HAL
open science

Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function

Michael Reimann, Max Nolte, Martina Scolamiero, Katharine Turner, Rodrigo Perin, Giuseppe Chindemi, Pawel Dlotko, Ran Levi, Kathryn Hess, Henry Markram

To cite this version:

Michael Reimann, Max Nolte, Martina Scolamiero, Katharine Turner, Rodrigo Perin, et al.. Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function. Frontiers in Computational Neuroscience, 2017, 11, pp.1-16. 10.3389/fncom.2017.00048 . hal-01706964

HAL Id: hal-01706964

https://hal.science/hal-01706964

Submitted on 12 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 12 June 2017
doi: 10.3389/fncom.2017.00048

Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function

Michael W. Reimann[1†], Max Nolte[1†], Martina Scolamiero[2], Katharine Turner[2], Rodrigo Perin[3], Giuseppe Chindemi[1], Paweł Dłotko[4‡], Ran Levi[5‡], Kathryn Hess[2‡] and Henry Markram[1,3*‡]*

[1] Blue Brain Project, École Polytechnique Fédérale de Lausanne, Geneva, Switzerland, [2] Laboratory for Topology and Neuroscience, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [3] Laboratory of Neural Microcircuitry, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [4] DataShape, INRIA Saclay, Palaiseau, France, [5] Institute of Mathematics, University of Aberdeen, Aberdeen, United Kingdom

The lack of a formal link between neural network structure and its emergent function has hampered our understanding of how the brain processes information. We have now come closer to describing such a link by taking the direction of synaptic transmission into account, constructing graphs of a network that reflect the direction of information flow, and analyzing these directed graphs using algebraic topology. Applying this approach to a local network of neurons in the neocortex revealed a remarkably intricate and previously unseen topology of synaptic connectivity. The synaptic network contains an abundance of cliques of neurons bound into cavities that guide the emergence of correlated activity. In response to stimuli, correlated activity binds synaptically connected neurons into functional cliques and cavities that evolve in a stereotypical sequence toward peak complexity. We propose that the brain processes stimuli by forming increasingly complex functional cliques and cavities.

Keywords: connectomics, topology, directed networks, structure-function, correlations, Betti numbers

OPEN ACCESS

Edited by:
Paul Miller,
Brandeis University, United States

Reviewed by:
Cees van Leeuwen,
KU Leuven, Belgium
Andreas Knoblauch,
Hochschule Albstadt-Sigmaringen, Germany

***Correspondence:**
Henry Markram
henry.markram@epfl.ch
Kathryn Hess
kathryn.hess@epfl.ch

[†] *These authors have contributed equally to this work.*
[‡] *Co-senior author.*

Received: *11 March 2017*
Accepted: *18 May 2017*
Published: *12 June 2017*

Citation:
Reimann MW, Nolte M, Scolamiero M, Turner K, Perin R, Chindemi G, Dłotko P, Levi R, Hess K and Markram H (2017) Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function. Front. Comput. Neurosci. 11:48. doi: 10.3389/fncom.2017.00048

1. INTRODUCTION

How the structure of a network determines its function is not well understood. For neural networks specifically, we lack a unifying mathematical framework to unambiguously describe the emergent behavior of the network in terms of its underlying structure (Bassett and Sporns, 2017). While graph theory has been used to analyze network topology with some success (Bullmore and Sporns, 2009), current methods are usually constrained to analyzing how local connectivity influences local activity (Pajevic and Plenz, 2012; Chambers and MacLean, 2016) or global network dynamics (Hu et al., 2014), or how global network properties like connectivity and balance of excitatory and inhibitory neurons influence network dynamics (Renart et al., 2010; Rosenbaum et al., 2017). One such global network property is small-worldness. While it has been shown that small-worldness optimizes information exchange (Latora and Marchiori, 2001), and that adaptive rewiring during chaotic activity leads to small world networks (Gong and Leeuwen, 2004), the degree of small-worldness cannot describe most local network properties, such as the different roles of individual neurons.

Algebraic topology (Munkres, 1984) offers the unique advantage of providing methods to describe quantitatively both local network properties and the global network properties that

emerge from local structure, thus unifying both levels. More recently, algebraic topology has been applied to functional networks between brain regions using fMRI (Petri et al., 2014) and between neurons using neural activity (Giusti et al., 2015), but the underlying synaptic connections (structural network) were unknown. Furthermore, all formal topological analyses have overlooked the direction of information flow, since they analyzed only undirected graphs.

We developed a mathematical framework to analyze both the structural and the functional topology of the network, integrating local and global descriptions, enabling us to establish a clear relationship between them. We represent a network as a directed graph, with neurons as the vertices and the synaptic connections directed from pre- to postsynaptic neurons as the edges, which can be analyzed using elementary tools from algebraic topology (Munkres, 1984). The structural graph contains all synaptic connections, while a functional graph is a sub-graph of the structural graph containing only those connections that are active within a specific time bin (i.e., in which a postsynaptic neuron fires within a short time of a presynaptic spike). The response to a stimulus can then be represented and studied as a time series of functional graphs.

Networks are often analyzed in terms of groups of nodes that are all-to-all connected, known as *cliques*. The number of neurons in a clique determines its size, or more formally, its *dimension*. In directed graphs it is natural to consider *directed cliques*, which are cliques containing a single source neuron and a single sink neuron and reflecting a specific *motif* of connectivity (Song et al., 2005; Perin et al., 2011), wherein the flow of information through a group of neurons has an unambiguous direction. The manner in which directed cliques bind together can be represented geometrically. When directed cliques bind appropriately by sharing neurons, and without forming a larger clique due to missing connections, they form *cavities* ("*holes*," "*voids*") in this geometric representation, with high-dimensional cavities forming when high-dimensional (large) cliques bind together. Directed cliques describe the flow of information in the network at the local level, while cavities provide a global measure of information flow in the whole network. Using these naturally arising structures, we established a direct relationship between the structural graph and the emergent flow of information in response to stimuli, as captured through time series of functional graphs.

We applied this framework to digital reconstructions of rat neocortical microcircuitry that closely resemble the biological tissue in terms of the numbers, types, and densities of neurons and their synaptic connectivity (a "microconnectome" model for a cortical microcircuit, **Figures 1A,B**; see Markram et al., 2015; Reimann et al., 2015). Simulations of the reconstructed microcircuitry reproduce multiple emergent electrical behaviors found experimentally in the neocortex (Markram et al., 2015). The microcircuit, formed by ∼8 million connections (edges) between ∼31,000 neurons (vertices), was reconstructed from experimental data, guided by biological principles of organization, and iteratively refined until validated against a battery of independent anatomical and physiological data obtained from experiments. Multiple instantiations of the reconstruction provide a statistical and biological range of microcircuits for analysis.

We found a remarkably high number and variety of high-dimensional directed cliques and cavities, which had not been seen before in neural networks, either biological or artificial, and in far greater numbers than those found in various null models of directed networks. Topological metrics reflecting the number of directed cliques and cavities not only distinguished the reconstructions from all null models, they also revealed subtle differences between reconstructions based on biological datasets from different animals, suggesting that individual variations in biological detail of neocortical microcircuits are reflected in the repertoire of directed cliques and cavities. When we simulated microcircuit activity in response to sensory stimuli, we observed that pairwise correlations in neuronal activity increased with the number and dimension of the directed cliques to which a pair of neurons belongs, indicating that the hierarchical structure of the network shapes a hierarchy of correlated activity. In fact, we found a hierarchy of correlated activity between neurons even within a single directed clique. During activity, many more high-dimensional directed cliques formed than would be expected from the number of active connections, further suggesting that correlated activity tends to bind neurons into high-dimensional active cliques.

Following a spatio-temporal stimulus to the network, we found that during correlated activity, active cliques form increasingly high-dimensional cavities (i.e., cavities formed by increasingly larger cliques). Moreover, we discovered that while different spatio-temporal stimuli applied to the same circuit and the same stimulus applied to different circuits produced different activity patterns, they all exhibited the same general evolution, where functional relationships among increasingly higher-dimensional cliques form and then disintegrate.

2. RESULTS

2.1. The Case for Directed Simplices

Networks of neurons connected by electrical synapses (gap junctions) can be represented as undirected graphs, where information can flow in both directions. Networks with chemical synapses, which impose a single direction of synaptic communication from the pre- to the postsynaptic neuron (**Figures 1B2,B3**), are more accurately represented as directed graphs. Sub-sampling networks of neurons experimentally has revealed small motifs of synaptic connectivity, but not large cliques of neurons (Song et al., 2005; Perin et al., 2011). Knowing the complete directed network of neurons, as we do in the case of the reconstructed microcircuit, enables us to detect all cliques, directed, and otherwise (**Figure 1**).

When the direction of connections is not taken into account, a great deal of information is lost. For example, in the undirected case, there is only one possible configuration for a clique of four fully connected neurons (**Figure 2A1**, left). However, in the directed case, there are $3^6 = 729$ possible configurations, as each of the six connections can be in one of three states ($i \rightarrow j$, $j \leftarrow i$, or $i \leftrightarrow j$ connection types; **Figure 2A1** right).

