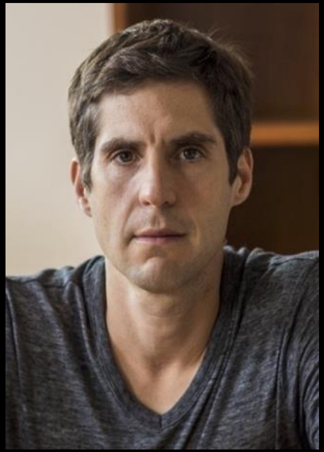
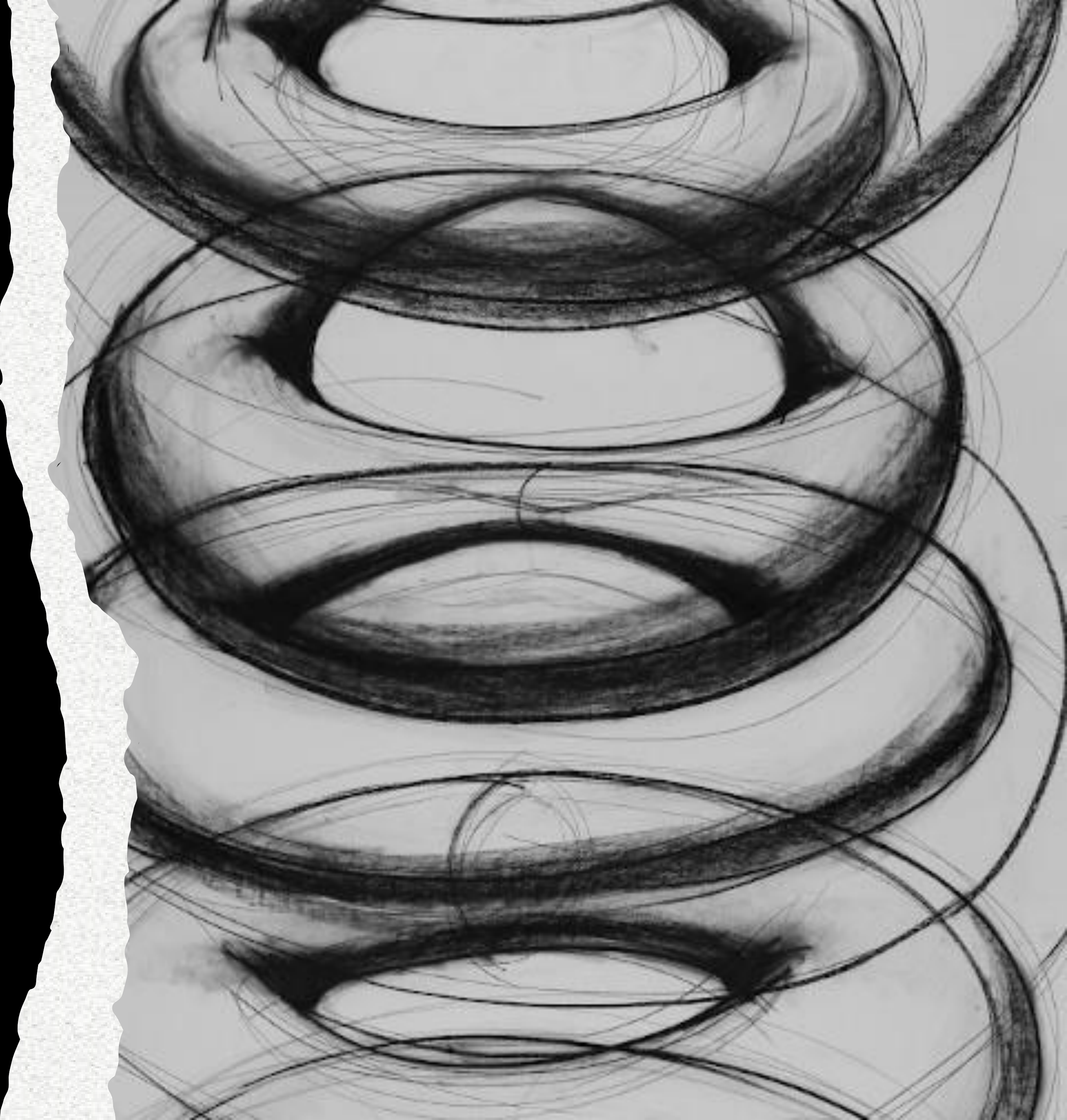
Stratifying Reinforcement Learning Games

Prof. Justin M. Curry @ SUNY Albany

Basic Stratification Theory

From Whitney to MacPherson to Lurie





Historical Stratification Theory for Varieties

The solution set V

$$\boxed{x^2 + y^2 - z^2 = 0}$$

naturally has

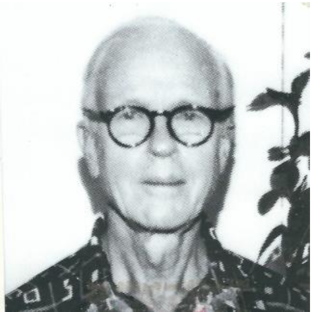
0-dimensional

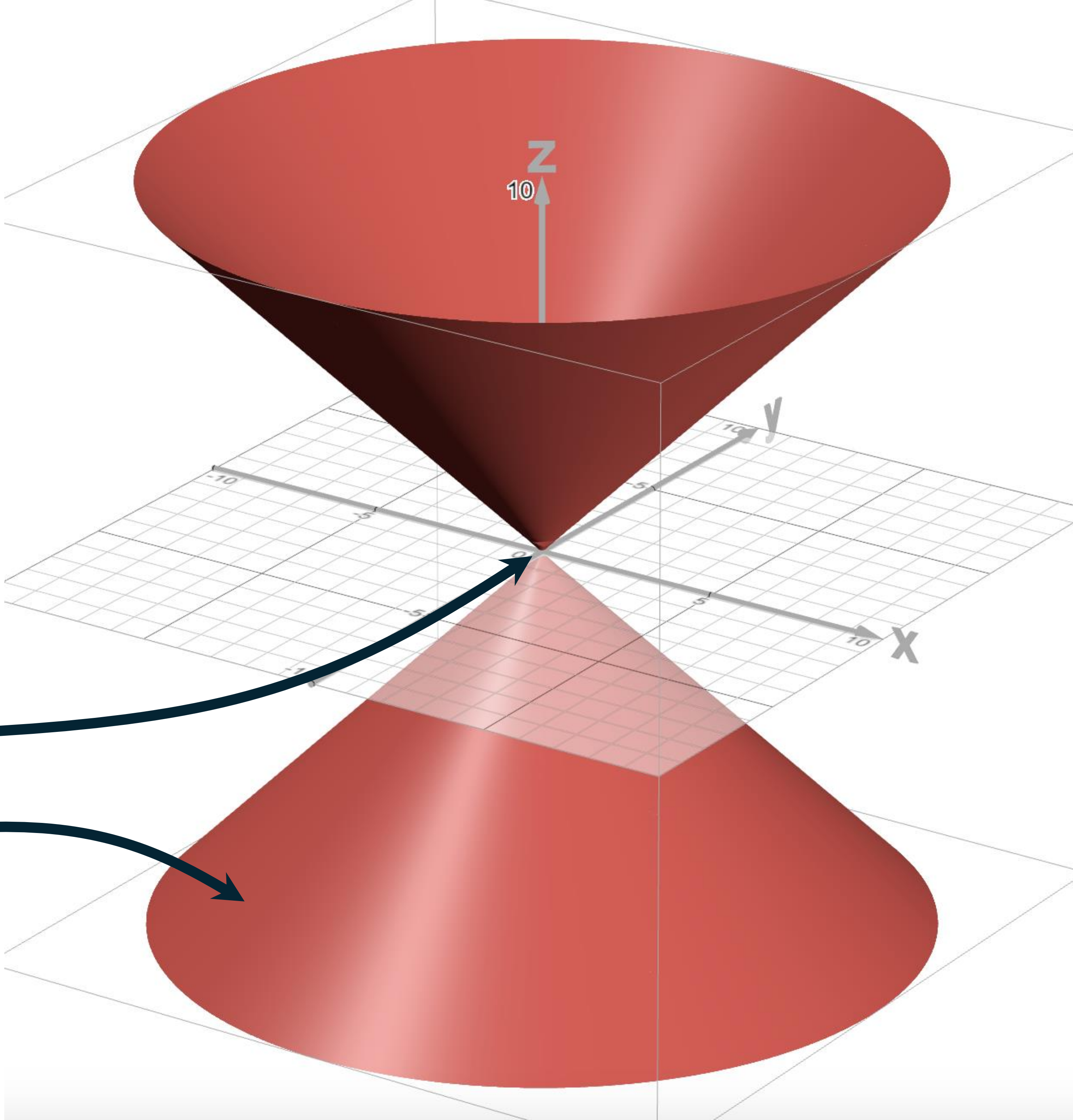
and

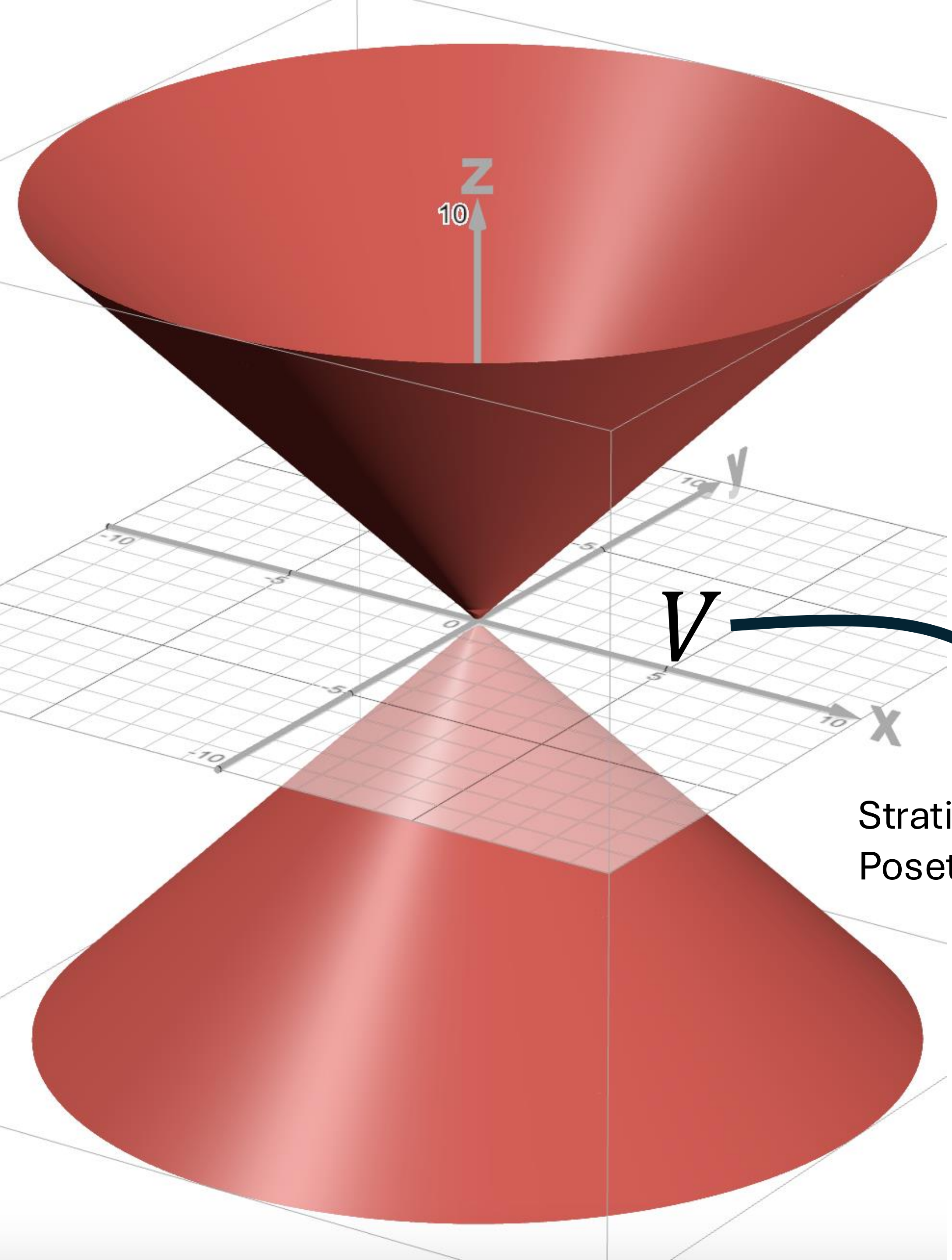
2-dimensional

“parts”

H. Whitney







J. Lurie

Modern Stratification Theory…

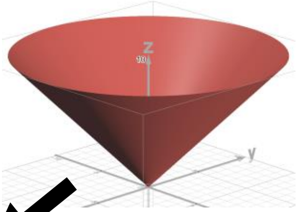
The solution to $\boxed{x^2 + y^2 - z^2 = 0}$

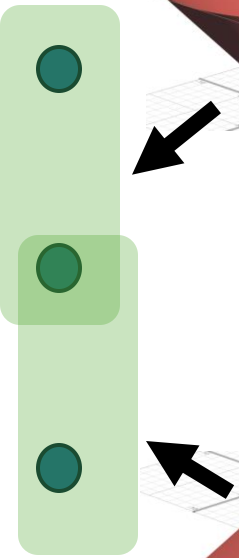
admits a *continuous* map to this poset!

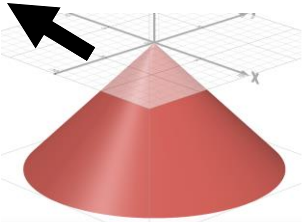
$$V \xrightarrow{\sigma} \mathcal{P}$$

Stratifying Poset

$$\boxed{\begin{array}{c}\beta_{>0} \\ \curlyvee \\ \alpha \\ \curlywedge \\ \beta_{<0}\end{array}}$$



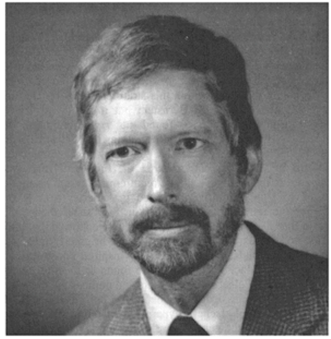




<u>Definition</u>

A map to a poset is **continuous** iff the pre-image of every down-set is closed

From Classical to Poset Stratifications



R. Macpherson

Historically, one defines a stratification from a filtration by closed sets:

$$\emptyset = X_{-1} \subset X_0 \subset X_1 \subset \cdots \subset X_n = X$$

Where one assumes that each $X^i := X_i - X_{i-1}$ is an i-dimensional manifold

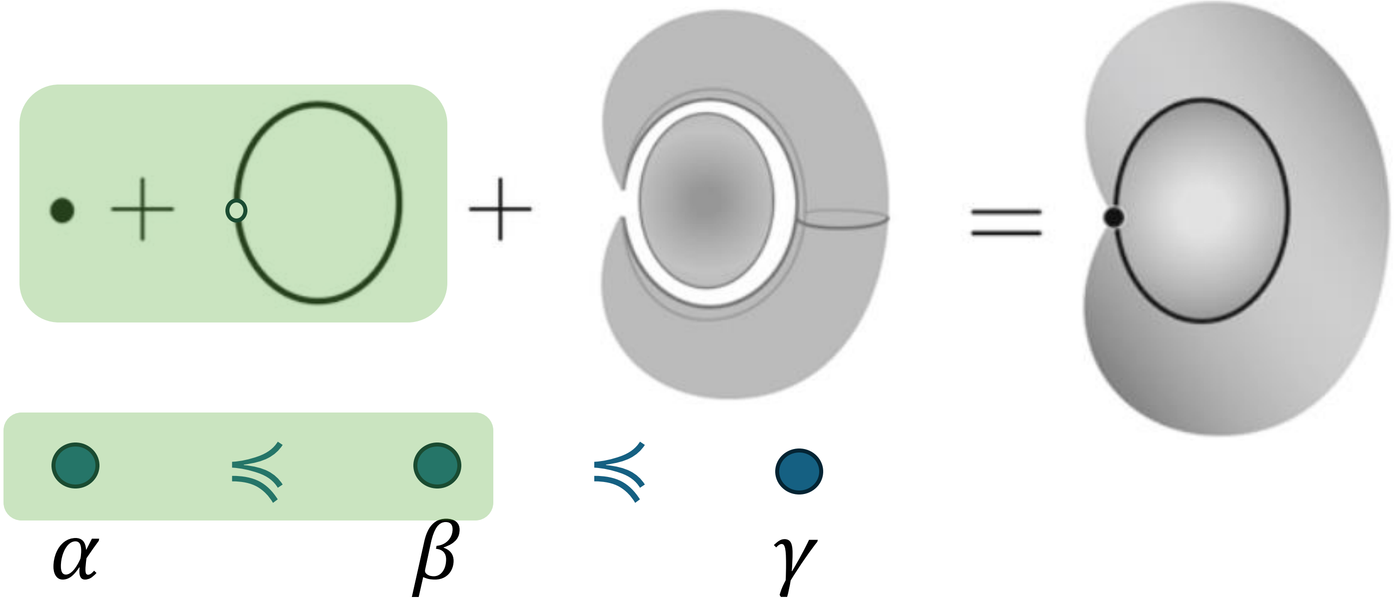
Use Map to a Poset!

