SIAM AG / Applications of Magnitude and Magnitude Homology to Network Analysis

Prospects for applications of magnitude homology

Steve Huntsman

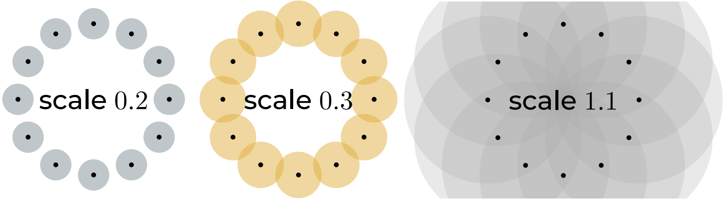
July 2023

Contents

1. **Context** from mainstream topological data analysis to discrete tools
2. **Magnitude homology** with the briefest nod towards cyber applications
3. **Examples** in the vein of neural stuff (feedforward nets and actual brains)

Mainstream TDA has a playbook

- Approximate $\{X_j\}_{j=1}^n \subset (\mathbb{R}^d)^n$ at various scales
- Compute a topological invariant of each approximation
- Highlight invariants that persist across scales
- E.g., 12 equispaced points on the unit circle (half-distance $\Delta \approx 0.2588$)
  - $\beta_0 = 12 \cdot 1_{[0,\Delta)} + 1_{[\Delta,1)} + 1_{[1,\infty)}$
  - $\beta_1 = 1_{[\Delta,1)}$



We are going to play a simpler game

CATEGORY CAPITAL

- Data that interests us is fundamentally discrete
  - Approximate **nothing** (via representation or model if need be)
  - Compute a (co)chain complex **directly from the data**: simplicial complexes are optional technical devices
  - Highlight any persistence **incidentally**: it often won't apply

We are going to play a simpler game

- Data that interests us is fundamentally discrete
  - Approximate **nothing** (via representation or model if need be)
  - Compute a (co)chain complex **directly from the data**: simplicial complexes are optional technical devices
  - Highlight any persistence **incidentally**: it often won't apply
- Archetype: Dowker homology
  - Relation corresponds to a pair of homotopy equivalent simplicial complexes
  - $\mathbb{F}_2$ homology readily computed directly from relation itself, without ever constructing either simplicial complex

We are going to play a simpler game

- Data that interests us is fundamentally discrete
  - Approximate **nothing** (via representation or model if need be)
  - Compute a (co)chain complex **directly from the data**: simplicial complexes are optional technical devices
  - Highlight any persistence **incidentally**: it often won't apply
- Archetype: Dowker homology
  - Relation corresponds to a pair of homotopy equivalent simplicial complexes
  - $\mathbb{F}_2$ homology readily computed directly from relation itself, without ever constructing either simplicial complex
- Many well known discrete tools are underutilized
  - Finite topological spaces *per se*
  - Simplicial complexes generated by posets/hypergraphs
  - Discrete Morse theory
  - ...

This simpler game has practical relevance

- Dowker applications:
    - Witness complexes
    - Systems/sociology
    - Robotics
    - Neuroscience
    - Privacy
    - Software engineering
    - ...
    - Exercises for audience: cyber interactions like user/computer; process/file; etc

This simpler game has practical relevance

- Dowker applications:
  - Witness complexes
  - Systems/sociology
  - Robotics
  - Neuroscience
  - Privacy
  - Software engineering
  - …
  - Exercises for audience: cyber interactions like user/computer; process/file; etc
- Magnitude homology for digraphs is suited for novel applications…
  - …that can give rise to new capabilities

This simpler game has practical relevance

CATEGORY CAPITAL

- Dowker applications:
  - Witness complexes
  - Systems/sociology
  - Robotics
  - Neuroscience
  - Privacy
  - Software engineering
  - …
  - Exercises for audience: cyber interactions like user/computer; process/file; etc
- Magnitude homology for digraphs is suited for novel applications…
  - …that can give rise to new capabilities
  - Cf. path homology, which is related but lacks *ab initio* source/target specificity

