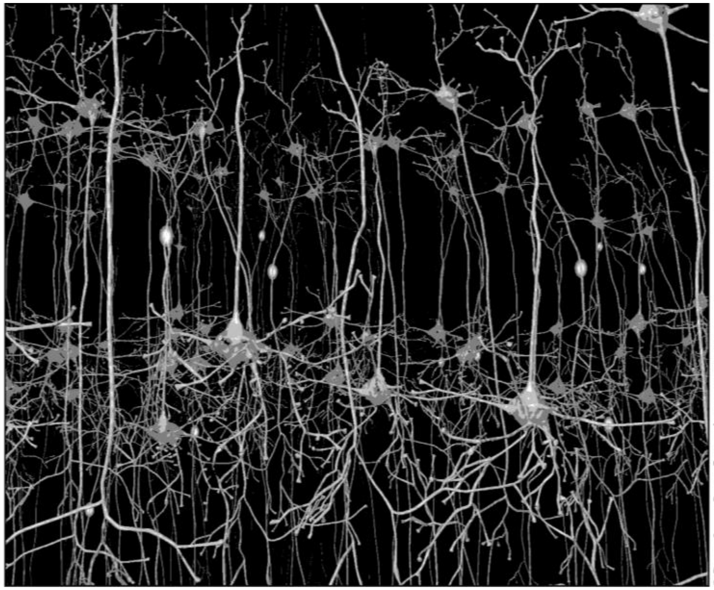
Dynamics and Computation of CANNs

吴思
心理与认知科学学院
IDG/McGovern 脑科学研究所
北大-清华联合生命科学中心
北京大学

Brain computes with neural networks

- The brain is extremely complex: $\sim 10^{11}$ neurons, each with $10^3 \sim 10^4$ connections
- The brain carries out computations relying on neural circuits.


© Digital Studio, Paris - France. All rights reserved.

<u>Connectionism</u>:

- The computation of a neuron is simple.
- Rich network dynamics implement complicated brain functions.

Attractors for robust information representation

- Neurons receive external inputs and interact with each other via synapses (connections), updating the state of a neural network.
- The states of a neural network convey, manipulate, and store information.
- Specifically, information is represented by the stationary states (attractors) of a neural network, such that the information can be retrieved reliably and repeatedly.

Hopfield model

$S_i = \pm 1$: the neuronal state

w_{ij} : the neuronal connection

The network dynamics:

$$S_i = \mathrm{sign}\left(\sum_j w_{ij} S_j - \theta\right)$$

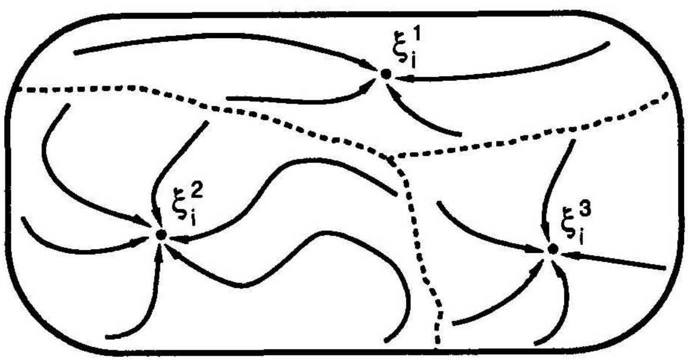
- Analogy to the Ising model in Physics
- The simplest model captures some computational characteristics of a neural network
- Should be the Amari-Hopfield model