Checking Continuity

$$X \qquad \sigma^{-1}(\boldsymbol{D}_{\beta})$$

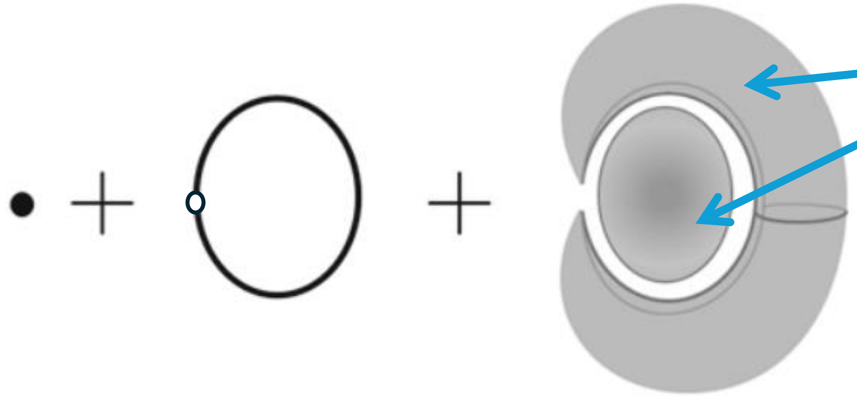
$$\sigma \downarrow$$

$$\mathcal{P} \qquad \boldsymbol{D}_{\beta}$$



Refining Poset Stratifications

Two connected components
in the top stratum!



$$\begin{tikzcd} X \arrow[d, "\sigma"'] \arrow[dd, bend right] \\ \mathcal{P} \\ Q \arrow[u] \end{tikzcd}$$

$$\alpha \preccurlyeq \beta \preccurlyeq \gamma$$

$$\alpha \preccurlyeq \beta \preccurlyeq \gamma_1 \qquad \beta \preccurlyeq \gamma_2$$

$$\begin{tikzcd} Y \arrow[r, "f"] \arrow[d, "\sigma_Y"'] & X \arrow[d, "\sigma_X"] \\ Q \arrow[r, "\sigma_f"'] & \mathcal{P} \end{tikzcd}$$

Refinement as a map
of stratified spaces

A. Einstein

M. Morse

Stratifications from Morse Theory

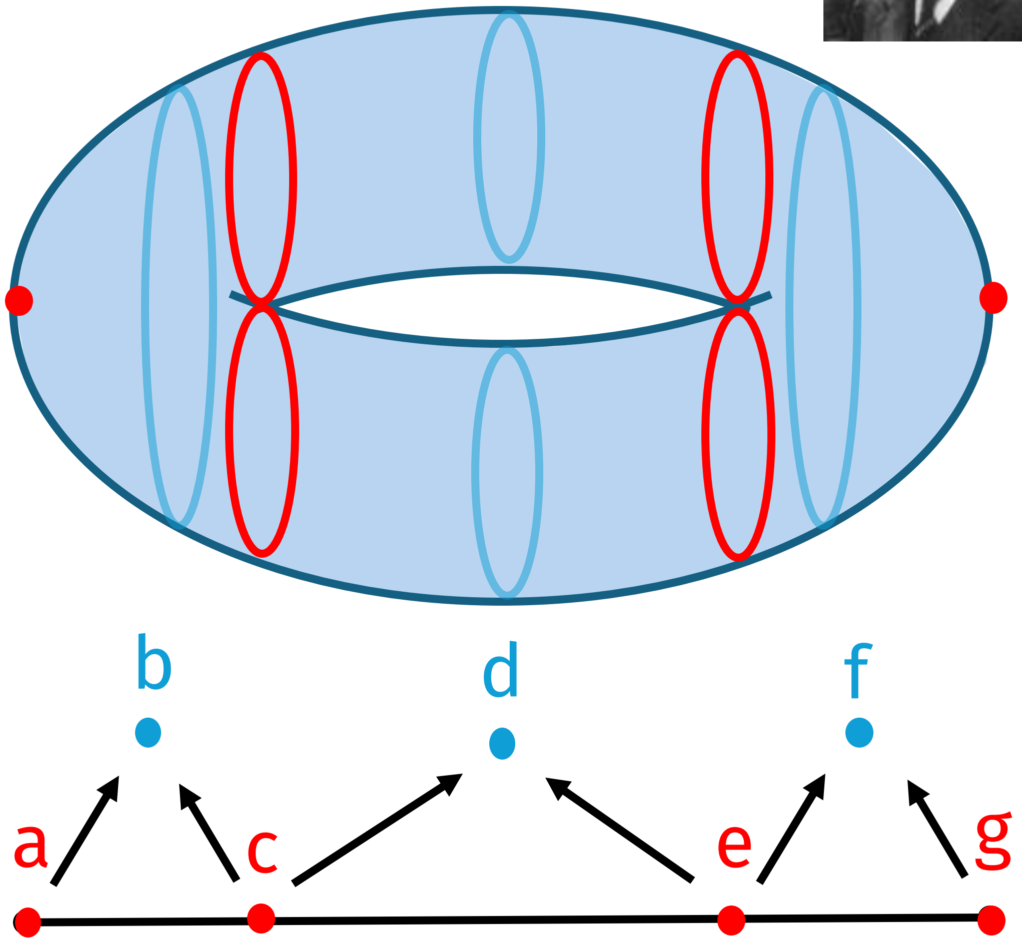
If $f : M \to \mathbb{R}$ is a Morse function with critical values

$$c_1, \dots, c_n$$

Then one obtains automatically a stratifying poset

$$(-\infty, c_1) \quad (c_1, c_2) \quad \dots \quad (c_n, \infty)$$

$$\{c_1\} \qquad \{c_2\} \qquad \{c_n\}$$

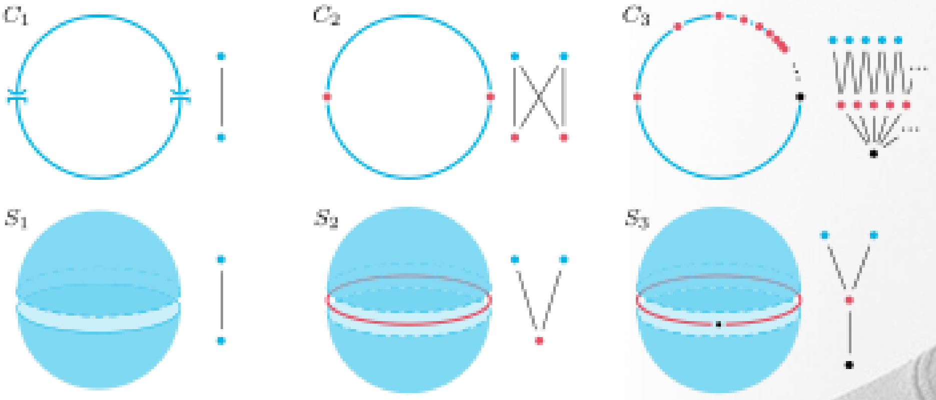


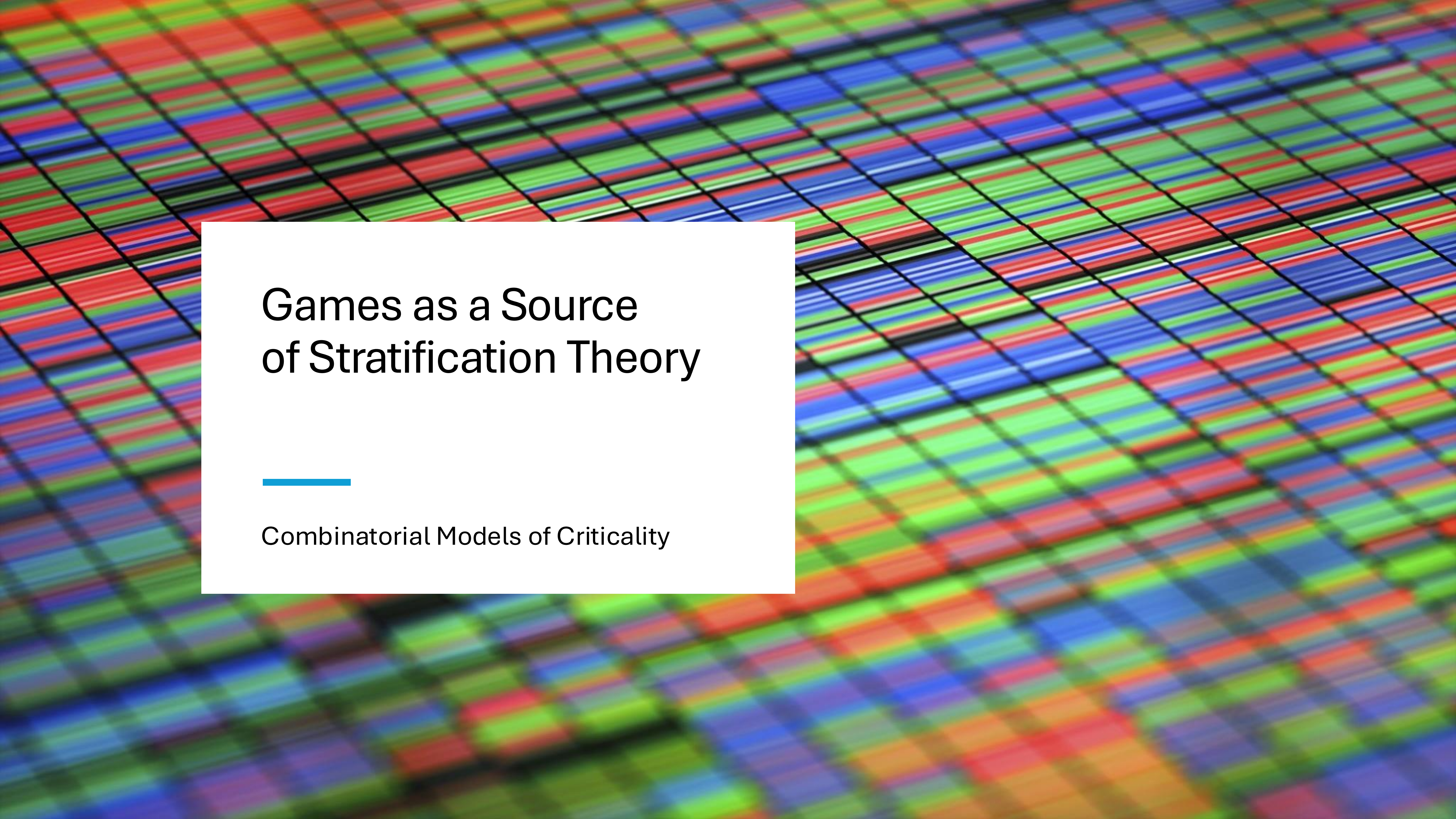
“Big” Stratifications in TDA

Stratifications on the Ran Space

Jānis Lazovskis[1]

Received: 20 December 2019 / Accepted: 28 April 2021

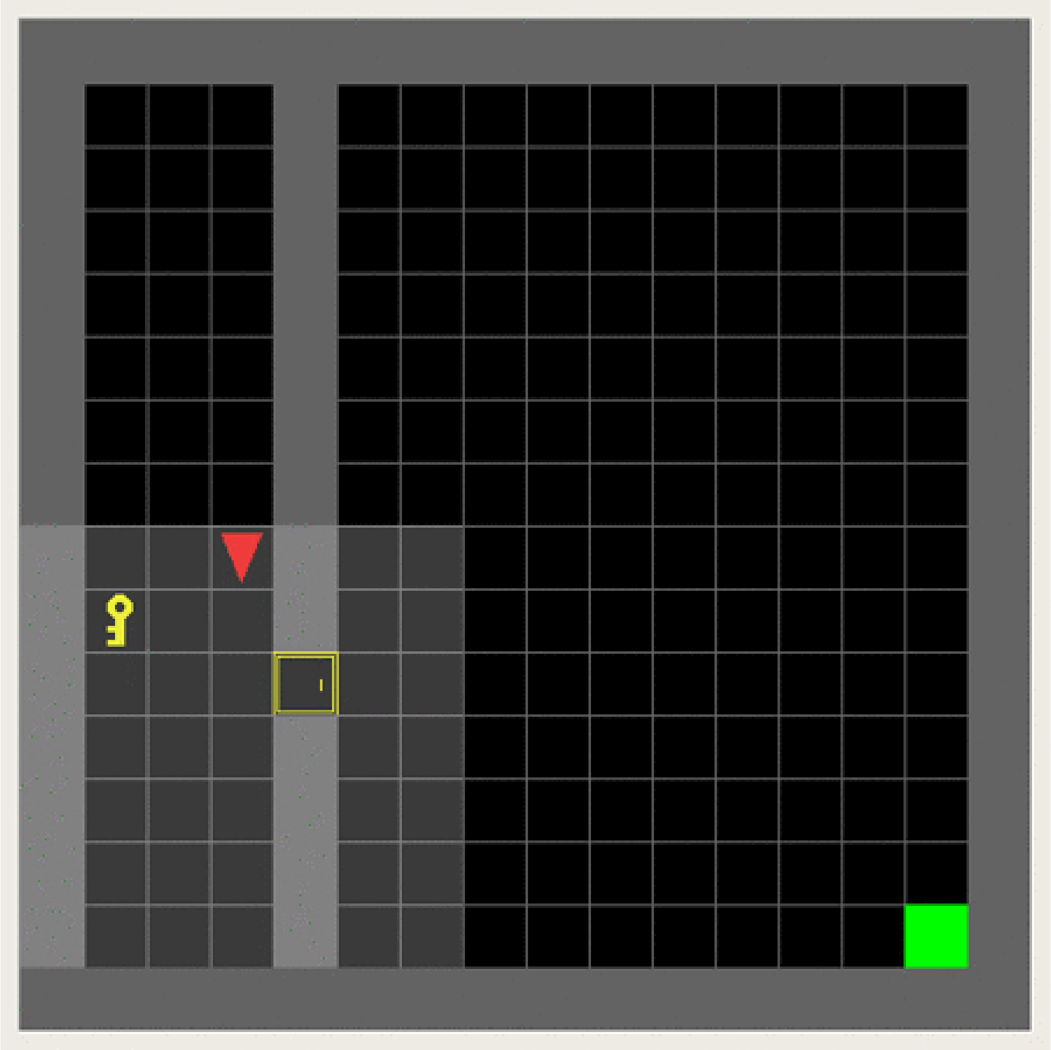




Games as a Source of Stratification Theory

Combinatorial Models of Criticality

A Digital Stratification Example

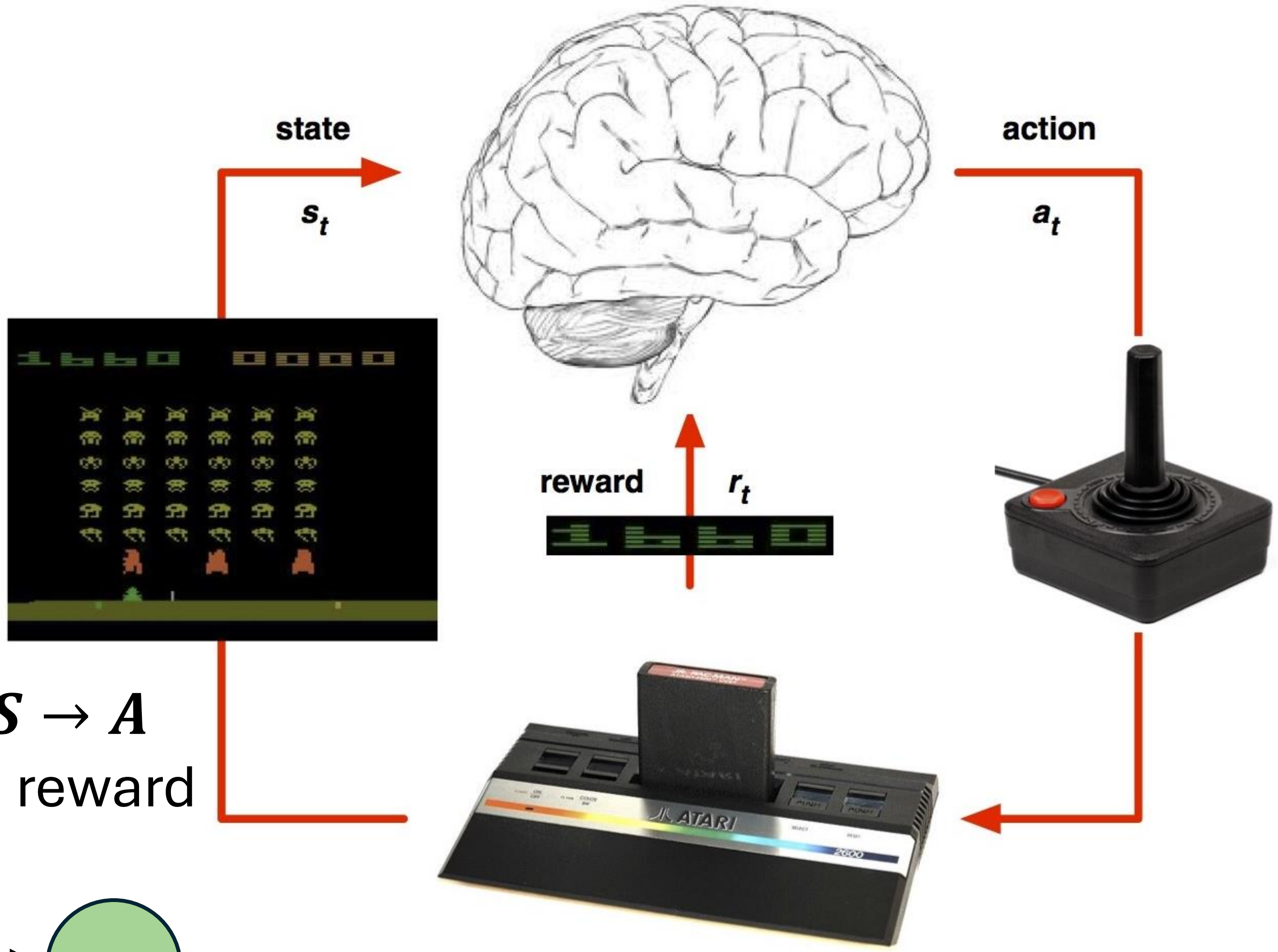


- What kinds of stratifications do we see here?
  - ➢Directionality: separation of inaccessible “past” regions. Describes a pre-order!
  - ➢Critical Events: reaching a key raises one’s energy level to access another connected component
  - ➢Dynamics: Goal state is a sink of some combinatorial vector field