Magnitude homology is readily constructed

CATEGORY CAPITAL

- Let (X, d) be a Lawvere metric space $\Leftrightarrow$ d = extended quasipseudometric
  - $d : X \times X \to [0, \infty]$ (extended)
  - d need not be symmetric (quasi-)
  - $d(x, x)$ need not be zero (pseudo-)
  - (X, d) best thought of here as a category enriched over the poset $([0, \infty], \geq)$

Magnitude homology is readily constructed

CATEGORY CAPITAL

- Let (X, d) be a Lawvere metric space $\Leftrightarrow$ $d =$ extended quasipseudometric
  - $d : X \times X \to [0, \infty]$ (extended)
  - d need not be symmetric (quasi-)
  - $d(x, x)$ need not be zero (pseudo-)
  - (X, d) best thought of here as a category enriched over the poset $([0, \infty], \geq)$
- A k-*simplex* in X is an ordered tuple $x^{(k)} := (x_0, \ldots, x_k) \in X^{k+1}$ s.t. $x_j \neq x_{j-1}$
- The *length* of $x^{(k)}$ is $\lambda(x^{(k)}) := \sum_{j=1}^{k} d(x_{j-1}, x_j)$

Magnitude homology is readily constructed

CATEGORY CAPITAL

- Let (X, d) be a Lawvere metric space $\Leftrightarrow d =$ extended quasipseudometric
  - $d : X \times X \to [0, \infty]$ (extended)
  - d need not be symmetric (quasi-)
  - $d(x, x)$ need not be zero (pseudo-)
  - (X, d) best thought of here as a category enriched over the poset $([0, \infty], \geq)$
- A k-*simplex* in X is an ordered tuple $x^{(k)} := (x_0, \dots, x_k) \in X^{k+1}$ s.t. $x_j \neq x_{j-1}$
- The *length* of $x^{(k)}$ is $\lambda(x^{(k)}) := \sum_{j=1}^{k} d(x_{j-1}, x_j)$

The $\mathbb{R}$-graded *magnitude chain complex* (as before R is a coefficient ring):

- k-chains: $MC_{k,L}(X) := R\left\{x^{(k)} \text{ a } k\text{-simplex in } X : \lambda(x^{(k)}) = L\right\}$
- differential: $\partial_k : MC_{k,L}(X) \to MC_{k-1,L}(X)$ given by $\partial_k := \sum_{j=1}^{k-1} (-1)^j \partial^{(j)}$
  - $\partial^{(j)}(x^{(k)}) := \nabla_j x^{(k)}$ if $d(x_{j-1}, x_{j+1}) = d(x_{j-1}, x_j) + d(x_j, x_{j+1})$ and $:= 0$ otherwise
- Appropriate notion of chain map is induced by d-nonincreasing maps

Magnitude homology is readily constructed

CATEGORY CAPITAL

- Let (X, d) be a Lawvere metric space $\Leftrightarrow$ d = extended quasipseudometric
  - $d : X \times X \to [0, \infty]$ (extended)
  - d need not be symmetric (quasi-)
  - $d(x, x)$ need not be zero (pseudo-)
  - (X, d) best thought of here as a category enriched over the poset $([0, \infty], \geq)$
- A k-*simplex* in X is an ordered tuple $x^{(k)} := (x_0, \dots, x_k) \in X^{k+1}$ s.t. $x_j \neq x_{j-1}$
- The *length* of $x^{(k)}$ is $\lambda(x^{(k)}) := \sum_{j=1}^{k} d(x_{j-1}, x_j)$

The $\mathbb{R}$-graded *magnitude chain complex* (as before R is a coefficient ring):

