Topological analysis of neural systems

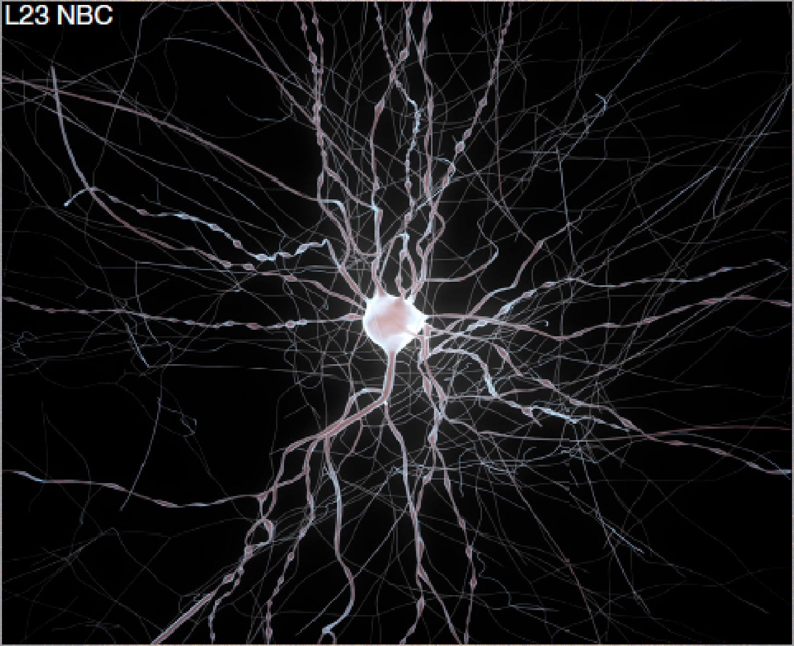
Ran Levi

University of Aberdeen

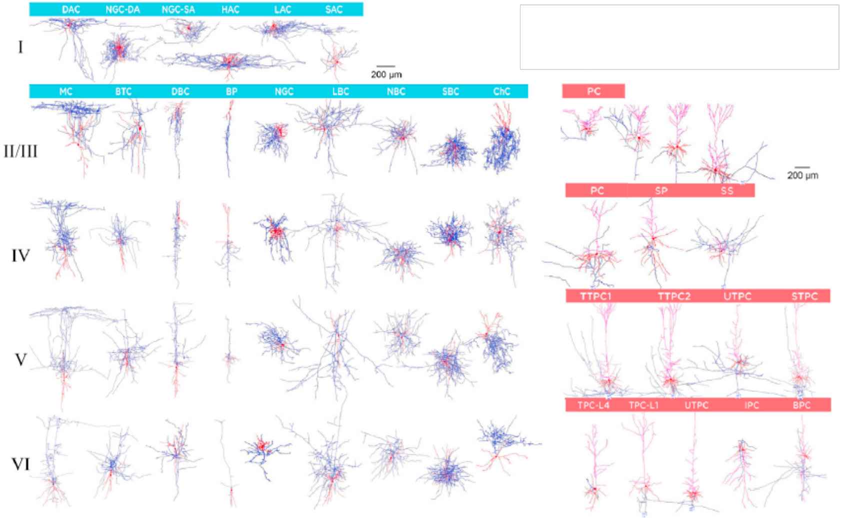
September 12, 2016

Joint project with

- Pawel Dłotko – INRIA
- Kathryn Hess, Martina Scolamiero and Katharine Turner – Laboratory for Topology and Neuroscience, EPFL
- Henry Markram, Eilif Muller, Max Nolte and Michael Reimann – Blue Brain Project, EPFL


L23 NBC

55 morphological types of neurons



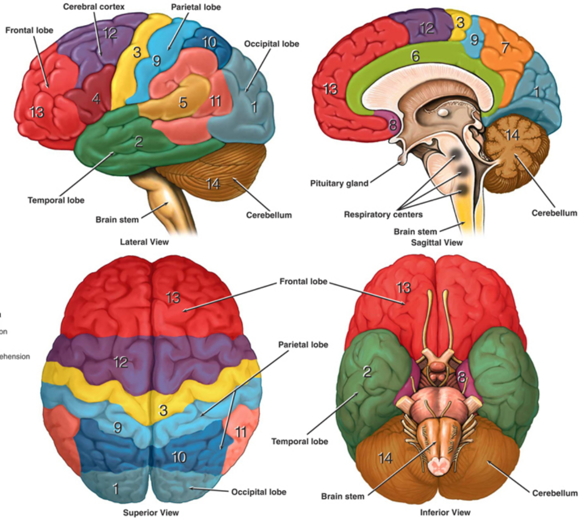
Anatomy and Functional Areas of the Brain