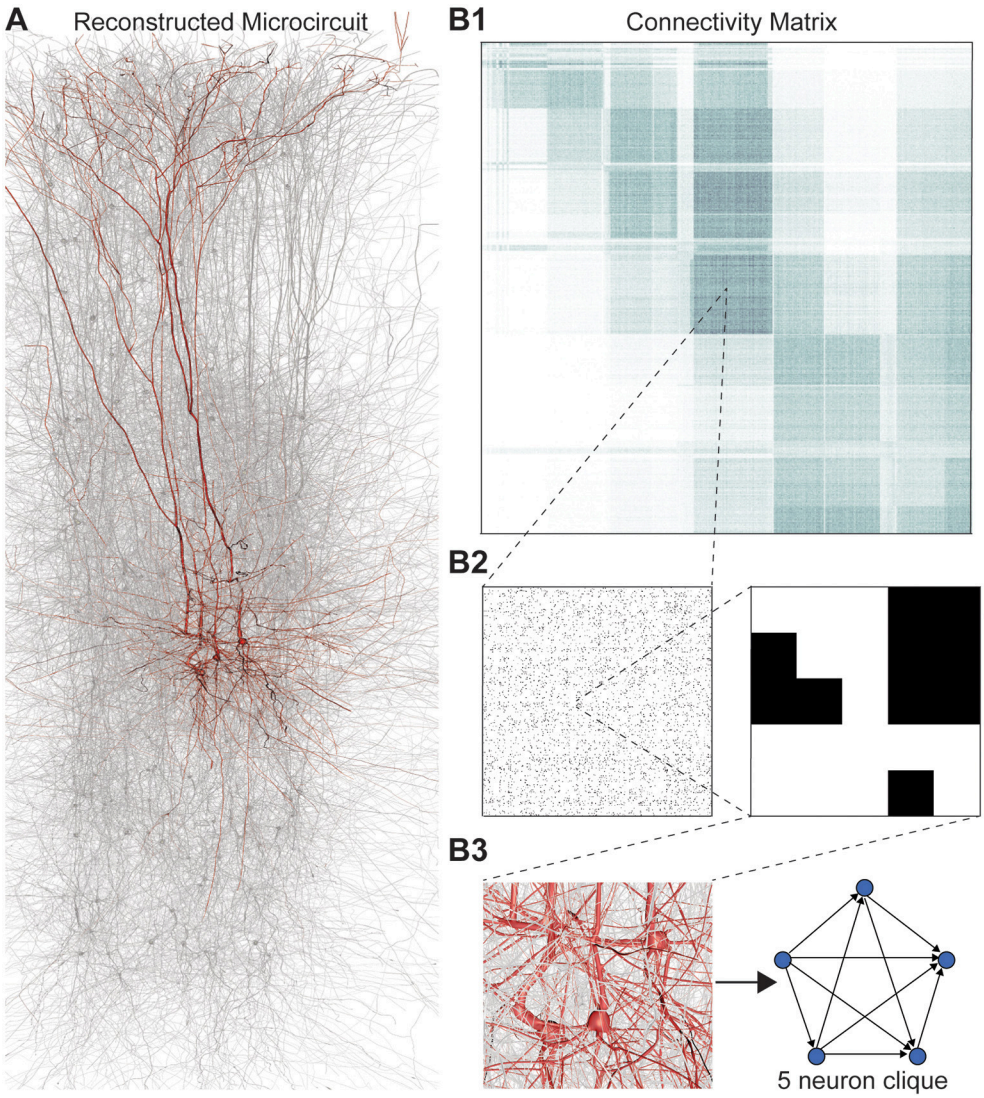

FIGURE 1 | (A) Thin (10 μm) slice of *in silico* reconstructed tissue. Red: A clique formed by five pyramidal cells in layer 5. **(B1)** Full connection matrix of a reconstructed microcircuit with 31,146 neurons. Neurons are sorted by cortical layer and morphological type within each layer. Pre-/postsynaptic neurons along the vertical/horizontal axis. Each grayscale pixel indicates the connections between two groups of 62 neurons each, ranging from white (no connections) to black ($\geq$8% connected pairs). **(B2)** Zoom into the connectivity between two groups of 434 neurons each in layer 5, i.e., 7 by 7 pixels in **(A)**, followed by a further zoom into the clique of 5 neurons shown in **(A)**. Black indicates presence, and white absence of a connection. **(B3)** Zoom into the somata of the clique in **(A)** and representation of their connectivity as a directed graph.

A clique with reciprocal connections contains two or more cliques consisting only of uni-directional connections (**Figure 2A2**). When only uni-directional connections are considered, there are 2^6 possible configurations of four fully connected neurons, which are of two types: those that contain cycles (40 configurations; **Figure 2A3** left; Section 4.1.3) and those that do not (24 configurations; **Figure 2A3** right). Directed cliques are exactly the acyclic cliques. The net *directionality* of information flow through any motif can be defined as the sum over all neurons of the squares of the differences between their in-degree and their out-degree (see Equation 2, Figure S1). Directed cliques have the highest net directionality among all cliques (Figure S1; Section 4.1.4). A clique that contains cycles always decomposes into directed cliques with the same number of neurons or fewer, at the very least any single connection between two neurons forms a 2-clique. A cyclical clique of three neurons therefore decomposes into three 2-cliques. Following the conventions in algebraic topology, we refer to directed cliques of *n* neurons as *directed simplices of dimension n-1* or *directed (n-1)-simplices* (which reflects their natural geometric representation as (n-1)-dimensional polyhedra) (see Figure S2; Section 4.1.3). Correspondingly, their sub-cliques are called *sub-simplices*.

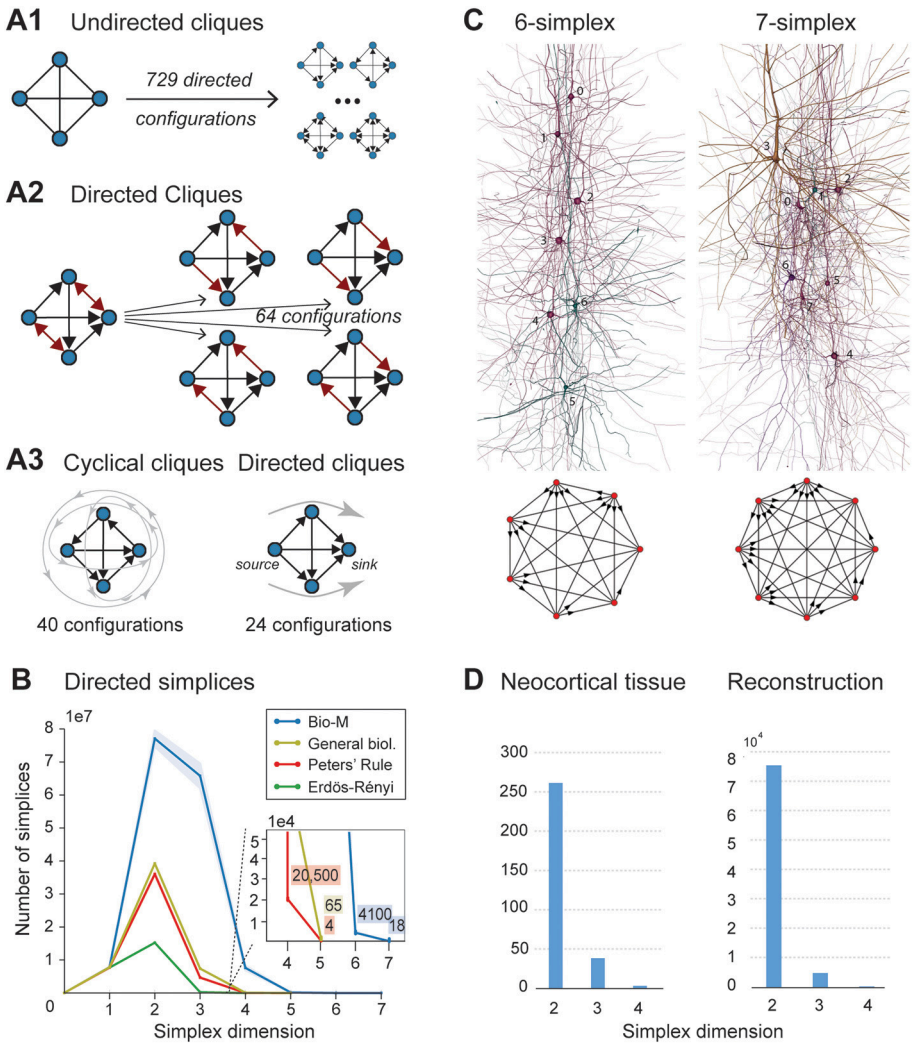

FIGURE 2 | (A1) A 4-clique in the undirected connectivity graph has one of 729 *configurations* in the directed graph. **(A2)** Configurations containing bidirectional connections are resolved by considering all sub-graphs without bidirectional connections. **(A3)** Without bidirectional connections, 64 possible configurations remain, 24 of which are acyclic, with a clear sink-source structure (*directed simplices*, in this case of dimension 3). **(B)** Number of simplices in each dimension in the Bio-M reconstruction (shaded area: standard deviation of seven statistical instantiations) and in three types of random control networks. **(C)** Examples of neurons forming high-dimensional simplices in the reconstruction. Bottom: Their representation as directed graphs. **(D)** (Left) Number of directed simplices of various dimensions found in 55 *in vitro* patch-clamp experiments sampling groups of pyramidal cells in layer 5. (Right) Number of simplices of various dimensions found in 100,000 *in silico* experiments mimicking the patch-clamp procedure of **(B)**.

2.2. An Abundance of Directed Simplices

2.2.1. Reconstructed Neocortical Microcircuitry

We analyzed 42 variants of the reconstructed microconnectome, grouped into six sets, each comprised of seven statistically varying instantiations (Markram et al., 2015; Section 4.3). The first five sets were based on specific heights of the six layers of the neocortex, cell densities, and distributions of different cell types experimentally measured in five different rats (Bio1-5), while the sixth represents the mean of these measurements (Bio-M). Individual instantiations within a set varied with the outcome of the stochastic portions of the reconstruction process. Surprisingly, we found that the reconstructions consistently contained directed simplices of dimensions up to 6 or 7, with as many as 80 million directed 3-simplices (**Figure 2B**; blue). This is the first indication of the existence of such a vast number of high-dimensional directed simplices in neocortical microcircuitry, or in any neural network.

2.2.2. Control Models

To compare these results with null models, we examined how the numbers of directed simplices in these reconstructions differed from those of artificial circuits and from circuits in which some of the biological rules of connectivity were omitted (see Section 4.4). For one control, we generated five Erdős-Rényi random graphs (ER) of equal size (∼31,000 vertices) and the same average connection probability as the Bio-M circuit (∼0.8%; ∼8 million

edges) (**Figure 2B**; dark green). For another, we constructed a circuit with the same 3D model neurons as the Bio-M circuit, but connected the neurons using a random connectivity rule ["Peters' Rule" (Peters and Feldman, 1976), PR; **Figure 2B**, red]. For the last control we connected the neurons in the Bio-M circuit according to the distance-dependent connection probabilities between the different morphological types of neurons. Since this control is similar to deriving connectivity from the average overlap of neuronal arbors (Shepherd et al., 2005), it retains the general biological (GB) features of connectivity between different types of neurons (Reimann et al., 2015), excluding only explicit pairwise connectivity between individual neurons, which is determined by the overlap of their *specific* arbors (**Figure 2B**, yellow). In all cases, the number of directed simplices of dimensions larger than 1 was far smaller than in the Bio-M circuit. In addition, the relative differences between the Bio-M and the null models increased markedly with dimension.

2.2.3. *In vitro*

Simplices of high dimensions (such as those depicted in **Figure 2C**) have not yet been observed experimentally, as doing so would require simultaneous intracellular recording of large numbers of neurons. To obtain an indication of the presence of many high-dimensional directed simplices in the actual neocortical tissue, we performed multi-neuron patch-clamp experiments with up to 12 neurons at a time in *in vitro* slices of the neocortex of the same age and brain region as the digitally reconstructed tissue (Section 4.5.1). Although limited by the number of neurons we could simultaneously record from, we found a substantial number of directed simplices up to dimension 3, and even one 4-dimensional simplex, in just 55 multi-neuron recording experiments (**Figure 2D**, left). We then mimicked these experiments on the reconstructed microcircuit by repeating the same multi-neuron patch-clamp recordings *in silico* (Section 4.5.2) and found a similar shape of the distribution of 4-, 3-, and 2-simplices, though in lower frequencies than in the actual tissue (**Figure 2D**, right). These findings not only confirm that high-dimensional directed simplices are prevalent in the neocortical tissue, they also suggest that the degree of organization in the neocortex is even greater than that in the reconstruction, which is already highly significant (see Section 3).

2.2.4. *C. elegans*

To test whether the presence of large numbers of high-dimensional directed simplices is a general phenomenon of neural networks rather than a specific phenomenon found in this part of the brain of this particular animal and at this particular age, we computed the numbers of directed simplices in the *C. elegans* connectome (Varshney et al., 2011) (Section 4.6). Again, we found many more high-dimensional simplices than expected from a random circuit with the same number of neurons (Figure S3).