- k-chains: $MC_{k,L}(X) := R\left\{x^{(k)} \text{ a } k\text{-simplex in } X : \lambda(x^{(k)}) = L\right\}$
- differential: $\partial_k : MC_{k,L}(X) \to MC_{k-1,L}(X)$ given by $\partial_k := \sum_{j=1}^{k-1} (-1)^j \partial^{(j)}$
  - $\partial^{(j)}(x^{(k)}) := \nabla_j x^{(k)}$ if $d(x_{j-1}, x_{j+1}) = d(x_{j-1}, x_j) + d(x_j, x_{j+1})$ and $:= 0$ otherwise
- Appropriate notion of chain map is induced by d-nonincreasing maps

The *magnitude homology* of X…

…is the homology of the magnitude chain complex

MH has attractive features

CATEGORY CAPITAL

- For a digraph D, there is a decomposition of the form

$$MC_{\bullet,L}(D) = \bigoplus_{s,t \in V(D)} MC^{(s,t)}_{\bullet,L}(D)$$

  - $MC^{(s,t)}_{\bullet,L}(D)$ generated by simplices with (initial, terminal) entries (s,t)

MH has attractive features

- For a digraph D, there is a decomposition of the form

$$MC_{\bullet,L}(D) = \bigoplus_{s,t \in V(D)} MC_{\bullet,L}^{(s,t)}(D)$$

  - $MC_{\bullet,L}^{(s,t)}(D)$ generated by simplices with (initial, terminal) entries (s,t)
  - This structure makes MH particularly convenient for unweighted digraphs

MH has attractive features

CATEGORY CAPITAL

- For a digraph D, there is a decomposition of the form

$$MC_{\bullet,L}(D) = \bigoplus_{s,t \in V(D)} MC_{\bullet,L}^{(s,t)}(D)$$

  - $MC_{\bullet,L}^{(s,t)}(D)$ generated by simplices with (initial, terminal) entries (s,t)
  - This structure makes MH particularly convenient for unweighted digraphs
- A "blurred" variant replaces $\lambda = L$ in the chain complex definition with $\lambda \leq L$

MH has attractive features

- For a digraph D, there is a decomposition of the form

$$MC_{\bullet,L}(D) = \bigoplus_{s,t \in V(D)} MC_{\bullet,L}^{(s,t)}(D)$$

  - $MC_{\bullet,L}^{(s,t)}(D)$ generated by simplices with (initial, terminal) entries (s,t)
  - This structure makes MH particularly convenient for unweighted digraphs
- A “blurred” variant replaces $\lambda = L$ in the chain complex definition with $\lambda \leq L$
- MH is related to both persistent homology and path homology

MH has attractive features

CATEGORY CAPITAL

- For a digraph D, there is a decomposition of the form

$$MC_{\bullet,L}(D) = \bigoplus_{s,t \in V(D)} MC_{\bullet,L}^{(s,t)}(D)$$

  - $MC_{\bullet,L}^{(s,t)}(D)$ generated by simplices with (initial, terminal) entries (s,t)
  - This structure makes MH particularly convenient for unweighted digraphs
- A "blurred" variant replaces $\lambda = L$ in the chain complex definition with $\lambda \leq L$
- MH is related to both persistent homology and path homology
- *Magnitude cohomology* has an analogue of the cup product
  - Ring structure determines the space for finite extended quasi-metric spaces
    - E.g., digraphs; finite metric spaces
  - Meanwhile, tree magnitude (co)homology *groups* only depend on #(vertices)

MH generalizes to enriched categories

CATEGORY CAPITAL

- Consider enriched category of sub-flow graphs of a given flow graph
  - For semicartesianness, take a specific category $\mathbf{F}$ of “two-terminal graphs”
    - Presents some technical issues for decomposition (probably surmountable)
  - The topological entropy is a “size function” over the max-plus semiring
    - Recall this is log rate of growth of possible paths as a function of length
  - This setup dovetails with the source-target direct sum decomposition
  - A categorification might be interesting and efficiently computable
    - Representatives might encode nice structure
  - Possibility to connect to useful cyber applications via compilers
- More ambitiously, consider $\mathbf{F} \times \mathbf{M}$ where $\mathbf{M}$ is a “matrix category” of data
  - Size function may be something like the zeta function of a Markov chain
- Suitable data obtainable from program analysis
  - Simpler initial goal: use hitting probabilities metric of Boyd *et al.* (this is vanilla)