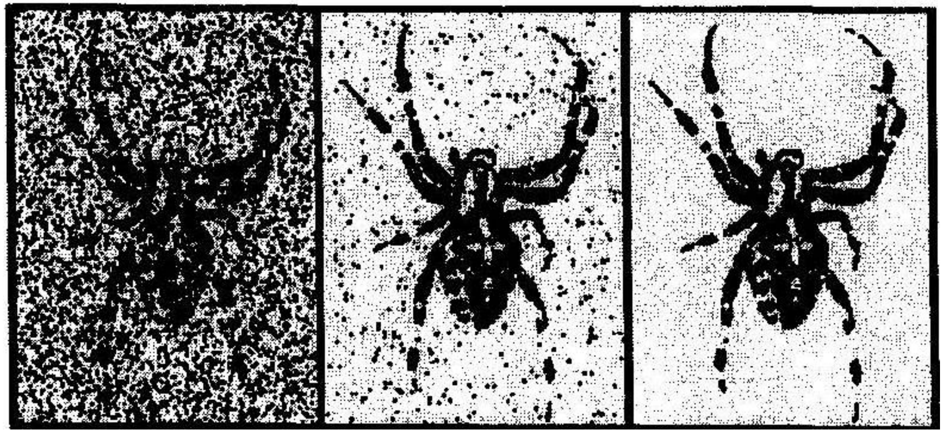
Hopfield model for associative memory

Energy function: $E = -\frac{1}{2}\sum_{i,j} w_{ij} S_i S_j + \theta \sum_i S_i$

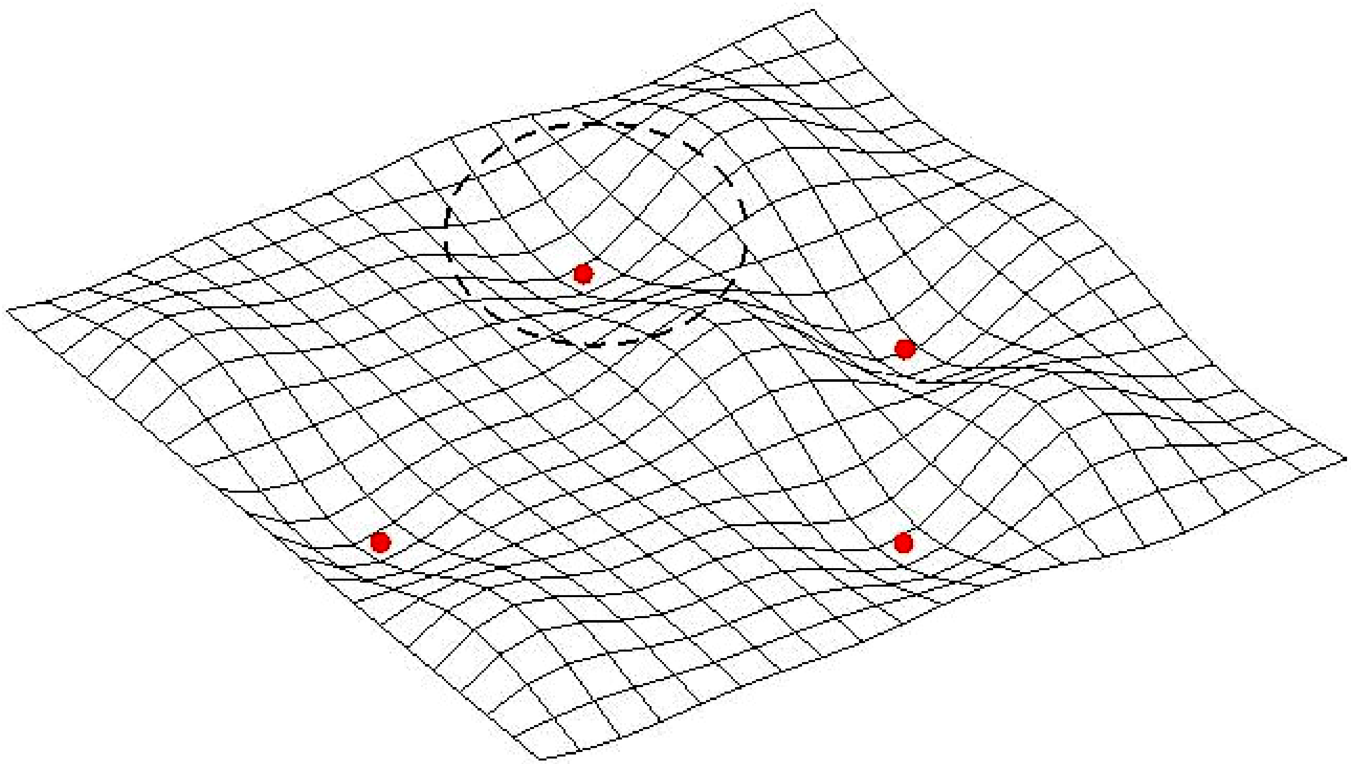
Consider S_i is updated, $S_i(t+1) = sign[\sum_j w_{ij} S_j(t) - \theta]$