2.2.5. Simplicial Architecture of Neocortical Microcircuitry

To understand the simplicial architecture of the microcircuit, we began by analyzing the sub-graphs formed only by excitatory neurons, only by inhibitory neurons, and only in individual layers by both excitatory and inhibitory neurons. Restricting to only excitatory neurons barely reduces the number of simplices in each dimension (**Figure 3A1**), while simplex counts in inhibitory sub-graphs are multiple orders of magnitude smaller (**Figure 3A2**), consistent with the fact that most neurons in the microcircuitry are excitatory. Analyzing the sub-graphs of the layers in isolation shows that layers 5 and 6, where most of the excitatory neurons reside (Markram et al., 2015), contain the most simplices and the largest number of high-dimensional simplices (**Figure 3A3**).

The large number of simplices relative to the number of neurons in the microcircuit implies that each neuron belongs to many directed simplices. Indeed, when we counted the number of simplices to which each neuron belongs across dimensions, we observed a long-tailed distribution such that a neuron belongs on average to thousands of simplices (**Figure 3B**). Both the mean maximal dimension and the number of simplices a neuron belongs to are highest in the deeper cortical layers (**Figure 3C**). Neurons in layer 5 belong to the largest number of simplices, many spanning multiple layers (**Figure 3D**), consistent with the abundance of neurons with the largest morphologies, which are connected to all layers. On the other hand, layer 6 has the largest number of simplices that are fully contained in the layer (**Figure 3A3**), consistent with the fact that layer 6 contains the most neurons. While the number of simplices that can form in the microcircuitry depends essentially on the number of neurons, the number of simplices to which a single neuron belongs depends fundamentally on its number of incoming and outgoing connections (its *degree*), which in turn depends on its morphological size (**Figure 3E**).

2.3. Topology Organizes Spike Correlations

The presence of vast numbers of directed cliques across a range of dimensions in the neocortex, far more than in null models, demonstrates that connectivity between these neurons is highly organized into fundamental building blocks of increasing complexity. Since the structural topology of the neural network takes into account the direction of information flow, we hypothesized that emergent electrical activity of the microcircuitry mirrors its hierarchical structural organization. To test this hypothesis, we simulated the electrical activity of the microcircuit under *in vivo*-like conditions (Markram et al., 2015).

Stimuli, configured as nine different spatio-temporal input patterns (**Figure 4A**), were injected into the reconstructed microcircuit through virtual thalamo-cortical fibers in which spike trains were induced using patterns recorded *in vivo* (Bale et al., 2015; Figure S4; Section 4.7). These stimuli differed primarily in the degree of synchronous input received by the neurons. As expected, the neurons in the microcircuit responded to the inputs with various spiking patterns (**Figures 4B1,B2,B4**). We then calculated for each connected pair of neurons the correlation of their spiking activity (**Figure 4C**) and found a broad distribution of correlation coefficients, with only ∼12% of

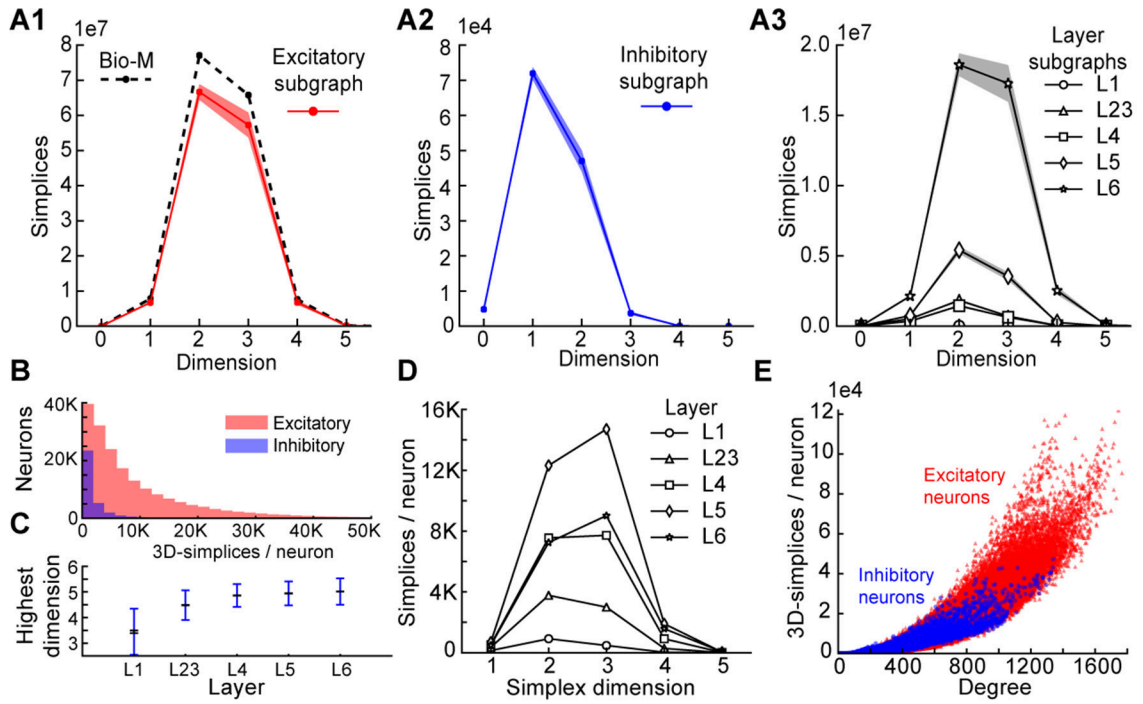

FIGURE 3 | (A1) Number of simplices in each dimension in the excitatory subgraph (shaded area: standard deviation across seven instantiations). **(A2)** Same, for the inhibitory subgraph. **(A3)** Same, for the subgraphs of individual layers. **(B)** Distribution across seven instantiations of the Bio-M graph of the number of 3- simplices an excitatory (red) or inhibitory (blue) neuron belongs to (simplices/neuron). **(C)** Mean over neurons in individual layers of the highest dimension of a simplex that they belong to. **(D)** Simplices/neuron by layer and dimension. **(E)** Correlation of 3-simplices/neuron and degree in the graph for all neurons.

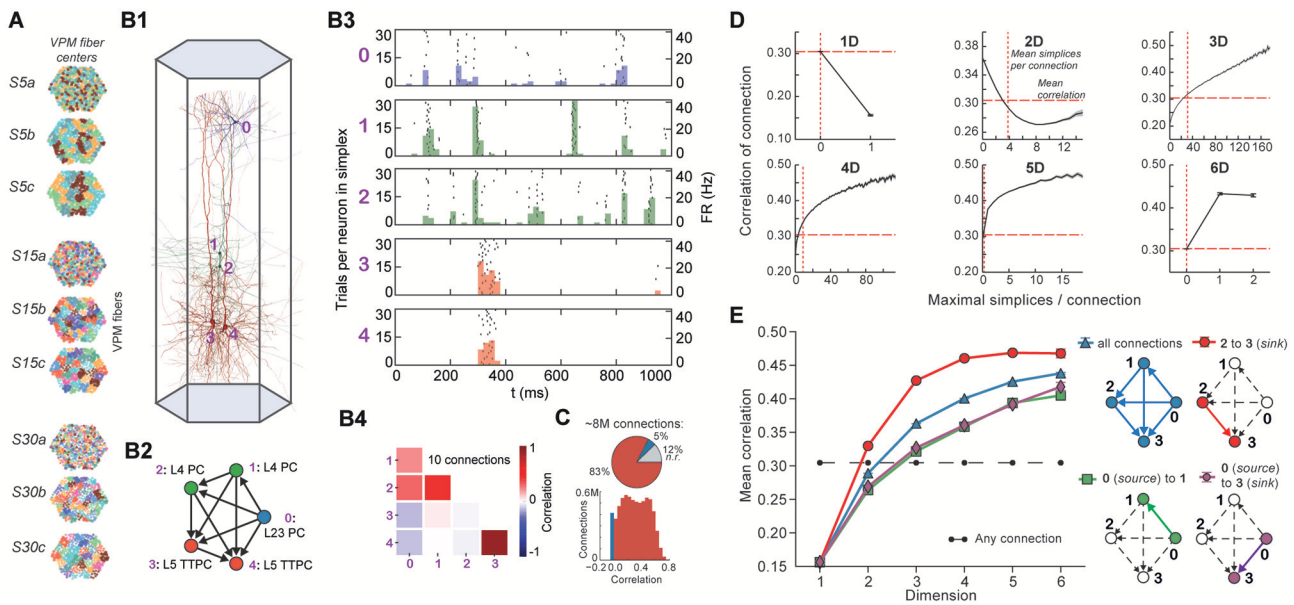

FIGURE 4 | (A) Patterns of thalamic innervation in the reconstruction. Each circle represents the center of innervation of a thalamic fiber. Each color represents a unique thalamic spike train assigned to that fiber. **(B1)** Exemplary directed simplex in a microcircuit. **(B2)** Connectivity and morphological types of neurons in the exemplary simplex. **(B3)** Raster plot and PSTH (Δt =10 ms) of spiking response of neurons in **(B1,B2)** to stimulus S30b. **(B4)** Correlation coefficients of all pairs of PSTHs in **(B3)**. **(C)** Correlation coefficients of PSTHs for all stimuli and all connected pairs of neurons in a microcircuit (Δt = 25 ms). **(D)** Mean correlation coefficients for connected pairs of neurons against the number of maximal simplices the edge between them belongs to, dimension by dimension. Means of fewer than 1,000 samples omitted. **(E)** Mean correlation coefficient of pairs of neurons, given their position within a simplex and its dimension.

connections where either the pre- or postsynaptic neuron failed to respond during all stimuli.

To avoid redundant sampling when testing the relationship between simplex dimension and activity, we restricted our analysis to *maximal simplices*, i.e., directed simplices that are not part of any higher-dimensional simplex (Section 4.1.2). A connection can be part of many higher-dimensional maximal simplices, unless it is itself a maximal 1-simplex. Despite the restriction to maximal simplices, we retained all information about the structure of the microcircuit because the complete structure is fully determined by its list of maximal simplices (Section 4.1.2). Correlations were calculated from histograms of the average spiking response (peri-stimulus time histogram, PSTH; bin size, 25 ms) to five seconds of thalamo-cortical input over 30 repetitions of a given input pattern (**Figure 4B3**). We then calculated the normalized cross-covariance of the histograms for all connections (**Figure 4C**; Section 4.8) and compared it to the number of maximal simplices associated with each connection in each dimension (see **Figure 4D**).

The neurons forming maximal 1-simplices displayed a significantly lower spiking correlation than the mean (**Figure 4D**), an indication of the fragility and lack of integration of the connection into the network. The mean correlation initially decreased with the number of maximal 2-simplices a connection belongs to, and then increased slightly. We observed that the greater the number of maximal 2-simplices a connection belongs to, the less likely it is to belong to higher-dimensional maximal simplices, with the minimum correlation occurring when the connection belongs to no simplices of dimension higher than 3. In higher dimensions, the correlation increased with the number of maximal simplices to which a connection belongs. While very high mean correlation can be attained for connections belonging to many maximal 3- or 4-simplices, the mean correlation of connections belonging to just one maximal 5- or 6-simplex was already considerably greater than the mean. These findings reveal a strong relationship between the structure of the network and its emergent activity and specifically that spike correlations depend on the level of participation of connections in high-dimensional simplices.