MH measures nonconvexity

Let $P_{p,n}$ be the DAG formed from p parallel paths of n arcs from 1 to $(n-1)p+2$

Then all the Betti numbers are zero except for

$$\beta(P_{p,n})_{0,0} = |V(P_{p,n})| = (n-1)p+2;$$

$$\beta(P_{p,n})_{1,1} = |A(P_{p,n})| = np,$$

and

$$\beta(P_{p,n})_{2,n} = p-1;$$

i.e., there are $p-1$ “convexity defects” of length n in homology dimension 2

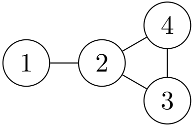
Measuring nonconvexity may be useful from the PoV of flows/routing

Besides information/transportation networks, neural stuff is also interesting

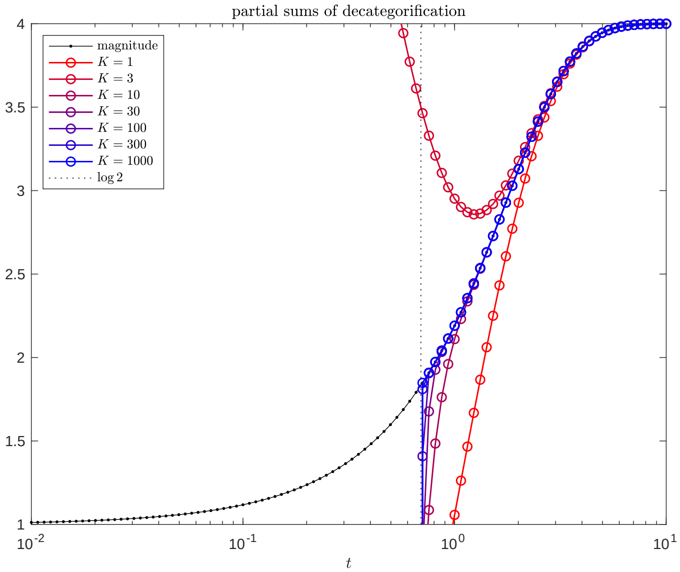
Categorification helps analyze magnitude

CATEGORY CAPITAL

Magnitude homology of



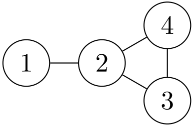
decategorifies like magnitude as shown



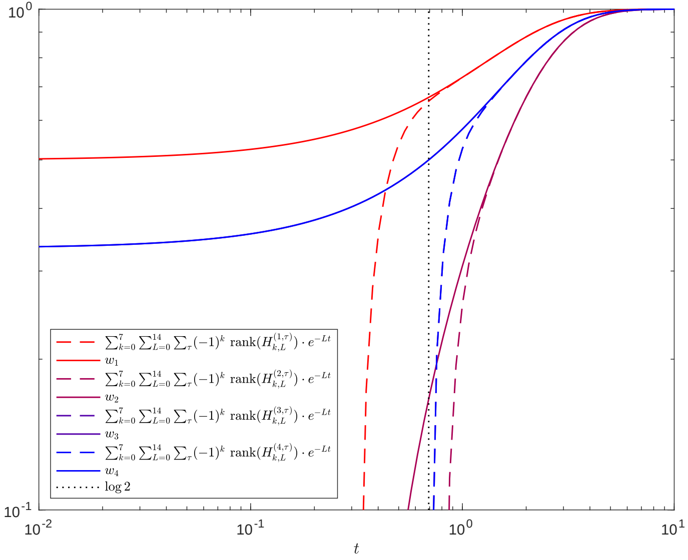
Categorification helps analyze weighting

