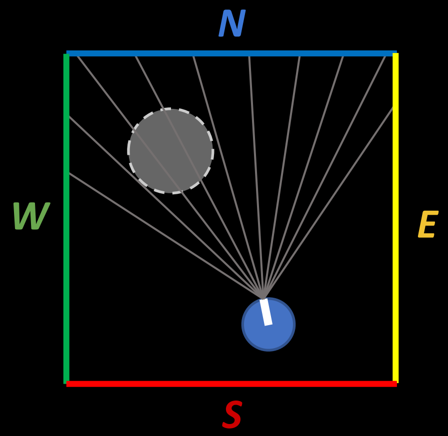
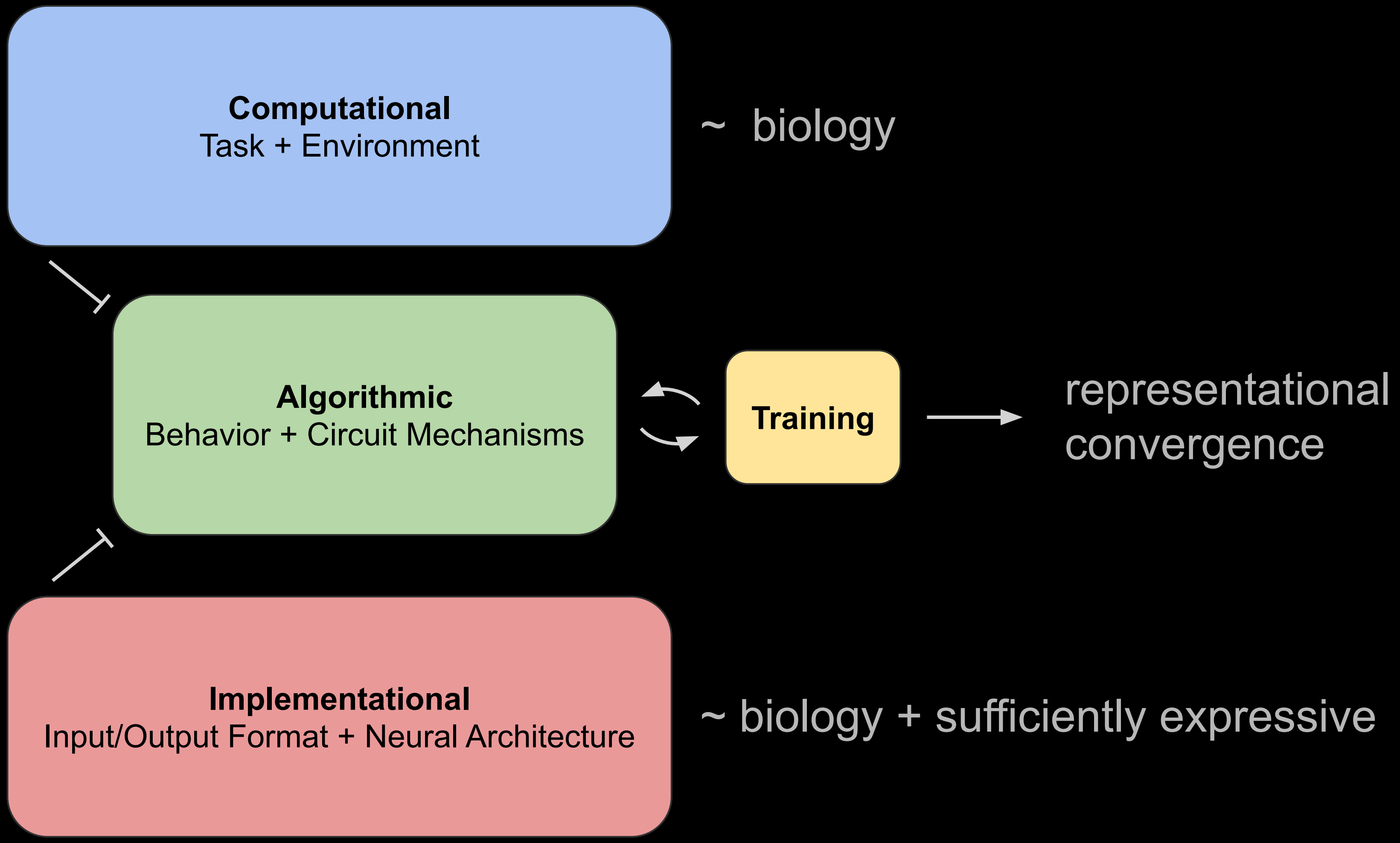
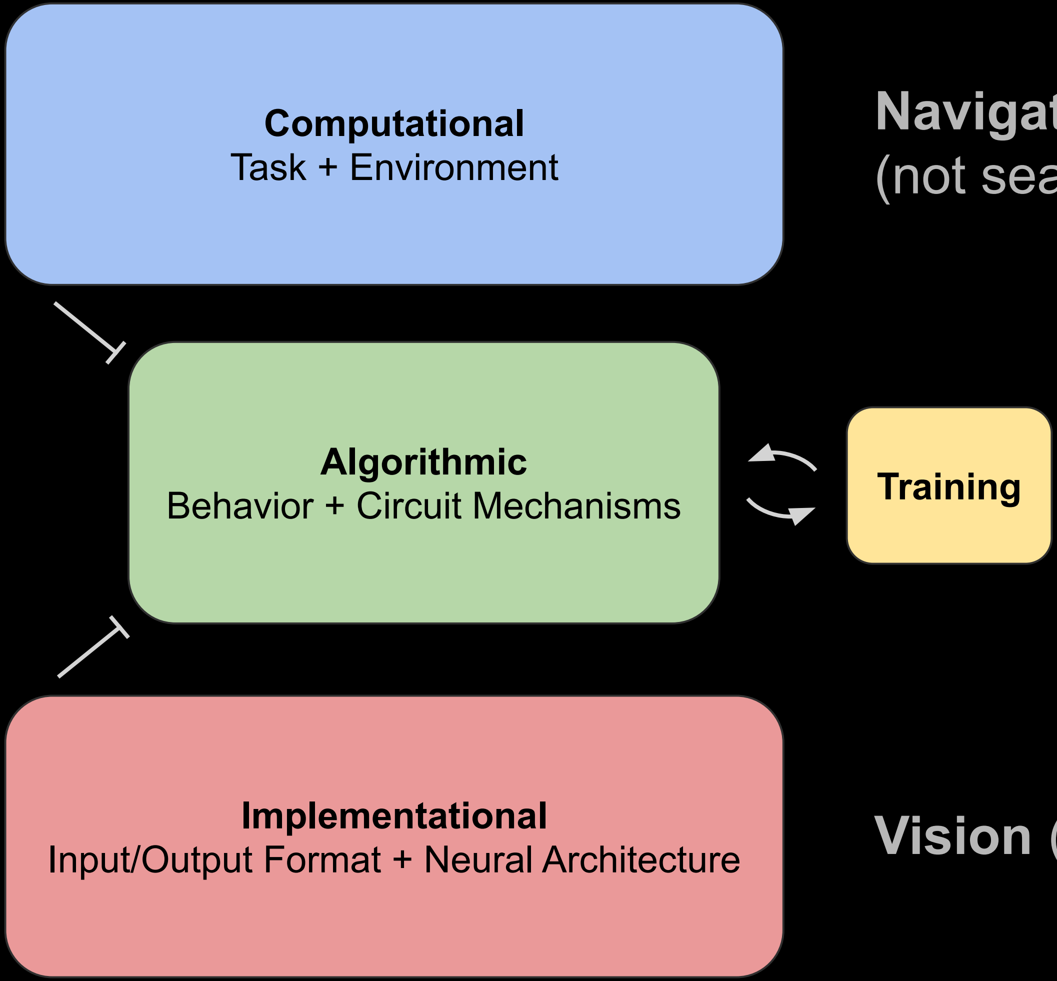
Fundamental Visual Navigation Algorithms: Indirect Sequential, Biased Diffusive, & Direct Pathing



Patrick Govoni
Romanczuk Lab
Humboldt-Universität zu Berlin


Computational
Task + Environment
~ biology
Algorithmic
Behavior + Circuit Mechanisms
Training
representational
convergence
Implementational
Input/Output Format + Neural Architecture
~ biology + sufficiently expressive



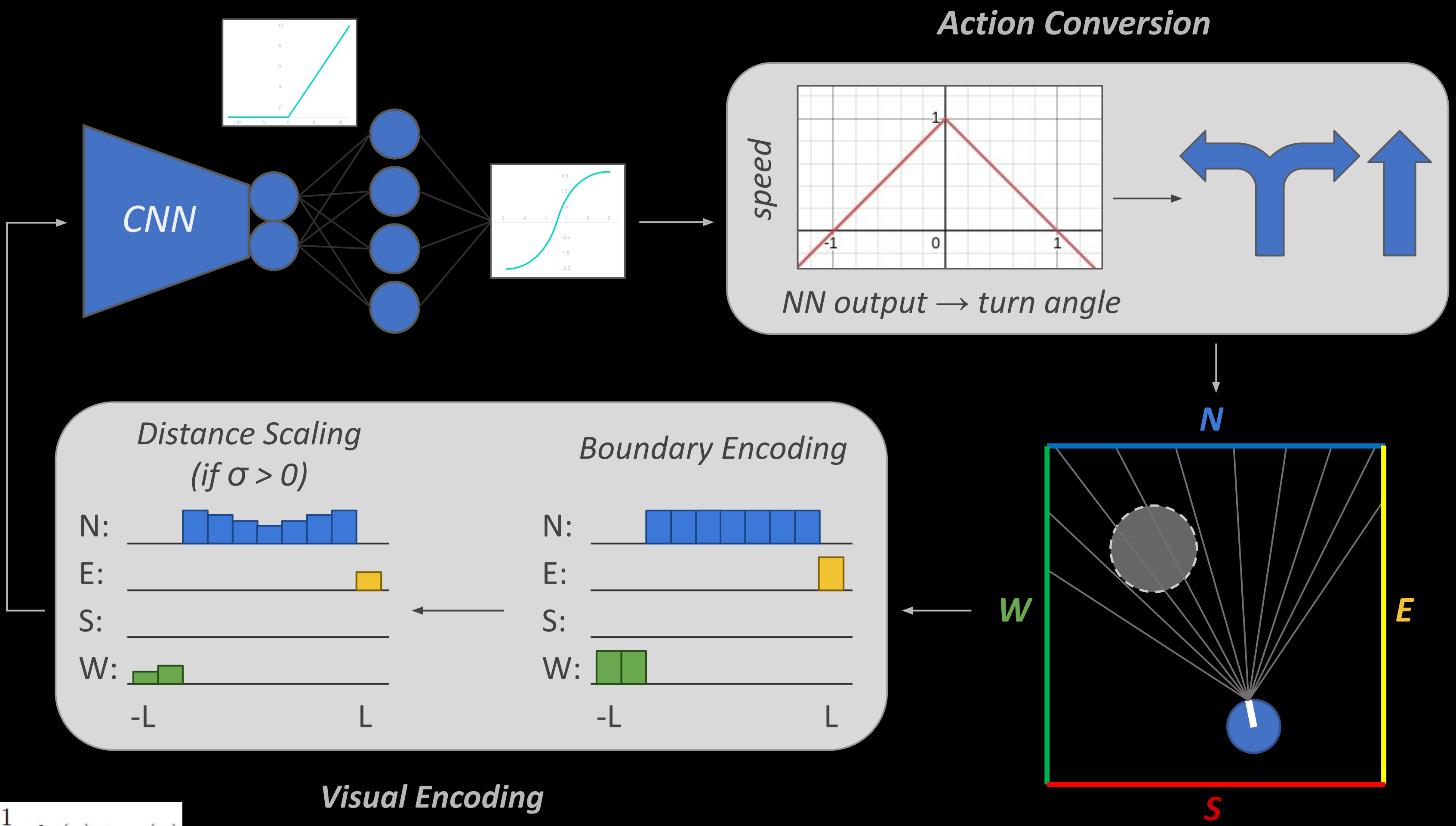
Navigation to a known destination (not search)

min(assumptions)

↓

max(generalizability)

Vision (CNN)

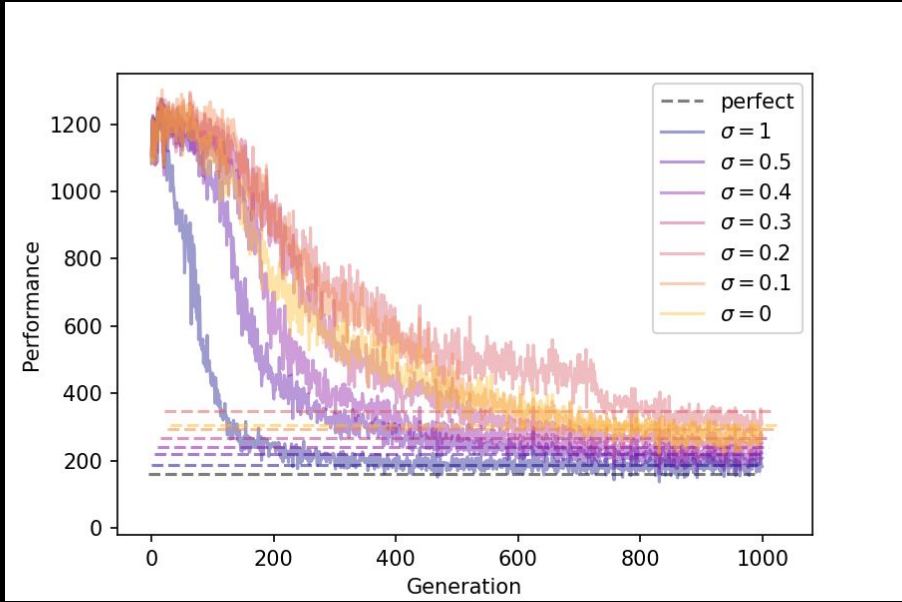


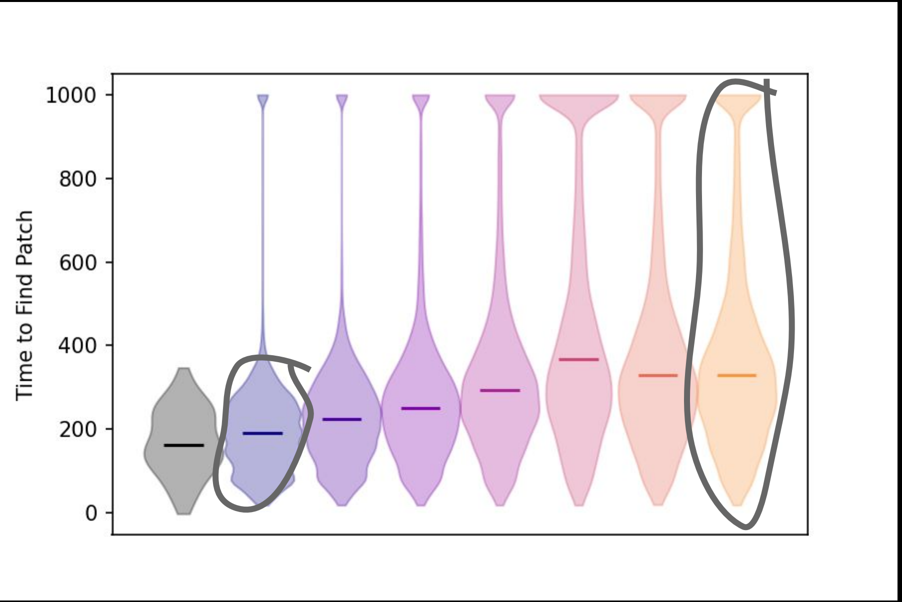
$$y = -\frac{1}{k(\sigma)}\ln(x) + m(\sigma)$$

Trained via Evolutionary Strategies

- Performance =
  time to patch +
  remaining distance

visual input =
angles (σ = 0) → learns task
+ distance (σ > 0) → near perfect





Local behavior
Angles Only

Navigational
Class
1

“Indirect
Sequential”

Local behavior

Angles Only

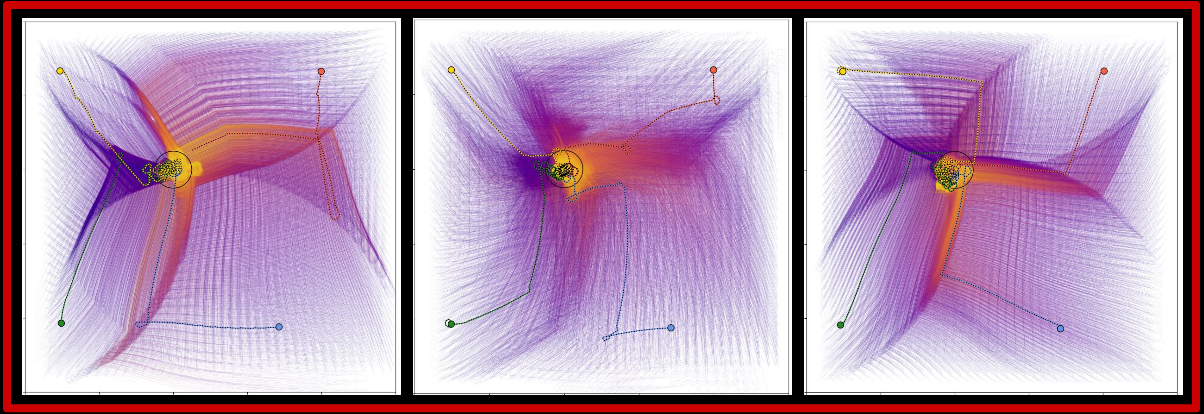
Navigational

Class

2

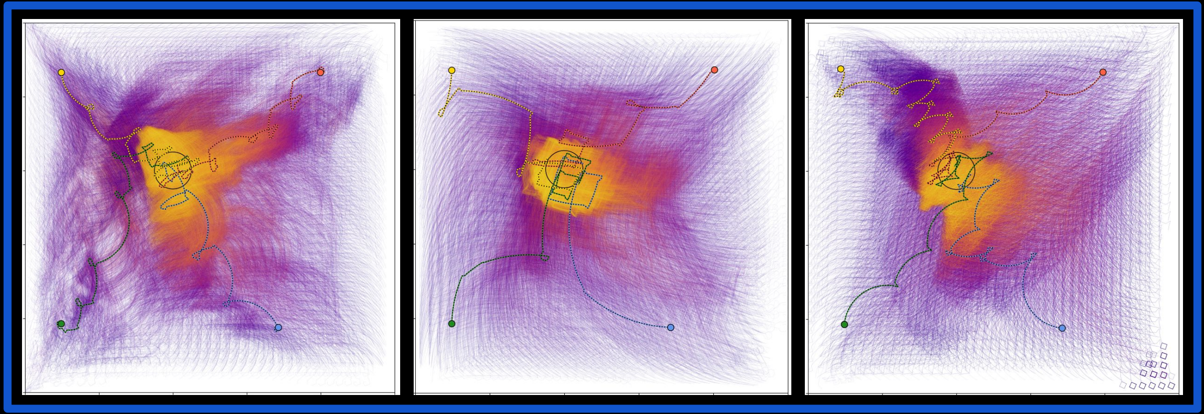
“Biased Diffusive”