Functional Areas of the Cerebral Cortex

1. **Visual Area:**
   Sight
   Image recognition
   Image perception
2. **Association Area**
   Short-term memory
   Equilibrium
   Emotion
3. **Motor Function Area**
   Initiation of voluntary muscles
4. **Broca's Area**
   Muscles of speech
5. **Auditory Area**
   Hearing
6. **Emotional Area**
   Pain
   Hunger
   "Fight or flight" response
7. **Sensory Association Area**
8. **Olfactory Area**
   Smelling
9. **Sensory Area**
   Sensation from muscles and skin
10. **Somatosensory Association Area**
    Evaluation of weight, texture, temperature, etc. for object recognition
11. **Wernicke's Area**
    Written and spoken language comprehension
12. **Motor Function Area**
    Eye movement and orientation
13. **Higher Mental Functions**
    Concentration
    Planning
    Judgment
    Emotional expression
    Creativity
    Inhibition

Functional Areas of the Cerebellum

14. **Motor Functions**
    Coordination of movement
    Balance and equilibrium
    Posture



The Blue Brain Project

- An intricate and biologically accurate digital reconstruction of the neocortical column of a 14 days old rat.
- Each microcircuit consists of roughly $3 \cdot 10^4$ neurons of 55 distinct morphological types and approximately 8×10^6 connections between neurons.
- The reconstruction allows recovering an abundance of information, including
  - the full adjacency matrix,
  - the type and spatial position of each individual neuron, and
  - full spiking data for each neuron under varying conditions.
- This gives rise to graphs (directed, potentially weighted, and time dependent).
- Graphs give rise to topological objects.

Abstract ordered simplicial complexes and the directed flag complex of a directed graph

An abstract *ordered* simplicial complex is a collection $\mathcal{S}$ of finite *ordered* sets, such that

$$\sigma \in \mathcal{S} \quad \Longrightarrow \quad \tau \in \mathcal{S}, \qquad \forall \tau \subset \sigma.$$

The subsets $\sigma \in \mathcal{S}$ are called the simplices of $\mathcal{S}$.

- A *directed* graph $\mathcal{G}$ is a pair (V, E) where V is the set of vertices and $E \subseteq V \times V$ is the set of directed edges.
- The *directed* flag complex of a directed graph $\mathcal{G}$ is the abstract directed simplicial complex $S = S(\mathcal{G})$, whose n-simplices are ordered $(n+1)$-tuples of vertices

$$S_n = \{(v_0, v_1, \dots, v_n) \mid (v_i, v_j) \in E, \ \forall 0 \leq i < j \leq n\}$$

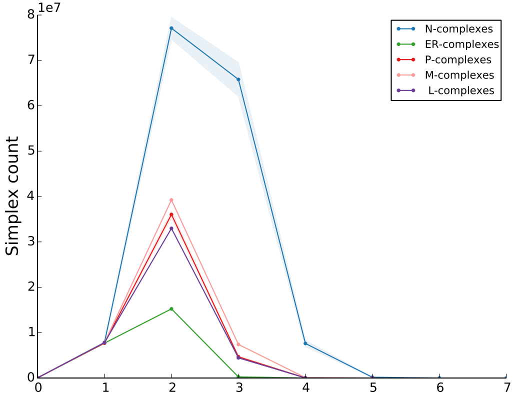
Analysis of the Blue Brain Project Reconstruction

Topological Analysis of Structure

At our disposal 42 reconstructed microcircuits based on the cortex of 5 individual rats. Adjacency matrices for each microcircuit - average size 31,000 with average connectivity of 0.8%. In addition we generated:

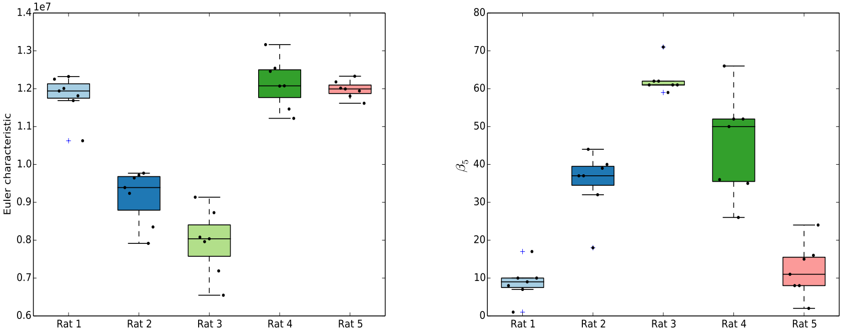
- Erdős-Rényi random connectivity matrices with the same size and average connectivity as the reconstruction.
- Two sets of controlled randomisations of an average microcircuit, preserving distance-dependent connection probability across **i)** all pairs of layers, and **ii)** all pair of morphological types, but otherwise random.
- A set of controlled randomisations of an average microcircuit according to Peter's rule: Connect two neurons if they contain arbors with distance at most $3\mu m$, and then prune uniformly until the required average connectivity is obtained.

Distribution of simplices



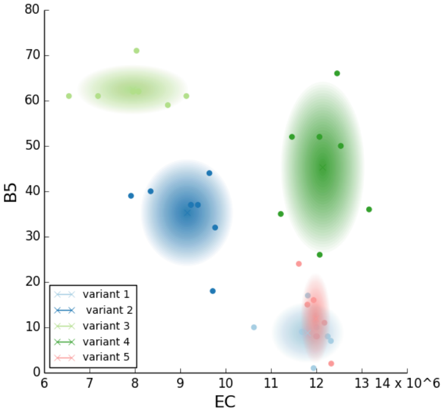
For each of the connectivity matrices we computed the directed flag complex. The complexes resulting from the Blue Brain reconstruction show dramatically different behaviour from the randomised matrices.

Euler characteristic and (mod 2) homology (mod 2)



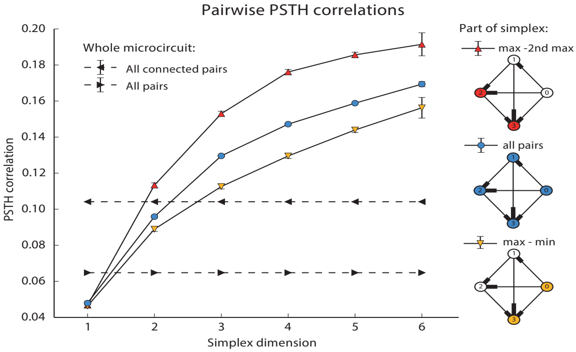
On the left, the Euler characteristic of the complex of each individual reconstruction, sorted by the animal that gave rise to the data. On the right the same with respect to betti 5.

Euler characteristic and (mod 2) homology



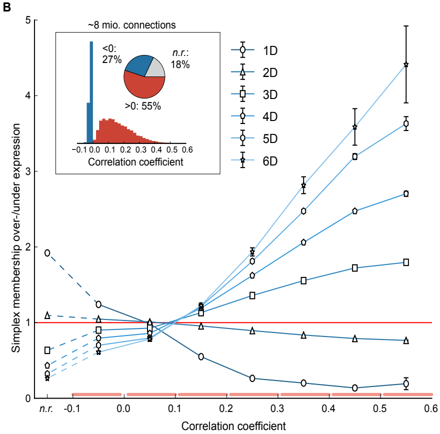
Rats 1 and 5 cannot be clearly distinguished, but the others form clearly distinguishable groups.

Dimension and directionality matter



Correlation is strongest between the last two neurons in a simplex, and is rising with dimension. This provides evidence to the importance of simplices, dimension and directionality.