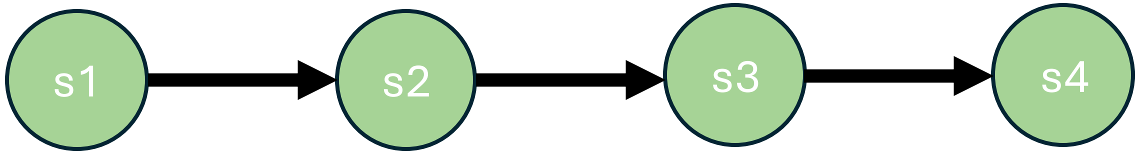
Markov Decision Processes (MDP)

Reinforcement Learning (RL) is often modeled using a MDP:

- $\boldsymbol{S}$ = set of **states**
- $\boldsymbol{A}$ = set of **actions**
- $\boldsymbol{T}: \boldsymbol{S} \times \boldsymbol{A} \rightarrow \boldsymbol{S}$ a **transition** map
- $\boldsymbol{R}: \boldsymbol{S} \times \boldsymbol{A} \times \boldsymbol{S} \rightarrow \mathbb{R}$ a **reward** for going from s to s' via a.



<u>GOAL:</u> Learn a **policy**, i.e., a map $\boldsymbol{P}: \boldsymbol{S} \rightarrow \boldsymbol{A}$ that maximizes cumulative expected reward

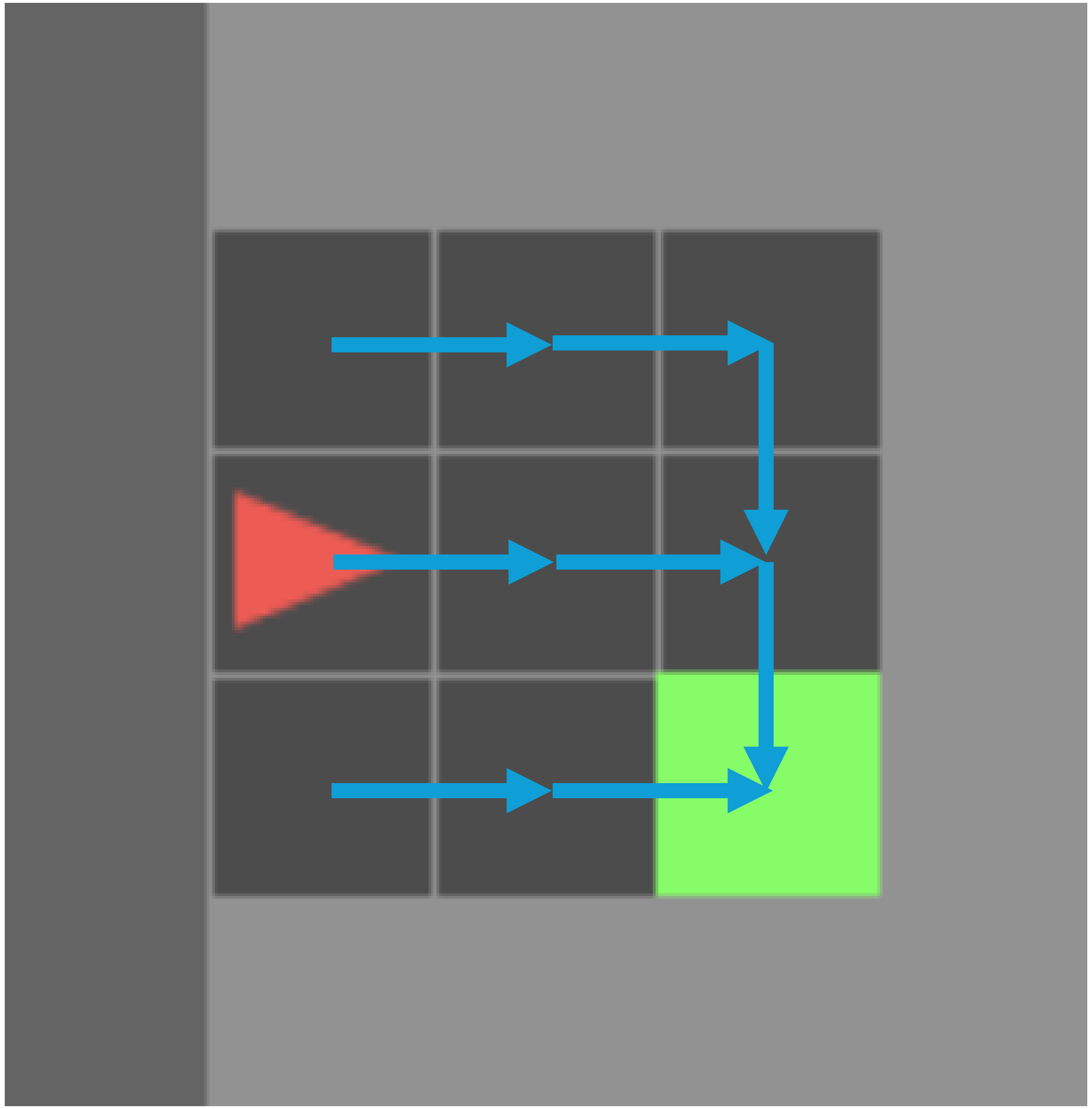


Policies and Tree Stratified Spaces

A policy that routes to the green "goal" state g is equivalent to picking a spanning tree of the dual graph rooted at g.

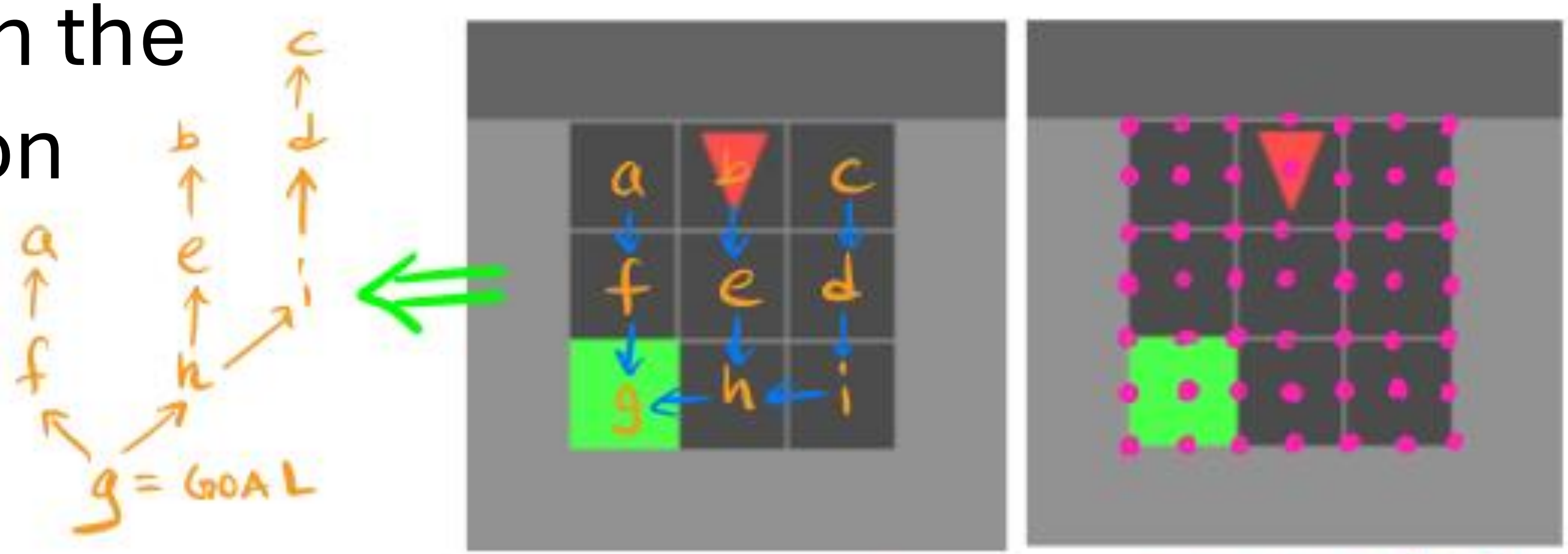
Proposition:

Equipping the cubical grid X with the finite topology where closed cells are a basis for closed sets, then the policy produces a tree-stratified space!



Subtleties in the Construction

Each top dim cell maps to unique node in the policy tree

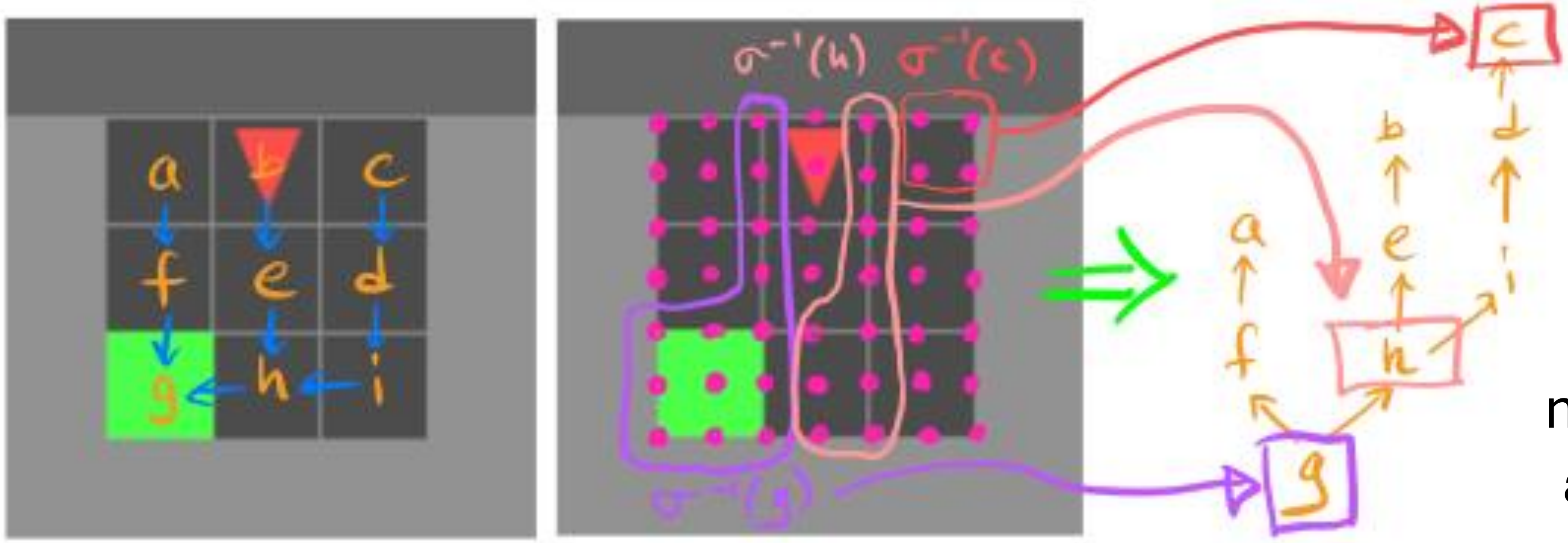


Finite Top. Space

49 pts!

What to do with boundary cells?

Map to meet of the paths in the policy tree!



Fibers of the stratification map can look a little weird!

Stratifying Trajectory Space

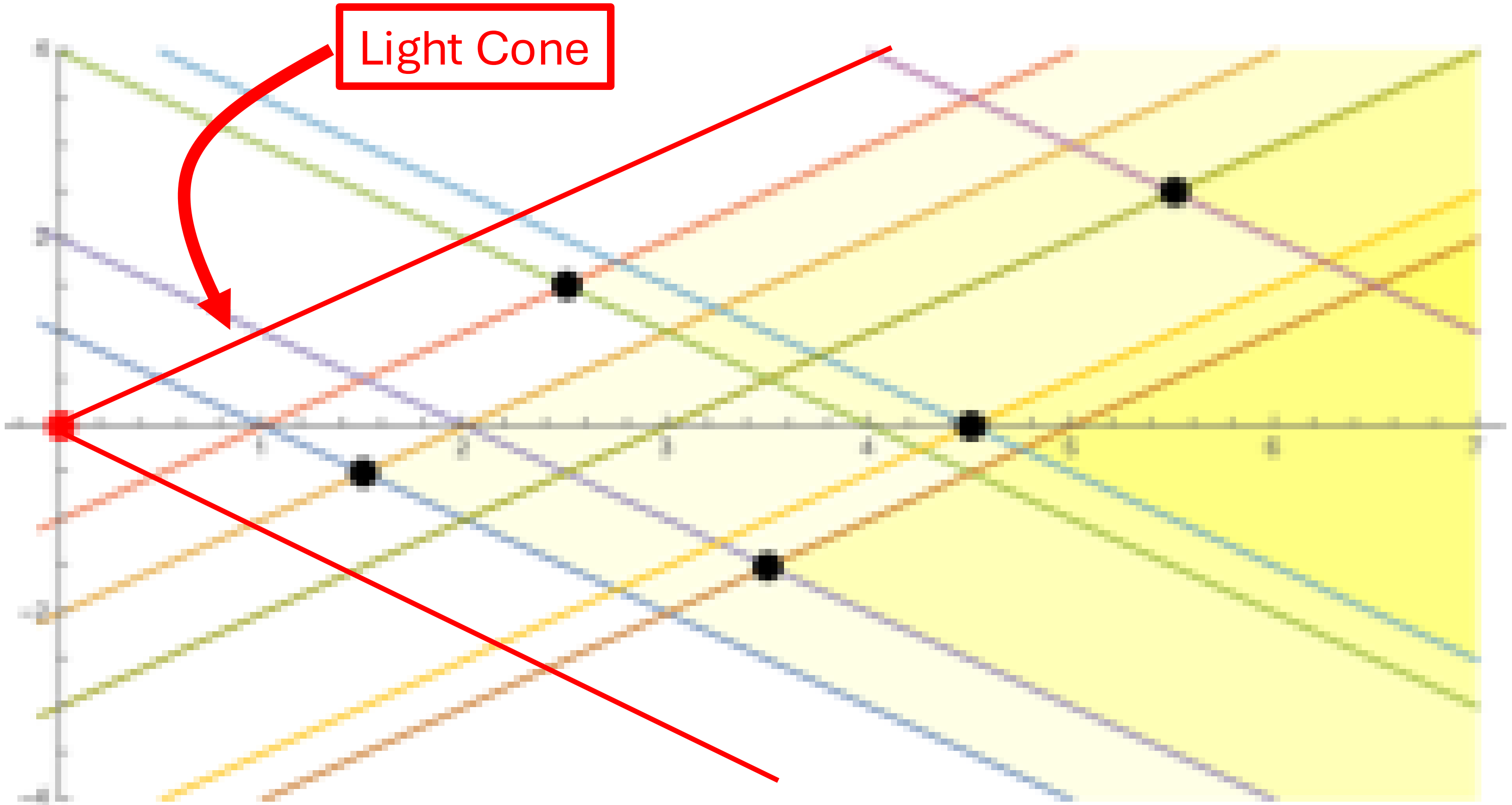
<u>New Game</u>

Start at the red dot

Collect black coins w/o going too fast

How many coins can we collect?