To determine the full extent to which the topological structure could organize activity of neurons, we examined spike correlations between pairs of neurons within individual simplices. These correlations increased with simplex dimension (**Figure 4E**, blue), again demonstrating that the degree of organization in the activity increases with structural organization. Spike correlation between pairs of neurons is normally an ambiguous measurement of connection strength because it is influenced by the local structure, specifically by indirect connections and/or shared inputs (Palm et al., 1988; Brody, 1999). However, since in our case the local structure is known and described in terms of directed simplices, we could infer how the local structural organization influences spike correlations. We compared the impact of indirect connections and of shared inputs on correlated activity by calculating the average correlation of pairs of neurons at different positions in a simplex when ordered from source to sink (**Figure 4E**, right panel). The number of indirect connections is highest for the pair consisting of the first (source) and last (sink) neurons (**Figure 4E**, purple), while the number of shared inputs is highest for the last and second-to-last neurons (**Figure 4E**, red). The first (source) and second neurons (**Figure 4E**, green) serve as a control because they have the smallest numbers of both indirect connections and shared inputs in the simplex.

We found that correlations were significantly higher for the last two neurons in the simplex, suggesting that shared input generates more of the pairwise correlation in spiking than indirect connections in directed simplices ($p < 8 \cdot 10^{-6}$, all dimensions except 1D). Moreover, the spiking correlation of the source and sink neurons was similar to the correlation of the first and second neurons (**Figure 4E**, purple and green), further suggesting that spike correlations tend to increase as shared input increases. These results hold for a range of histogram time bin sizes (Figure S5). The specific positions of neurons in local structures such as directed simplices therefore shape the emergence of correlated activity in response to stimuli.

2.4. Cliques of Neurons Bound into Cavities

Simplices are the mathematical building blocks of the microcircuitry. To gain insight into how its global structure shapes activity, it is necessary to consider how simplices are bound together. This can be achieved by analyzing the *directed flag complex*, which is the set of all directed simplices together with the set of all sub-simplices for each simplex (Figure S6, Section 4.1.2). The directed flag complex is a complete representation of the graph, including in particular the cycles neglected when examining directed simplices in isolation. The relationship between any two directed simplices depends on how they share sub-simplices. Just as any simplex can be realized as a polyhedron, a directed flag complex can be realized as a geometric object, built out of these polyhedra. If two simplices share a sub-simplex, the corresponding polyhedra are glued together along a common face (**Figure 5A**). The "shape" (or, more precisely, the topology) of this geometric object fully describes the global structure of the network.

To analyze directed flag complexes we computed two descriptors, the *Euler characteristic* and *Betti numbers* (Section 4.1.5). The Euler characteristic of a flag complex is given by the alternating sum of the number of simplices in each dimension, from zero through the highest dimension (**Figure 5A**). The Betti numbers together provide an indication of the number of cavities (or more precisely, *homology classes*) fully enclosed by directed simplices in the geometric object realizing the directed flag complex, where the dimension of a cavity is determined by the dimension of the enclosing simplices. The n-th Betti number, denoted β_n, indicates the number of n-dimensional cavities. For example, in **Figure 5A**, there is one 2-dimensional cavity (and therefore $\beta_2 = 1$) enclosed by the eight triangles; if an edge were added between any two non-connected nodes, then the geometric object realizing the corresponding flag complex would be filled in with solid tetrahedra, and the cavity would disappear. In the flag complexes of the reconstructions, it was not possible to compute more than the zeroth and top nonzero Betti numbers, as lower dimensions were computationally too expensive (Section 4.2.2). We could easily compute all Betti numbers for the *C. elegans*

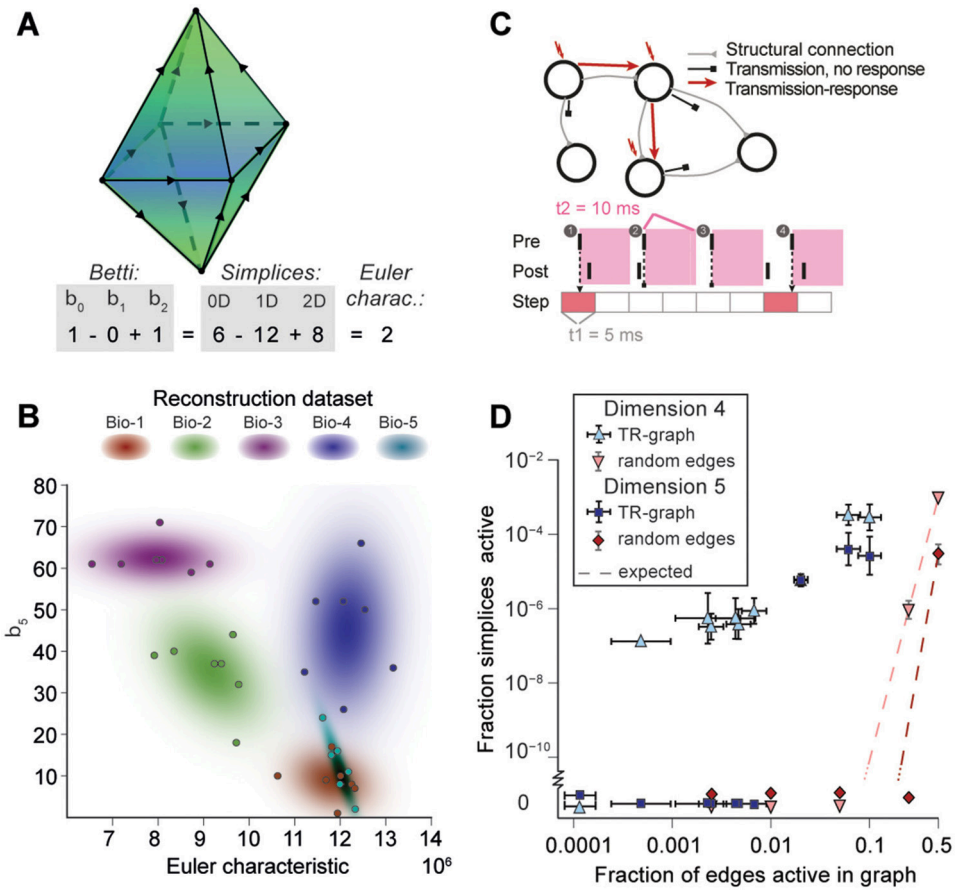

FIGURE 5 | (A) Example of the calculation of the Euler characteristic of a directed flag complex as an alternating sum of Betti numbers or simplex counts. **(B)** Euler characteristic against the highest non-zero Betti number (β_5) for seven instances of reconstructed microcircuits based on five different biological datasets (Bio 1-5). **(C)** Top: The transmission-response (TR) graph of the activity of a microcircuit is a subgraph of its structural connectivity containing all nodes, but only a subset of the edges (connections). Bottom: An edge is contained if its presynaptic neuron spikes in a defined time bin and its postsynaptic neurons spikes within 10 ms of the presynaptic spike. **(D)** Fraction of edges active against fraction of high-dimensional simplices active in TR graphs for various time bins of a simulation. Error bars indicate the standard deviation over 10 repetitions of the simulation. Blue triangles: 4-dimensional simplices, blue squares: 5-dimensional simplices. Red symbols and dashed lines indicate the results for choosing edges randomly from the structural graph and the number expected for random choice, respectively.

connectome, however, as it has many fewer nodes and edges (Figure S3).

The Betti number computations showed that there are cavities of dimension 5 (cavities completely enclosed by 5-simplices/6-neuron directed cliques) in all seven instances of each of the reconstructions (Bio1-Bio5, **Figure 5B**; Bio-M not shown). In contrast, the ER- and PR-control models have no cavities of dimension higher than 3, and the GB-model has no cavities of dimension higher than 4, demonstrating that there are not only non-random building blocks in the reconstruction, but also non-random relationships among them. We found as well that the information encoded in β_5 and the Euler characteristic together captures enough of the structure of the flag complex of a reconstruction to reveal subtle differences in their connectivity arising from the underlying biological data (**Figure 5B**, different colors).

2.5. Cliques and Cavities in Active Sub-Graphs

Thus far we have shown that the structural network guides the emergence of correlated activity. To determine whether this correlated activity is sufficiently organized to bind neurons together to form active cliques and to bind cliques together to form active cavities out of the structural graph, we represented the spiking activity during a simulation as a time series of sub-graphs for which we computed the corresponding directed flag complexes. Each sub-graph in this series comprises the same nodes (neurons) as the reconstruction, but only a subset of the edges (synaptic connections), which are considered *active*, i.e., the presynaptic neuron spikes in a time bin of size Δt_1 and the postsynaptic neuron spikes within a time Δt_2 after the presynaptic spike (**Figure 5C** and Figure S7, Section 4.9). By considering subsequent, non-overlapping time bins of constant size Δt_1, we obtain a time series of *transmission-response* (TR)

graphs reflecting correlated activity in the microcircuitry. We converted the time series of TR graphs in response to the different patterns of thalamo-cortical inputs (see **Figure 4A**) into time series of directed flag complexes. We found significantly more simplices in the TR graphs ($\Delta t_1 = 5$ ms, $\Delta t_2 = 10$ ms) than would be expected based on the number of edges alone (**Figure 5D**), indicating that correlated activity becomes preferentially concentrated in directed simplices.