Dimension and directionality matter

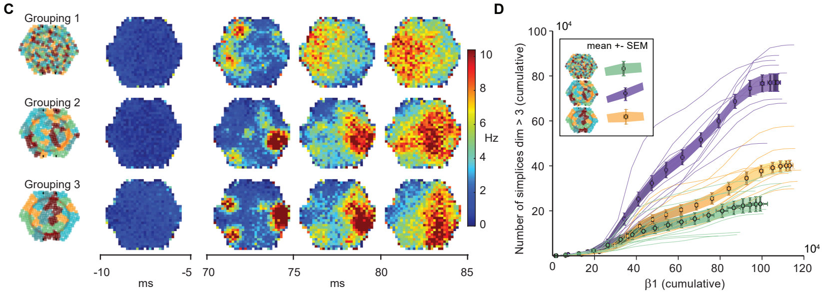


Correlation against normalised maximal simplex membership.

The Transmission-Response method

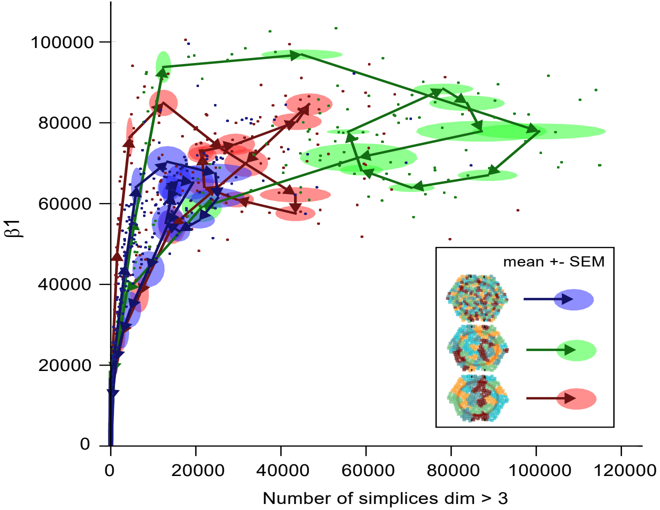
- The microcircuit is stimulated in time intervals of 50ms for a whole second. The reaction is recorded in time bins $t = 0 \dots 199$ of 5 ms each (size optimised by experimentation).
- Let A denote the structural connectivity matrix for the given microcircuit.
- In each time bin k consider the "successful transmission" connectivity matrix A^k where $A^k_{i,j} = 1$ if and only if the following three conditions are satisfied:
  - $A_{i,j} = 1$, i.e., there is a structural connection from neuron i to neuron j,
  - neuron i fired in time bin k, and
  - neuron j fired within 7.5ms after neuron i did (optimised by experimentation).

Patterns of activity

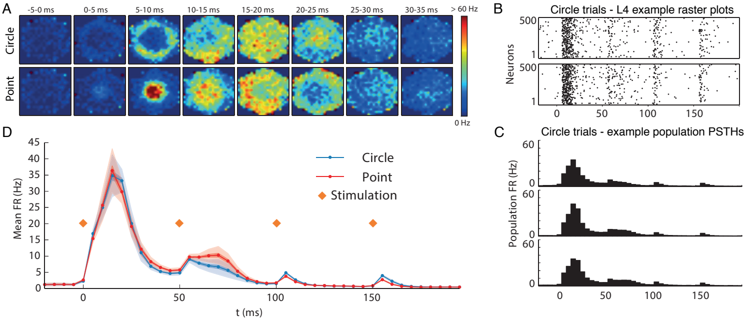


Trail of the number of high-dimensional (≥ 3) simplices in a transmission-response graph against Betti 1 for three thalamocortical input patterns.