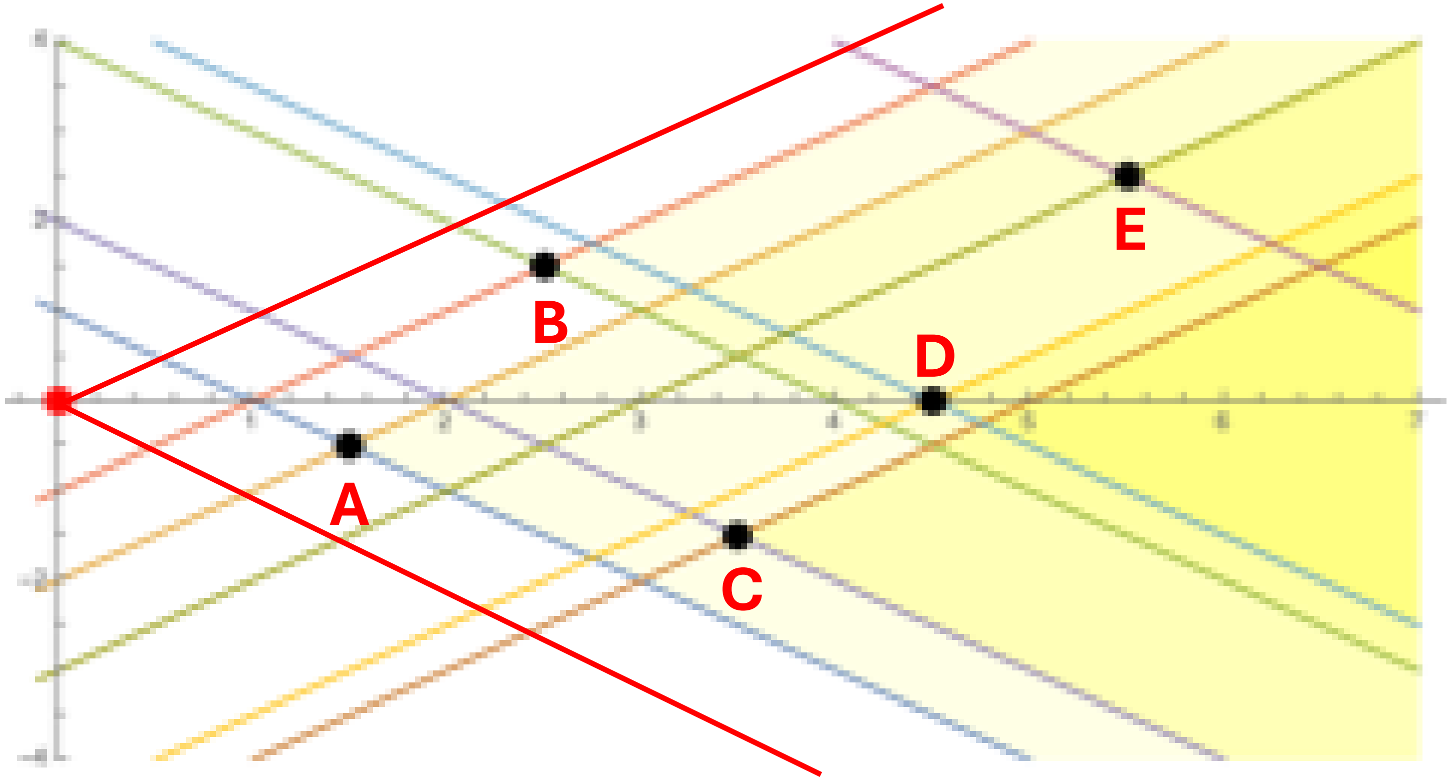
2 coins!



Yuliy Baryshnikov's "Linear Obstacles and How to Avoid Them"

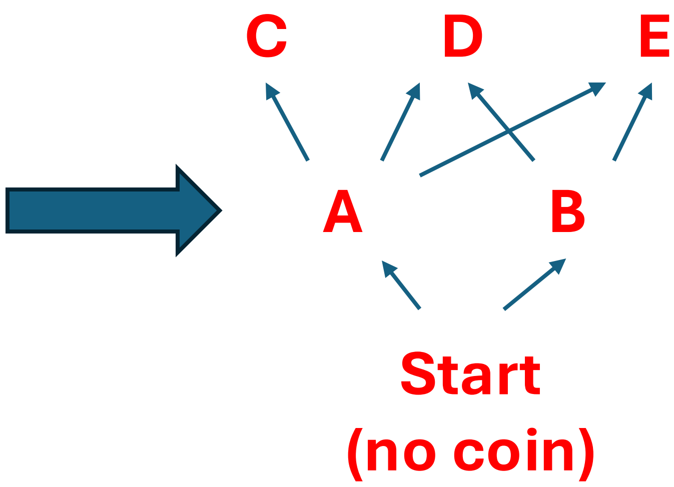
State Space

Yuliy Baryshnikov's "Linear Obstacles and How to Avoid Them"



- State space is space-time $Y \times \mathbb{R}$ where Y is the y-axis.
- Stratifying poset is a sub-poset and reward function is intersection depth of these cones

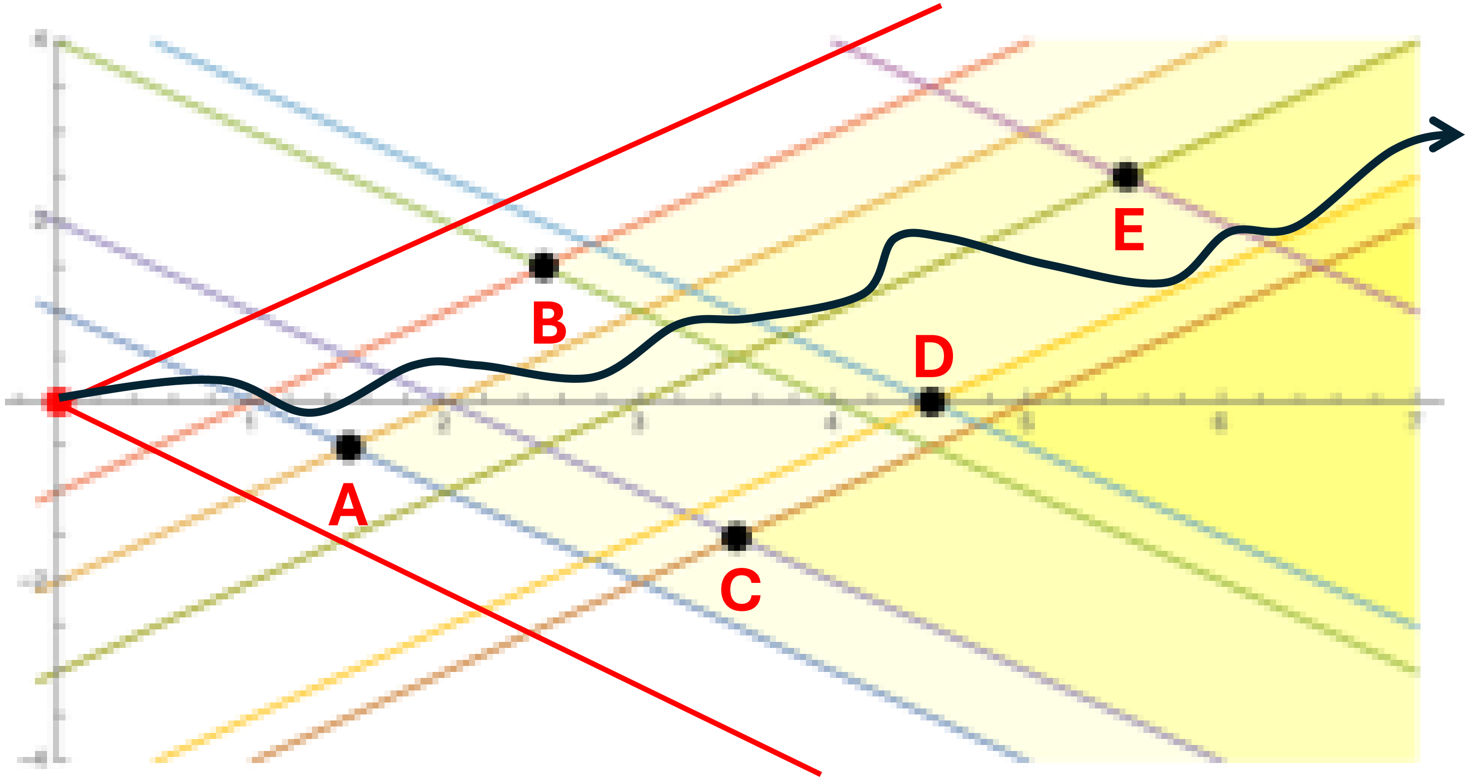
Stratifying Poset



Notice that the maximal chains have length at most 2, indicating the max 2 coins you can collect

Trajectory Space

Yuliy Baryshnikov's "Linear Obstacles and How to Avoid Them"



N.B. The reward function induces this stratification, but lots of different reward functions could induce this same poset. This is a "common core" among them all...

Poset of Chains

2-coin level

SAC SAD SAE SBD SBE

1-coin level

SA SC SD SE SB

0-coin level

S

Length of chain = Reward

This provides a coarse, topological notion of a reward function.

Stratifying Function Space

Lemma: Let $\mathcal{C}(X,Y)$ = space of cts maps from X to Y

If $T \subseteq Y$ is a closed set of "target values"

$$F_{S,T} = \{f \in \mathcal{C}(X,Y) \mid f^{-1}(T) \supseteq S\}$$

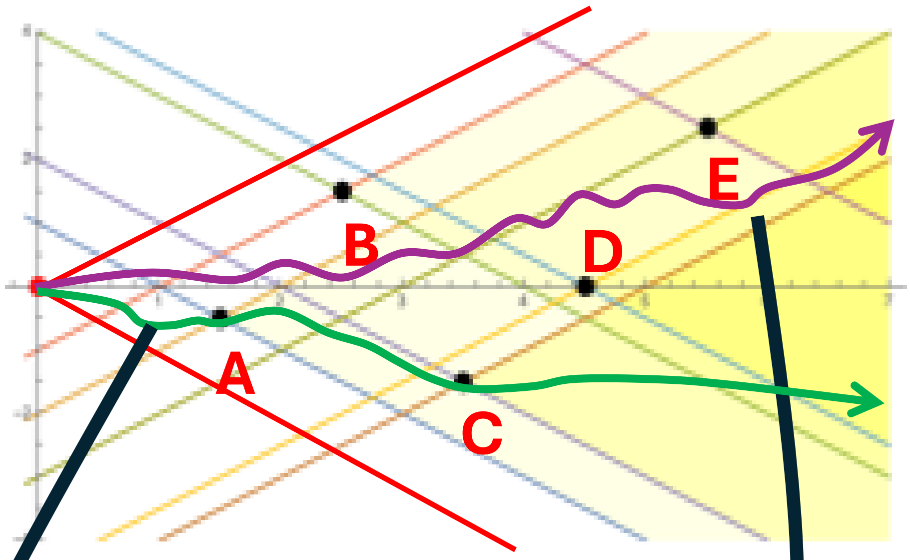
is **closed** for any choice of $S \subseteq X$

Theorem:

Let $\mathcal{P} = Closed(X)$ be the poset of closed sets.
Topologize with down-sets being closed.

$$\sigma_T : \mathcal{C}(X,Y) \to Closed(X)^{op}$$
$$f \longmapsto f^{-1}(T)$$

For any closed $T \subseteq Y$, this is a poset stratification of the space of maps!



$\boxed{T = \{A,B,C,D,E\} \subseteq \mathbb{R}^2}$

Possible $f^{-1}(T)$

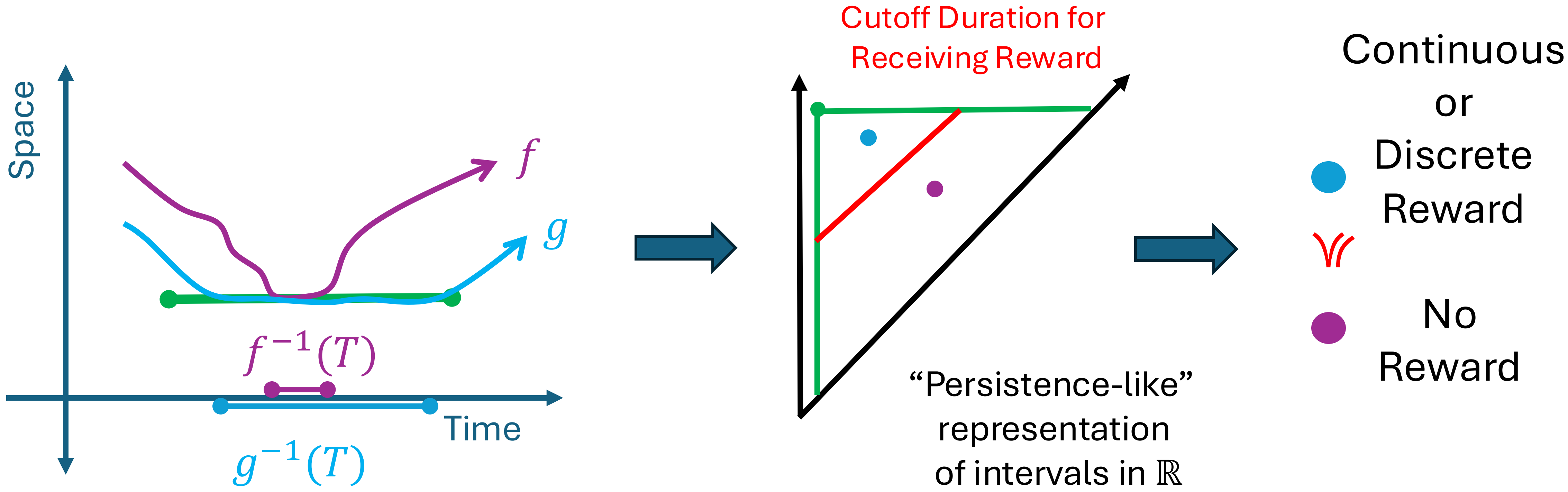
"Generic" Trajectory = No Reward! $\boxed{\emptyset}$

$\boxed{x_A}$ $\boxed{x_C}$ $\boxed{x_D}$ $\boxed{x_E}$ $\boxed{x_B}$

$\boxed{x_A, x_C}$ $\boxed{x_A, x_D}$ $\boxed{x_A, x_E}$ $\boxed{x_B, x_D}$ $\boxed{x_B, x_E}$

"Skillful" Trajectory = Reward!

Stratifying by Time on Target



The Policy Stratification Hypothesis

ALL SKILLFUL BEHAVIOR IS SINGULAR

Journey into the Net

Detecting Stratifications inside Latent Representations

Quick Review of Neural Nets

A neural net is a generalized regression machine

A neural net learns a map $f_\theta\colon \mathbb{R}^m \to \mathbb{R}^n$
that arises as a composition of simpler maps
$$\mathbb{R}^{\ell_0} \to \mathbb{R}^{\ell_1} \to \cdots \to \mathbb{R}^{\ell_{k-1}} \to \mathbb{R}^{\ell_k}$$
where each map has a tunable parameter to “do better”.

The Tokenizer Playground

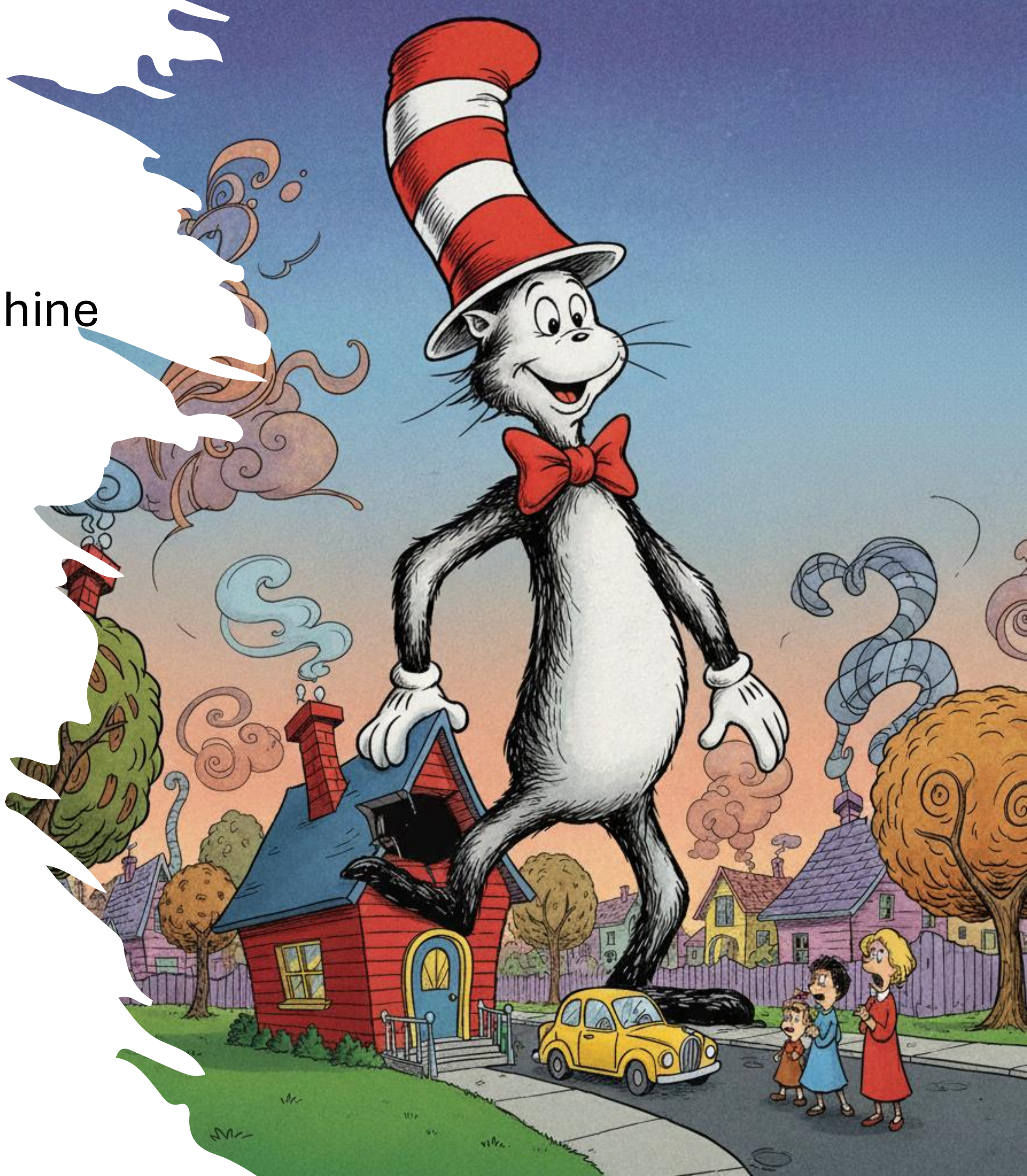
Experiment with different tokenizers (running locally in your browser).