The nine stimuli generated different spatio-temporal responses and different numbers of active edges (**Figure 6A**). The variation in Betti numbers and Euler characteristic over time indicates that neurons become bound into cliques and cavities by correlated activity (**Figure 6A** and Figure S8). When we plotted the number of cavities of dimension 1 (β_1) against the number of those of dimension 3 (β_3) (the highest dimension in which cavities consistently occur), the trajectory over the course of ∼100 ms (**Figure 6B**) began ∼50 ms after stimulus onset with the formation of a large number of 1-dimensional cavities, followed by the emergence of 2-dimensional (not shown) and 3-dimensional cavities. The

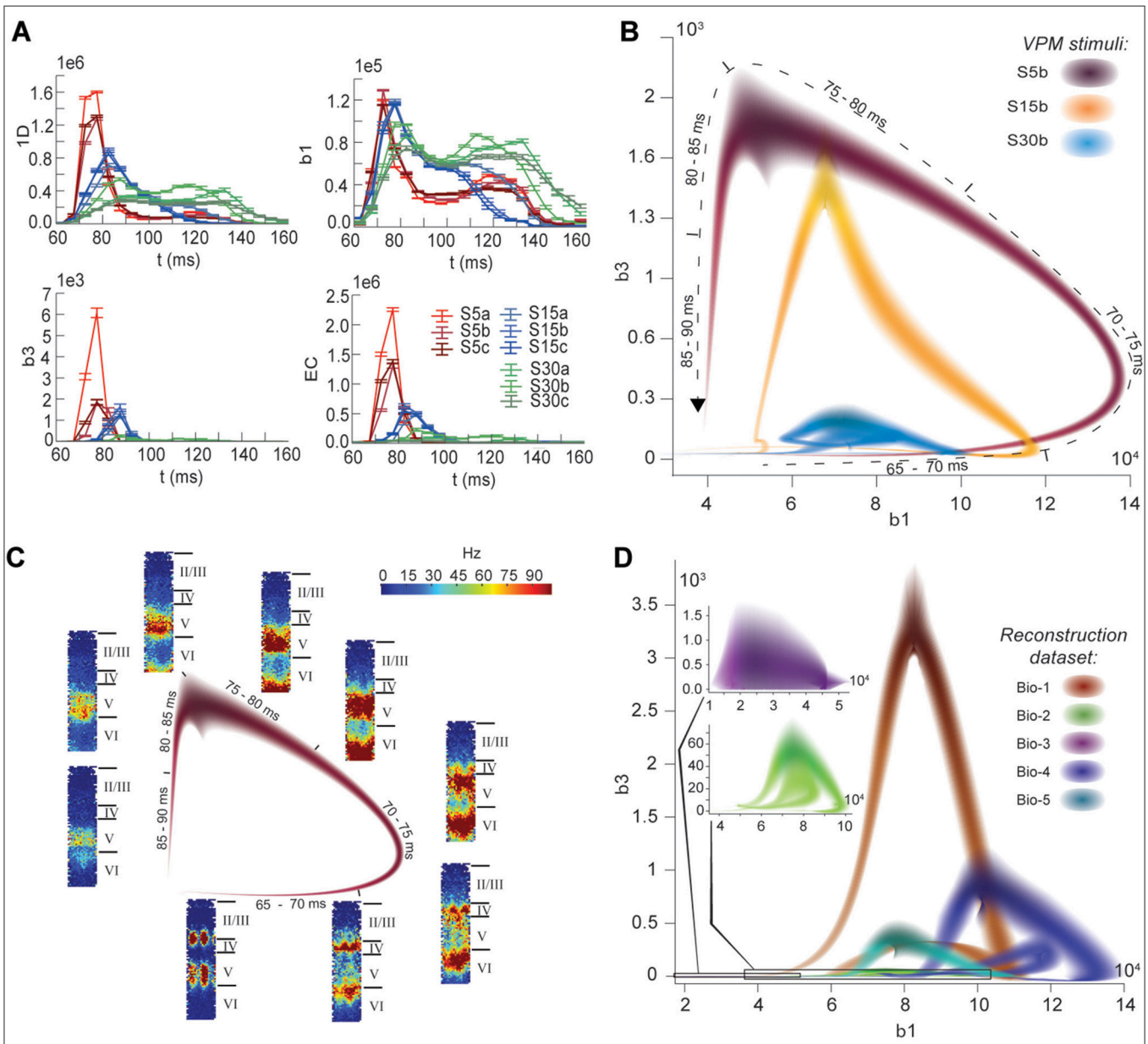

FIGURE 6 | (A) Number of edges, β_1, β_3, and Euler characteristic of the time series of TR graphs in response to the stimulus patterns shown in **Figure 4** (mean and SEM of 30 repetitions of each stimulus). **(B)** Trace of the time series of β_1 against β_3 for three of the stimuli. Shading of colors indicates Gaussian profiles at each time step with means and standard deviations interpolated from 30 repetitions of each stimulus. **(C)** Trace for one of the stimuli in B, along with the mean firing activity at different locations of the microcircuit during time steps of 2 ms. **(D)** Like **(B)**, but for TR graphs of Bio 1-5, in response to stimulus S15b.

decrease in β_1 began while β_3 was still increasing and continued until β_3 reached its peak, indicating that higher-dimensional relationships between directed simplices continued to be formed by correlated activity as the lower dimensional relationships subside.

Different stimuli led to Betti number trajectories of different amplitudes, where higher degrees of synchrony in the thalamic input produced higher amplitudes. The trajectories all followed a similar progression of cavity formation toward a peak level of functional organization followed by relatively rapid disintegration. The center of the projection of each trajectory onto the β_1-axis (its β_1-center) was approximately the same. Together, these characteristics of the trajectories reveal a stereotypical evolution of cliques and cavities in response to stimuli. These observations are consistent with experimentally recorded *in vivo* responses to sensory stimuli in terms of onset delay, response duration, and the presence of distinct phases of the response (Luczak et al., 2015).

To determine the neurons involved in this robust evolution of functional organization, we recorded the mean levels of spiking activity at different spatial locations within the microcircuit for one exemplary stimulus (**Figure 6C**). The activity started at depths that correspond to the locations of the thalamo-cortical input (Meyer et al., 2010; Markram et al., 2015), increasing in layer 4 and at the top of layer 6, before propagating downwards, reaching the top of layer 5 and the center of layer 6 as β_1 peaks, consistent with the finding that most directed simplices are in these layers. The transition from increasing β_1 to increasing β_3 coincided with the spread of the upper activity zone deeper into layer 5 and the top of layer 6, consistent with the presence of the highest dimensional directed simplices in these layers. The bottom activity zone also continued moving deeper, until it eventually subsided. As the top activity zone reached the bottom of layer 5, β_3 attained its peak. The zones of activity at the peaks of β_1 and β_3 are highly complementary: zones active at the peak of β_1 were generally inactive at the peak of β_3 and vice versa. The activity zone then remained in layer 5 until the cavities collapsed.

Finally, we applied the same stimulus to the reconstructions based on variations in the underlying biological data (see **Figure 5B**, Bio-1 to 5) and found similar Betti number trajectories, indicating that the general sequence of cavity formation toward peak functional organization followed by disintegration is preserved across individuals. On the other hand, we observed markedly different amplitudes, indicating that biological variability leads to variation in the number of high-dimensional cavities formed by correlated activity (**Figure 6D**). We also found that, unlike the case of different stimuli applied to the same microcircuit (**Figure 6B**), trajectories arising from different biological variations have different β_1-centers. In some cases, we observed reverberant trajectories that also followed a similar sequence of cavity formation, though smaller in amplitude. The general sequence of cavity formation and disintegration, however, appears to be stereotypic across stimuli and individuals.

3. DISCUSSION

This study provides a simple, powerful, parameter-free, and unambiguous mathematical framework for relating the activity of a neural network to its underlying structure, both locally (in terms of simplices) and globally (in terms of cavities formed by these simplices). Using this framework revealed an intricate topology of synaptic connectivity containing an abundance of cliques of neurons and of cavities binding the cliques together. The study also provides novel insight into how correlated activity emerges in the network and how the network responds to stimuli.

Such a vast number and variety of directed cliques and cavities had not been observed before in any neural network. The numbers of high-dimensional cliques and cavities found in the reconstruction are also far higher than in null models, even in those closely resembling the biology-based reconstructed microcircuit, but with some of the biological constraints released. We verified the existence of high-dimensional directed simplices in actual neocortical tissue. We further found similar structures in a nervous system as phylogenetically different as that of the worm *C. elegans* (Varshney et al., 2011), suggesting that the presence of high-dimensional topological structures is a general phenomenon across nervous systems.

We showed that the spike correlation of a pair of neurons strongly increases with the number and dimension of the cliques they belong to and that it even depends on their specific position in a directed clique. In particular, spike correlation increases with proximity of the pair of neurons to the sink of a directed clique, as the degree of shared input increases. These observations indicate that the emergence of correlated activity mirrors the topological complexity of the network. While previous studies have found a similar link for motifs built from 2-dimensional simplices (Pajevic and Plenz, 2012; Chambers and MacLean, 2016), we generalize this to higher dimensions. The fact that each neuron belongs to many directed cliques of various dimensions explains *in vivo* observations that neurons can "flexibly join multiple ensembles" (Miller et al., 2014). Braids of directed simplices connected along their appropriate faces could possibly act as synfire chains (Abeles, 1982), with a superposition of chains (Bienenstock, 1995) supported by the high number of cliques each neuron belongs to.

Topological metrics reflecting relationships among the cliques revealed biological differences in the connectivity of reconstructed microcircuits. The same topological metrics applied to time-series of transmission-response sub-graphs revealed a sequence of cavity formation and disintegration in response to stimuli, consistent across different stimuli and individual microcircuits. The size of the trajectory was determined by the degree of synchronous input and the biological parameters of the microcircuit, while its location depended mainly on the biological parameters. Neuronal activity is therefore organized not only within and by directed cliques, but also by highly structured relationships between directed cliques, consistent with a recent hypothesis concerning the relationship between structure and function (Luczak et al., 2015).

The higher degree of topological complexity of the reconstruction compared to any of the null models was found to depend on the morphological detail of neurons, suggesting that the local statistics of branching of the dendrites and axons is a crucial factor in forming directed cliques and cavities, though the exact mechanism by which this occurs remains to be determined (but see Stepanyants and Chklovskii, 2005). The number of directed 2-, 3-, and 4-simplices found per 12-patch *in vitro* recording was higher than in the digital reconstruction, suggesting that the level of structural organization we found is a conservative estimate of the actual complexity. Since the reconstructions are stochastic instantiations at a specific age of the neocortex, they do not take into account rewiring driven by plasticity during development and learning. Rewiring is readily triggered by stimuli as well as spontaneous activity (Le Be and Markram, 2006), which leads to a higher degree of organization (Chklovskii et al., 2004; Holtmaat and Svoboda, 2009) that is likely to increase the number of cliques. The difference may also partly be due to incomplete axonal reconstructions that would lead to lower connectivity, but such an effect would be minor because the connection rate between the specific neurons recorded for this comparison is reasonably well constrained (Reimann et al., 2015).

The digital reconstruction does not take into account intracortical connections beyond the microcircuit. The increase in correlations between neurons with the number of cliques to which they belong should be unaffected when these connections are taken into account because the overall correlation between neurons saturates already for a microcircuit of the size considered in this study, as we have previously shown (Markram et al., 2015). However, the time course of responses to stimuli and hence the specific shape of trajectories may be affected by the neighboring tissue.

In conclusion, this study suggests that neocortical microcircuits process information through a stereotypical progression of clique and cavity formation and disintegration, consistent with a recent hypothesis of common strategies for information processing across the neocortex (Harris and Shepherd, 2015). We conjecture that a stimulus may be processed by binding neurons into cliques of increasingly higher dimension, as a specific class of cell assemblies, possibly to represent *features* of the stimulus (Hebb, 1949; Braitenberg, 1978), and by binding these cliques into cavities of increasing complexity, possibly to represent the *associations* between the features (Willshaw et al., 1969; Engel and Singer, 2001; Knoblauch et al., 2009).

4. MATERIALS AND METHODS

4.1. The Topological Toolbox

Specializing basic concepts of algebraic topology, we have formulated precise definitions of cliques (*simplices*) and cavities (as counted by *Betti numbers*) associated to directed networks. What follows is a short introduction to directed graphs, simplicial complexes associated to directed graphs, and homology, as well as to the notion of directionality in directed graphs used in this study. We define, among others, the following terms and concepts.