Global behavior
Angles Only



Indirect Sequential

- Grid-like trajectories
- Compositional route segments
- Elliptical decision manifolds



Biased Diffusive

- Spin/ratchet-like trajectories
- Directed towards patch

Local behavior

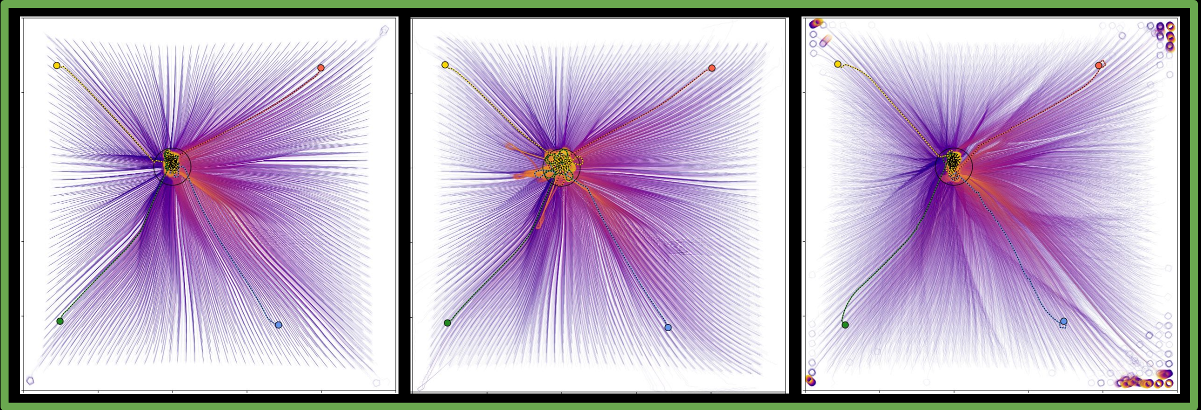
Angles + Distance

Navigational
Class
3

“Direct Pathing”

Global behavior

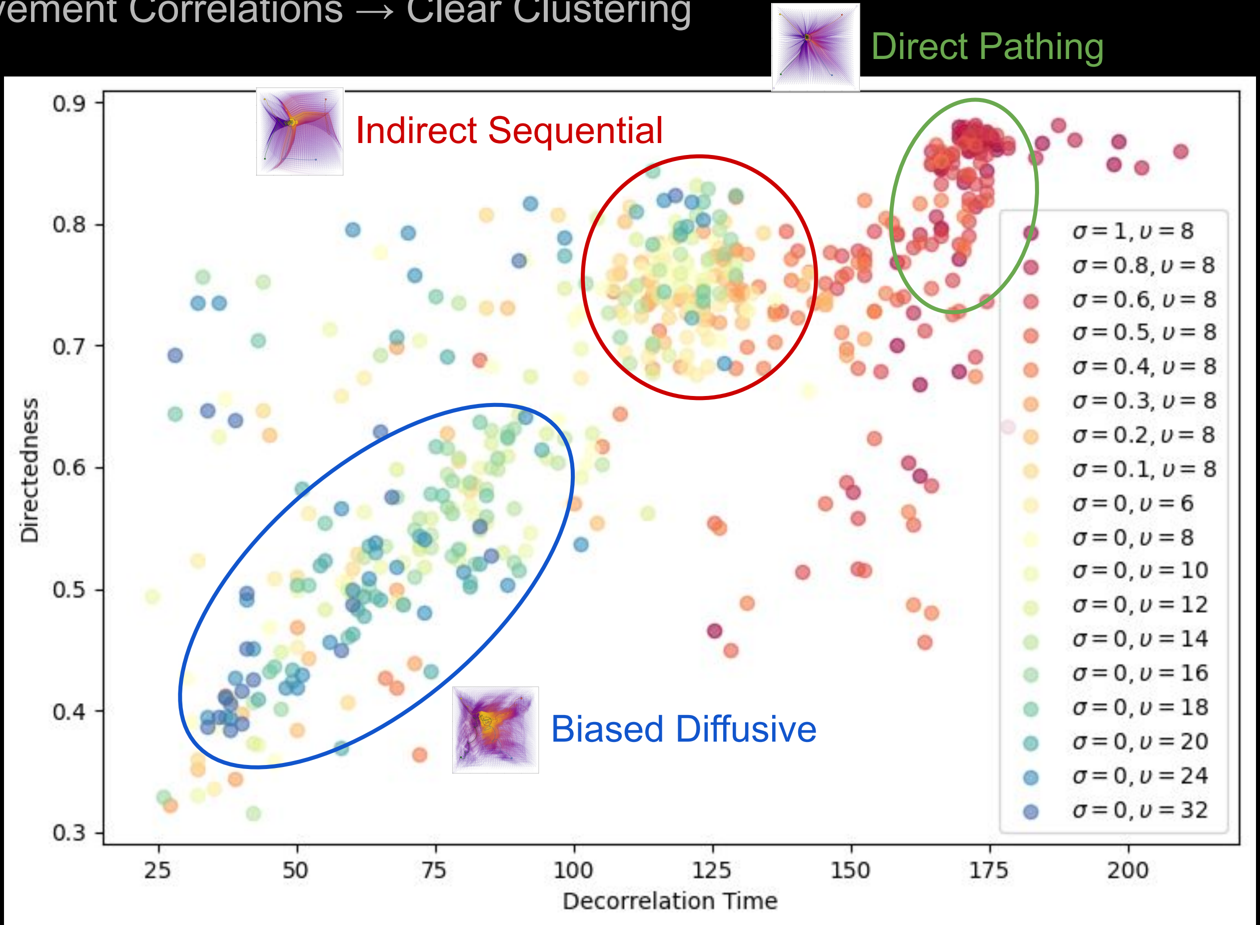
Angles + Distance



Direct Pathing

- Strong unidirectional persistence
- Highly accurate patch estimation

Movement Correlations → Clear Clustering



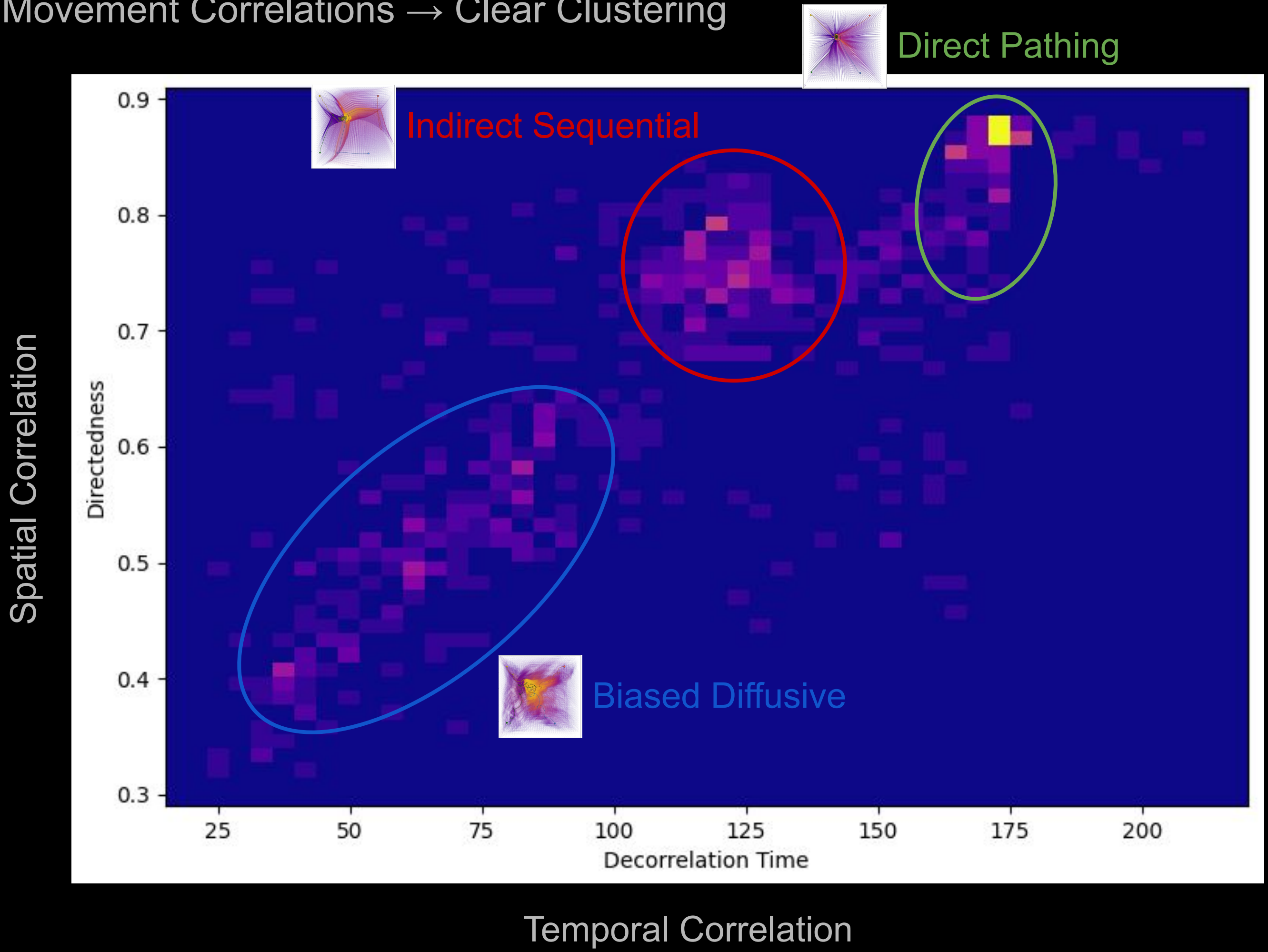
Spatial Correlation

Temporal Correlation

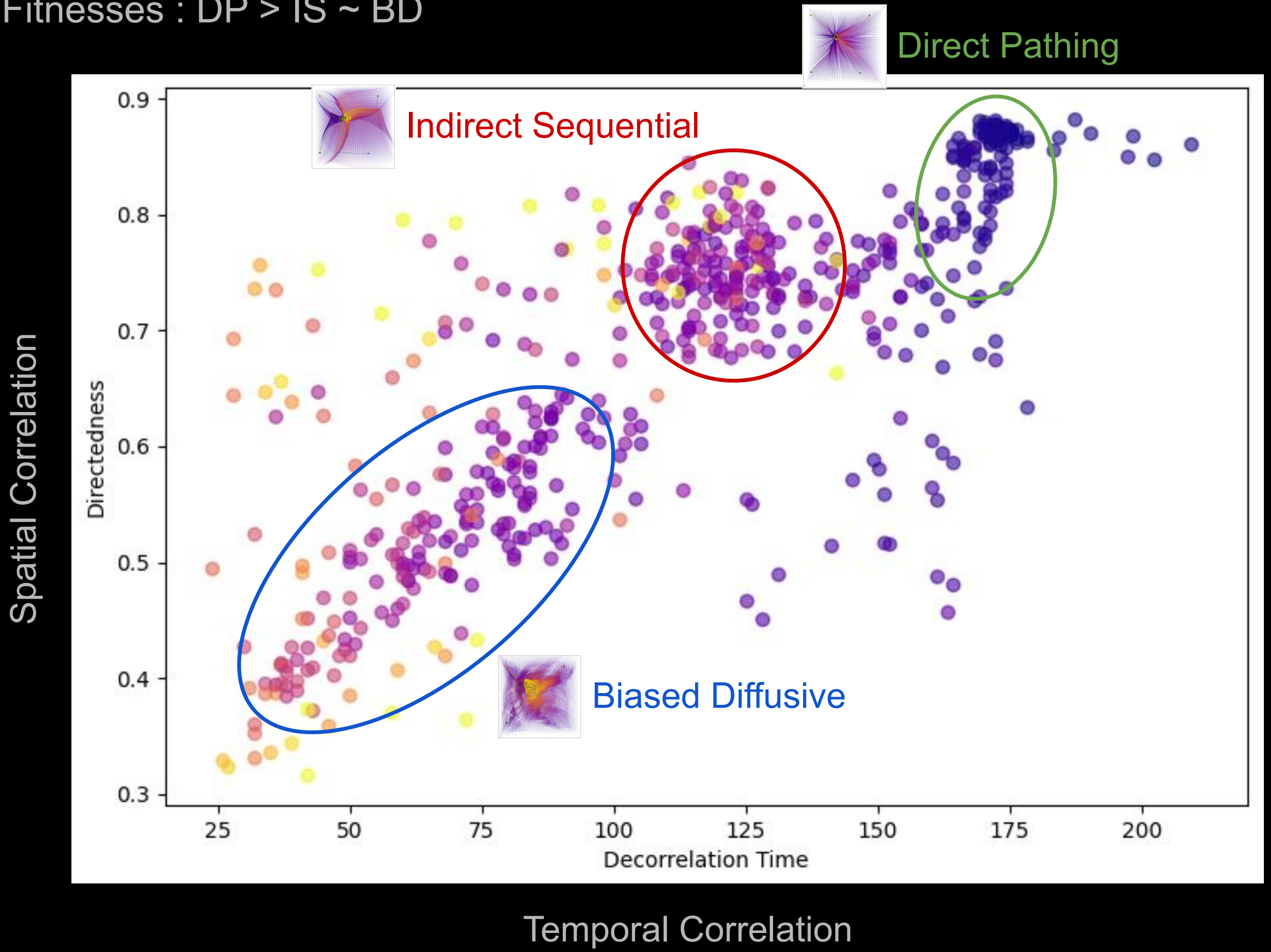
σ = distance scaling factor

υ = visual resolution

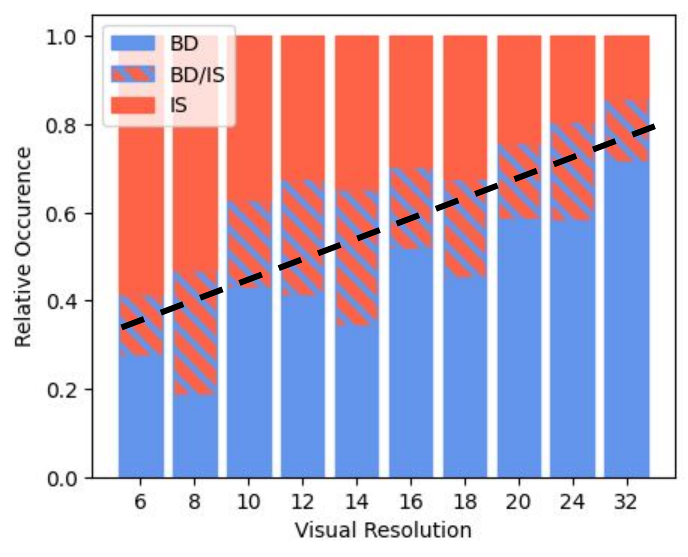
Movement Correlations → Clear Clustering