gpt-4 / gpt-3.5-turbo / text-embedding-ada-002

```
The Cat in the Hat grew ten feet tall.
```

TOKENS	CHARACTERS
10	38

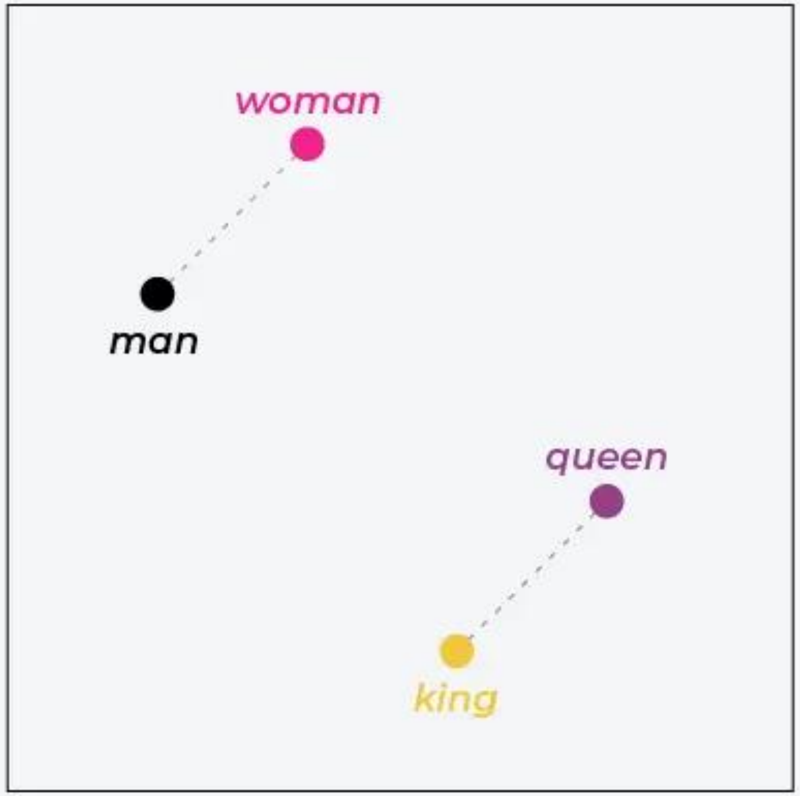
```
The Cat in the Hat grew ten feet tall.
[791, 17810, 304, 279, 22050, 14264, 5899, 7693, 16615, 13]
```



word	living being	feline	human	gender	royalty	verb	plural
man	0.6	-0.2	0.8	0.9	-0.1	-0.9	-0.7
woman	0.7	0.3	0.8	-0.7	0.1	-0.5	-0.4
king	0.5	-0.4	0.7	0.8	0.9	-0.7	-0.6
queen	0.8	-0.1	0.8	-0.9	0.8	-0.5	-0.9

word

Word embedding



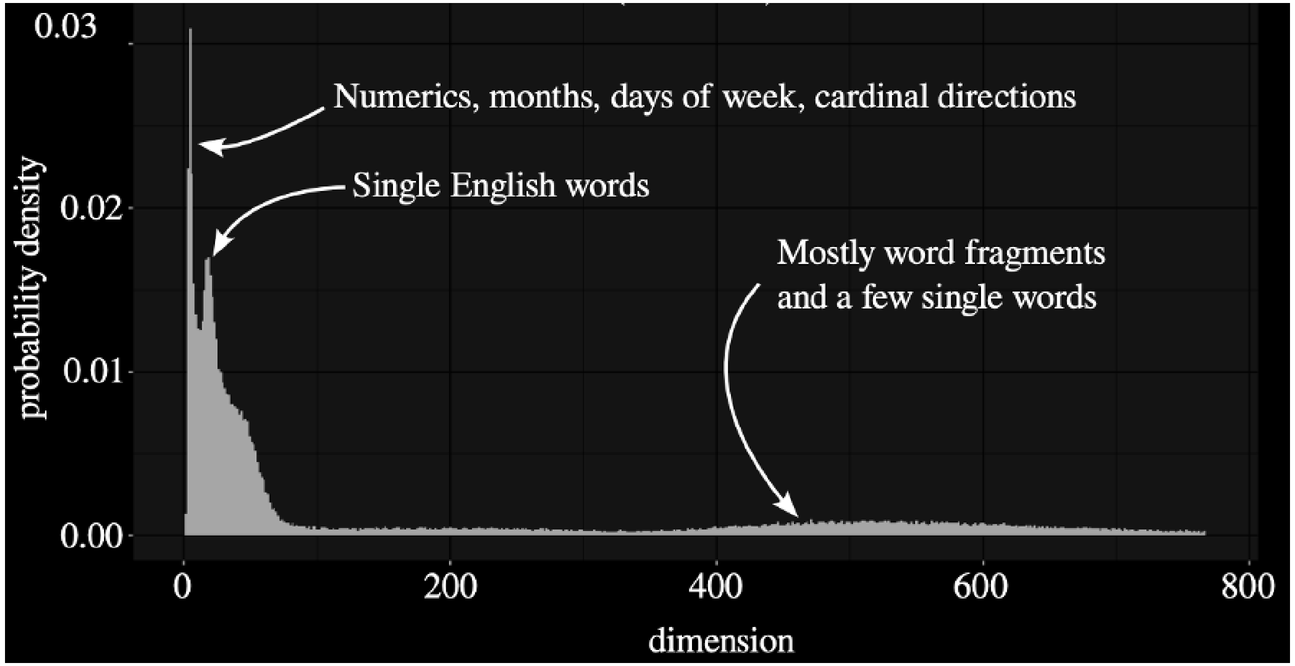
Visualization of word embedding

“Latent Representation”

https://arize.com/blog-course/embeddings-meaning-examples-and-how-to-compute/

Language Models and Stratified Spaces

Contemporaneous with the previous insights, Michael Robinson and others developed completely different arguments for why neural nets are stratified!



[1] *M. Robinson, S. Dey, and S. Sweet. "The structure of the token space for large language models"*, https://arxiv.org/abs/2410.08993, 2024.

[2] *M. Robinson, S. Dey, and T. Chiang. "Token embeddings violate the manifold hypothesis",* https://arxiv.org/abs/2504.01002, 2025.

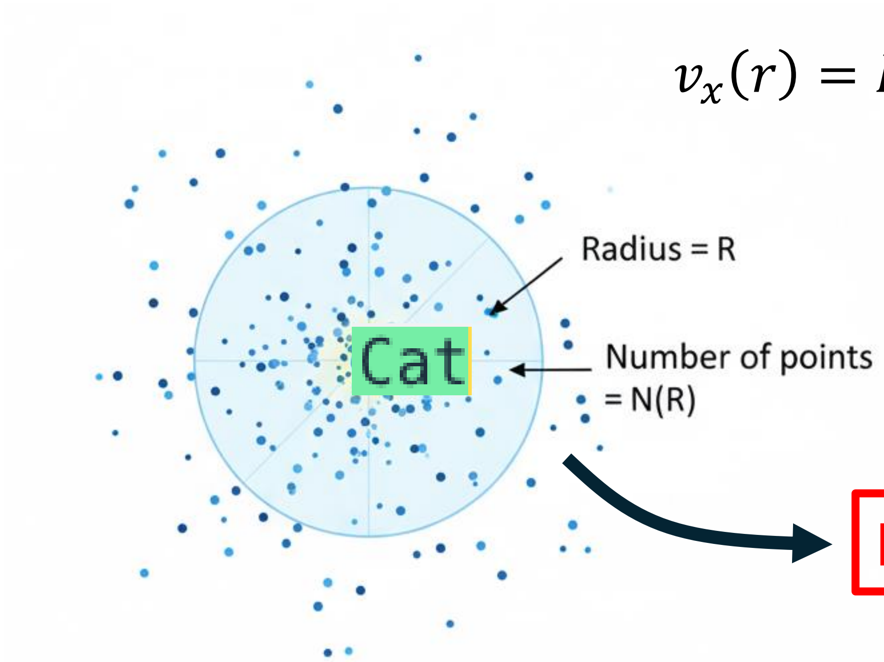
Volume-Radius Relations

Key insight of [1,2] was that investigating GPT-type models are *way* too big for typical TDA methods. Instead, investigate [3], e.g., **volume growth laws**!

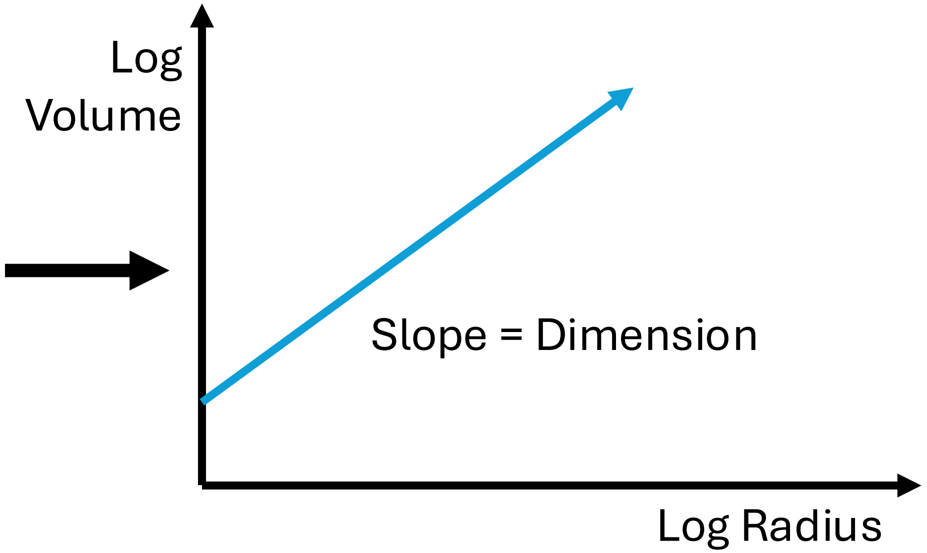
$$v_x(r) = Kr^n\left(1 - \frac{Ric}{6(n+2)}r^2 + O(r^4)\right)$$

Assuming uniform sampling from an n-manifold

$$\log v_x \approx \log K + n \log r + O(r^2)$$

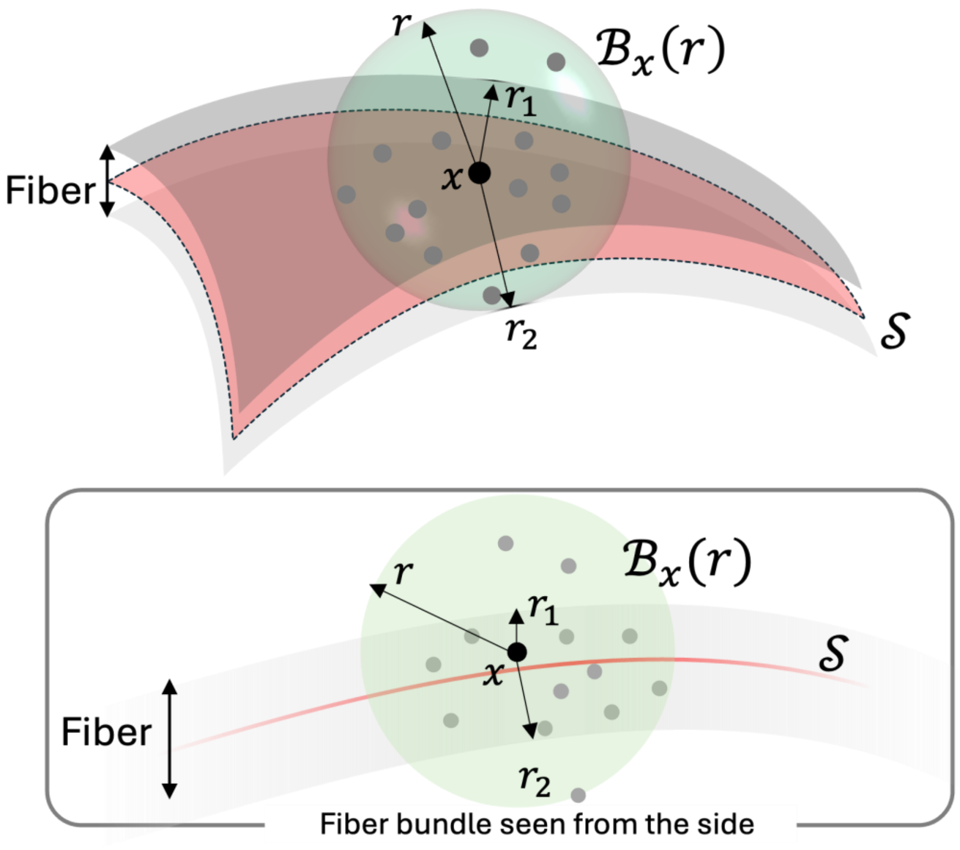


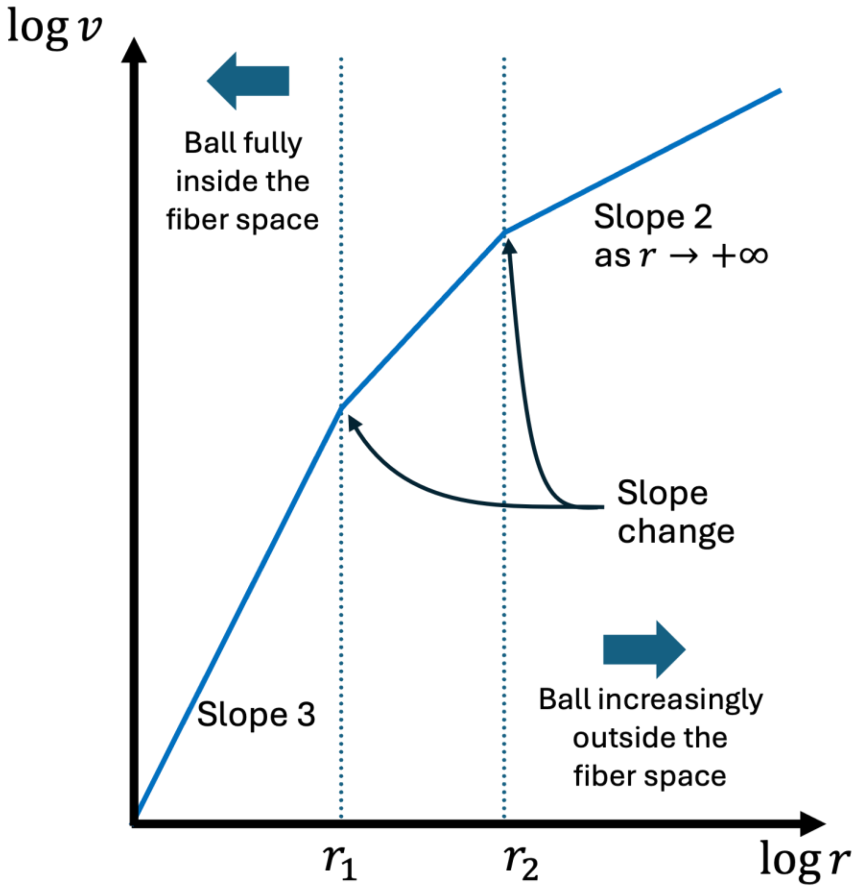
Do Regression!



[3] Alfred Gray. The volume of a small geodesic ball of a Riemannian manifold. Michigan Mathematical Journal, 20(4):329–344, 1974.