CATEGORY CAPITAL

Magnitude homology of

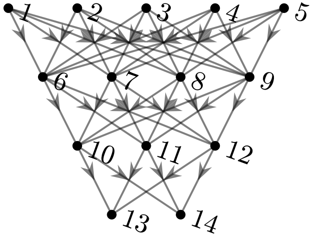


decategorifies like weighting as shown

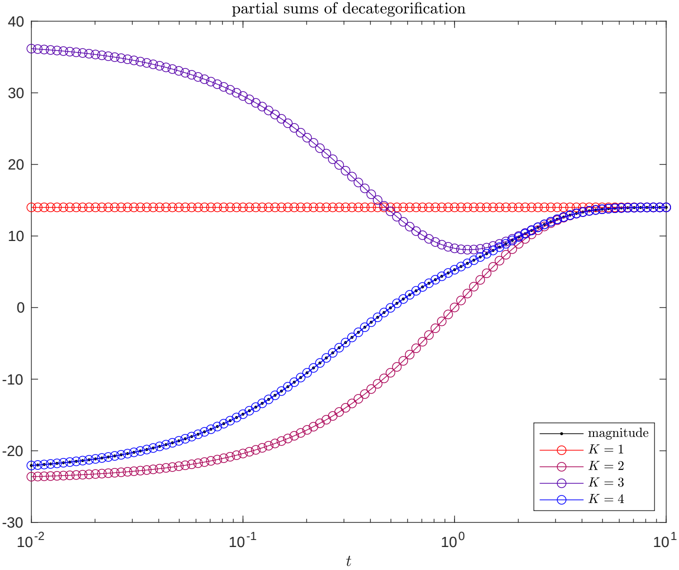


DAGs are very nicely handled

Magnitude homology of *multilayer perceptron* (MLP) $K^{\rightarrow}_{5,4,3,2}$

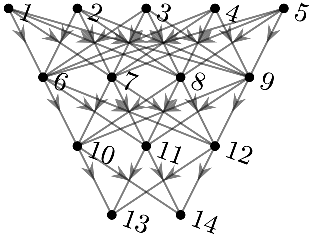


decategorifies as shown

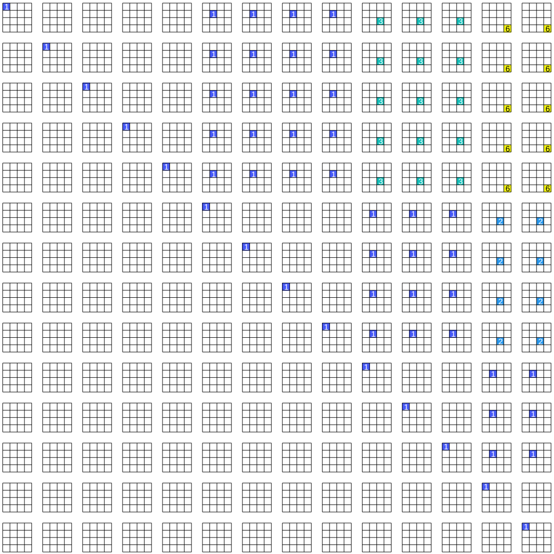


DAGs are very nicely handled

Magnitude homology of *multilayer perceptron* (MLP) $K^{\rightarrow}_{5,4,3,2}$



has Betti numbers shown



Fully connected MLPs have simple MH

CATEGORY CAPITAL

- Let $N_\ell := \sum_{i=1}^{\ell} n_i$, $e_{(\ell)} := \sum_{j=N_{\ell-1}+1}^{N_\ell} e_j$, and $B[x_1, \ldots, x_{L-k}; n] := \sum_{\ell=1}^{L-k} x_\ell e_{(\ell)}^T e_{(\ell+k)}$
  - E.g. $A = B[1_1, \ldots, 1_{L-1}; n]$

The preceding slide's mechanics suggest the conjecture (exercise)