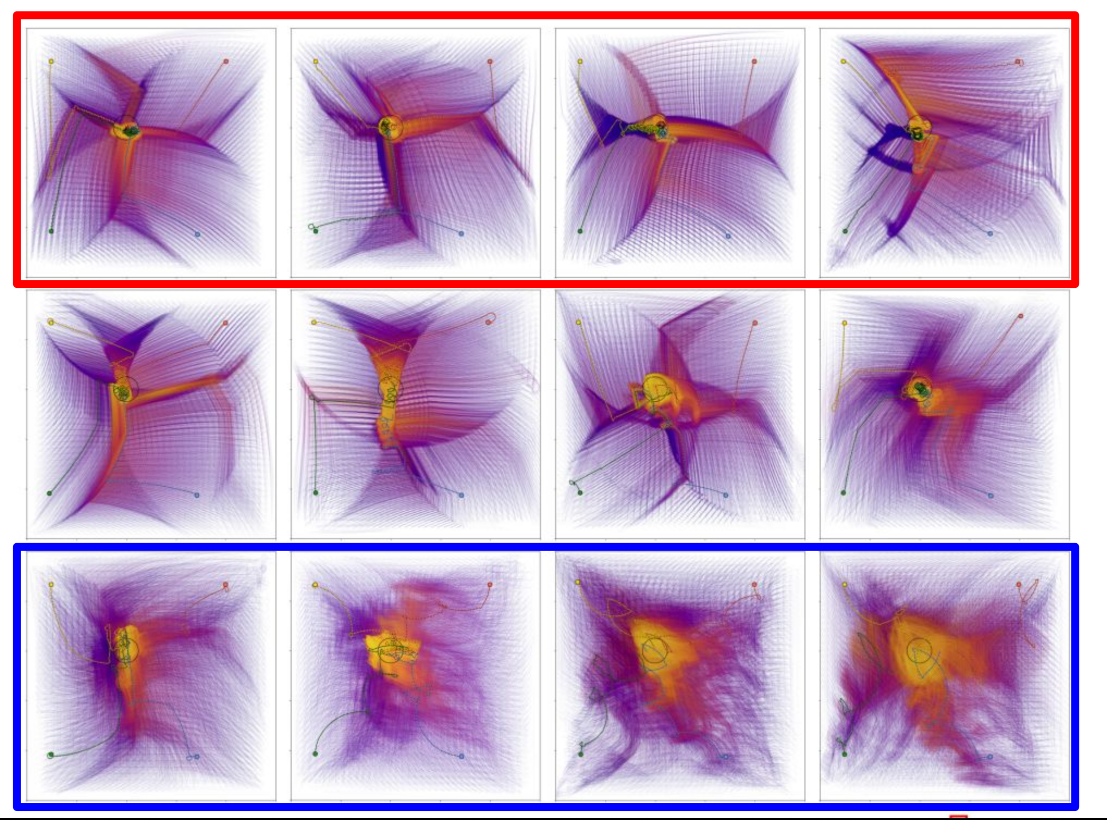


Fitnesses : DP > IS ~ BD

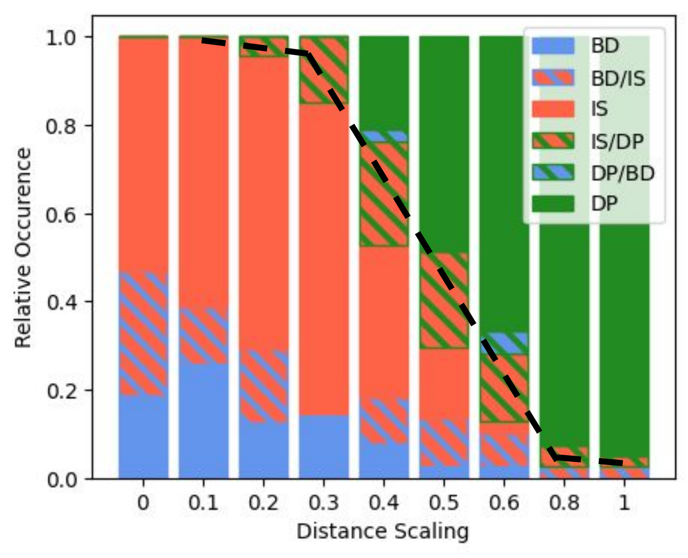


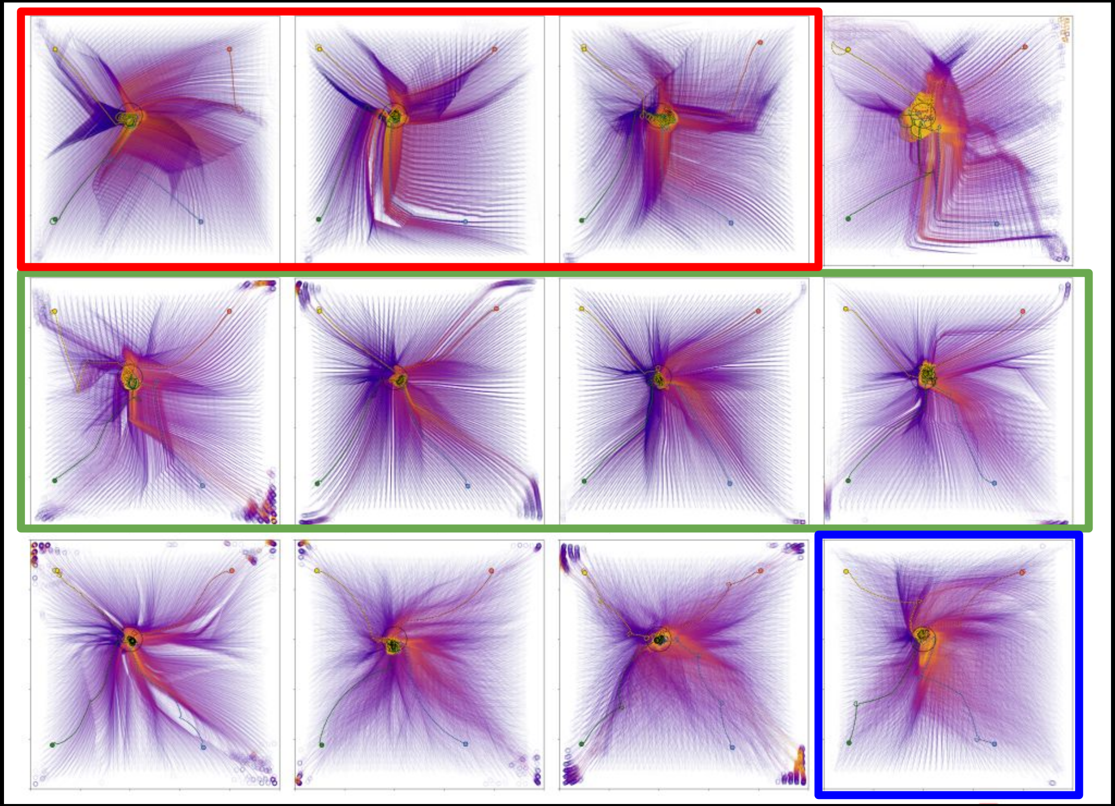
Visual resolution gradually modulates BD/IS class preference/evolvability





Sigmoidal appearance of DP

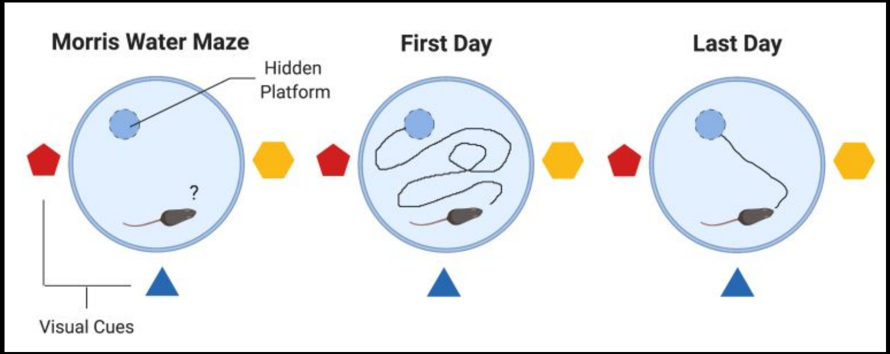


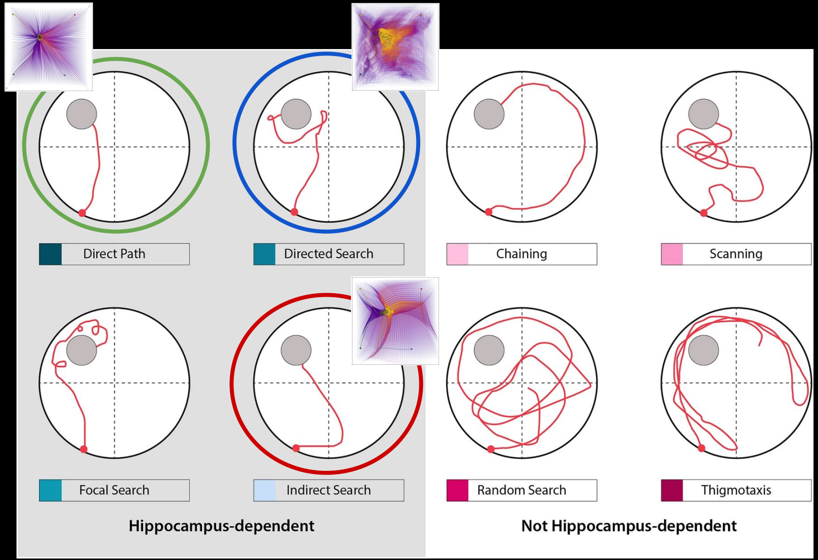


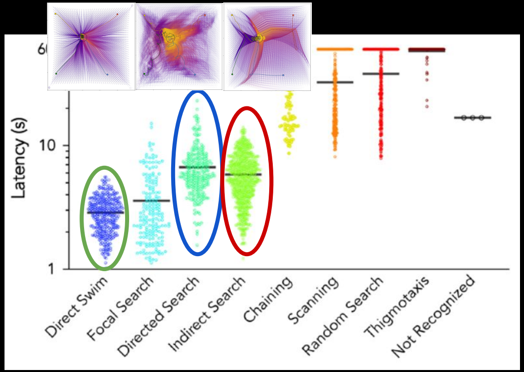
Classes ~ Morris Water Maze strategies

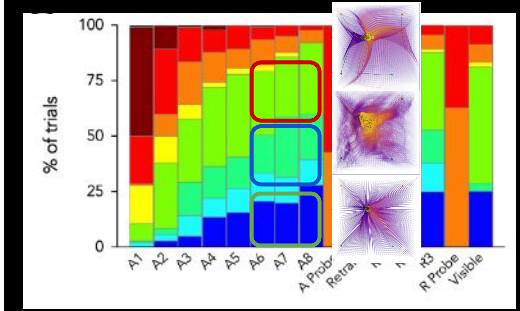
- Indirect Sequential ~ Indirect Search
- Biased Diffusive ~ Directed Search + Self-Orienting Behavior
- Direct Pathing ~ Direct Path/Swim

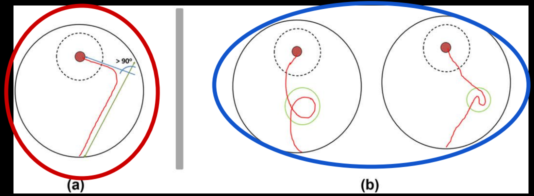
- Similar relative fitness + occurrence metrics



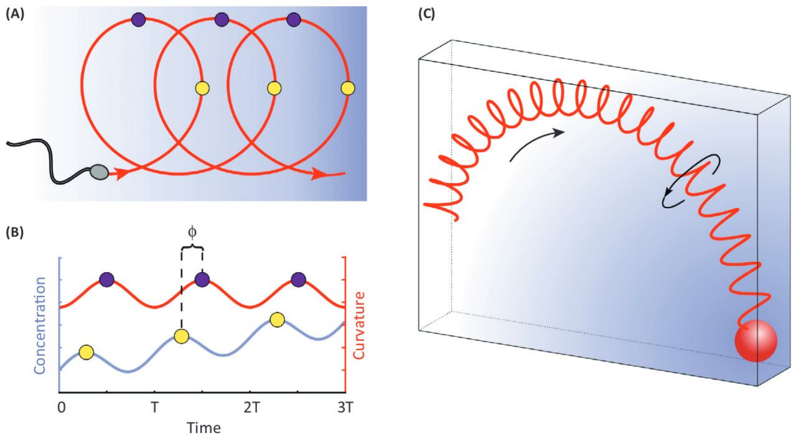








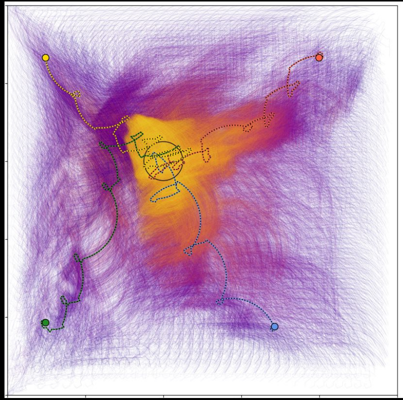
Cooke 2020
Curdt 2022
Villarreal-Silva 2022

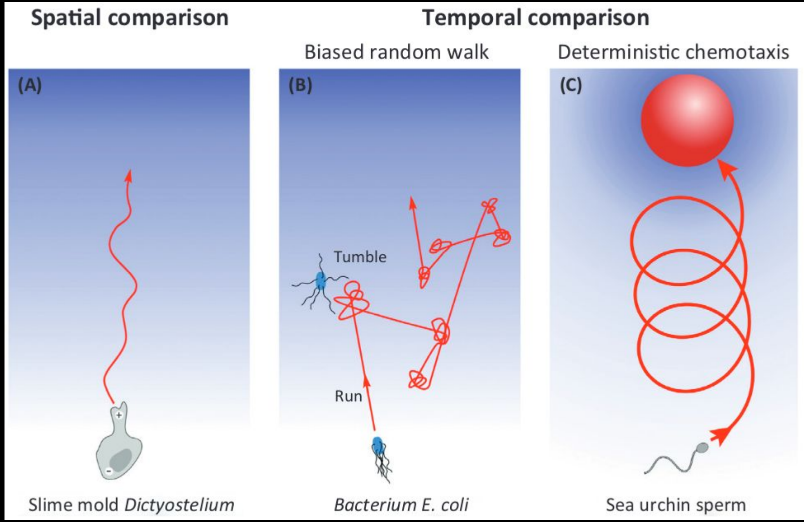


Biased Diffusive ~ sperm chemotaxis

- Dynamic curvature modulation
  → up-gradient helical bias

- concentration sampling vs. landmark observations
- gradient in signal vs. indirect inference

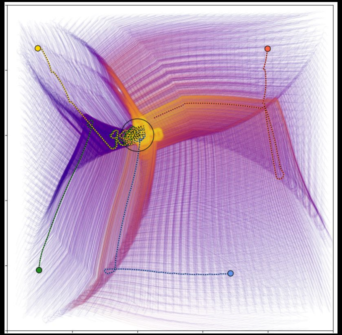


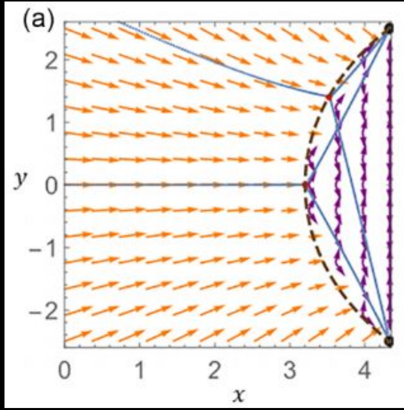


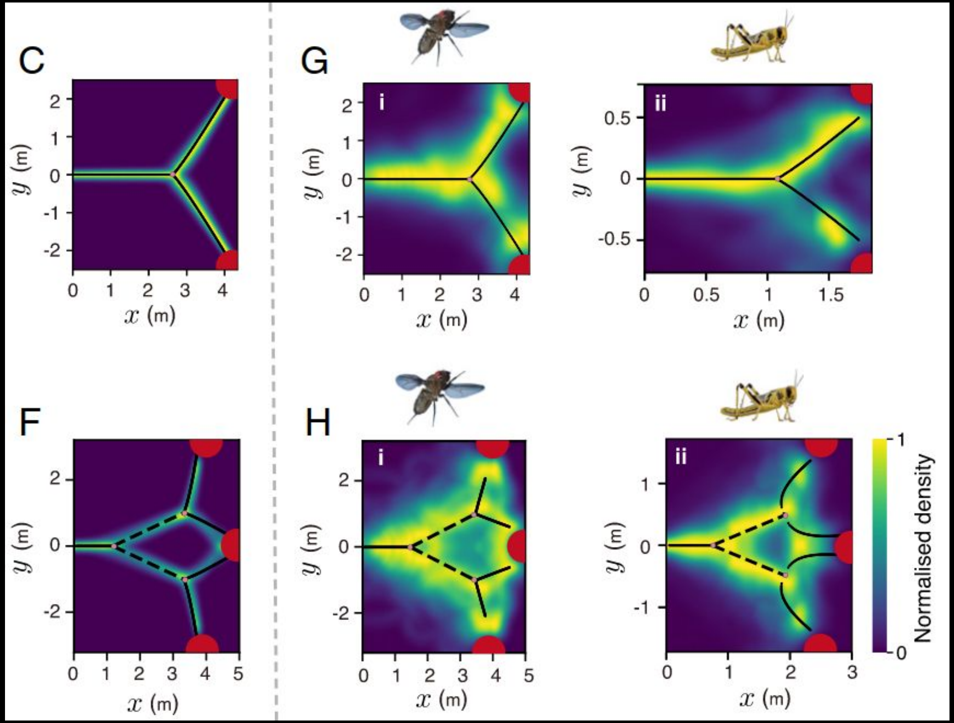
Armon 2012
Alvarez 2014

Indirect Sequential ~ decision-making thresholds from binary spatial choice
→ observed in flies + locusts + fish (+ ring attractor model / set of spin systems)

→ goal choice vs. indirect spatial inference
+ route composition
→ set of spin systems vs. CNN
→ emergent vs. learned transitions

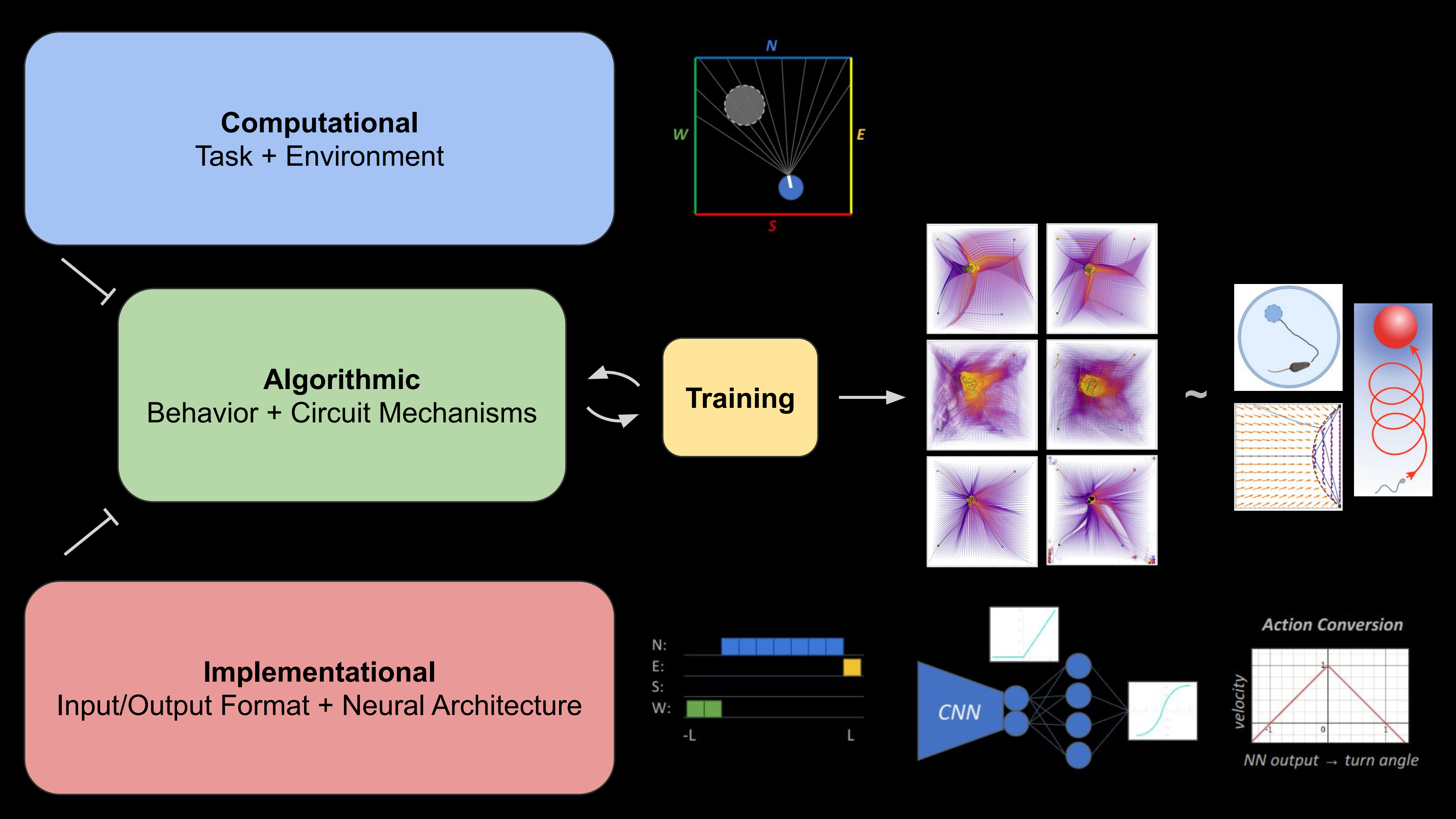






Visual space scaffolded by elliptical geometry
→ it's not all Euclidean cognitive maps

Sridhar 2021
Gorbonos 2024

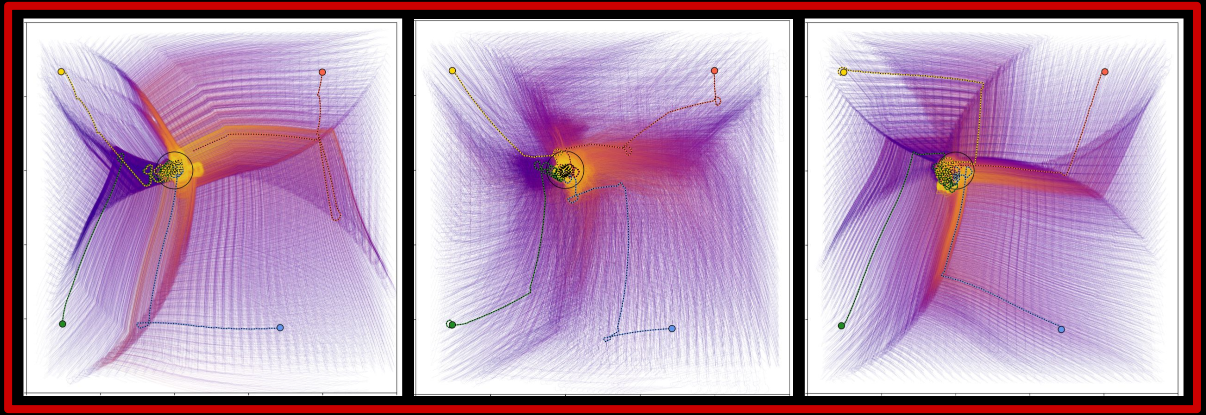

Computational
Task + Environment
N
W
E
S
Algorithmic
Behavior + Circuit Mechanisms
Training
~
Implementational
Input/Output Format + Neural Architecture
N:
E:
S:
W:
-L
L
CNN
Action Conversion
velocity
NN output → turn angle

Govoni, P., & Romanczuk, P. (2024).
arXiv preprint arXiv:2407.13535.