Term	Description	Section
Directed graph	Network where each edge has a *source* and a *target*	4.1.1
Simplex	Clique of all-to-all connected nodes	4.1.2
Directed simplex	Simplex in a directed graph, with a *source* and a *sink*	4.1.3
Source (of simplex)	The node that is only a source of edges in a directed simplex	4.1.3
Sink (of simplex)	The node that is only a target of edges in a directed simplex	4.1.3
Face (of simplex)	Obtained by leaving out one or more nodes of a simplex	4.1.2
Simplicial complex	A collection of simplices "glued" together along common faces	4.1.2
Maximal simplex	Not a face of any larger simplex	4.1.2
Directionality	Formalized, intuitive measure of directionality in a graph	4.1.4
Betti numbers	Description of a graph in terms of the number of *cavities*	4.1.5.1
Euler characteristic	Alternating sum of number of simplices	4.1.5.2

4.1.1. Directed Graphs

A *directed graph* $\mathcal{G}$ consists of a pair of finite sets (V, E) and a function $\tau = (\tau_1, \tau_2)\colon E \to V \times V$. The elements of the set V are the *vertices* of $\mathcal{G}$, the elements of E are the *edges* of $\mathcal{G}$, and the function τ associates with each edge an ordered pair of vertices. The *direction* of an edge e with $\tau(e) = (v_1, v_2)$ is taken to be from $\tau_1(e) = v_1$, the *source vertex*, to $\tau_2(v) = v_2$, the *target vertex*. The function τ is required to satisfy the following two conditions.

1. There are no (self-) loops in the graph (i.e., for each $e \in E$, if $\tau(e) = (v_1, v_2)$, then $v_1 \neq v_2$).
2. For any pair of vertices (v_1, v_2), there is at most one edge directed from v_1 to v_2 (i.e., the function τ is injective).

Notice that a directed graph may contain pairs of vertices that are reciprocally connected, i.e., there may exist edges $e, e' \in E$ such that $\tau(e) = (v_1, v_2)$ and $\tau(e') = (v_2, v_1)$ (Figure S6A1ii).

A vertex $v \in \mathcal{G}$ is said to be a *sink* if there exists no $e \in E$ such that $v = \tau_1(e)$, but there is at least one edge $e' \in E$ such that $\tau_2(e') = v$. Similarly v is said to be a *source* is if there exists no $e \in E$ such that $v = \tau_2(e)$, but there is at least one $e' \in E$ such that $\tau_1(e') = v$ (Figures S6A1i,iii). A *path* in a directed graph $\mathcal{G}$ consists of a sequence of edges $(e_1, ..., e_n)$ such that for all $1 \leq k < n$, the target of e_k is the source of e_{k+1}, i.e., $\tau_2(e_k) = \tau_1(e_{k+1})$ (Figure S6A1iii). The *length* of the path $(e_1, ..., e_n)$ is n. If, in addition, the target of e_n is the source of e_1, i.e., $\tau_2(e_n) = \tau_1(e_1)$, then $(e_1, ..., e_n)$ is an *oriented cycle*. A graph that contains no oriented cycles is said to be *acyclic* (Figure S6A1i).

A directed graph is said to be *fully connected* if for every pair of distinct vertices, there exists an edge from one to the other, in at least one direction.

4.1.2. Simplices, Simplicial Complexes, and Flag Complexes

An *abstract directed simplicial complex* is a collection $\mathcal{S}$ of finite, *ordered* sets with the property that if $\sigma \in \mathcal{S}$, then every subset τ of σ, with the natural ordering inherited from σ, is also a member of $\mathcal{S}$. A *subcomplex* of an abstract directed simplicial complex is a sub-collection $\mathcal{S}' \subseteq \mathcal{S}$ that is itself an abstract directed simplicial complex. Abstract directed simplicial complexes are a variation on the more common ordinary *abstract simplicial complexes*, where the sets forming the collection $\mathcal{S}$ are not assumed to be ordered. To be able to study directed graphs, we use this slightly more subtle concept. Henceforth, we always refer to abstract directed simplicial complexes as *simplicial complexes*.

The elements σ of a simplicial complex $\mathcal{S}$ are called its *simplices*. We define the *dimension* of σ (denoted $\dim(\sigma)$) to be the cardinality of the set σ minus one. If σ is a simplex of dimension n, then we refer to σ as an *n-simplex* of $\mathcal{S}$. The set of all n-simplices of $\mathcal{S}$ is denoted $\mathcal{S}_n$. A simplex τ is said to be a *face* of σ if τ is a subset of σ of a strictly smaller cardinality. The *front face* of an n-simplex $\sigma = (v_0, ..., v_n)$ is a face $\tau = (v_0, ..., v_m)$ for some $m < n$. Similarly, the *back face* of σ is a face $\tau' = (v_i, \ldots, v_n)$ for some $0 < i < n$. If $\sigma = (v_0, \ldots, v_n) \in \mathcal{S}_n$ then, for each $0 \leq i \leq n$, the ith *face* of σ is the $(n-1)$-simplex σ^i obtained from σ by removing the vertex v_{n-i}. A simplex that is not a face of any other simplex is said to be *maximal*. The set of all maximal simplices of a simplicial complex determines the entire simplicial complex, since every simplex is either maximal itself or a face of a maximal simplex.

A simplicial complex gives rise to a topological space by *geometric realization*. A 0-simplex is realized by a single point, a 1-simplex by a line segment, a 2-simplex by a (filled in) triangle, and so on for higher dimensions. (see Munkres, 1984, Section 1). To form the geometric realization of the simplicial complex, one then glues the geometrically realized simplices together along common faces. The intersection of two simplices in $\mathcal{S}$, neither of which is a face of the other, is a proper subset, and hence a face, of both of them. In the geometric realization this means that the geometric simplices that realize the abstract simplices intersect on common faces, and hence give rise to a well-defined geometric object.

If $\mathcal{S}$ is a simplicial complex, then the union $\mathcal{S}^{(n)} = \mathcal{S}_n \cup \cdots \cup \mathcal{S}_0$, called the *n-skeleton* of $\mathcal{S}$, is a subcomplex of $\mathcal{S}$. We say that $\mathcal{S}$ is *n-dimensional* if $\mathcal{S} = \mathcal{S}^{(n)}$, and n is minimal with this property. If $\mathcal{S}$ is n-dimensional, and $k \leq n$, then the collection $\mathcal{S}_k \cup \ldots \cup \mathcal{S}_n$ is not a subcomplex of $\mathcal{S}$ because it is not closed under taking subsets. However, if one adds to that collection all the faces of all simplices in $\mathcal{S}_k \cup \ldots \cup \mathcal{S}_n$, one obtains a subcomplex of $\mathcal{S}$ called the *k-coskeleton* of $\mathcal{S}$, which we will denote by $\mathcal{S}_{(k)}$. Coskeleta are important for computing homology (see Section 4.2.2).

4.1.3. Simplicial Complexes of Directed Graphs

Directed graphs give rise to directed simplicial complexes in a natural way. The directed simplicial complex associated to a directed graph $\mathcal{G}$ is called the *directed flag complex* of $\mathcal{G}$ (Figure S6A2). This concept is a variation on the more common construction of a *flag complex* associated with an undirected graph (Aharoni et al., 2005). If $\mathcal{G} = (V, E, \tau)$ is a directed graph, then the directed flag complex associated to $\mathcal{G}$ is the abstract directed simplicial complex $\mathcal{S} = \mathcal{S}(\mathcal{G})$, with $\mathcal{S}_0 = V$ and whose *directed n-simplices* $\mathcal{S}_n$ for $n \geq 1$ are $(n+1)$-tuples $(v_0, \ldots, v_n)$, of vertices such that for each $0 \leq i < j \leq n$, there is an edge in $\mathcal{G}$ directed from v_i to v_j. The vertex v_0 is called the *source* of the simplex $(v_0, \ldots, v_n)$, as there is an edge directed from v_0 to v_i for all $0 < i \leq n$. Conversely, the vertex v_n is called the *sink* of the simplex $(v_0, \ldots, v_n)$, as there is an edge directed from v_i to v_n for all $0 \leq i < n$.

Notice that because of the assumptions on τ, an n-simplex in $\mathcal{S}$ is characterized by the (ordered) sequence $(v_0, \ldots, v_n)$, but not by the underlying set of vertices. For instance (v_1, v_2, v_3) and (v_2, v_1, v_3) are distinct 2-simplices with the same set of vertices.

4.1.4. Directionality of Directed Graphs

We give a mathematical definition of the notion of directionality in directed graphs, and prove that directed simplices are fully connected directed graphs with maximal directionality. Let $\mathcal{G} = (V, E, \tau)$ be a directed graph. For each vertex $v \in \mathcal{G}$, define the *signed degree* of v to be

$$\mathrm{sd}(v) = \mathrm{Indeg}(v) - \mathrm{Outdeg}(v). \tag{1}$$

Note that for any finite graph $\mathcal{G}$, $\sum_{v \in \mathcal{G}} \mathrm{sd}(v) = 0$. We define the *directionality* of $\mathcal{G}$, denoted $\mathrm{Dr}(\mathcal{G})$, to be the sum over all vertices of the square of their signed degrees (Figure S1),

$$\mathrm{Dr}(\mathcal{G}) = \sum_{v \in V} \mathrm{sd}(v)^2. \tag{2}$$

Let $\mathcal{G}_n$ denote a directed n-simplex, i.e., a fully connected directed graph on $n+1$ vertices such that every complete subgraph has a unique source and a unique sink. Note that a directed n-simplex has no reciprocal connections. If $\mathcal{G}$ is any directed graph on $n+1$ vertices, then $\mathrm{Dr}(\mathcal{G}) \leq \mathrm{Dr}(\mathcal{G}_n)$. If additionally $\mathcal{G}$ is a fully connected directed graph without reciprocal connections, then equality holds if and only if $\mathcal{G}$ is isomorphic to $\mathcal{G}_n$ as a directed graph. A full proof of these statements is given in the Supplementary Methods.

4.1.5. Homology

Betti numbers and *Euler characteristic* are numerical quantities associated to simplicial complexes that arise from an important and very useful algebraic object one can associate with any simplicial complex, called *homology*. Homology serves to measure the "topological complexity" of simplicial complexes, leading us to refer to Betti numbers and Euler characteristic as *topological metrics*. In this study we use only *mod 2 simplicial homology*, computationally the simplest variant of homology, which is why it is very commonly used in applications (Bauer et al., 2017). What follows is an elementary description of homology and its basic properties.

4.1.5.1. Betti numbers

Let $\mathbb{F}_2$ denote the field of two elements. Let $\mathcal{S}$ be a simplicial complex. Define the *chain complex* $C_*(\mathcal{S}, \mathbb{F}_2)$ to be the sequence $\{C_n = C_n(\mathcal{S}, \mathbb{F}_2)\}_{n \geq 0}$, such that C_n is the $\mathbb{F}_2$-vector space whose basis elements are the n-simplices $\sigma \in \mathcal{S}_n$, for each $n \geq 0$. In

other words, the elements of C_n are formal sums of n-simplices in $\mathcal{S}$.