$$\begin{aligned}
\Delta E &= E(t+1) - E(t) \\
&= -[S_i(t+1) - S_i(t)]\sum_j w_{ij} S_j(t+1) + \theta[S_i(t+1) - S_i(t)] \\
&= -[S_i(t+1) - S_i(t)][\sum_j w_{ij} S_j(t) - \theta] \\
&\le 0
\end{aligned}$$



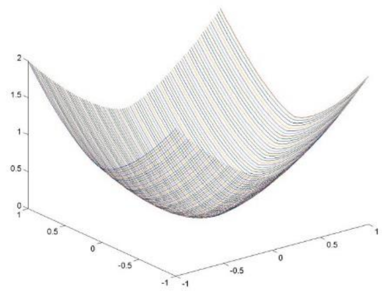


Discrete attractors in the Hopfield model

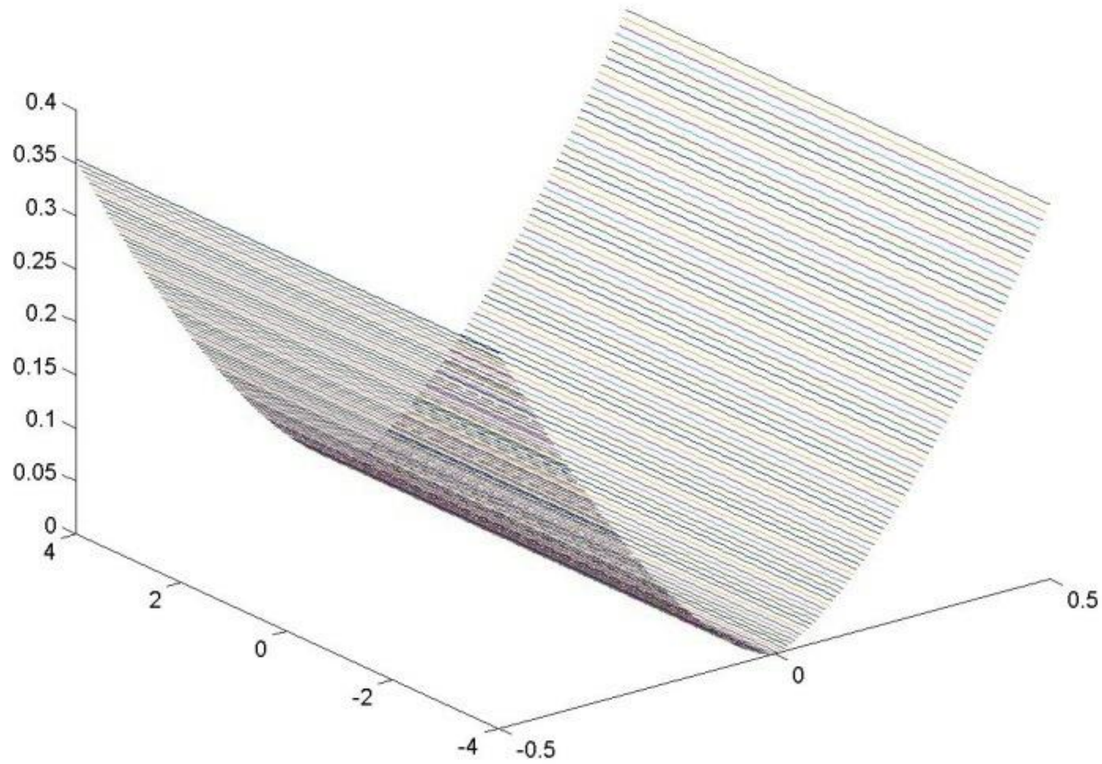


Isolated attractors are too “rigid”, which are good for associative memory, but not appealing for some computations, such as information navigation, information search et al.