Patterns of activity

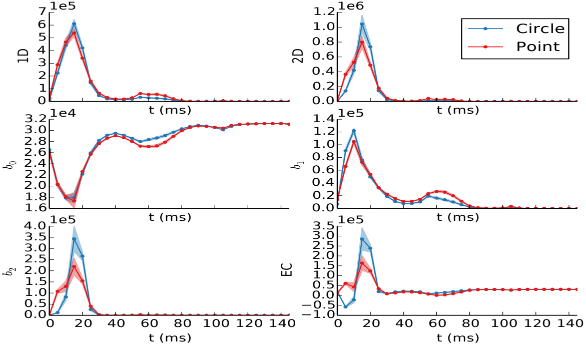


Point vs. Circle Experiment



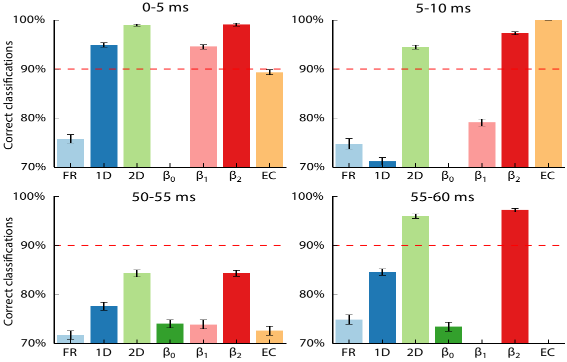
"Point vs. Circle" stimulation. Top left shows the evolution of the signal. The other panels show the reaction as measured by standard methods in neuroscience.

Time series of topological invariants



Average values of the various metrics for the dot and circle stimuli. Response after the first 100ms is much smaller than the initial response due to recovery time required.

Correct classification



In the indicated time bins topological invariants give the highest percentage of correct classifications performed by Gaussian Bayes classifier based on each of the metrics.

Segregation, Clustering, Integration, Small worldness

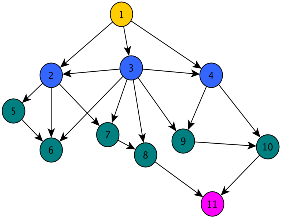
Based on a survey paper by Rubinov and Sporns: There are many graph theoretic invariants which proved useful in neuroscience. We restrict to a few such invariants.

- The degree k_i of a node i: the number of nodes connected to i.
- A basic measure of segregation at a node i: The number t_i of triangles with i as a vertex.
- The clustering coefficient of a node i: $C_i = \frac{2t_i}{k_i(k_i-1)}$ = the number of triangles divided by the number of possible triangles. The clustering coefficient of the network: $C = \frac{1}{n}\sum_i C_i$.
- Measure of integration: $L = \frac{1}{n}\sum_i L_i$, where L_i is the average path length from i to any other node.
- Small worldness: Higher than random segregation, close to random integration.

Topological metrics - k-valence

Let X be a directed abstract simplicial complex.

- For $v \in X_0$, let $m_k^{\text{out}}(v)$ denote the outgoing k-valence of v - the number of simplices $\sigma \in X_k$, such that v is an initial vertices in σ.
- Similarly, define $m_k^{\text{in}}(v)$ - the incoming k-valence of v.
- Define $m_k(v)$ - the k-valence of v to be the number of k-simplices in X which contain v.



$m_1^{\text{out}}(3) = 6,\ m_1^{\text{in}}(3) = 1, m_1(3) = 7,$
$m_2^{\text{out}}(3) = 4,\ m_2^{\text{in}}(3) = 0,\ m_2(3) = 6.$

Topological metrics - Degree polynomials

- Define the *local* degree polynomial of $v \in X_0$ by

$$M_v(t) = \sum_{k \geq 0} m_k(v) t^k.$$

- Similarly define incoming and outgoing *local* degree polynomials $M_v^{\text{in}}(t)$ and $M_v^{\text{out}}(t)$ resp.
- Define the *global* degree polynomial of X by

$$M_X(t) = \frac{1}{|X_0|} \sum_{v \in X_0} M_v(t).$$

- Similarly define incoming and outgoing *global* degree polynomials $M_X^{\text{in}}(t)$ and $M_X^{\text{out}}(t)$ resp. Notice:

$$M_X^{\text{out}}(t) = M_X^{\text{in}}(t) = \frac{1}{|X_0|} \sum_{k \geq 0} |X_k| t^k.$$