For each $n \geq 1$, there is a linear transformation called a *differential*

$$\partial_n \colon C_n \to C_{n-1} \tag{3}$$

specified by $\partial_n(\sigma) = \sigma^0 + \sigma^1 + \cdots + \sigma^n$ for every n-simplex σ, where σ^i is the i-th face of σ, as defined above. Having defined ∂_n on the basis, one then extends it linearly to the entire vector space C_n. The n-th Betti number $\beta_n(\mathcal{S})$ of a simplicial complex $\mathcal{S}$ is the $\mathbb{F}_2$-vector space dimension of its n-th mod 2 *homology group*, which is defined by

$$H_n(\mathcal{S}, \mathbb{F}_2) = \mathrm{Ker}(\partial_n)/\mathrm{Im}(\partial_{n+1}) \tag{4}$$

for $n \geq 1$ and

$$H_0(\mathcal{S}, \mathbb{F}_2) = C_0/\mathrm{Im}(\partial_1). \tag{5}$$

For all $n \geq 1$, there is an inclusion of vector subspaces $\mathrm{Im}(\partial_{n+1}) \subseteq \mathrm{Ker}(\partial_n) \subseteq C_n$, and thus the definition of homology makes sense.

Computing the Betti numbers of a simplicial complex is conceptually very easy. Let $|\mathcal{S}_n|$ denote the number of n-simplices in the simplicial complex $\mathcal{S}$. If one encodes the differential ∂_n as a $\left(|\mathcal{S}_{n-1}| \times |\mathcal{S}_n|\right)$-matrix D_n with entries in $\mathbb{F}_2$, then one can easily compute its *nullity*, $\mathrm{null}(\partial_n)$, and its *rank*, $\mathrm{rk}(\partial_n)$, which are the $\mathbb{F}_2$-dimensions of the null-space and the column space of D_n, respectively. The *Betti numbers* of $\mathcal{S}$ are then a sequence of natural numbers defined by

$$\begin{aligned} \beta_0(\mathcal{S}) &= \dim_{\mathbb{F}_2}(C_0) - \mathrm{rk}(\partial_1), \text{ and} \\ \beta_n(\mathcal{S}) &= \mathrm{null}(\partial_n) - \mathrm{rk}(\partial_{n+1}). \end{aligned} \tag{6}$$

Since $\mathrm{Im}(\partial_{n+1}) \subseteq \mathrm{Ker}(\partial_n)$ for all $n \geq 1$, the Betti numbers are always non-negative. The n-th Betti number β_n gives an indication of the number of "n-dimensional cavities" in the geometric realization of $\mathcal{S}$.

4.1.5.2. Euler characteristic

If $\mathcal{S}$ is a simplicial complex, and $|\mathcal{S}_n|$ denotes the cardinality of the set of n-simplices in $\mathcal{S}$, then the Euler characteristic of $\mathcal{S}$ is defined to be

$$\chi(\mathcal{S}) = \sum_{n \geq 0} (-1)^n |\mathcal{S}_n|. \tag{7}$$

There is a well-known, close relationship between Euler characterstic and Betti numbers (Munkres, 1984, Theorem 22.2), which is expressed as follows. If $\left\{\beta_n(\mathcal{S})\right\}_{n \geq 0}$ is the sequence of Betti numbers for $\mathcal{S}$, then

$$\chi(\mathcal{S}) = \sum_{n \geq 0} (-1)^n \beta_n(\mathcal{S}). \tag{8}$$

4.2. Computation of Simplices and Homology

4.2.1. Generating Directed Flag Complexes with Hasse Diagrams

To obtain the simplices, Betti numbers and Euler characteristic of a directed graph, we first generate the directed flag complex associated to the graph. Our algorithm encodes a directed graph and its flag complex as a *Hasse diagram*. The Hasse diagram then gives immediate access to all simplices and simplex counts. The algorithm to generate the Hasse diagrams is fully described in the Supplementary Methods Section 2.2, and the C++ implementation of the code is publicly available at http://neurotop.gforge.inria.fr/.

4.2.2. Homology Computations

Betti numbers and Euler characteristic are computed from the directed flag complexes. All homology computations carried out for this paper were made with $\mathbb{F}_2$ coefficients, using the boundary matrix reduced by an algorithm from the PHAT library (Bauer et al., 2017).

The complexity of computing the n-th Betti numbers scales with the number of simplices in dimensions $n-1$, n, and $n+1$. In particular, it requires the computation of rank and nullity of matrices with shapes $(n-1) \times n$ and $n \times (n+1)$. Due to the millions of simplices in dimensions 2 and 3 in the reconstructed microcircuits (see Results), the calculation of Betti numbers above 0 or below 5 was computationally not viable, while the computation of the 5th Betti number was possible using the 5-coskeleton for each of the complexes. Nevertheless, our Euler characteristic computations imply that at least one of β_2 or β_4 must be nonzero, and it is highly likely the β_k is nonzero for all $k \leq 5$.

4.3. Model of Neocortical Microcircuitry

Analyses of connectivity and simulations of electrical activity are based on a previously published model of neocortical microcircuitry and related methods (Markram et al., 2015). We analyzed microcircuits that were reconstructed with layer height and cell density data from five different animals (Bio-1-5), with seven microcircuits per animal forming a mesocircuit (35 microcircuits in total). In addition, we analyzed microcircuits that were reconstructed using average data (Bio-M, seven microcircuits). Simulations were run on one microcircuit each of Bio-1-5 and Bio-M. Each microcircuit contains ∼31,000 neurons and ∼8 million connections. Data about the microcircuit and the neuron models used in the simulations, as well as the connection matrices, are available on https://bbp.epfl.ch/nmc-portal/ (Ramaswamy et al., 2015).

4.4. Control Networks

Additional control models of connectivity were constructed by removing different biological constraints on connectivity. We created three types of random matrices of sizes and connection probabilities identical to the connectivity matrices of the reconstructed microcircuits.

4.4.1. ER-Model (Random-Independent Graph)

An empty square connection matrix of the same size as the connection matrix of the reconstruction was instantiated and then randomly selected off-diagonal entries were activated. Specifically, entries were randomly selected with equal probabilities until the same number of entries as in the reconstruction were active. The directed graph corresponding to such a matrix is the directed analog of an Erdős-Rényi random graph (Erdos and Rényi, 1960).

4.4.2. PR-Model (Morphology-Only, "Peters' Rule")

A square connection matrix was generated based on the existence of spatial appositions between neurons in the reconstruction, i.e., instances where the axon of one neuron is within 1 μm of a dendrite of the other neuron. Appositions were then randomly removed from the matrix with equal probabilities until the same number of connections as in the reconstruction remained.

4.4.3. GB-Model (Shuffled, Preserving Distance Dependance)

The connection matrix of a reconstructed microcircuit was split into 55^2 submatrices based on the morphological types of pre- and postsynaptic neurons. Each submatrix was then randomized by shuffling its connections as follows. Connections in a submatrix were first grouped into bins according to the distance between the somata of their pre- and postsynaptic cells. Next, for each connection a new postsynaptic target was randomly selected from the same distance bin. We selected a distance bin size of 75 μm, which was the largest bin size that preserved the distribution of soma-distances of connected pairs of neurons in all sub-matrices (no statistically significant difference; $p > 0.05$, KL-test).

4.5. Patch Clamp Experiments

4.5.1. *In vitro*

Connectivity between layer 5 thick-tufted pyramidal cells was analyzed using multiple somatic whole-cell recordings (6–12 cells simultaneously) on 300 μm slices of primary somatosensory cortex of 14- to 16-day-old rats. Monosynaptic, direct excitatory connections were identified by stimulation of a presynaptic cell with a 20-70 Hz train of 5-15 strong and brief current pulses (1–2 nA, 2–4 ms). Experiments were carried out according to the Swiss national and institutional guidelines. Further details are explained in the Supplementary Methods.

4.5.2. *In silico*

In order to obtain *in silico* cell groups comparable to their patched *in vitro* counterparts, we designed a cell selection procedure approximating several of the experimental constraints of the *in vitro* patch-clamp setup used in this study and explained above. In brief, layer 5 thick-tufted pyramidal cells were selected from a volume with dimensions of $200 \times 200 \times 20\,\mu m$. The size of the volume was chosen to match the field of view usually available in the *in vitro* patch-clamp setup and to account for the tendency to patch nearby cells, which increases the probability of finding connected cells. The total number of cells was then reduced by randomly discarding a fraction of them, approximating the limited number of patching pipettes available *in vitro* (12) and the failure rate of the patching. This filtering step was optimized to match the *in silico* and *in vitro* cluster size distributions.

4.6. *C. elegans* Connectome

We analyzed part of the *C. elegans* connectome (Varshney et al., 2011), consisting of 6,393 directed chemical synapses, obtained from www.wormatlas.org/neuronalwiring.html.

4.7. Simulation of Electrical Activity

We performed simulations of neuronal electrical activity during stimulation with spatio-temporal patterns of thalamic input at the *in vivo*-like state (as in Markram et al., 2015), in the central microcircuit of Bio-M. Additionally, we repeated the same simulations in the central microcircuits of the Bio-1-5 reconstructions. We ran simulations using nine different organizations of thalamic input spike trains (see below).

4.7.1. Thalamic Stimulation

We used spike trains of 42 VPM neurons extracted from extracellular recordings of the response to texture-induced whisker motion in anesthetized rats, with up to nine cells in the same barreloid recorded simultaneously (Bale et al., 2015). Each reconstructed microcircuit is innervated by 310 virtual thalamo-cortical fibers (Markram et al., 2015). To generate sets of stimuli with different degrees of synchronous input, we assigned to each fiber one of 5 (SS5), 15 (SS15), or 30 (SS30) spike trains, recorded from distinct VPM neurons. In addition, we used k-means clustering to form clusters of fibers of size 1 (SSa), 5 (SSb), and 10 (SSc) (scikit-learn, sklearn.cluster.KMeans, Pedregosa et al., 2011) that were assigned the same spike train. This leads to different spatial arrangements of the identical thalamic inputs, and therefore to different degrees of synchronous input to individual neurons in the microcircuit.

4.8. Spike Train Correlations

We constructed postsynaptic time histograms (PSTHs) for each neuron for each stimulus, using the mean response to 30 trials of 5 s of thalamic stimulation (with bin size of 25 ms; for additional control, bin sizes of 10, 50, 100, 250, and 500 ms were also used). We then computed the normalized covariance matrix of the PSTHs of all neurons

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}, \tag{9}$$

where C_{ij} is the covariance of the PSTHs of neurons i and j. PSTHs of simulations with different thalamic stimuli were concatenated for each neuron to yield an average correlation coefficient for all stimuli. In total, correlations are based on the response of all neurons during 30 trials of nine stimuli for 5 s of activity (22.5 min).

4.9. Transmission-Response Matrices

The temporal sequence of transmission-response matrices associated to a simulation of neuronal activity of duration T is defined as

$$TR(\Delta t_1, \Delta t_2) := \{A(n) = A(n, \Delta t_1, \Delta t_2)\}_{n=1}^{N}, \tag{10}$$

where the n-th matrix, $A(n)$, is a binary matrix describing spiking activity in the time interval $[n \cdot \Delta t_1, (n+1) \cdot \Delta t_1 + \Delta t_2]$, and where $N = T/\Delta t_1$. The (j, k)-coefficient of $A(n)$ corresponding to the n-th time bin is 1 if and only if the following three conditions are satisfied, where s_i^j denotes the time of the i-th spike of neuron j.