Discrete vs. continuous attractors



A discrete attractor

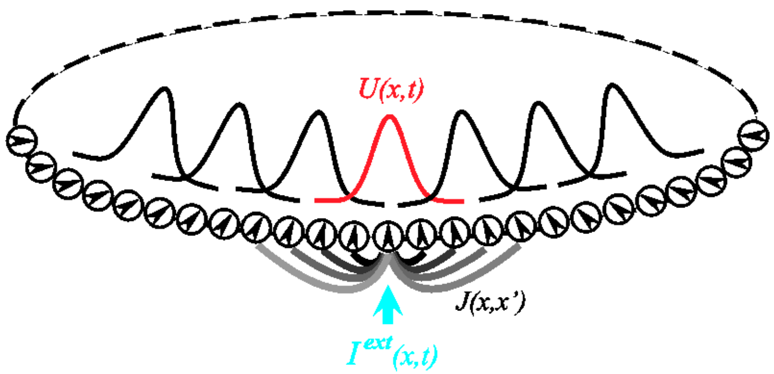


A continuous family of attractors forming a valley, on which the network state is neutrally stable.

Continuous Attractor Neural Network (CANN)

$$\tau \frac{\partial U(x,t)}{\partial t} = -U(x,t) + \rho \int dx' J(x-x') r(x',t) + I^{ext}(x,t)$$

$$r(x,t) = \frac{U(x,t)^2}{1 + k\rho \int dx' U(x',t)^2}; \quad J(x-x') = \frac{J}{\sqrt{2\pi a}} \exp\left[-\frac{(x-x')^2}{2a^2}\right]$$



Key Structure:

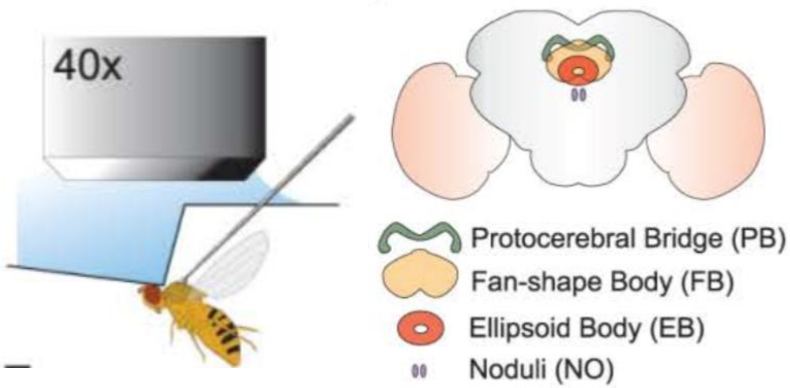
- Bell-shaped recurrent connection strength
- Translation-invariant connection pattern
- Global divisive normalization

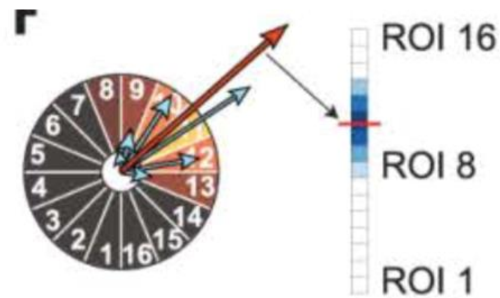
Key Mathematic Properties:

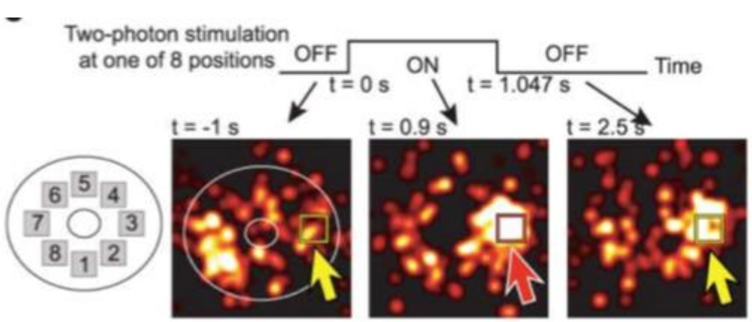
- Recurrent positive-feedback generates attractor, retaining input information
- Divisive normalization avoids exploration
- Translation-invariance ensures many attractors