What comes next?

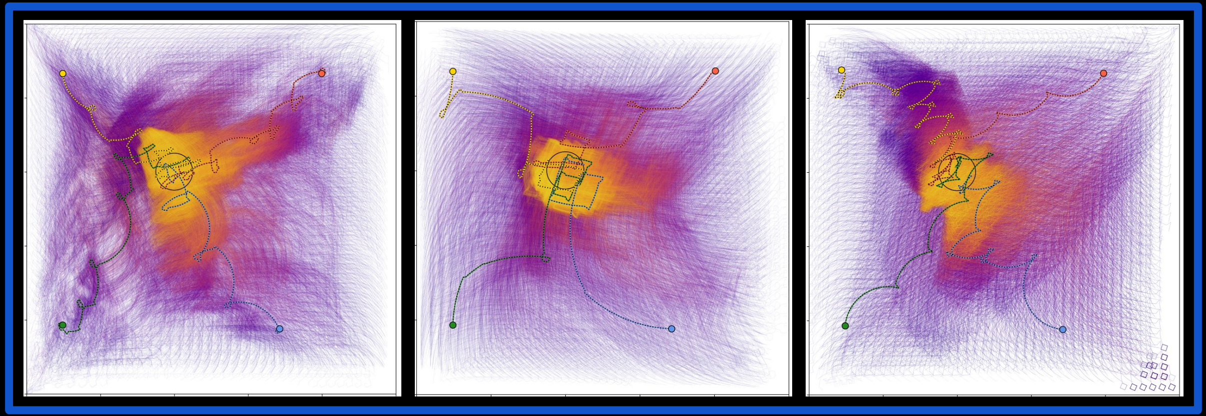
- Circuit mechanisms: 🚧
- Collective navigation:
  social vs. individual strategies +
  metabolic/attention-based constraint

Global behavior
Angles Only



Indirect Sequential

- Grid-like trajectories
- Compositional route segments
- Elliptical decision manifolds



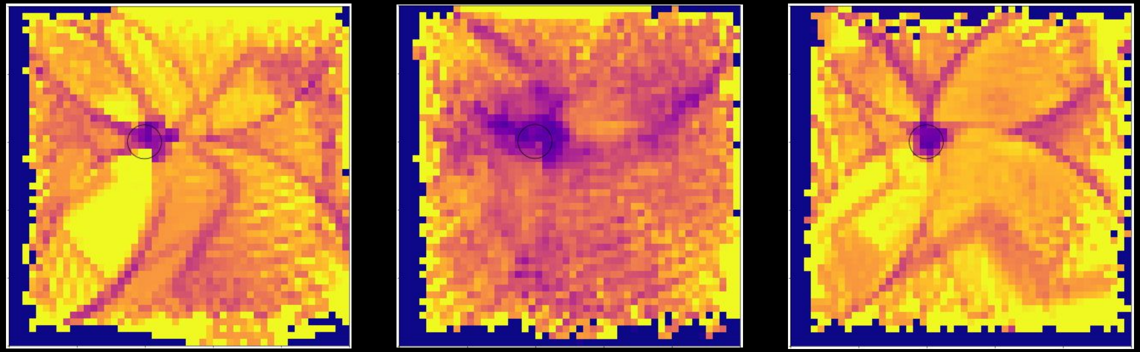
Biased Diffusive

- Spin/ratchet-like trajectories
- Directed towards patch

Entropic Directedness - Spatial Correlation - Allocentric Bias

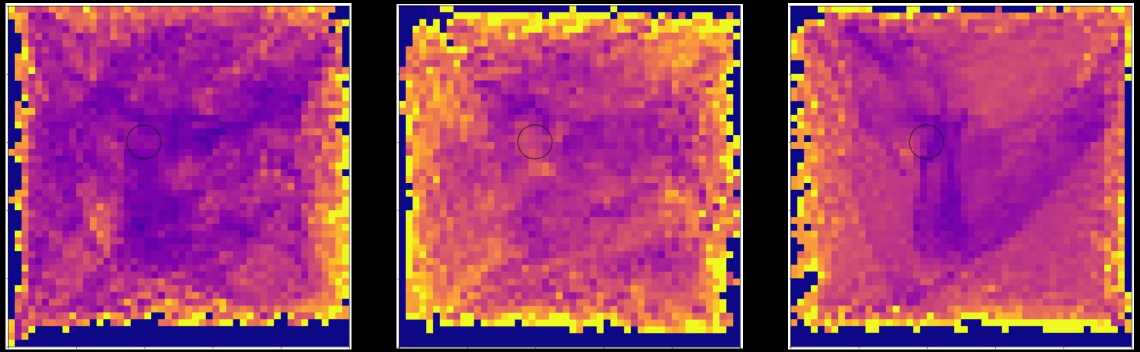
$$H(\Phi, b) = -\sum_{\phi \in \Phi} p(\phi) \log(p(\phi))$$

$$D(H, \Phi, b) = \frac{H_{max}(\Phi) - H(\Phi, b)}{H_{max}(\Phi) - H_{min}(\Phi)}$$



Indirect Sequential

→ Higher

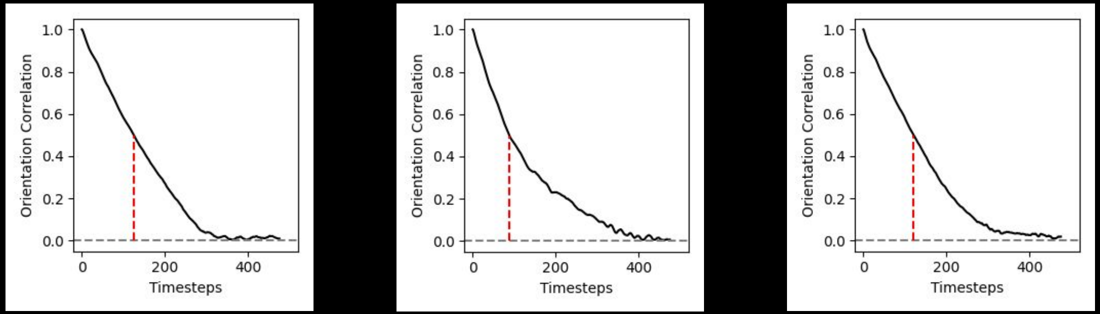


Biased Diffusive

→ Lower

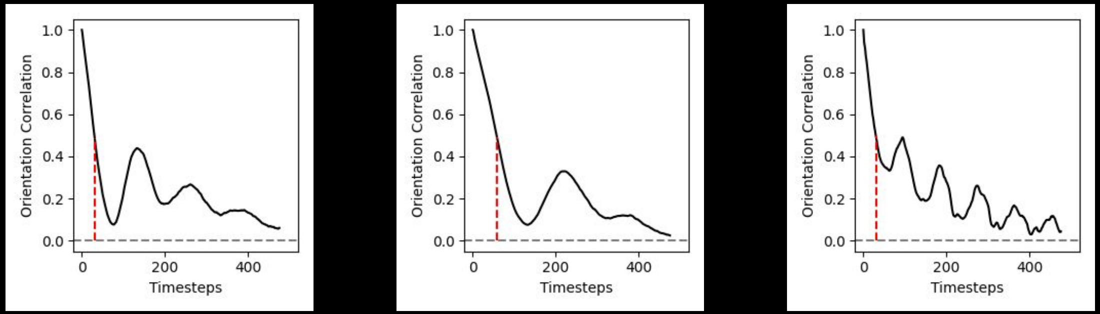
Decorrelation Time - Temporal Correlation

$$C(t) = < \cos[\phi(t_0 + t) - \phi(t_0)] >$$



Indirect Sequential

→ Slow



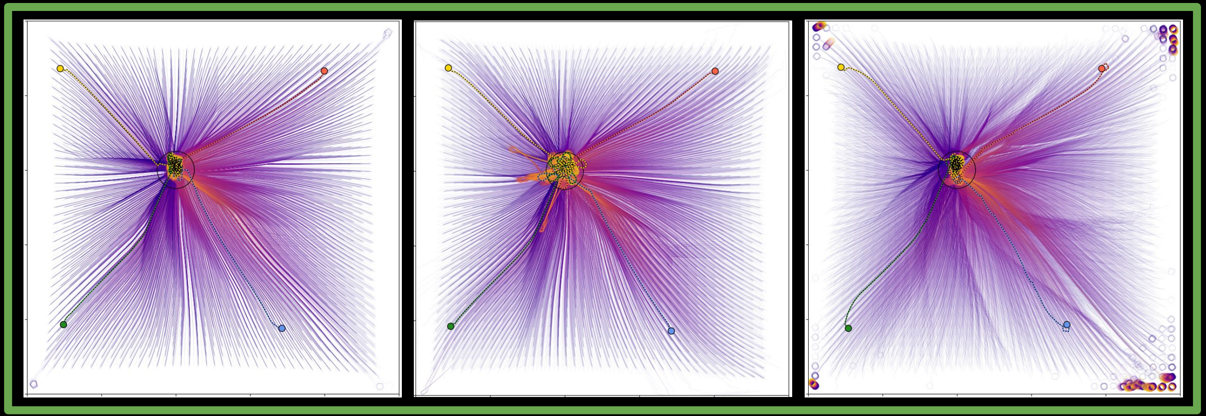
Biased Diffusive

→ Faster

+ Correlated oscillations

Global behavior

Angles + Distance

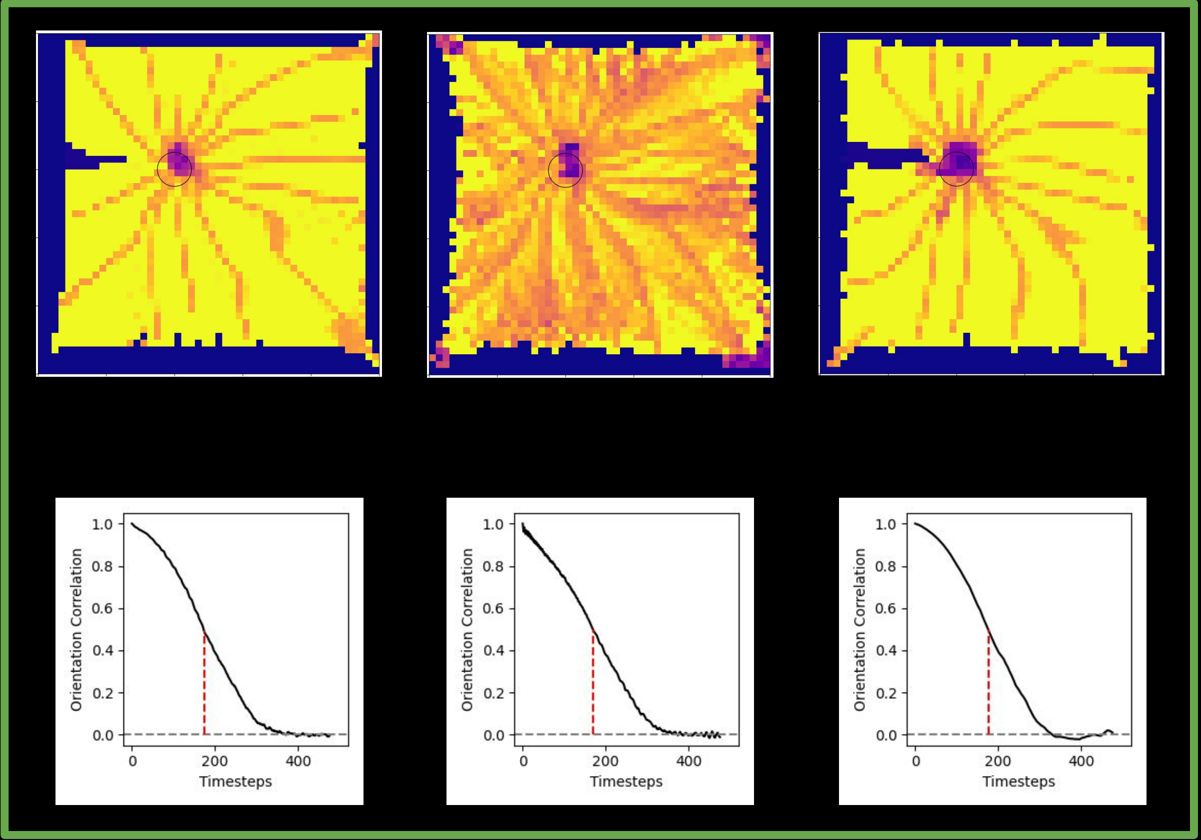


Direct Pathing

- Strong unidirectional persistence
- Highly accurate patch estimation

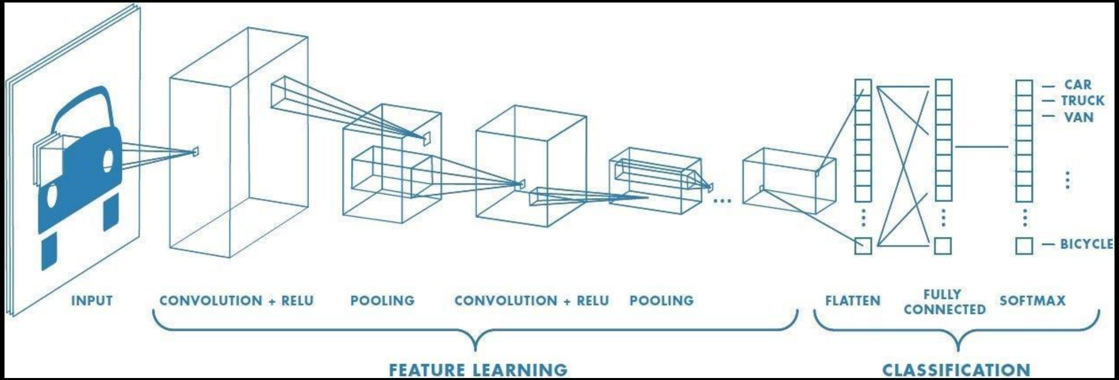
Global behavior

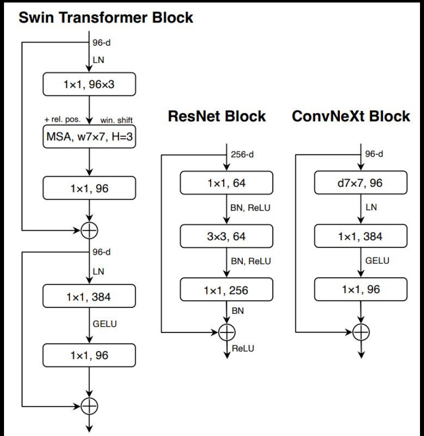
Angles + Distance

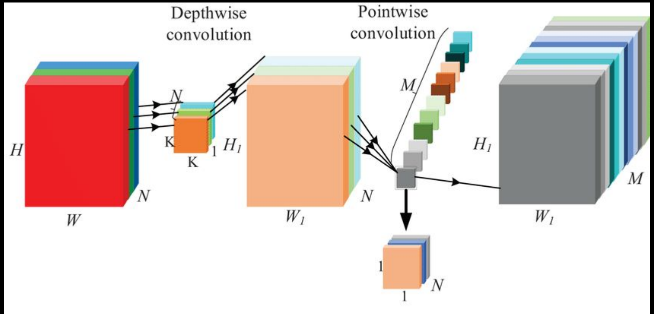


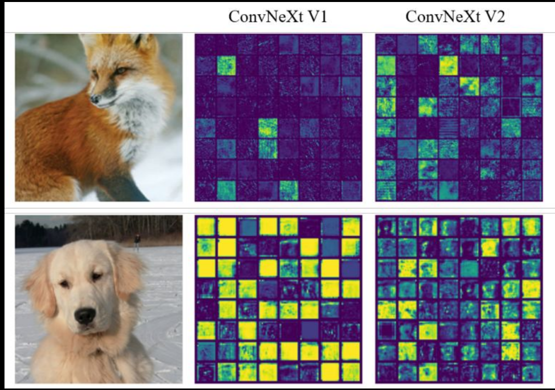
Direct Pathing

- Highest directedness
- Slowest decorrelation



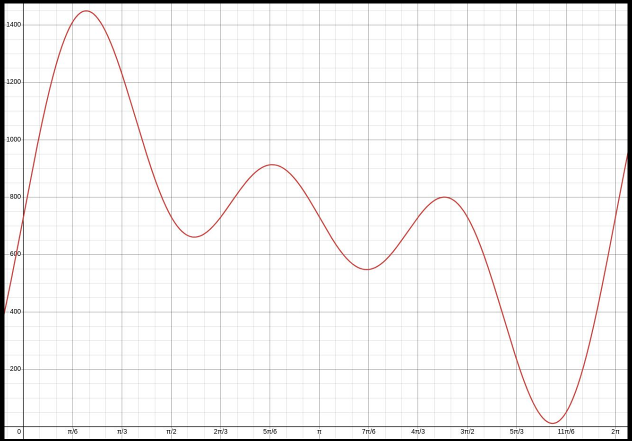




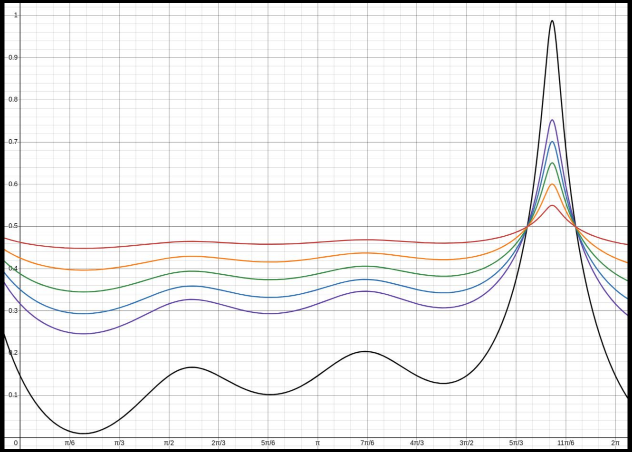


Liu 2022
Wu 2023

Distance
(toy data)