Volume Growth Laws for Fiber Bundles

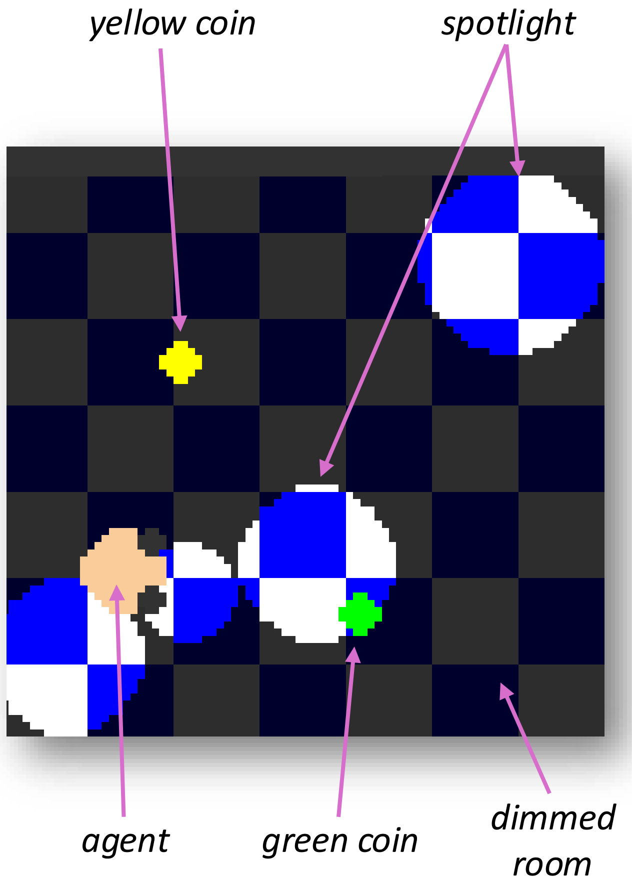


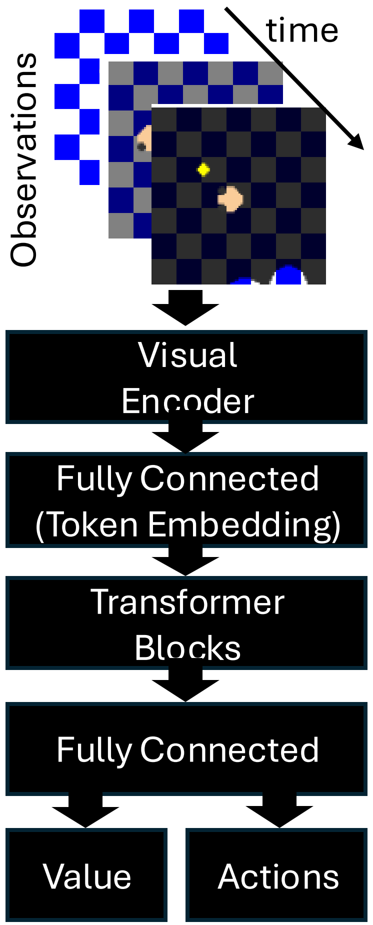


[1] *M. Robinson, S. Dey, and S. Sweet. "The structure of the token space for large language models"*, https://arxiv.org/abs/2410.08993, 2024.
[2] *M. Robinson, S. Dey, and T. Chiang. "Token embeddings violate the manifold hypothesis"*, https://arxiv.org/abs/2504.01002, 2025.

The Searing Spotlight Video Game

- **In our work [4],** we modify the *Searing Spotlight* game from [5]. An agent:
  - ➢ Earns Reward for collecting coins,
  - ➢ Earns Damage if caught in a spotlight.
- Game starts with the lights on then fades to black.
- Spotlights are introduced gradually, but move randomly and with increasing speed and radius.

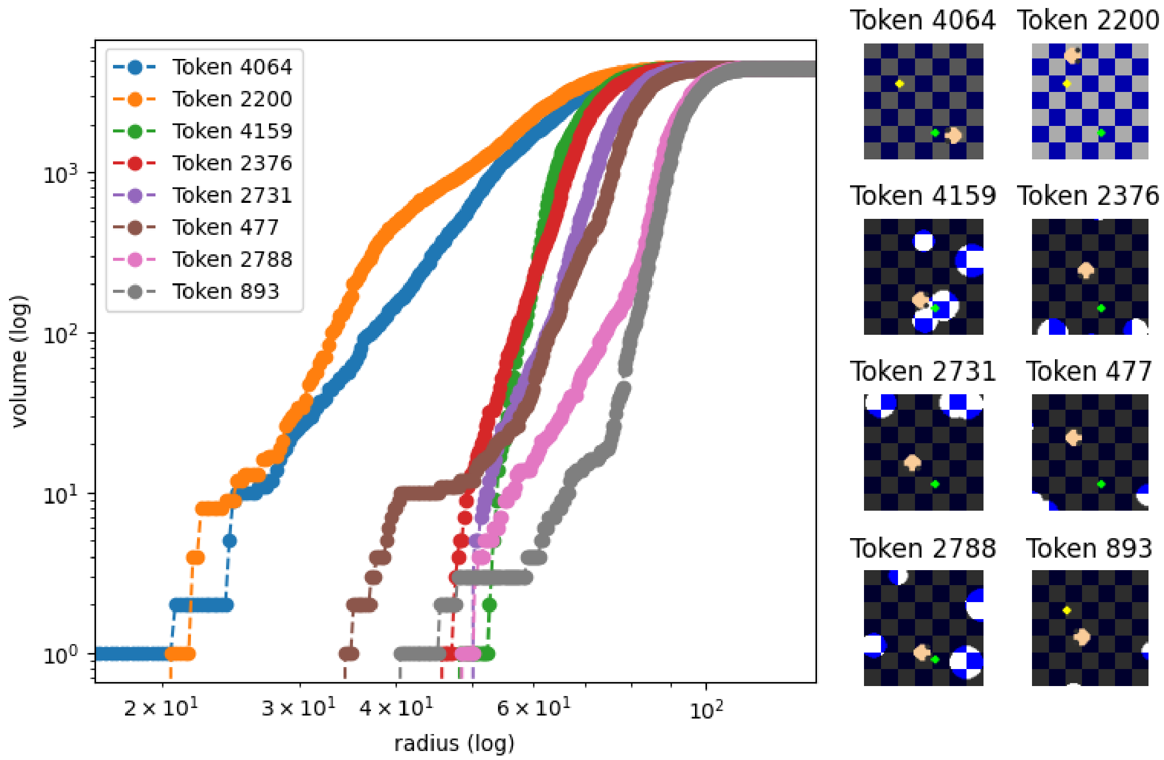


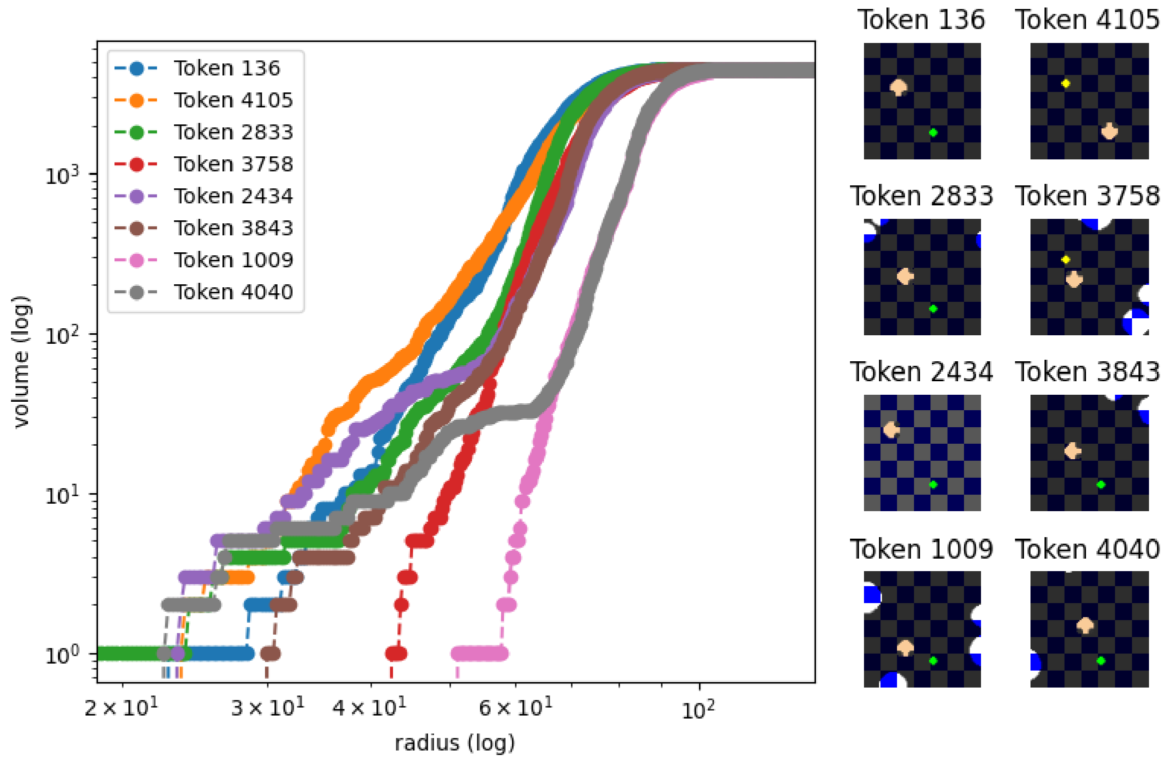


[4] *J. Curry, B. Lagasse, N.B. Lam, G. Cox, D. Rosenbluth, and A. Speranzon.* ***"Exploring the Stratified Space Structure of an RL Game with the Volume Growth Transform",*** https://www.arxiv.org/pdf/2507.22010, 2025.

[5] *M. Pleines, M. Pallasch, F. Zimmer and M. Preuss. "Memory Gym: Towards Endless Tasks to Benchmark Memory Capabilities of Agents",* https://arxiv.org/abs/2504.01002, 2025.

Token Embeddings and Volume Growth Laws





[4] *J. Curry, B. Lagasse, N.B. Lam, G. Cox, D. Rosenbluth, and A. Speranzon.* ***"Exploring the Stratified Space Structure of an RL Game with the Volume Growth Transform"*** https://www.arxiv.org/pdf/2507.22010, 2025.

Realization Result

Can a stratified space exhibit volume growth laws with sharp increases?

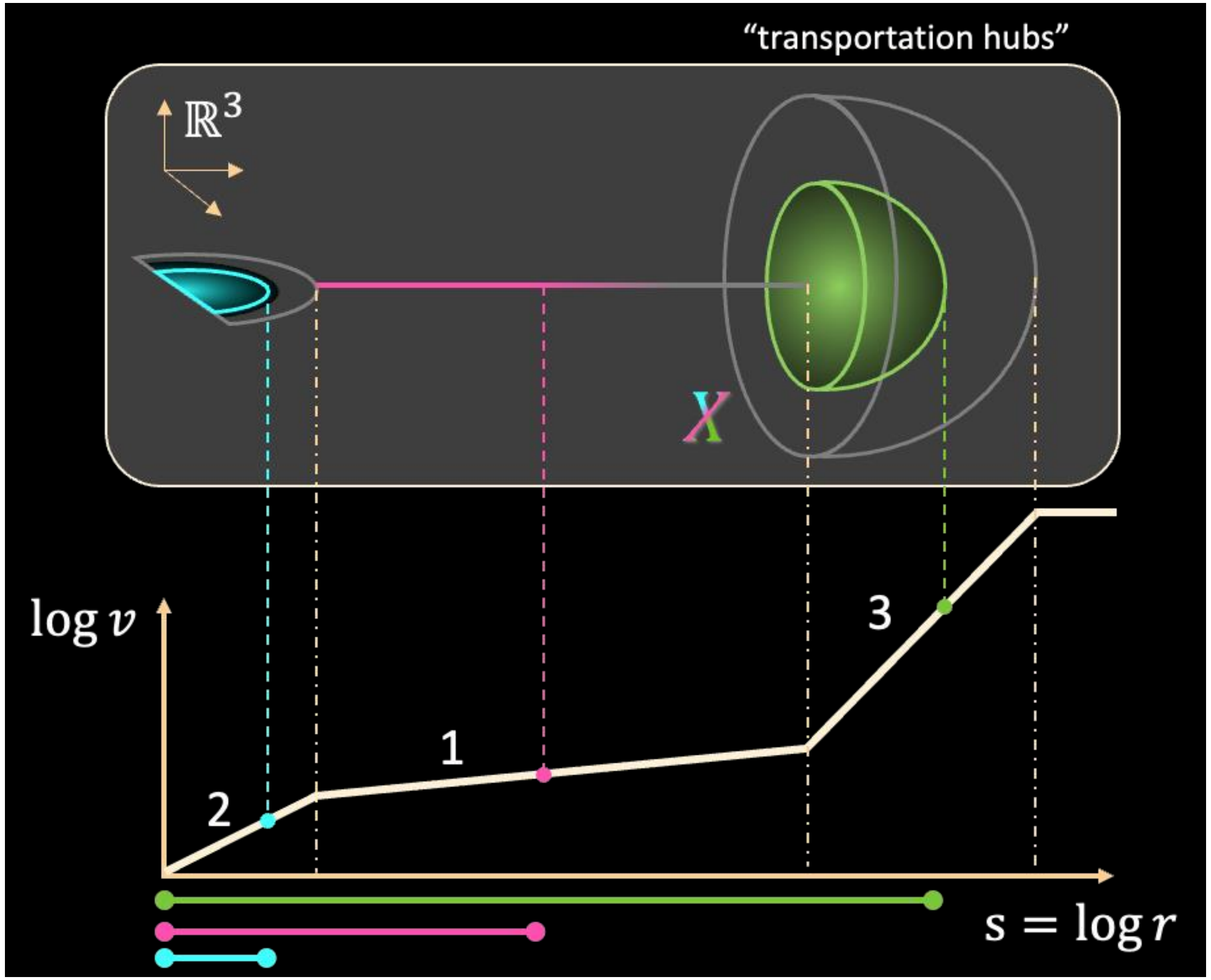
Theorem:

If $f: [0, \infty) \to [0, \infty)$ is a non-decreasing piecewise linear function such that:

- $f|_{[0,s_1]} = n_1 s$
- $f|_{[s_i,s_{i+1}]} = n_{i+1} s + f(s_i), \quad i \in \{1, \dots, k-1\}$
- $f|_{[s_k,\infty)} = f(s_k)$ is a constant

for some "critical scales" $\{s_1 < \cdots < s_k\}$ and natural numbers $n_1, \dots, n_k$ then there exists a *stratified space* X with a point $x \in X$ such that

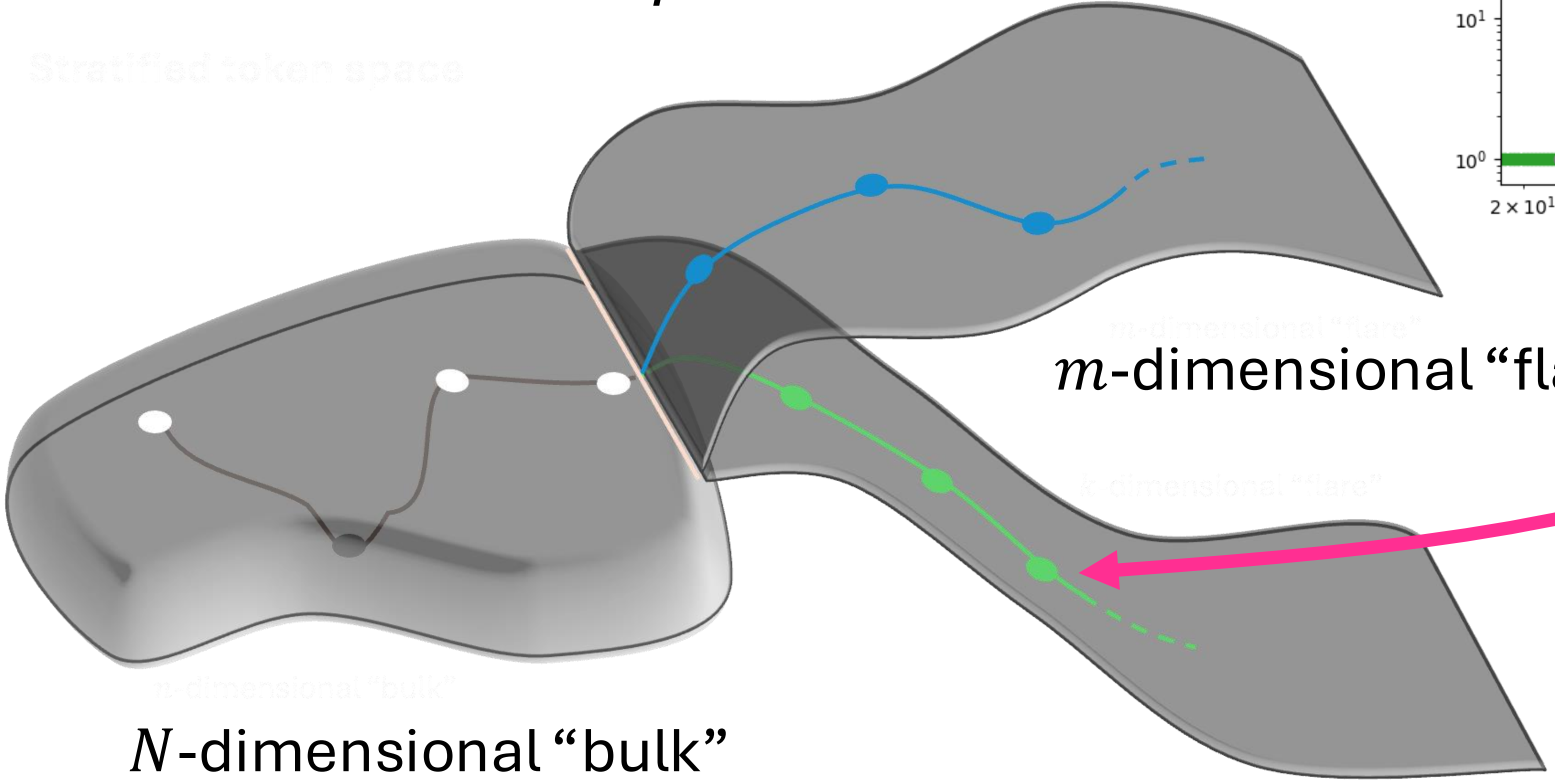
$$\text{Volume Growth Transform}(x) = f(s)$$



[4] *J. Curry, B. Lagasse, N.B. Lam, G. Cox, D. Rosenbluth, and A. Speranzon.* ***"Exploring the Stratified Space Structure of an RL Game with the Volume Growth Transform"*** https://www.arxiv.org/pdf/2507.22010, 2025.