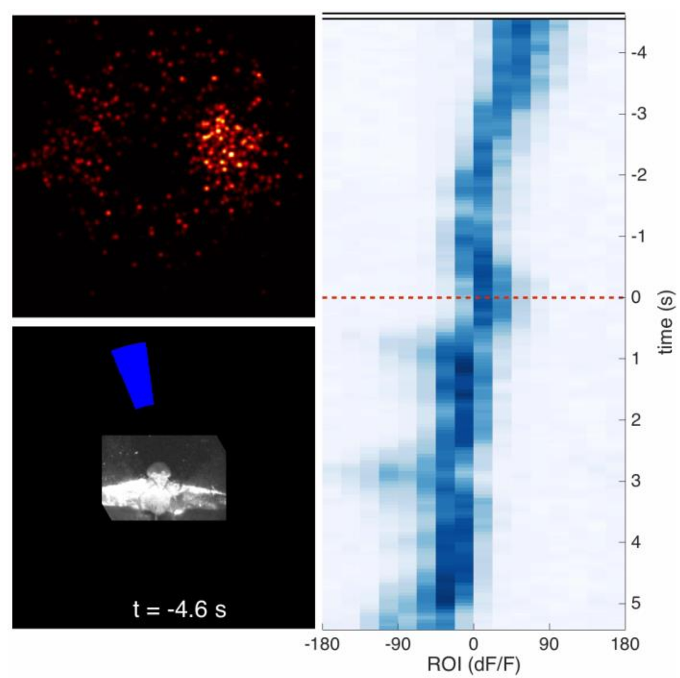
References: 1. Amari, 1977, 2. Ben-Yishai et al., 1995, 3. Zhang, 1996, 4. Seung, 1996, 5. Deneve et al, 1999, 6. Wu et al, 2002, 2005, 2008, 2010, 2012

1D CANN: head-direction system in Drosophila









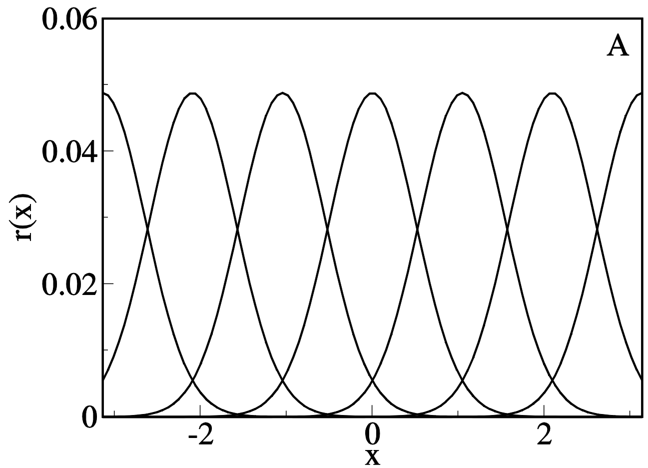
(Kim, Nature, 2017)

A Continuous Family of Stationary States

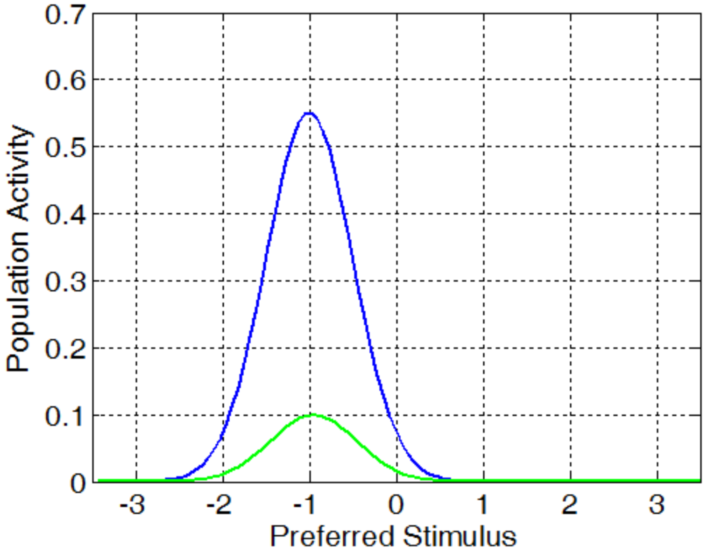
- ◆ Stationary states

$$\bar{U}(x|z) = \frac{A\rho J}{\sqrt{2}} \exp\left[-\frac{(x-z)^2}{4a^2} \right]$$