Percept

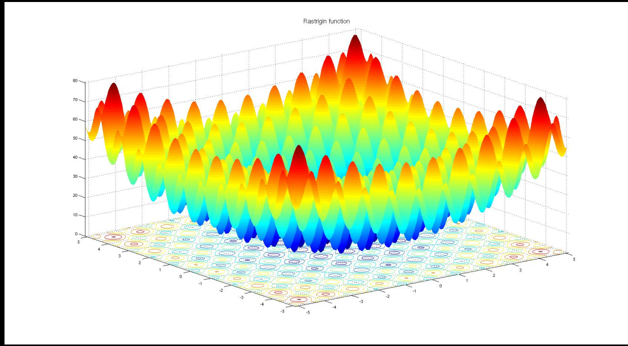


Field of View

Distance input scaled via Weber Fechner law

$y = -\ln(x) / k + m$

→ k,m fit to min/max output ranges



Evolutionary Strategies (ES)

Less complexity/assumptions

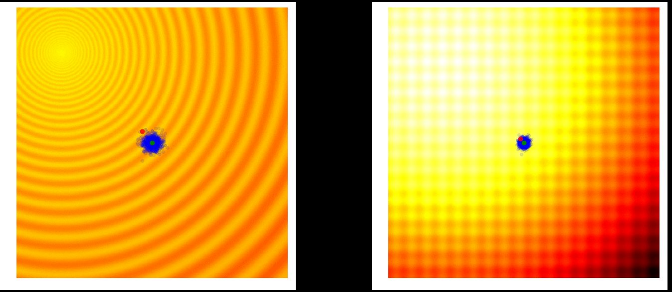
- No value function approximation
- No temporal reward dimension (no credit assignment problem)

Suitably expressive for my question

- Intrinsic exploration helps navigate rugged fitness landscapes
- Global convergence more likely

Main Issue

- Poor data efficiency (shared with RL)



Hansen 2003
Ha 2017

CMA-ES

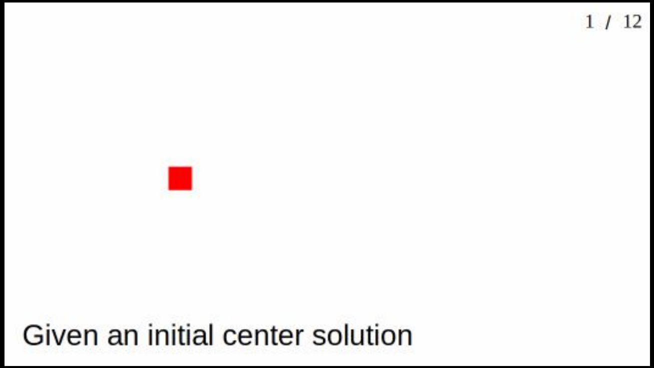
(Covariance-Matrix Adaptation)

PGPE

(Policy Gradient Parameter Exploration)

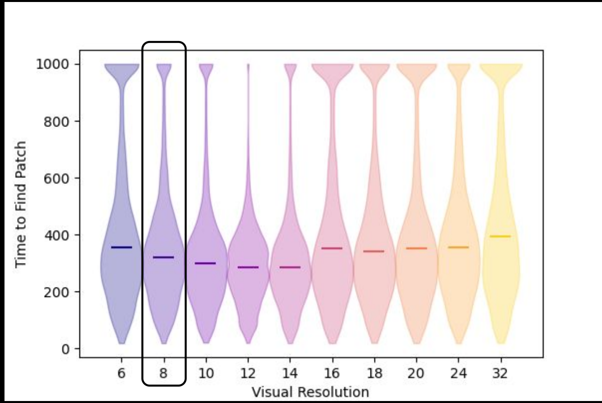
+

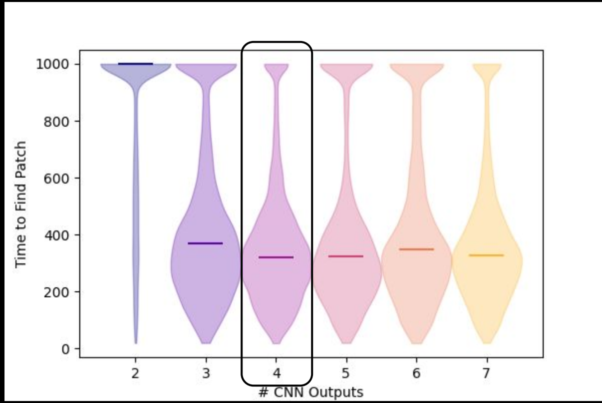
ClipUp optimizer

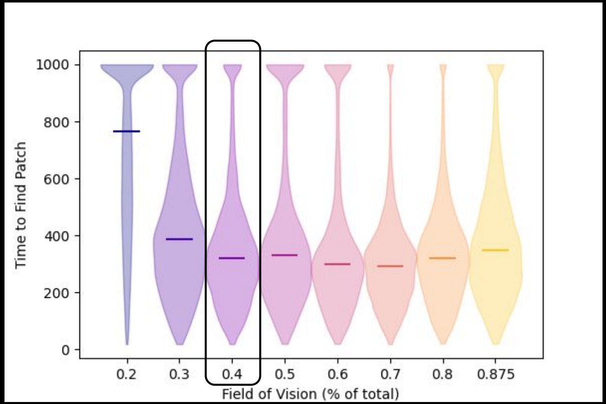


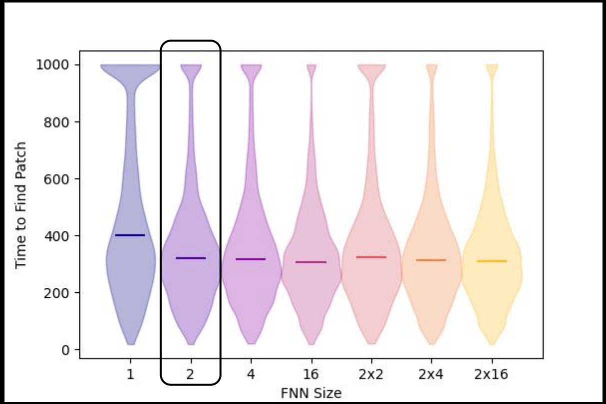
Sehnke 2010
Toklu 2020

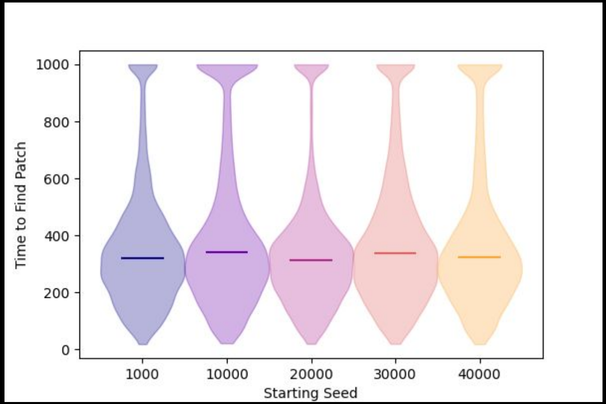
Performance distributions robust to perceptual/cognitive hyperparameters + seed



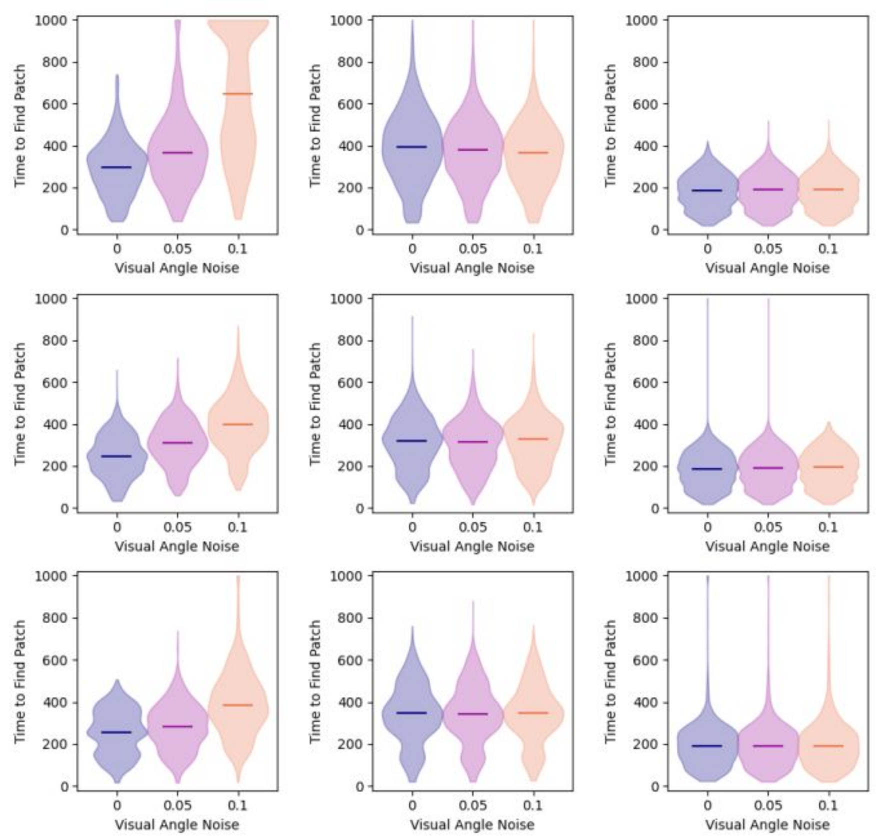




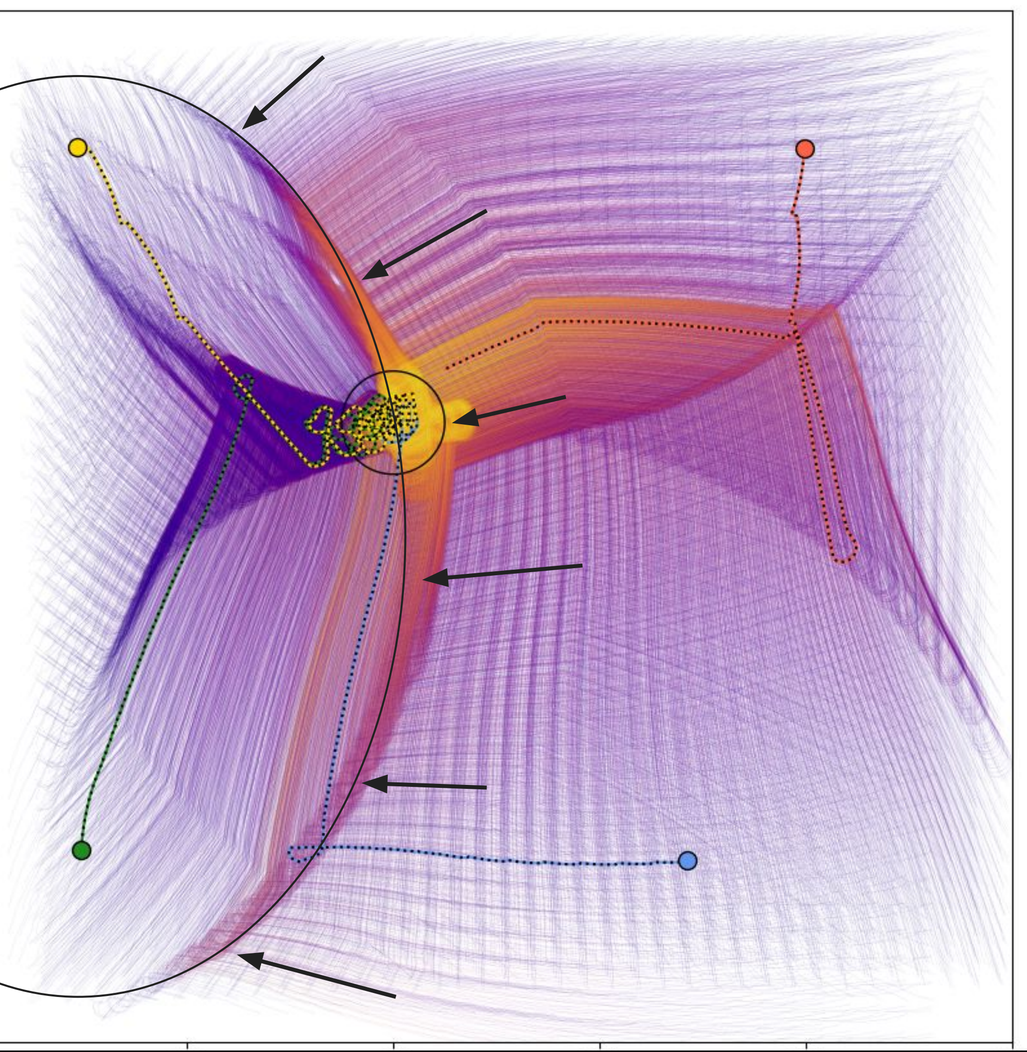




Noise to visual angle → neg affect to IS, BD/DP robust



(a) Indirect Sequential (b) Biased Diffusive (c) Direct Pathing

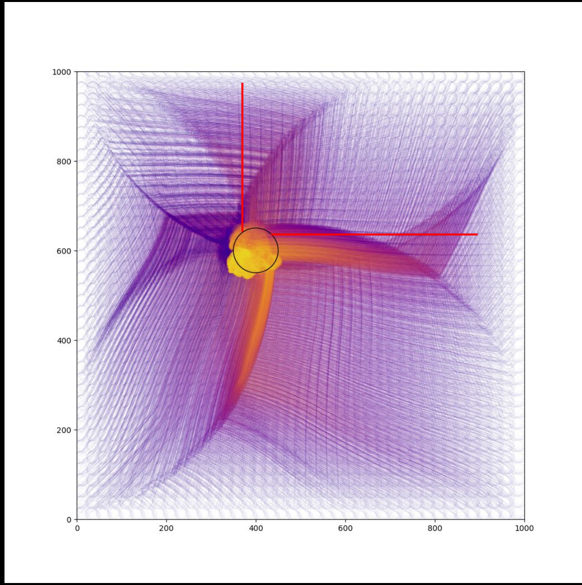


How are IS ellipses calculated?

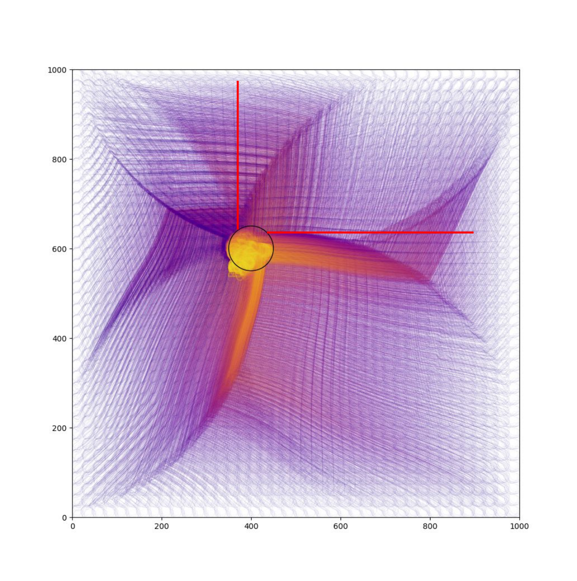
- Artifact from discretized vision?
- From CNN processing?
- Inherent geometrical affordance?

Ellipses : fn (FOV)

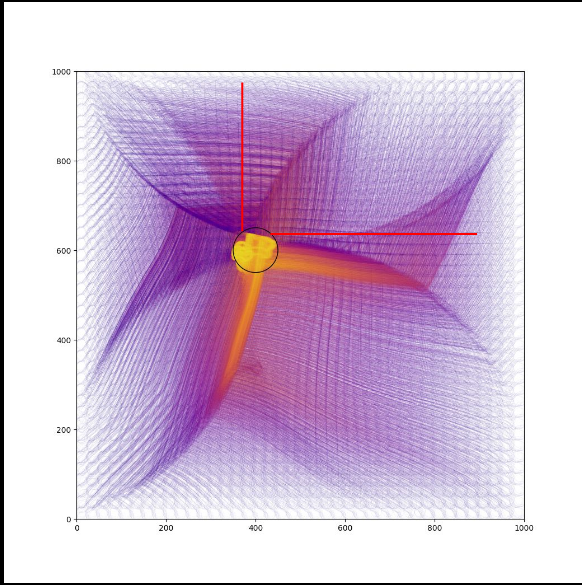
FOV = 0.39



FOV = 0.4

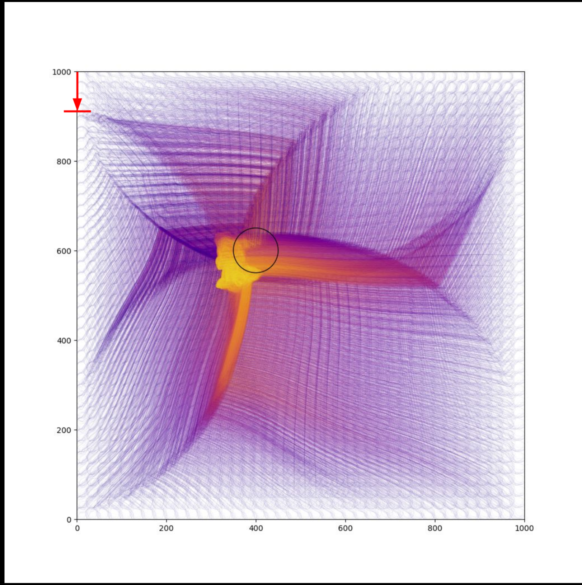


FOV = 0.41

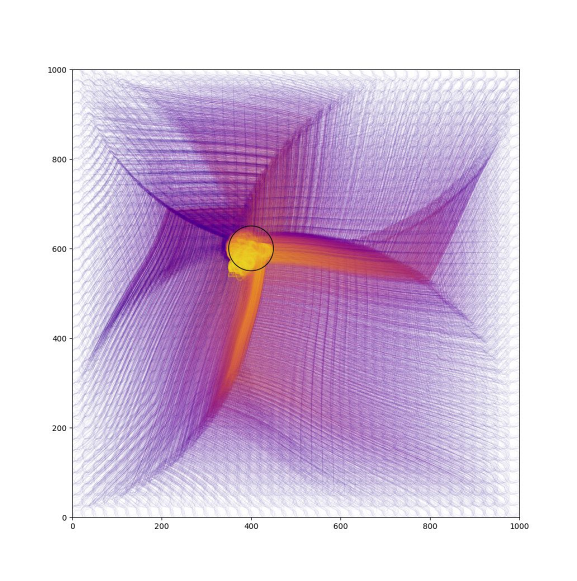


Ellipses : fn (FOV, corner locations)