Interpretation of the Data

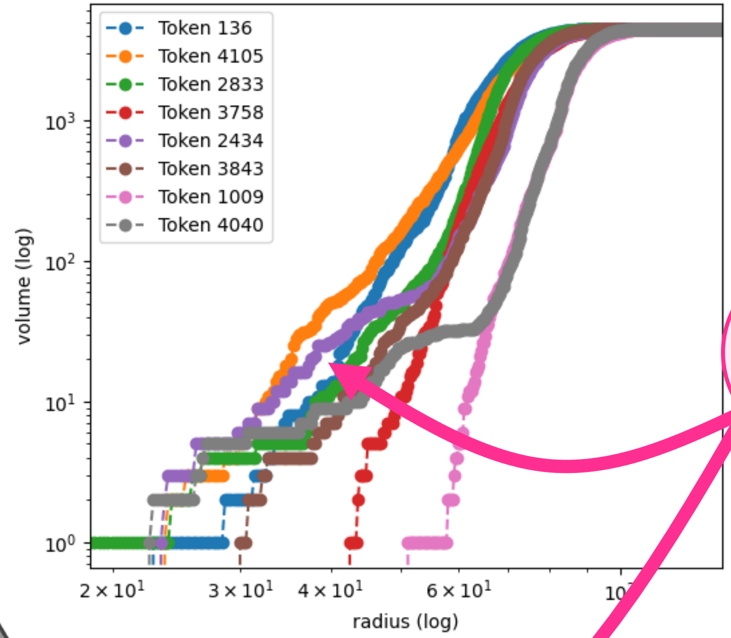
A Stratified Token Space



m-dimensional "flare"

N-dimensional "bulk"

k-dimensional "flare"

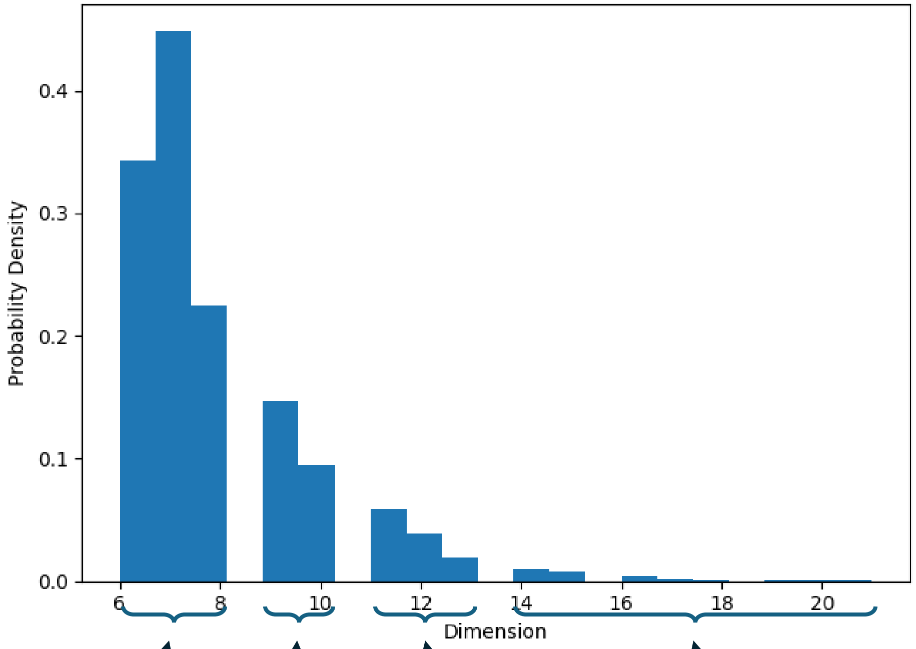


[4] *J. Curry, B. Lagasse, N.B. Lam, G. Cox, D. Rosenbluth, and A. Speranzon.* ***"Exploring the Stratified Space Structure of an RL Game with the Volume Growth Transform"*** https://www.arxiv.org/pdf/2507.22010, 2025.

Distribution of Local Dimensions

State Space Model:
- 2D for position
- 1D for rotation
- 3D color channel

= 6D State Space!



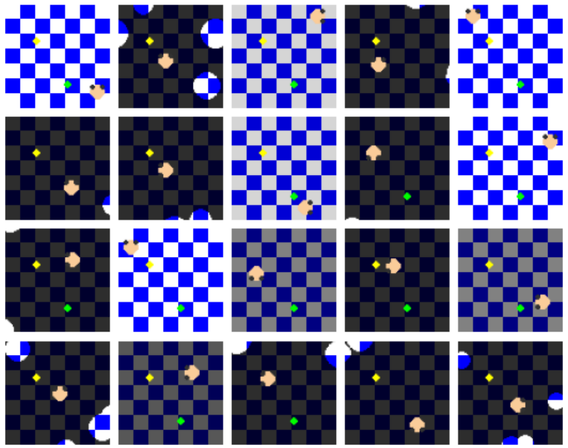
State Space Model:
- 2D for position
- 1D for rotation
- 3D color channel

= 6D State Space

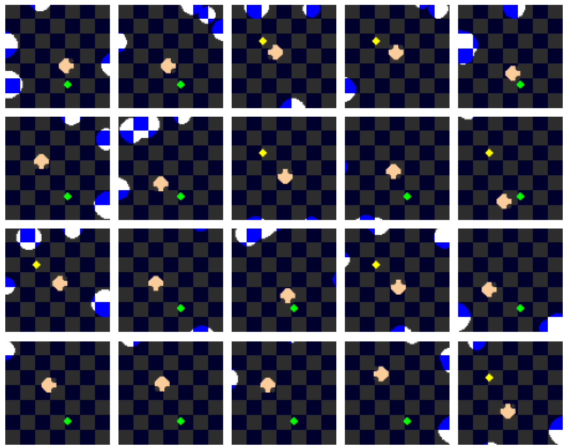
+ 4 Spotlights w/ 2D per spotlight

= 14D State Space!

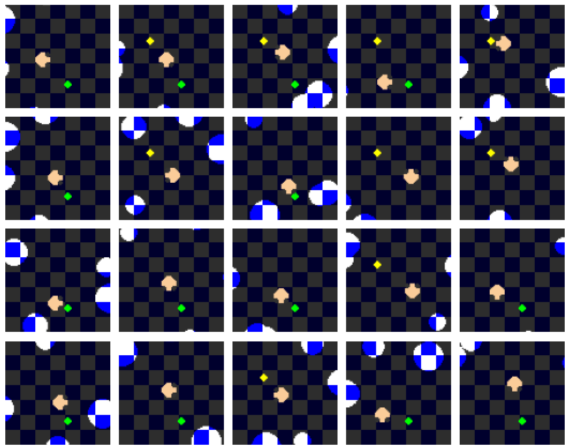
$dim \in [6, 8]$



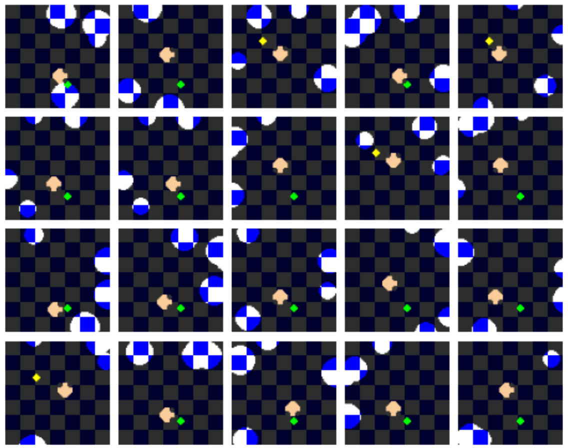
$dim \in [9, 10]$



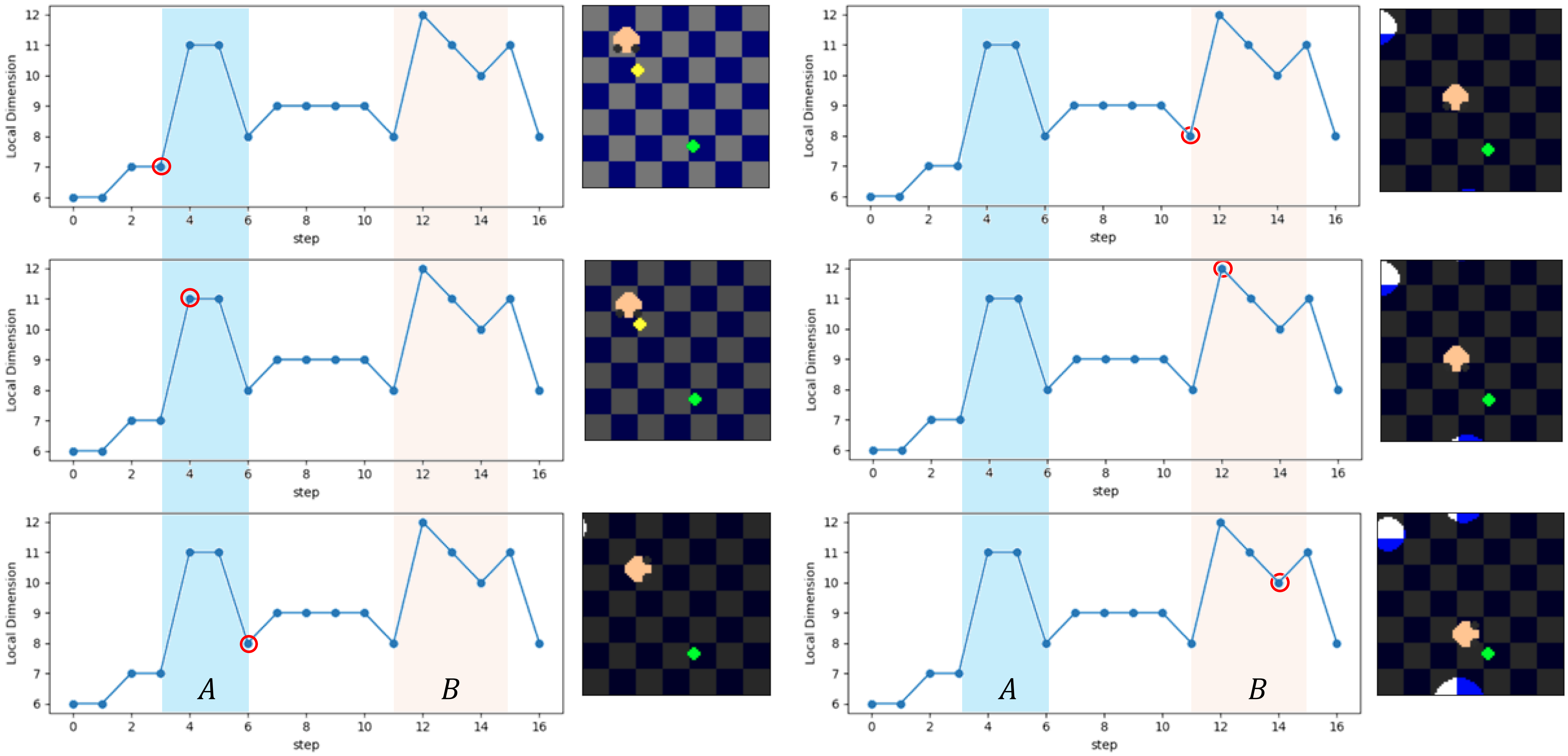
$dim \in [11, 13]$

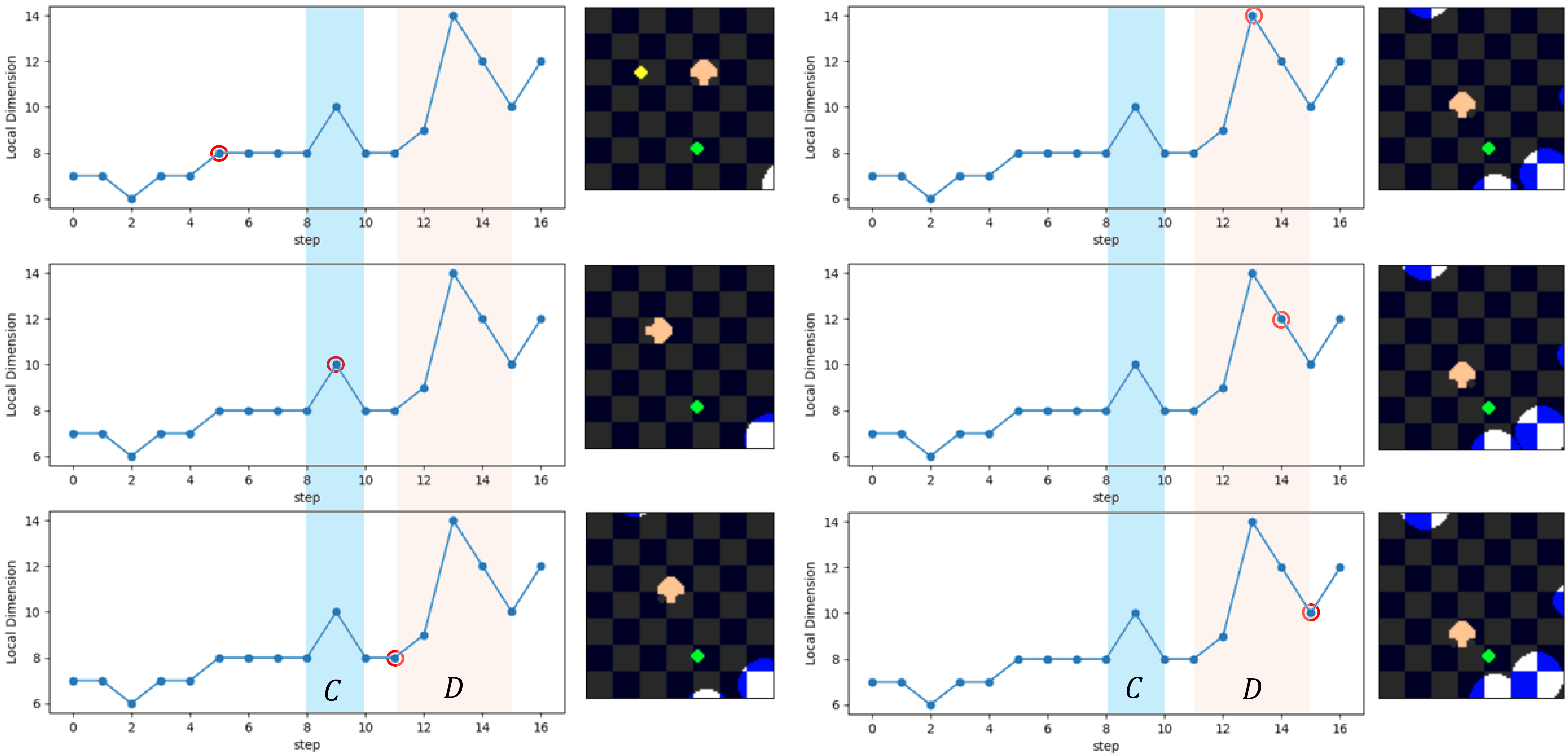


$dim \in [14, 21]$



20 randomly sampled tokens/observations from each “cluster”