$$\bar{r}(x|z) = A \exp\left[-\frac{(x-z)^2}{4a^2} \right]$$



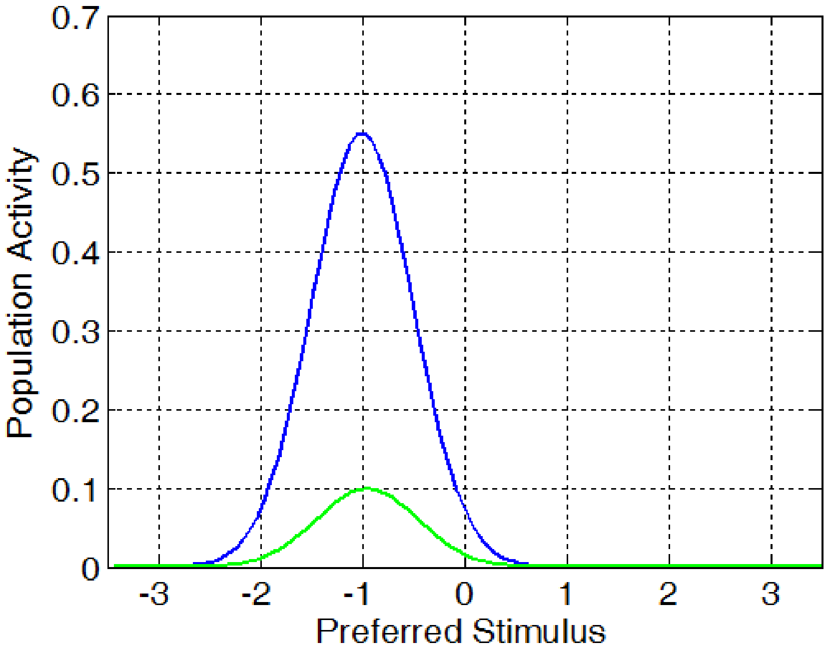
- ◆ Smooth tracking



Mobility of a CANN is still not enough

- In response to noises, the bump movement is Brownian motion on the representation space—not optimal information searching!

- In response to a moving input, the tracking of the bump is delayed.



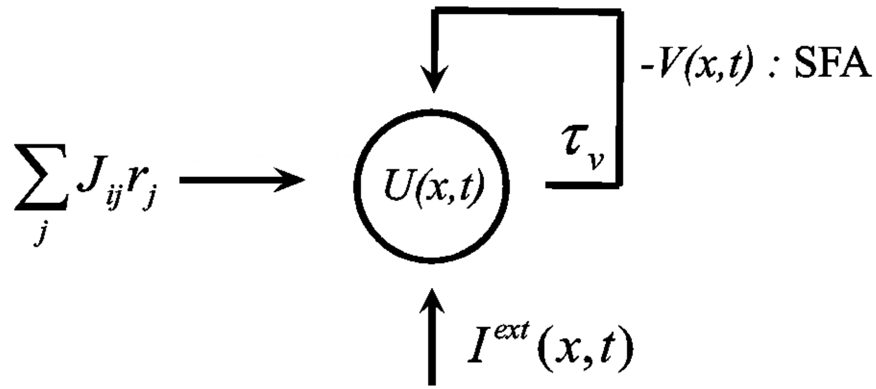
A CANN with Adaptation

$$\tau \frac{dU(x,t)}{dt} = -U(x,t) + \rho \int dx' J(x-x') r(x',t) - V(x,t) + I^{ext}(x,t)$$

$$\tau_v \frac{dV(x,t)}{dt} = -V(x,t) + mU(x,t)$$

V(x,t) represents the SFA effect,

$$V(x,t) = \frac{m}{\tau_v} \int_{-\infty}^{t} e^{-\frac{t-t'}{\tau_v}} U(x,t')dt'$$