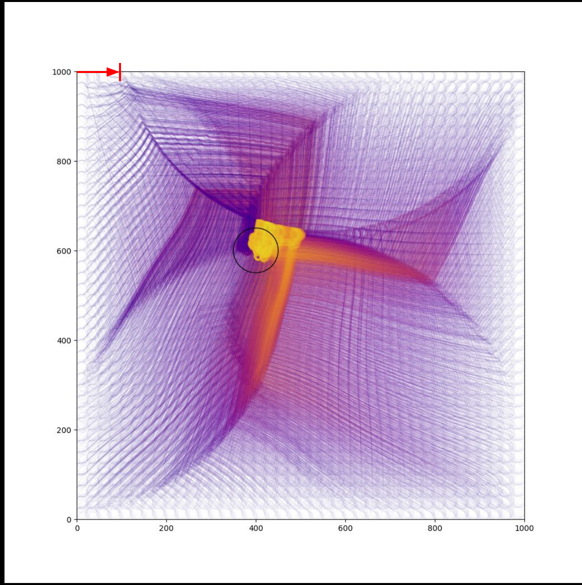
y(TL) - 100



unperturbed

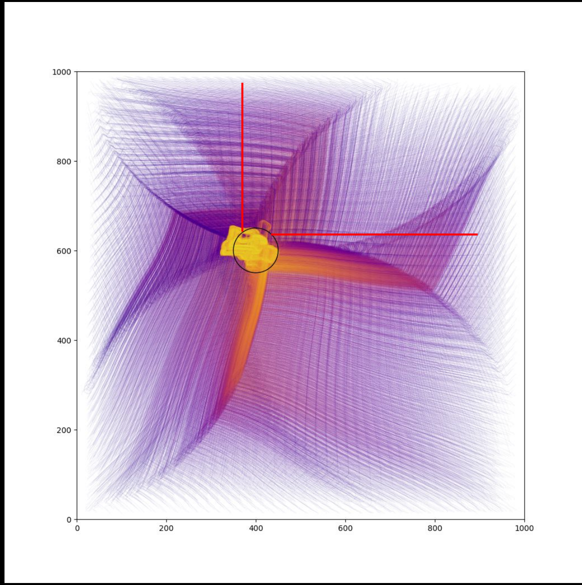


x(TL) + 100

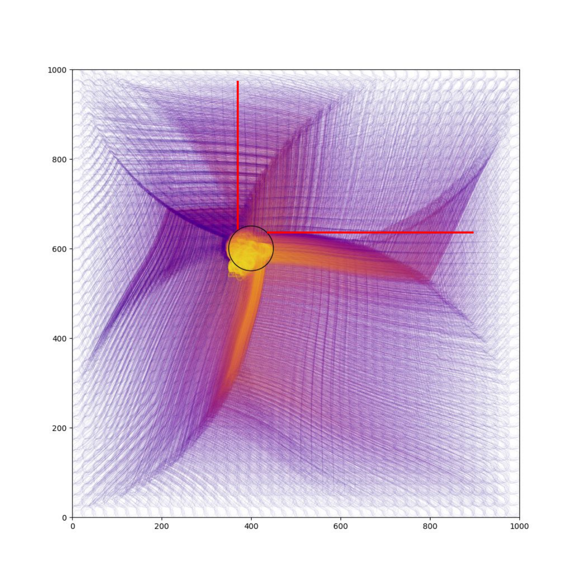


Ellipses : fn (FOV, corner locations, speed)

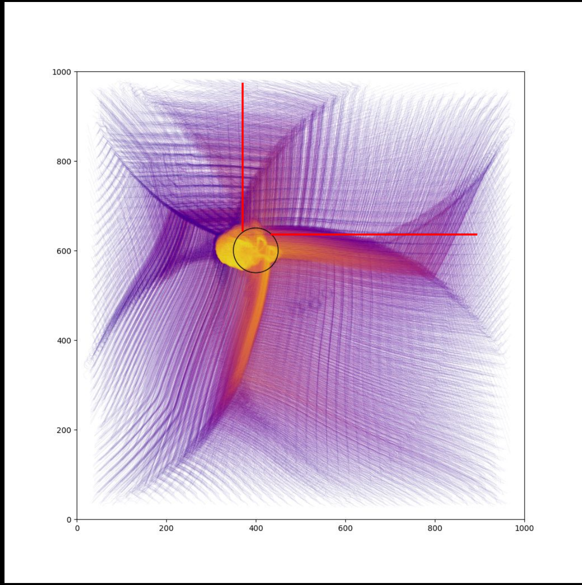
*speed * .75*



unperturbed



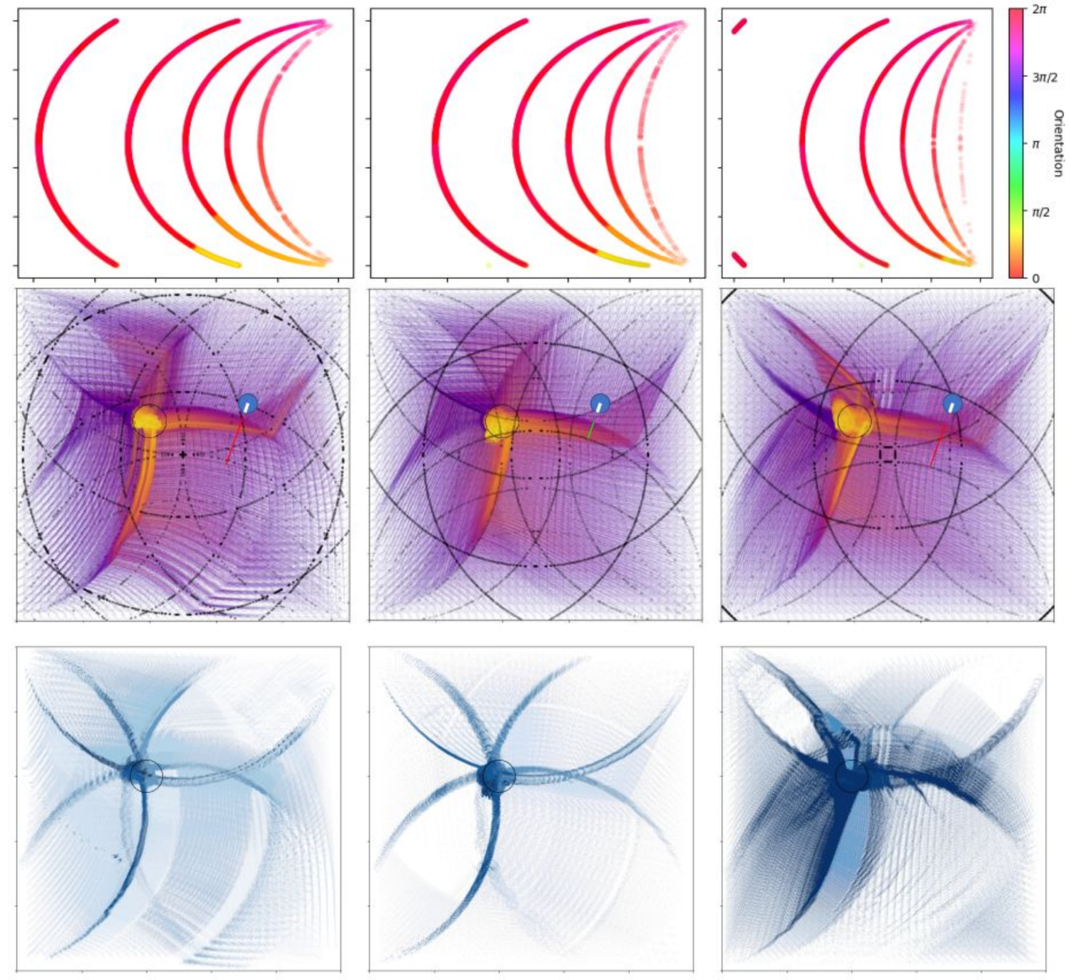
*speed * 1.25*



Ellipses mechanism: dual-corner detection

training at different FOV should shift ellipses
→ they do not

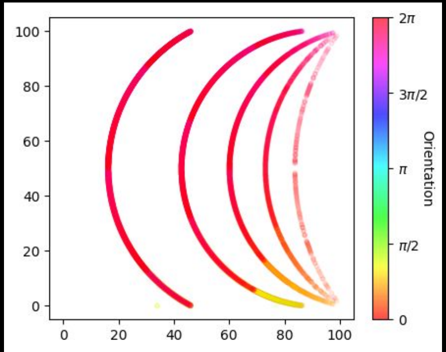
greater turning speed modulates ellipse position to align with patch position



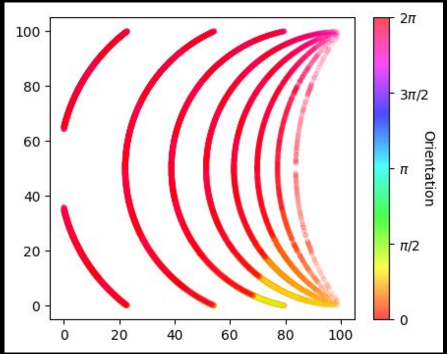
(a) FOV = 0.35 (b) FOV = 0.40 (c) FOV = 0.45

Ellipses :

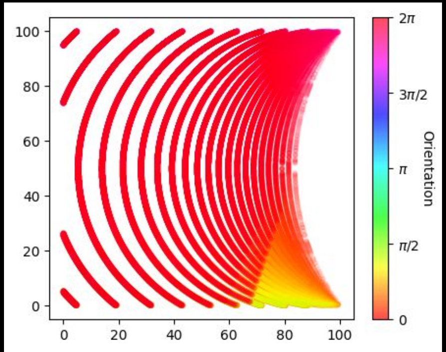
fn (FOV, corner locations, speed, visual resolution)



res = 8



res = 12



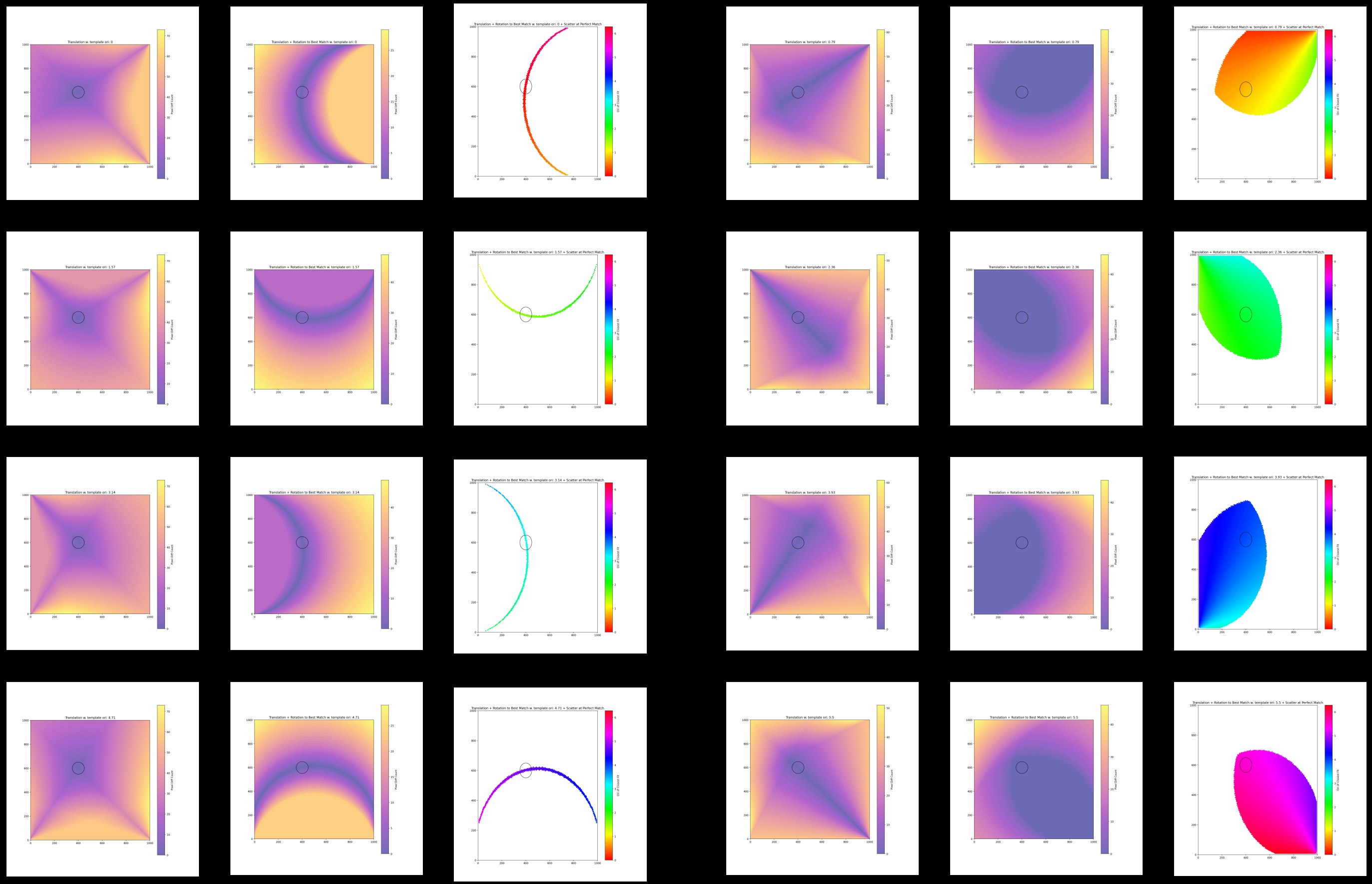
res = 32

Image Difference
Relative to Patch Views

Vis res: 128
Space step: 25
Ori step: pi / 256

2 corners → ellipse

1 corner → leaf



Trans
(# pixels match view)

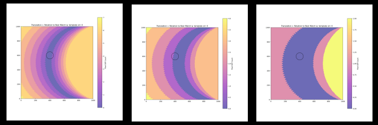
Trans+Rot

Trans+Rot
(ori of perfect match)

Rot + Trans → 2 DoF

Perceiving external geometry → degeneracy

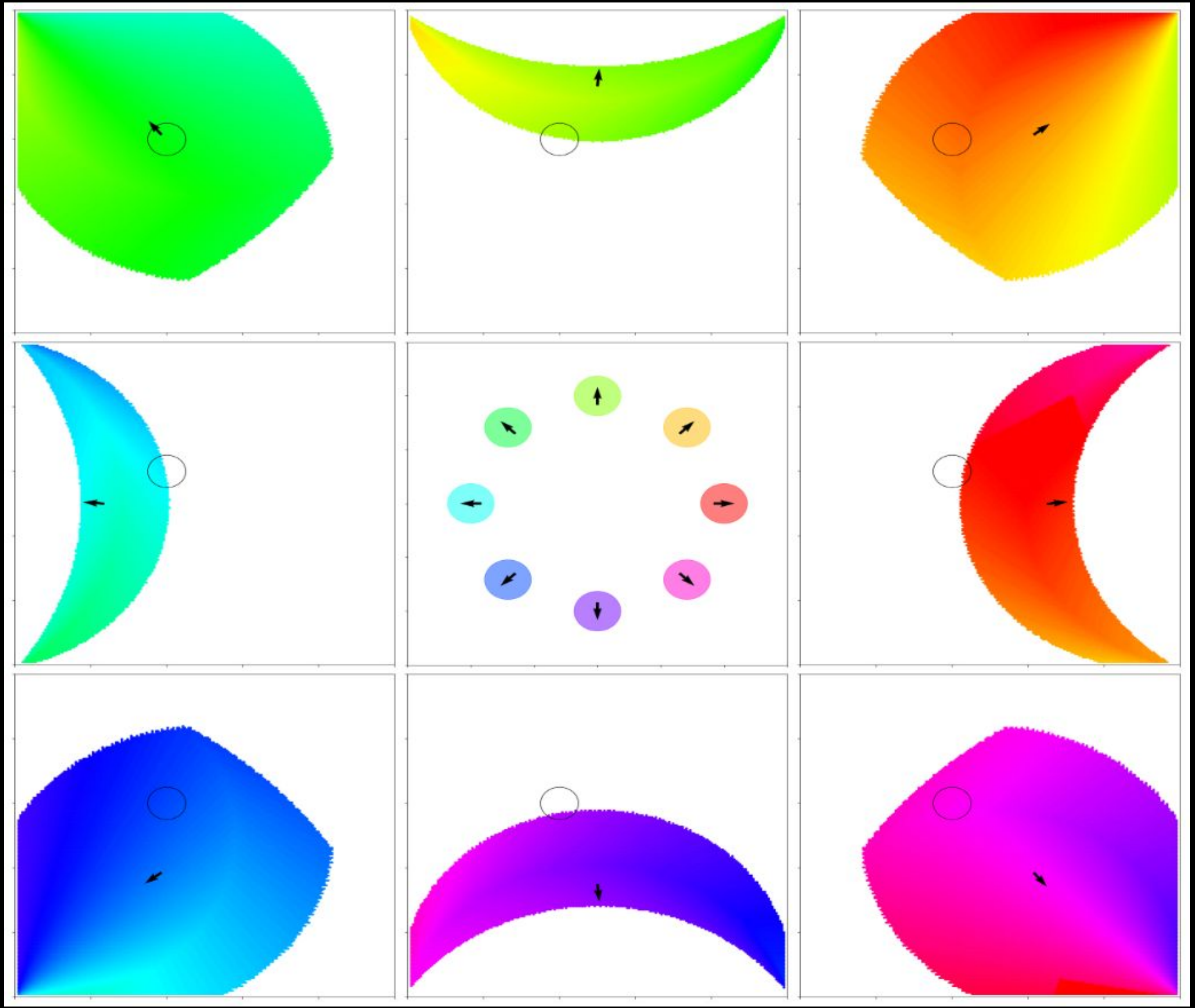
Varying vis res varies areas only, not shapes