Local Dimension
step
A
B



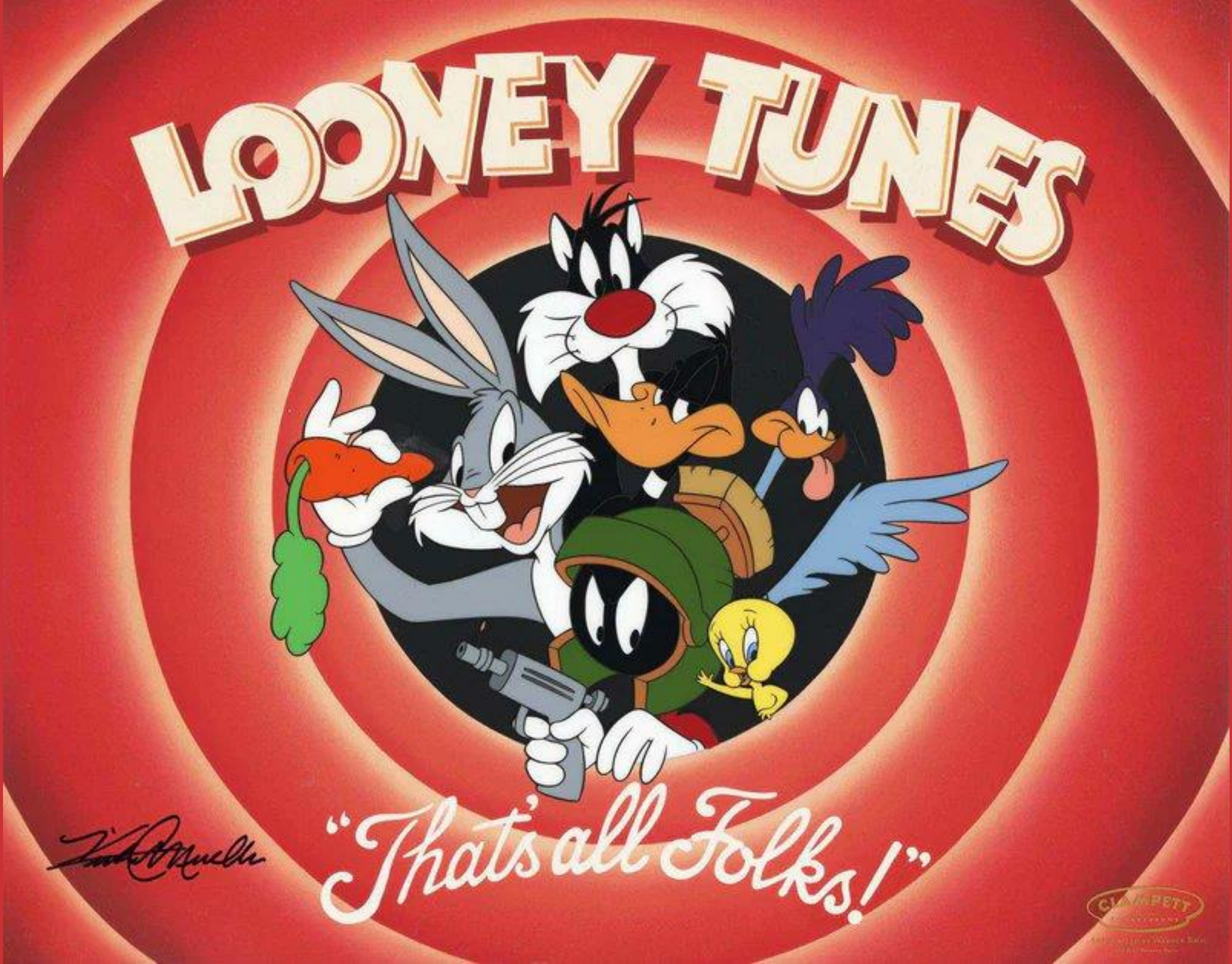
Local Dimension
step
C
D

Final Thoughts and Future Directions

- Why does the dimension spike near goal states?
  - ➢Over-Sampling?
  - ➢Complex decision making?
  - ➢Picking up *codimension* via a Steiner Polynomial?

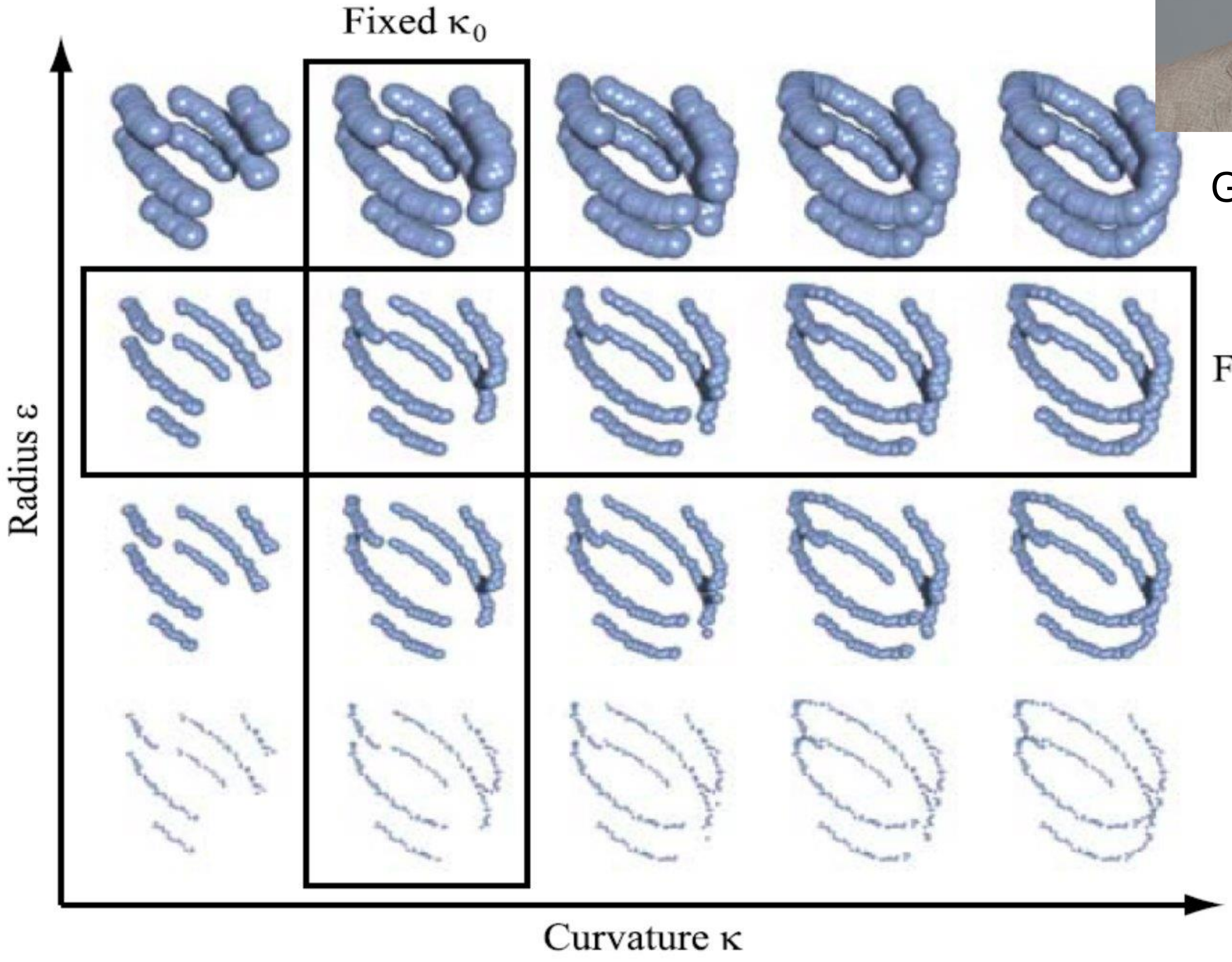
$$\mathcal{H}^d(X^r) =: \sum_{i=0}^{d} \omega_i V_{d-i}(X) r^i$$

- How does histogram of dimensions evolve during training?
  - ➢Novel measure of learning!
- Reward Machines?
  - ➢Reward hacking
  - ➢Stratification guides reward initialization
  - ➢Moduli space of equivalent reward functions


LOONEY TUNES
"That's all Folks!"

“Bizarre” Continuous Stratifications

- The Morse property is not necessary
- Every continuous function $f: M \to \mathbb{R}$ is automatically Alexandrov continuous
- Much of persistence can be thought of as stratification theory



G. Carlsson

Stratification Theory for the TDA Community

Discrete & Computational Geometry
https://doi.org/10.1007/s00454-020-00206-y

Sheaf-Theoretic Stratification Learning from Geometric and Topological Perspectives

Adam Brown[1,2] · Bei Wang[3]

Received: 21 December 2018 / Revised: 11 February 2020 / Accepted: 31 March 2020

Abstract
We investigate a sheaf-theoretic interpretation of stratification learning from geometric and topological perspectives. Our main result is the construction of stratification learning algorithms framed in terms of a sheaf on a partially ordered set with the Alexandroff topology. We prove that the resulting decomposition is the unique minimal stratifi-

Foundations of Computational Mathematics (2020) 20:195–222
https://doi.org/10.1007/s10208-019-09424-0

Local Cohomology and Stratification

Vidit Nanda[1]

Received: 14 August 2017 / Revised: 9 March 2019 / Accepted: 29 April 2019 / Published online: 13 June 2019

Abstract
We outline an algorithm to recover the canonical (or, coarsest) stratification of a given finite-dimensional regular CW complex into cohomology manifolds, each of which is a union of cells. The construction proceeds by iteratively localizing the poset of cells about a family of subposets; these subposets are in turn determined by a collection of cosheaves which capture variations in cohomology of cellular neighborhoods across

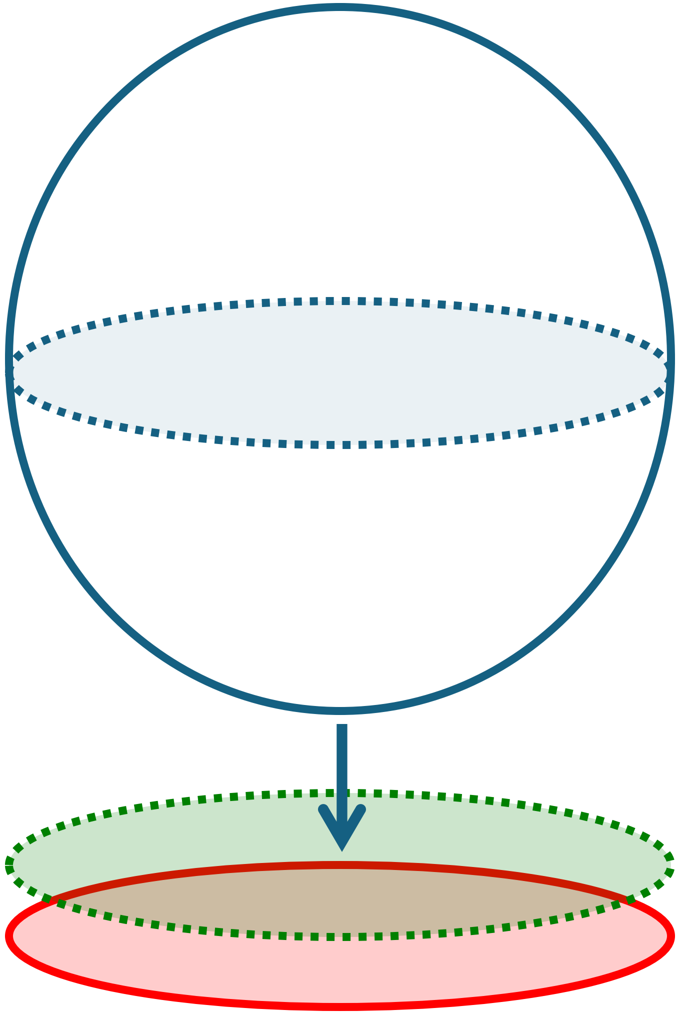
M. Lesnick, J. Curry, R. MacPherson, M. Goresky, A. Patel, G. Henselman, B. Fasy

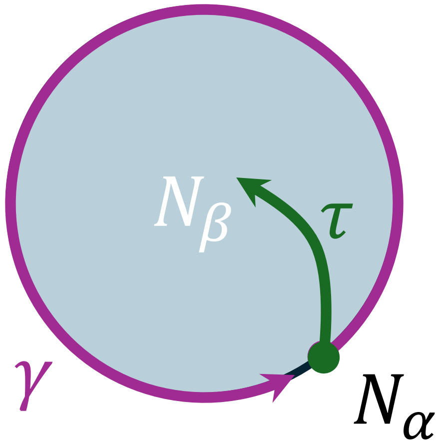
Posets as Shadows of Categories

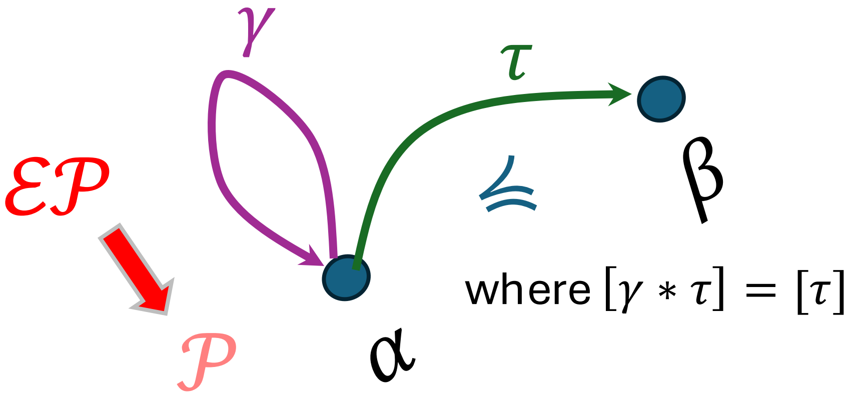
The **Exit Path Category** has points for objects and homotopy classes of exit paths for morphisms; paths that “exit” a stratum for a higher dimensional one.

There is a functor $\mathcal{EP} \to \mathcal{P}$ that forgets homotopy information

R. MacPherson

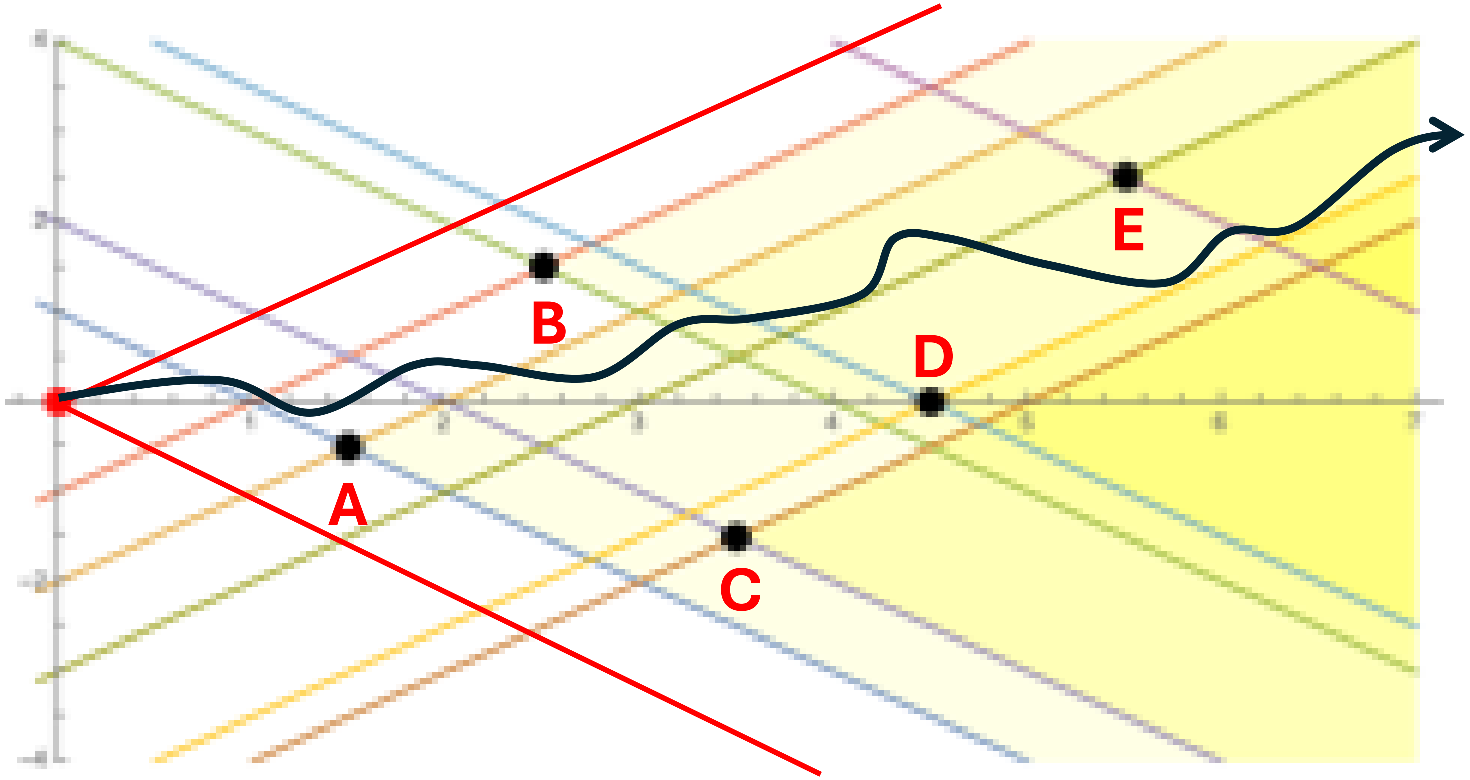




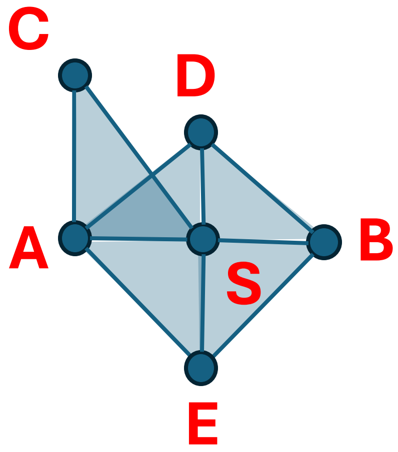


Trajectory Space

Yuliy Baryshnikov's "Linear Obstacles and How to Avoid Them"



Order Complex



N.B. Really need a directed order complex here...

Big Idea: How do we find signatures of these structures inside the latent space of neural networks?