(1) The (j, k)-coefficient of the structural matrix is 1, i.e., there is a structural connection from neuron j to neuron k, so that they form a pre-post synaptic pair.
(2) There is some i such that $n\Delta t_1 \leq s_i^j < (n+1)\Delta t_1$, i.e., neuron j spikes in the n-th time bin.
(3) There is some l such that $0 < s_l^k - s_i^j < \Delta t_2$, i.e., neuron k spikes after neuron j, within a Δt_2 interval.

In other words, a non-zero entry in a transmission-response matrix denotes a presynaptic spike, closely followed by a postsynaptic spike, maximizing the possibility of a causal relationship between the spikes. Based on firing data from spontaneous activity in the reconstructed microcircuit, we optimized Δt_i such that the resulting transmission-response matrices best reflect the actual sucessful transmission of signals between the neurons in the microcircuit (see Supplementary Methods). Unless noted otherwise, $\Delta t_1 = 5$ and $\Delta t_2 = 10$ ms were used throughout the study.

4.10. Data Analysis and Statistical Tests

Analysis of the model and simulations was performed on a Linux computing-cluster using Python 2.7, including the numpy and scipy libraries (Jones et al., 2001), and custom Python scripts. We calculated p-values using Welch's t-test (scipy.stats), unless noted otherwise.

AUTHOR CONTRIBUTIONS

HM and RL developed and initially conceived the study over 10 years of discussions. HM, RL, and KH conceived and directed the final study. KH and RL directed the applicability of concepts in algebraic topology to neuroscience. HM directed the relevance of algebraic topology in neuroscience. The Blue Brain Project team reconstructed the microcircuit and developed the capability to simulate the activity. MN performed the simulations. MN, MR, and PD generated the directed flag complexes from the connection matrices for analysis. KH and RL developed the theory for directed cliques and directed simplicial complexes. MR and RL developed the definition of directionality within motifs and directed cliques. MR developed the definition for transmission response matrices. PD developed the code to isolate simplices and directed simplices and performed initial computations. MS performed topological and statistical analyses on the flag complexes and on the *C. elegans* connectome. KT helped with initial statistical analysis of network responses to stimuli. MR and MN analyzed the simulation data, mapped it onto the topological data and generated the figures. RP performed the patch-clamp experiments. GC and MR performed the corresponding *in silico* experiments. HM, KH, RL, MR, and MN wrote the paper.

ACKNOWLEDGMENTS

This work was supported by funding from the ETH Domain for the Blue Brain Project and the Laboratory of Neural Microcircuitry. The Blue Brain Project's IBM BlueGene/Q system, BlueBrain IV, is funded by the ETH Board and hosted at the Swiss National Supercomputing Center (CSCS). MS was supported by the NCCR Synapsy grant of the Swiss National Science Foundation. Partial support for PD was provided by the GUDHI project, supported by an Advanced Investigator Grant of the European Research Council and hosted by INRIA. We thank Eilif Muller for providing input on the analysis, Magdalena Kedziorek for help with proving maximality in directed cliques, Gard Spreemann for help with the analysis of the *C. elegans* connectome, and Taylor H. Newton for helpful discussions about statistical methods.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fncom.2017.00048/full#supplementary-material

REFERENCES

Abeles, M. (1982). *Local Cortical Circuits: An Electrophysiological Study*. Berlin; Heidelberg; New York, NY: Springer.

Aharoni, R., Berger, E., and Meshulam, R. (2005). Eigenvalues and homology of flag complexes and vector representations of graphs. *Geom. Funct. Anal.* 15, 555–566. doi: 10.1007/s00039-005-0516-9

Bale, M. R., Ince, R. A. A., Santagata, G., and Petersen, R. S. (2015). Efficient population coding of naturalistic whisker motion in the ventro-posterior medial thalamus based on precise spike timing. *Front. Neural Circuits* 9:50. doi: 10.3389/fncir.2015.00050

Bassett, D. S., and Sporns, O. (2017). Network neuroscience. *Nat. Neurosci.* 20, 353–364. doi: 10.1038/nn.4502

Bauer, U., Kerber, M., Reininghaus, J., and Wagner, H. (2017). PHAT-Persistent Homology Algorithms Toolbox. *J. Symb. Comput.* 78, 76–90. doi: 10.1016/j.jsc.2016.03.008

Bienenstock, E. (1995). A model of neocortex. *Network Comput. Neural Syst.* 6, 179–224. doi: 10.1088/0954-898X 6 2 004

Braitenberg, V. (1978). "Cell assemblies in the cerebral cortex," in *Theoretical Approaches to Complex Systems*, eds R. Heim and G. Palm (Berlin; Heidelberg: Springer), 171–188. Available online at: http://www.springer.com/cn/book/9783540087571

Brody, C. D. (1999). Correlations without synchrony. *Neural Comput.* 11, 1537–1551. doi: 10.1162/089976699300016133

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575

Chambers, B., and MacLean, J. N. (2016). Higher-order synaptic interactions coordinate dynamics in recurrent networks. *PLOS Comput. Biol.* 12:e1005078. doi: 10.1371/journal.pcbi.1005078

Chklovskii, D. B., Mel, B. W., and Svoboda, K. (2004). Cortical rewiring and information storage. *Nature* 431, 782–788. doi: 10.1038/nature03012

Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25. doi: 10.1016/S1364-6613(00)01568-0

Erdős, P., and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* 5, 17.

Giusti, C., Pastalkova, E., Curto, C., and Itskov, V. (2015). Clique topology reveals intrinsic geometric structure in neural correlations. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13455–13460. doi: 10.1073/pnas.1506407112

Gong, P., and Leeuwen, C. v. (2004). Evolution to a small-world network with chaotic units. *EPL* 67:328. doi: 10.1209/epl/i2003-10287-7

Harris, K. D., and Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nat. Neurosci.* 18, 170–181. doi: 10.1038/nn.3917

Hebb, D. (1949). *The Organization of Behaviour*. New York, NY: Wiley & Sons.

Holtmaat, A., and Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* 10, 647–658. doi: 10.1038/nrn2699

Hu, Y., Trousdale, J., Josi, K., and Shea-Brown, E. (2014). Local paths to global coherence: cutting networks down to size. *Phys. Rev. E* 89:032802. doi: 10.1103/PhysRevE.89.032802

Jones, E., Oliphant, T., Peterson, P., and others (2001). *SciPy: Open Source Scientific Tools for Python*. Available online at: https://www.scipy.org/citing.html (Accessed November 11, 2016).

Knoblauch, A., Palm, G., and Sommer, F. T. (2009). Memory capacities for synaptic and structural plasticity. *Neural Comput.* 22, 289–341. doi: 10.1162/neco.2009.08-07-588

Latora, V., and Marchiori, M. (2001). Efficient behavior of small-world networks. *Phys. Rev. Lett.* 87:198701. doi: 10.1103/PhysRevLett.87.198701

Le Be, J.-V., and Markram, H. (2006). Spontaneous and evoked synaptic rewiring in the neonatal neocortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13214–13219. doi: 10.1073/pnas.0604691103

Luczak, A., McNaughton, B. L., and Harris, K. D. (2015). Packet-based communication in the cortex. *Nat. Rev. Neurosci.* 16, 745–755. doi: 10.1038/nrn4026

Markram, H., Muller, E., Ramaswamy, S., Reimann, M., Abdellah, M., Sanchez, C., et al. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell* 163, 456–492. doi:10.1016/j.cell.2015.09.029

Meyer, H. S., Wimmer, V. C., Hemberger, M., Bruno, R. M., de Kock, C. P., Frick, A., et al. (2010). Cell type-specific thalamic innervation in a column of rat vibrissal cortex. *Cereb. Cortex* 20, 2287–2303. doi: 10.1093/cercor/bhq069

Miller, J.-E. K., Ayzenshtat, I., Carrillo-Reid, L., and Yuste, R. (2014). Visual stimuli recruit intrinsically generated cortical ensembles. *Proc. Natl. Acad. Sci. U.S.A.* 111, E4053–E4061. doi: 10.1073/pnas.1406077111

Munkres, J. R. (1984). *Elements of Algebraic Topology*. Menlo Park, CA: Addison-Wesley Publishing Company.

Pajevic, S., and Plenz, D. (2012). The organization of strong links in complex networks. *Nat. Phys.* 8, 429–436. doi: 10.1038/nphys2257

Palm, G., Aertsen, A. M., and Gerstein, G. L. (1988). On the significance of correlations among neuronal spike trains. *Biol. Cybern.* 59, 1–11. doi: 10.1007/BF00336885

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. Available online at: http://www.jmlr.org/papers/v12/pedregosa11a.html

Perin, R., Berger, T. K., and Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5419–5424. doi:10.1073/pnas.1016051108

Peters, A., and Feldman, M. L. (1976). The projection of the lateral geniculate nucleus to area 17 of the rat cerebral cortex. I. General description. *J. Neurocytol.* 5, 63–84. doi: 10.1007/BF01176183

Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P. J., et al. (2014). Homological scaffolds of brain functional networks. *J. R. Soc. Interface* 11:20140873. doi: 10.1098/rsif.2014.0873

Ramaswamy, S., Courcol, J.-D., Abdellah, M., Adaszewski, S. R., Antille, N., Arsever, S., et al. (2015). The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Front. Neural Circuits* 9:44. doi: 10.3389/fncir.2015.00044

Reimann, M. W., King, J. G., Muller, E. B., Ramaswamy, S., and Markram, H. (2015). An algorithm to predict the connectome of neural microcircuits. *Front. Comput. Neurosci.* 9:120. doi: 10.3389/fncom.2015.00120

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., et al. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590. doi: 10.1126/science.1179850

Rosenbaum, R., Smith, M. A., Kohn, A., Rubin, J. E., and Doiron, B. (2017). The spatial structure of correlated neuronal variability. *Nat. Neurosci.* 20, 107–114. doi: 10.1038/nn.4433

Shepherd, G. M. G., Stepanyants, A., Bureau, I., Chklovskii, D., and Svoboda, K. (2005). Geometric and functional organization of cortical circuits. *Nat. Neurosci.* 8, 782–790. doi: 10.1038/nn1447

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly nonrandom features Of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3:e68. doi: 10.1371/journal.pbio.0030068

Stepanyants, A., and Chklovskii, D. (2005). Neurogeometry and potential synaptic connectivity. *Trends Neurosci.* 28, 387–394. doi: 10.1016/j.tins.2005.05.006

Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., and Chklovskii, D. B. (2011). Structural properties of the caenorhabditis elegans neuronal network. *PLoS Comput. Biol.* 7:e1001066. doi: 10.1371/journal.pcbi.1001066

Willshaw, D. J., Buneman, O. P., and Longuet-Higgins, H. C. (1969). Non-holographic associative memory. *Nature* 222, 960–962.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Reimann, Nolte, Scolamiero, Turner, Perin, Chindemi, Dłotko, Levi, Hess and Markram. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.