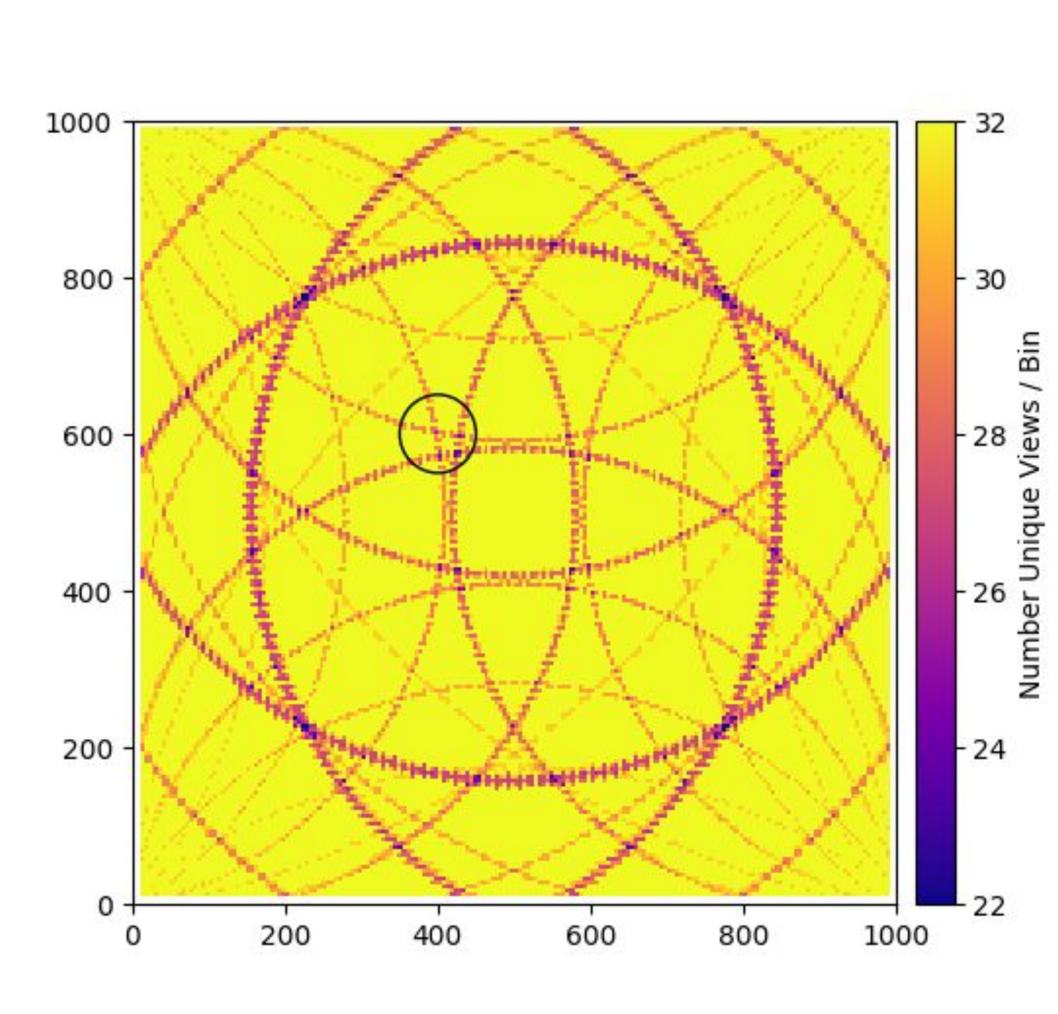


e.g… 32 16 8

9 representative views
(resolution = 8)

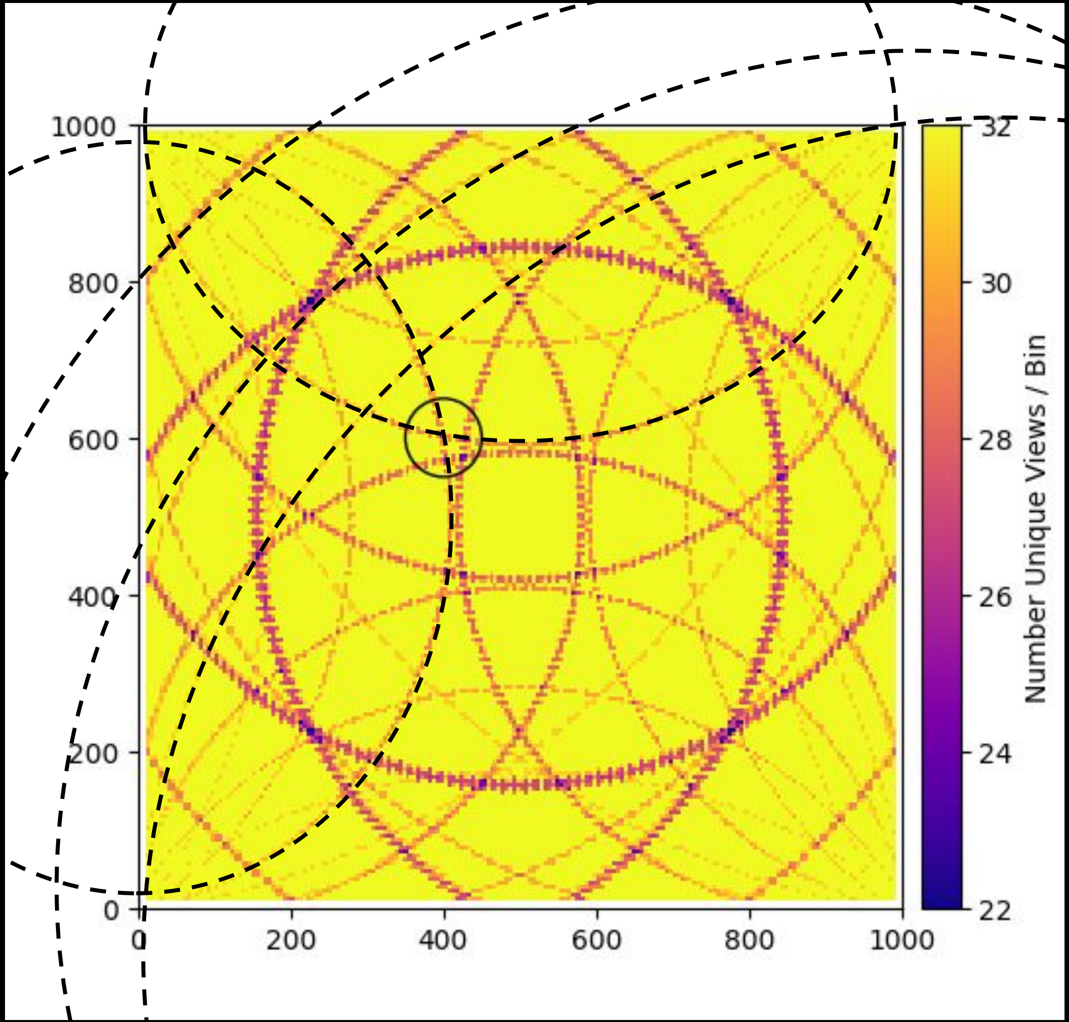
T/B/L/R : 2-corner views → ellipse
Corner : 1-corner views → leaf





Heatmap
for all perfect matches
for all views
(resolution = 8)

- thresholds : “perceptual affordance space”
- mark where agents can switch control regimes
- generated via symmetry breaking, where L+R copies collapse to one

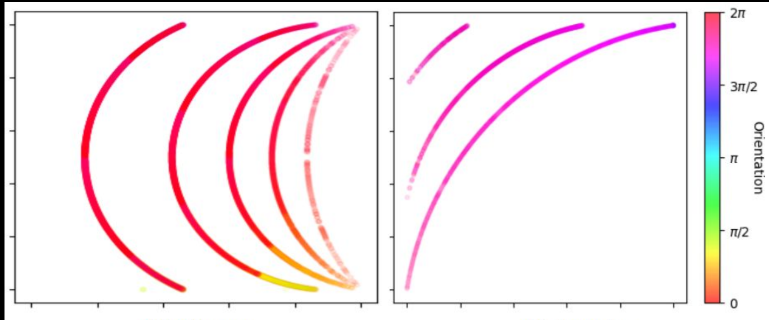


Major ellipses :

- adjacent 2-corner detection
- both predict 5 * 4 ellipses possible

Minor diagonal ellipses:

- opposite 2-corner detection
- both predict 3 * 4 ellipses possible



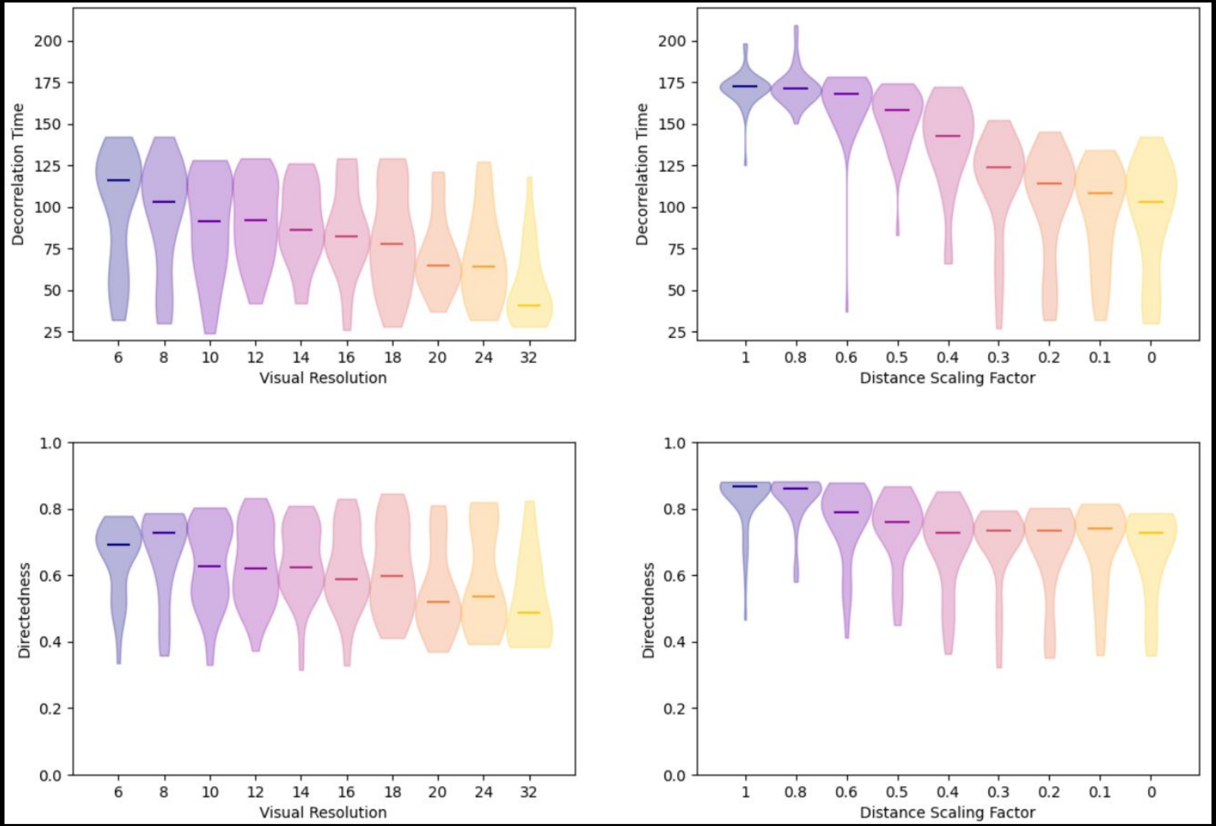
(a) Adjacent

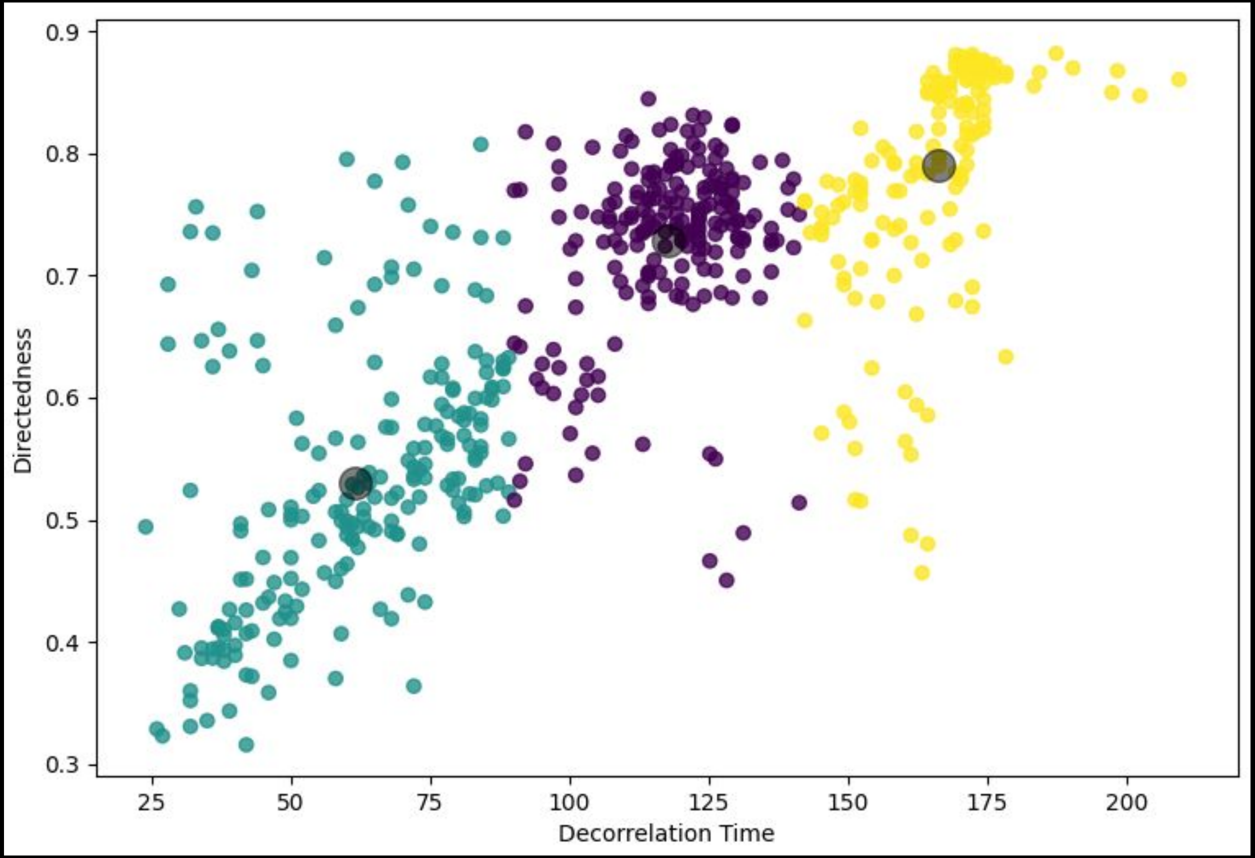
(b) Opposite

Classification parameters x (visual resolution, distance scaling factor)
→ fairly good 1D separation
→ bimodality apparent in visual resolution (IS + BD classes)

Temporal Correlation

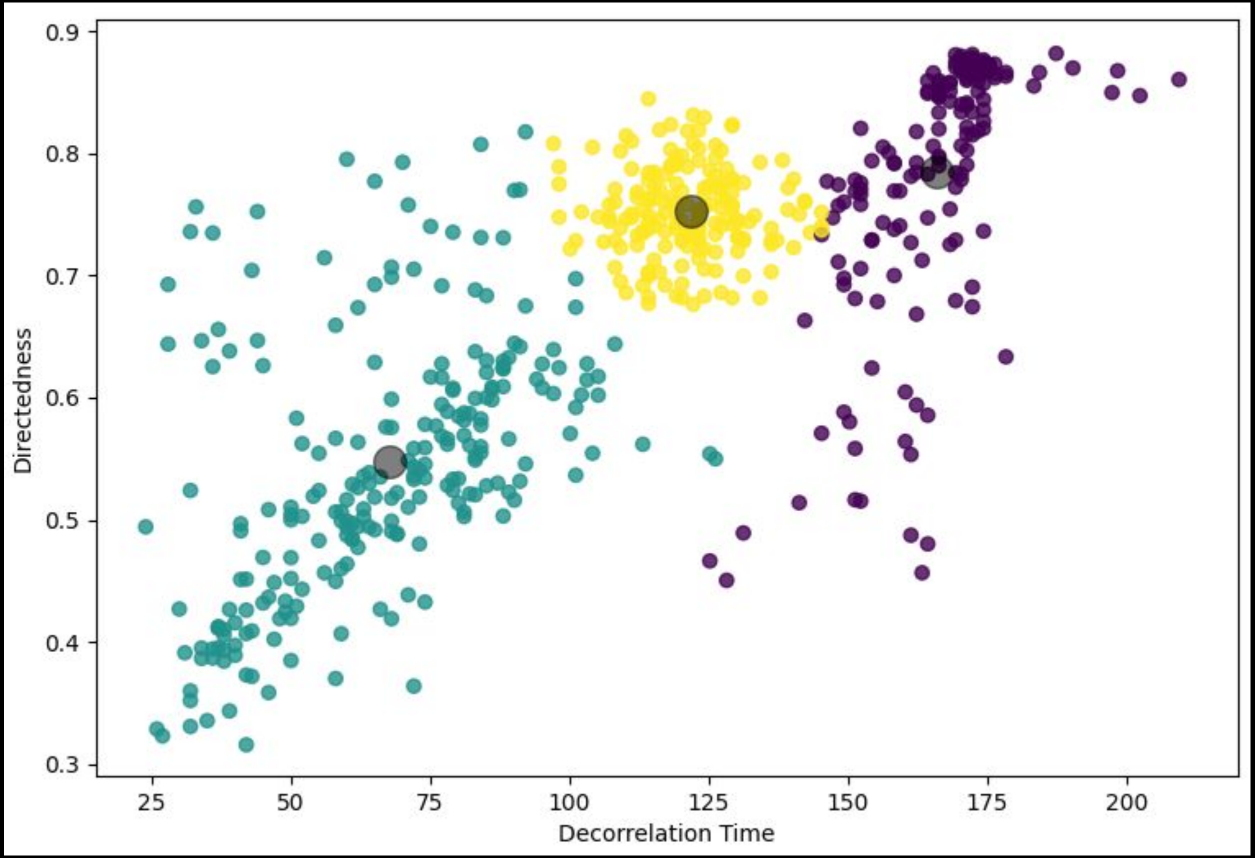
Spatial Correlation





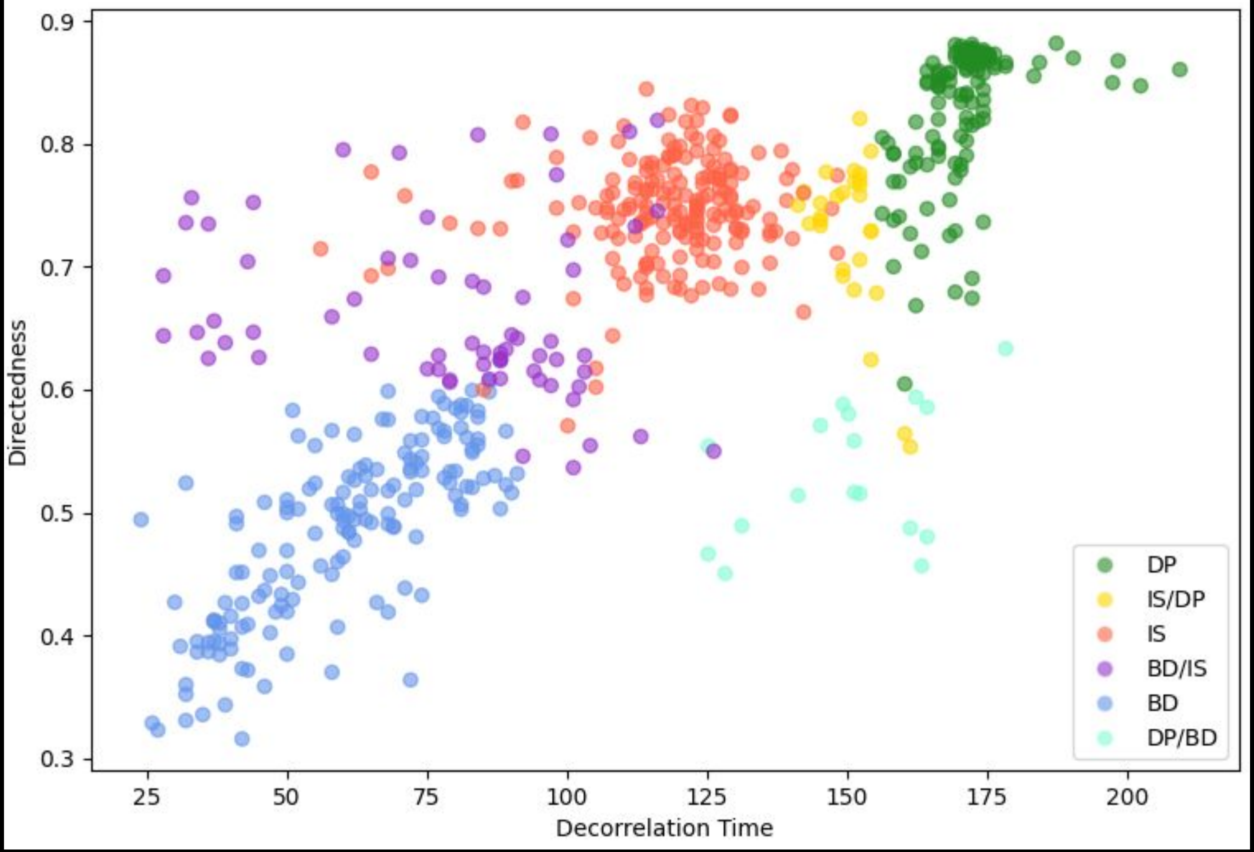
K-Means

Unsupervised clustering works, but solely data-driven



Gaussian Mix

Unsupervised clustering works, but solely data-driven



Class labels clustered via:

- heuristics based on clustering algos + heatmaps
- outliers qualitatively labeled
  - to single / hybrid class
  - some IS better labeled as hybrids

→Persistent random walks cannot reliably reach patch

RW

PRW
D = 0.5

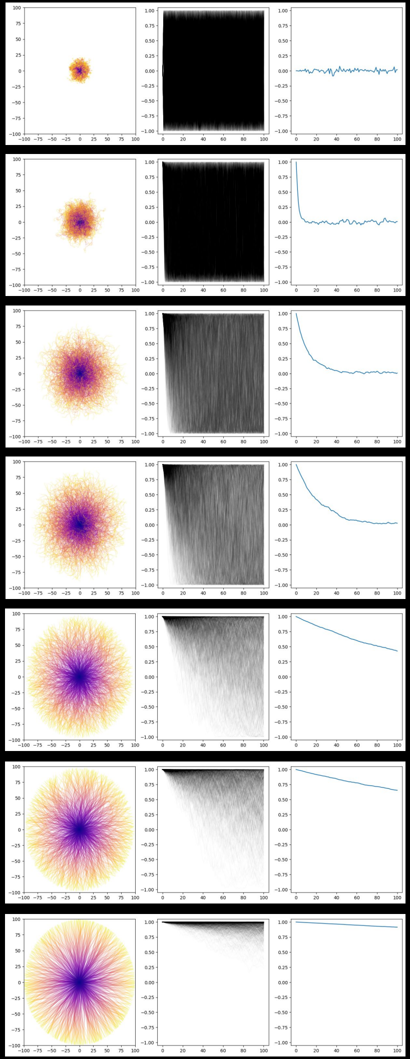
PRW
D = 0.1

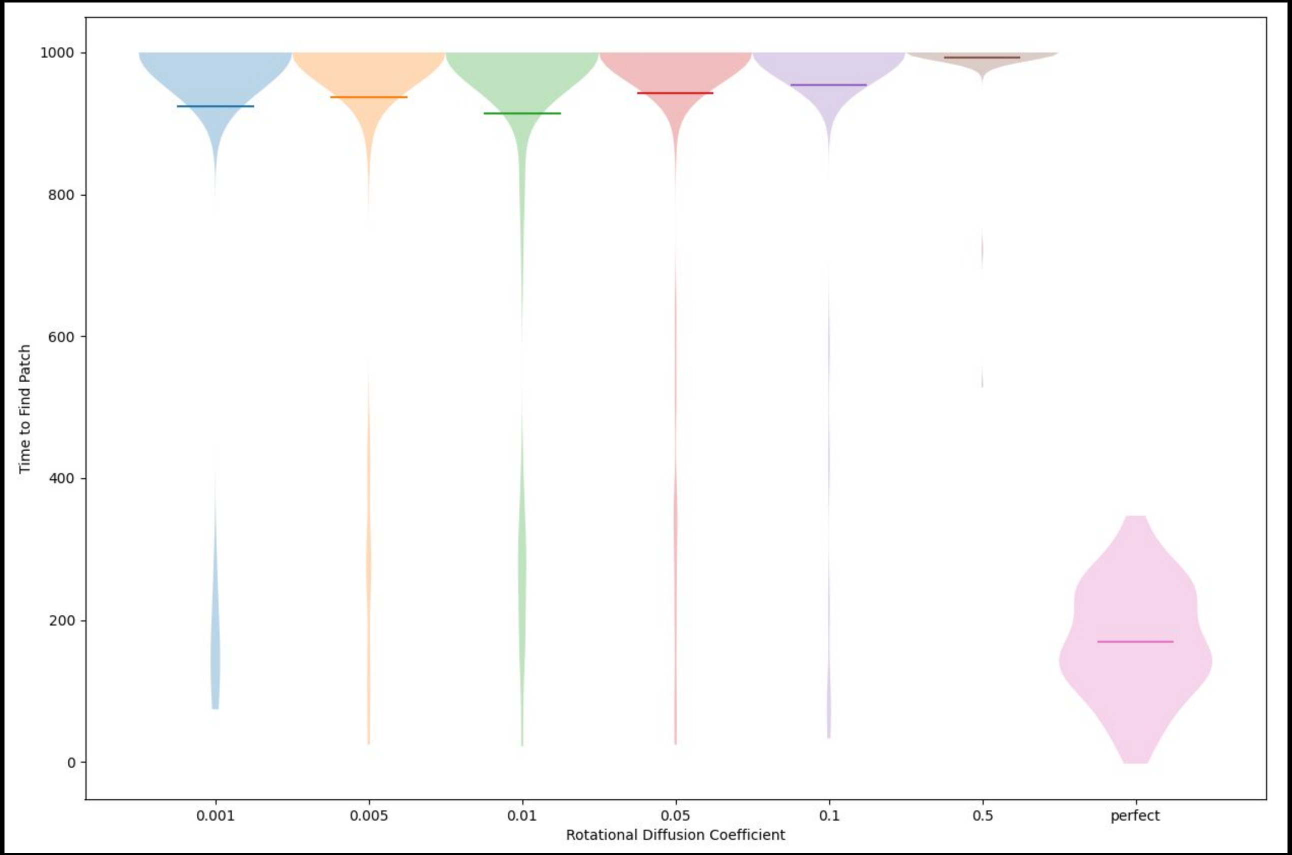
PRW
D = 0.05

PRW
D = 0.01

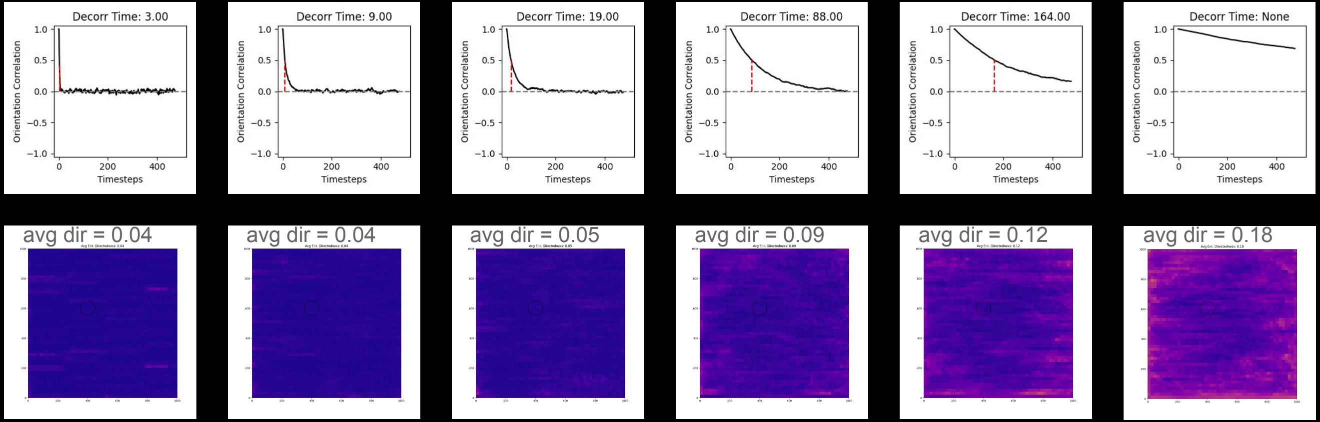
PRW
D = 0.005

PRW
D = 0.001





Persistent Random Walks



D = 0.5 | D = 0.1 | D = 0.05 | D = 0.01 | D = 0.005 | D = 0.001

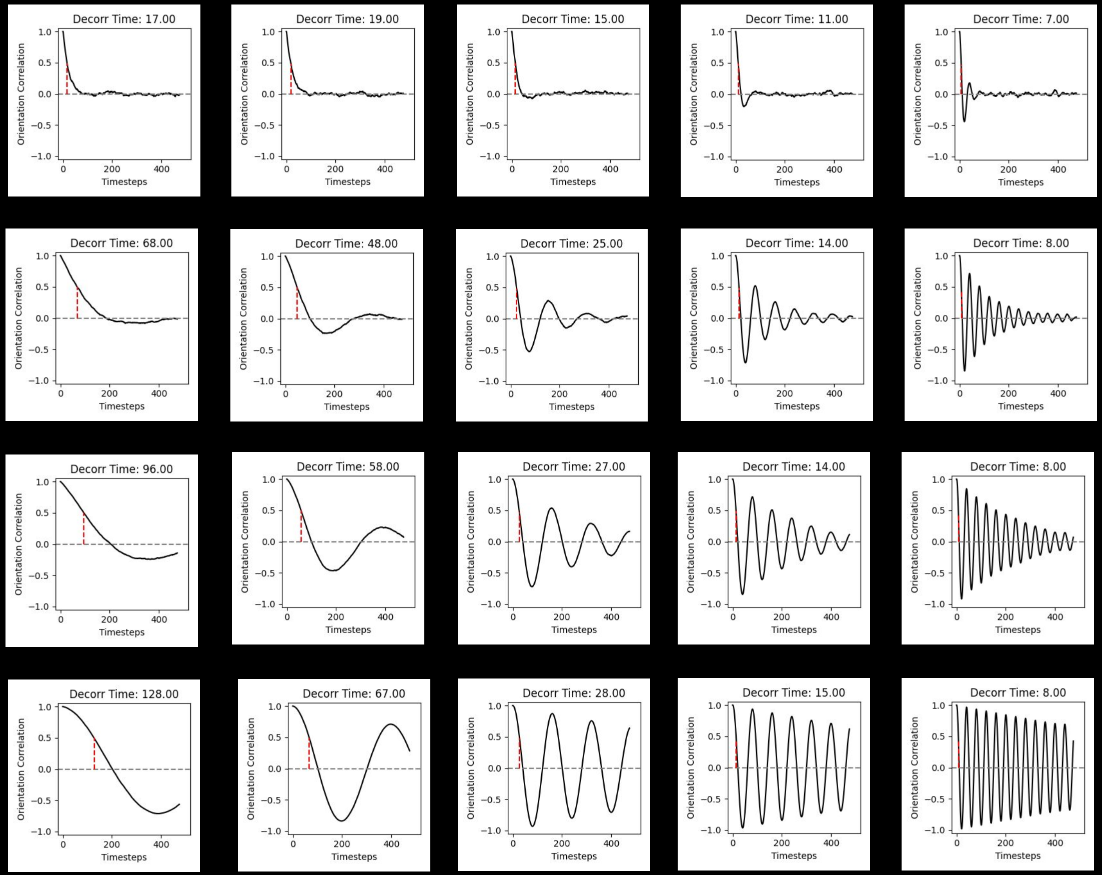
D = rotation diffusion coefficient

Decorrelation times (Temporal)
→ **encompasses observed range**

Directedness (Spatial)
→ much lower
→ random spatial direction, no allocentric bias

Persistent Random Walks
+ Fixed Rate Spin

→
correlated oscillations



D = 0.05

D = 0.01

D = 0.005

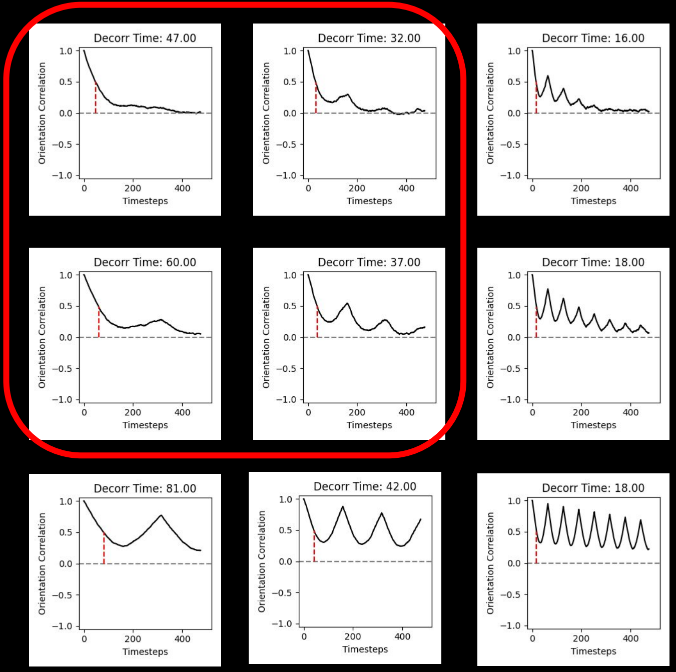
D = 0.001

S = 0.005 S = 0.01 S = 0.025 S = 0.05 S = 0.1

D =
rotation
diffusion
coefficient

S =
fixed spin rate

T = pi * 1/2



S = 0.005 S = 0.01 S = 0.025

D = 0.01

D = 0.005

D = 0.001

Ratcheting back once spins reach threshold (if sufficiently low)

→

Sustained positive correlation

→

Similar plots to Biased Diffusive class (if params in the highlighted regime)

Persistent Random Walks
+ Fixed Rate Spin
+ Ratchet Threshold

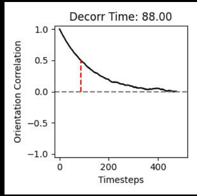
D = rotation diffusion coefficient

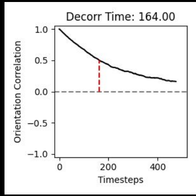
S = fixed spin rate

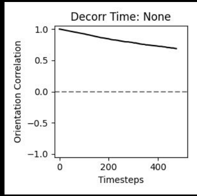
T = spin threshold (before ratcheting back)

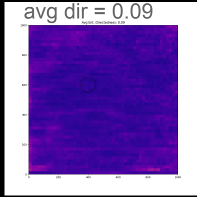
Persistent Random Walks
+ Patch-Oriented Bias

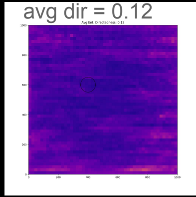
B = 0

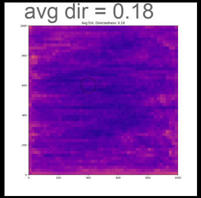






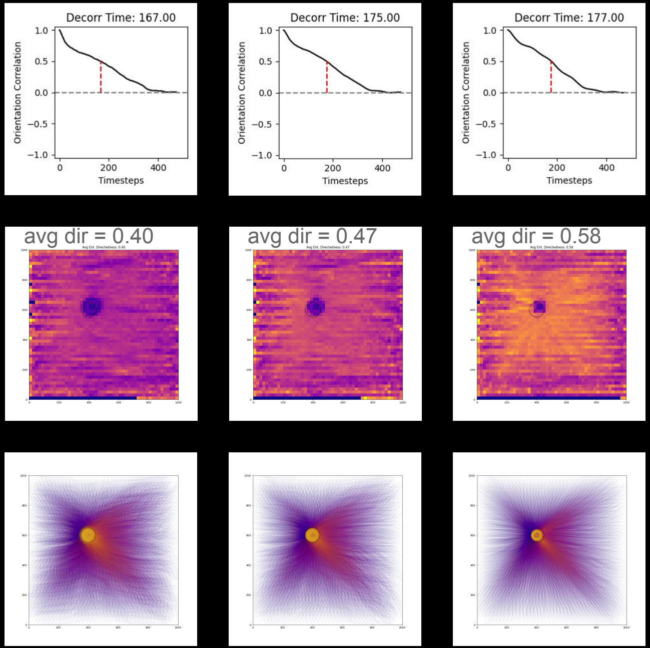






D = 0.01 D = 0.005 D = 0.001

B = bias
(scaling factor on
relative patch
direction info)



D = 0.01 D = 0.005 D = 0.001

B = 0.1

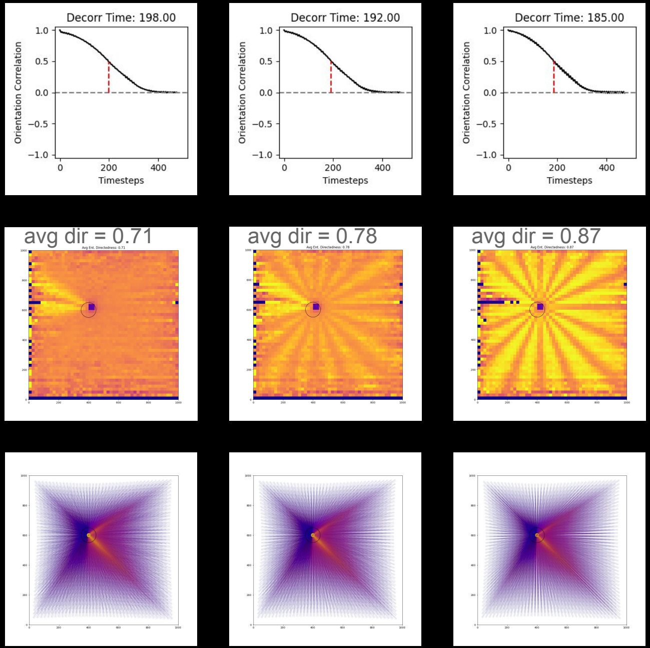
Persistent Random Walks
+ Patch-Oriented Bias

→ correlation concavity

B = bias
(scaling factor on
relative patch
direction info)

Decorrelation times (Temporal)
→ encompasses observed range
(for Direct Pathing agents)

Directedness (Spatial)
→ encompasses observed range
→ allocentric bias → predictable spatial direction
→ star-like plots similar to Direct Pathing



D = 0.01 D = 0.005 D = 0.001

B = 0.75

Persistent Random Walks
+ Patch-Oriented Bias

→ **correlation concavity**

B = bias
(scaling factor on relative patch direction info)

Decorrelation times (Temporal)
→ encompasses observed range
(for Direct Pathing agents)

Directedness (Spatial)
→ encompasses observed range
→ **allocentric bias** → **predictable spatial direction**
→ star-like plots similar to Direct Pathing