For $k > 1$, $\beta_{k+1,m}^{(s,t)}(K_{n_1,\ldots,n_L}^{\rightarrow}) = \delta_{k+1,m} \cdot \left(B\left[\prod_{\ell=2}^{k+1}(n_\ell - 1), \ldots, \prod_{\ell=L-k}^{L-1}(n_\ell - 1); n\right]\right)_{st}$

Note that we automatically have $\beta_{0,m}^{(s,t)} = \delta_{0m}\delta_{st}$ and $\beta_{1,m}^{(s,t)} = \delta_{1m}A_{st}$

Examples:

$\beta(K_{5,4,3,2}^{\rightarrow}) = \text{diag}(14, 38, 61, 60, 0, 0, \ldots)$
$\beta(K_{2,11,3,7,5}^{\rightarrow}) = \text{diag}(28, 111, 304, 940, 1200, 0, 0, \ldots)$
$\beta(K_{10,2,8,4,6}^{\rightarrow}) = \text{diag}(30, 92, 280, 532, 1260, 0, 0, \ldots)$

Sparsely connected MLPs are more relevant

Sparsity arises via (e.g.) pruning small edge weights in trained networks

"Lottery tickets" exhibit hierarchical modularity [Patil, Michael, and Dovrolis]

Magnitude homology might usefully indicate structure...

...insofar as it manifests as "convexity defects" that differ from a null model

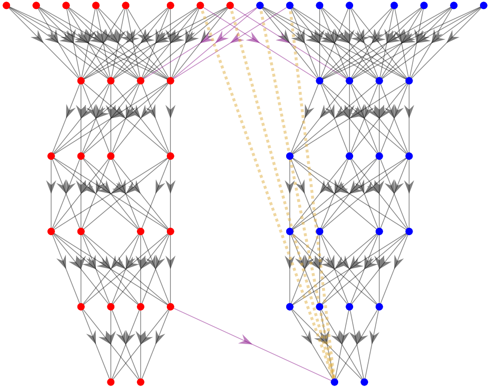
We provide some *very preliminary* supporting evidence

Sparse interconnections of MLPs and sparsification of a single MLP both induce off-diagonal magnitude homology, but differently

Sparse interconnections ⇒ off-diagonals

CATEGORY CAPITAL

$$\beta = \begin{pmatrix} 52 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 182 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 436 & 0 & 0 & 4 & 0 & \cdots \\ 0 & 0 & 0 & 1020 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 2196 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 2646 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$



Off-diagonal contributions indicated by dotted lines from sources to targets

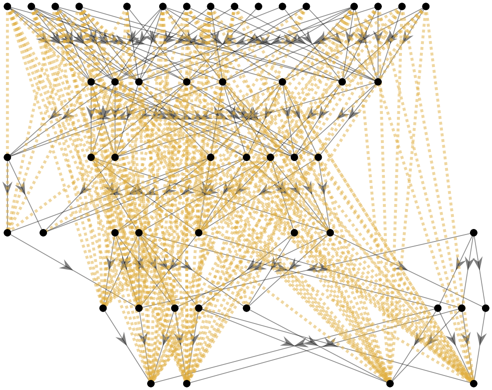
Vertices colored by undirected Fiedler vector (i.e., spectral) bipartition; traversing arcs are purple

Sparsification ⇒ off-diagonals (differently)

$$\beta = \begin{pmatrix} 52 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 143 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 164 & 44 & 5 & 0 & 0 & \dots \\ 0 & 0 & 0 & 71 & 71 & 38 & 0 & \dots \\ 0 & 0 & 0 & 0 & 2 & 47 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Off-diagonal contributions indicated by dotted lines from sources to targets