Topological metrics - Bottleneck polynomials

- Define the k-th bottleneck coefficient of $v \in X_0$ by

$$b_k(v) = \frac{m_k^{\text{out}}(v)}{m_k^{\text{in}}(v)}.$$

- Define the bottleneck polynomial of $v \in X_0$ to be

$$B_v(t) = \sum_{k \geq 0} b_k(v) t^k.$$

- The bottleneck polynomial of X by

$$B_X(t) = \frac{1}{|X_0|} \sum_{v \in X_0} B_v(t).$$

Topological metrics - clustering and segregation

- Define

$$N_k^{\text{in}}(v) = \begin{cases} \frac{m_1^{\text{in}}(v)!}{(m_1^{\text{in}}(v)-k)!} & m_1^{\text{in}}(v) \geq k \\ 0 & m_1^{\text{in}}(v) < k \end{cases},$$

$$N_k^{\text{out}}(v) = \begin{cases} \frac{m_1^{\text{out}}(v)!}{(m_1^{\text{out}}(v)-k)!} & m_1^{\text{out}}(v) \geq k \\ 0 & m_1^{\text{out}}(v) < k \end{cases},$$

  and

$$N_k(v) = \sum_{i=0}^{k} N_i^{\text{in}}(v) \cdot N_{k-i}^{\text{out}}(v).$$

- The number $N_k(v)$ is the largest number of k -simplices that can be formed with v as a vertex given the edges incident to it.

Topological metrics - clustering and segregation

For example, when $k = 2$ and $m_1^{\text{in}}(v), m_1^{\text{out}}(v) \geq 2$,

$$N_2(v) = m_1(v)(m_1(v) - 1) - m_1^{\text{in}}(v)m_1^{\text{out}}(v),$$

which is the maximal number of directed triangles to which a vertex that is the source of $m_1^{\text{out}}(v)$ edges and the target of $m_1^{\text{in}}(v)$ edges can belong.

- For every $k \geq 2$, the k-clustering coefficient of $v \in X_0$ is the ratio
$$C_k(v) = \frac{m_k(v)}{N_k(v)}.$$
- The clustering polynomial of $v \in X_0$ is defined by
$$S_v(t) = 1 + t + \sum_{k \geq 2} C_k(v)t^k.$$

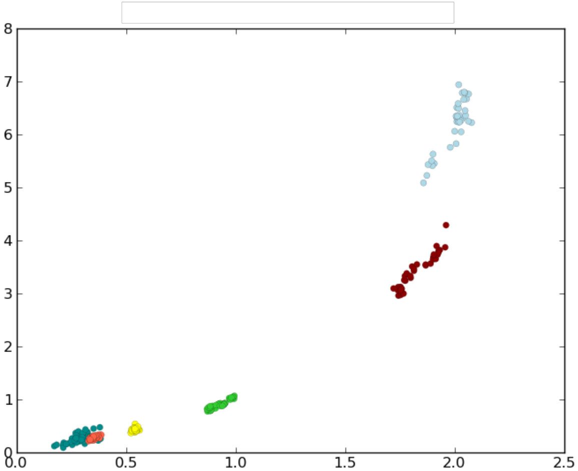
Topological metrics

- The segregation polynomial of a simplicial complex $\mathcal{S}$ is defined by
$$S_X(t) = \frac{1}{|X_0|} \sum_{v \in X_0} S_v(t).$$

- Notice that the coefficient of t^2 in $S_X(t)$ is the classical clustering coefficient of the graph corresponding to the 1-skeleton of X.

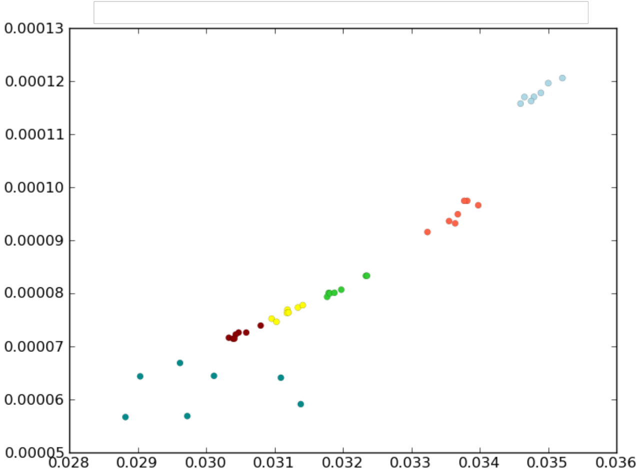
Example 1: Bottleneck polynomial - By layer

Coefficients of t and t^2.



Example 2: Segregation polynomial - by layer

Coefficients of t^2 and t^3.



