. Vigneaux, *Topology of statistical systems. A cohomological approach to information theory*, PhD Thesis, Institut de mathématiques de Jussieu, Université de Paris Diderot, 2019.

[28] G. Winskel, M. Nielsen, *Categories in concurrency*, in "Semantics and logics of computation (Cambridge, 1995)", pp. 299–354, Publ. Newton Inst., 14, Cambridge Univ. Press, 1997.

California Institute of Technology, USA
E-mail address: `matilde@caltech.edu`