Keeping $\approx 2/5$ of arcs in $K^{\rightarrow}_{16,8,8,8,8,4}$



This may be useful in practice

CATEGORY CAPITAL

Deep neural networks are functionally sparsified and modular

Weight filtrations (for sparsity) and quantizations (for efficiency) are common

A "null model" of IID arc weights may actually be fairly realistic

Neural networks are initialized randomly; training finds one of many extrema

For a DAG the dimensions of simplices and lengths are bounded

In this case numerics largely suffice: e.g. the (de)categorification formula

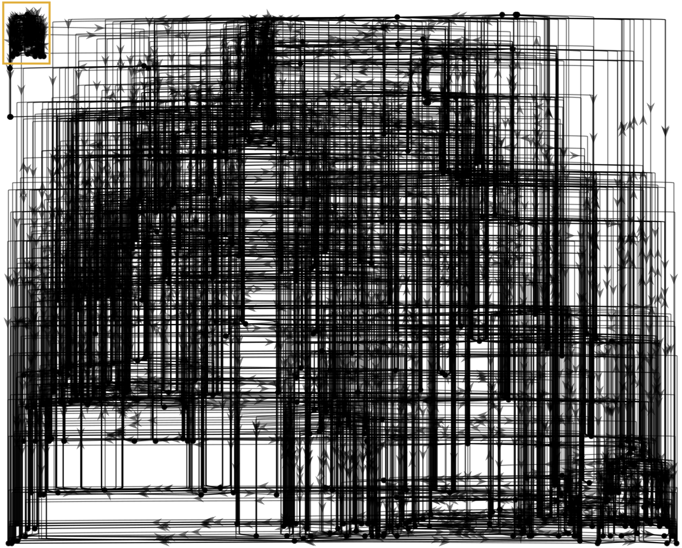
$$((\exp[-\tau d])^{-1})_{st} \overset{\tau \gg 0}{=} \sum_{k,L} (-1)^k \text{rank}\left(MH_{k,L}^{(s,t)}\right) \exp(-\tau L) = \sum_L \chi_{\bullet,L}^{(s,t)} \exp(-\tau L)$$

can exactly produce $\chi_{\bullet,L}^{(s,t)}$ via Laplace transform *without ever computing* MH

Consider a whole-animal connectome

Whole-animal chemical connectome of hermaphroditic *Caenorhabditis elegans* (nematode) is a digraph with 454 vertices and 4879 arcs (see https://wormwiring.org)

The pharynx subconnectome with 50 vertices and 242 arcs is in the box at upper left



The pharynx has off-diagonal 2-homology

CATEGORY CAPITAL

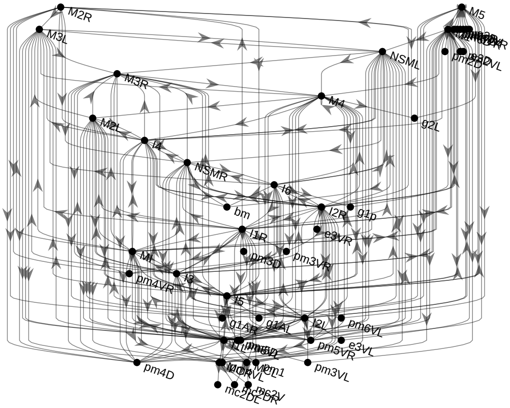
$$\beta = \begin{pmatrix} 50 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 241 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1290 & 11 & 0 & 9 & 13 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Off-diagonal 2-homology is intricate

The subconnectome is not Menger convex or geodetic; it has 4-cuts

Exercise: ∃ biological significance?

E.g., 10/11 reps for $(k, L) = (2, 3)$ either have source neurons M2R, M3R, or NSML, or have target neuron MCR



Takeaways

CATEGORY CAPITAL

Magnitude homology might be considered too abstract for real applications

But this would be a mistake

Takeaways

CATEGORY CAPITAL

Magnitude homology might be considered too abstract for real applications

But this would be a mistake

Magnitude homology is amenable to computation *in silico*

It can enable new capabilities for fundamentally discrete data and problems that are ubiquitous in network science, machine learning, and cybersecurity

Thanks

MATLAB code at
https://github.com/SteveHuntsman/MagnitudeHomology