Highways and Flow

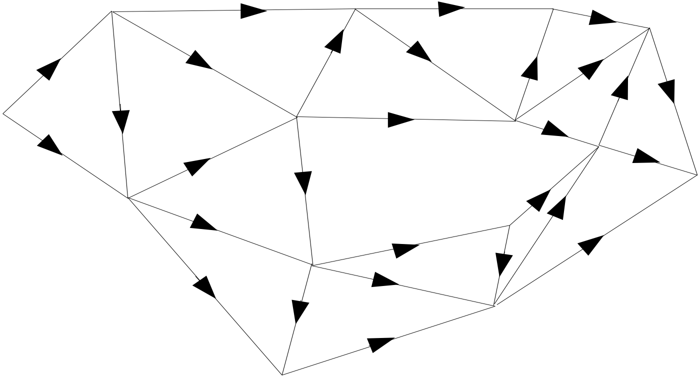
- Let X be an oriented simplicial complex, and let $x, y \in X_0$ be any two vertices.

- A *d-dimensional highway* from x to y is either a d-simplex $(x, x_1, \ldots x_{d-1}, y)$ in X_d or a sequence of $(d+1)$-simplices

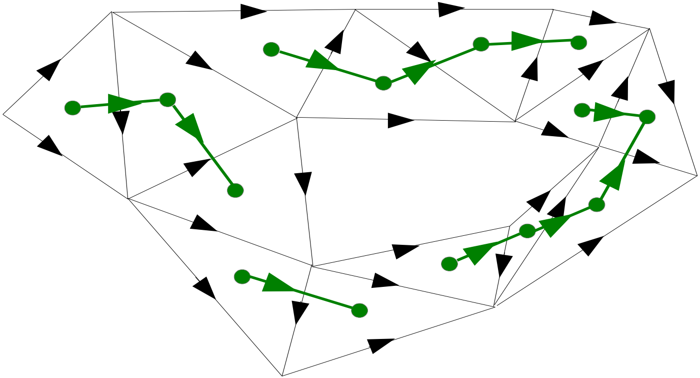
$$\sigma_0, \ldots, \sigma_m$$

  in X, such that $\sigma_i \cap \sigma_{i+1}$ is a back d-face of σ_i and a front d-face of σ_{i+1}, for all $i \geq 0$, and such that x is an initial vertex in σ_0 and y is a final vertex in σ_m.

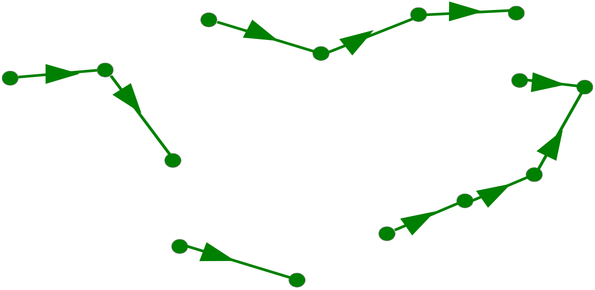
Example: 1-highways



Example: 1-highways



Example: 1-highway graph



Highways and Flow

- For vertices $v, w \in X_0$, and $d \leq 0$, let $i_d(v, w)$ denote the integration coefficient of the pair (v, w), i.e. the minimal length of a d-dimensional highway from v to w if one exists, and set $i_d(v, w) = 0$ if it doesn't or if $v = w$ and $d > 0$. Also set $i_0(v, v) = 1$.

- If the edges of a simplicial complex are weighted, then one can assign a max flow capacity on a d-dimensional highway $f_d(v, w)$ to each pair or vertices (v, w) (0 if such a highways does not exist).

- These are harder (more expensive in time and memory) to compute, but there are good approximation algorithms.

Highways and Flow

- The integration polynomial of a (weighted) oriented simplicial complex X is defined by

$$I_X(t) = \frac{1}{|X_0|(|X_0 - 1|)} \cdot \sum_{(x,y)\in X_0\times X_0} \sum_{d\geq 0} i_d(x,y)t^d.$$

- The flow polynomial of a (weighted) oriented simplicial complex X is defined by

$$F_X(t) = \frac{1}{|X_0|(|X_0 - 1|)} \cdot \sum_{(x,y)\in X_0\times X_0} \sum_{d\geq 0} f_d(x,y)t^d.$$