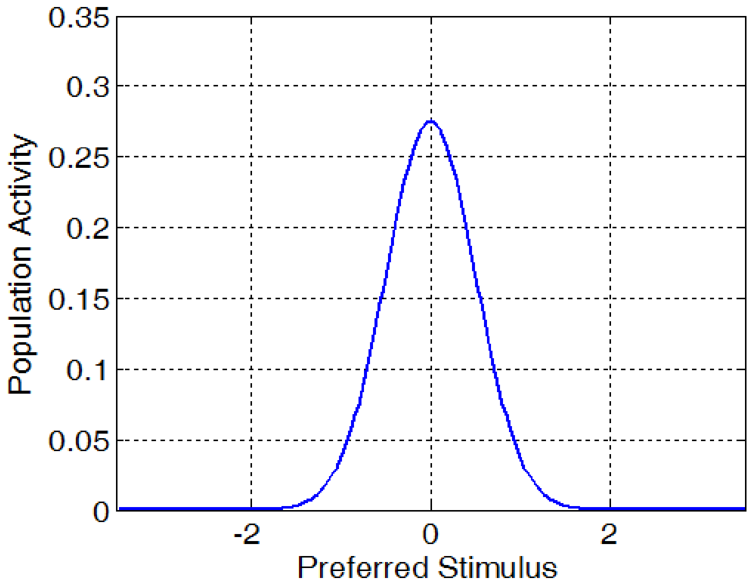
Spike frequency Adaptation (SFA):

- Neuronal response attenuates after experiencing prolonged firing.
- Slow negative feedback modulation to neuronal response.

Mi et al., NIPS 2014

Traveling wave state

<u>Traveling Wave</u>: a moving bump in the network without relying on external drive.
<u>The mechanism</u>: SFA suppresses localized neural activity and triggers the bump to move.



$$v_{\text{int}} = \frac{2a}{\tau_v}\sqrt{\ln \frac{m\tau_v}{\tau}}$$

Dynamics of a CANN with Adaptation

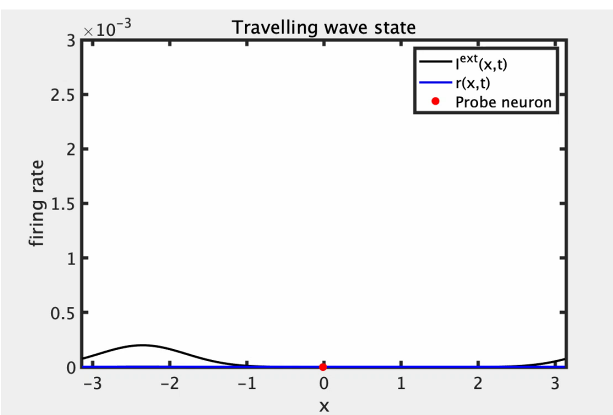
External moving input:

$$I^{ext}(x,t) = \alpha\, exp\left\{-\frac{[x - v_{ext}t]^2}{4a^2}\right\} \quad (6)$$

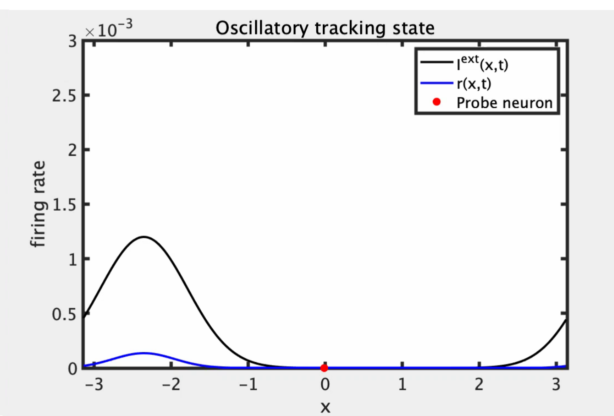
Spike frequency adaptation:

$$\tau \frac{\partial \mathrm{V}(x,t)}{\partial t} = -\mathrm{V} + \mathrm{m}U \quad (4)$$

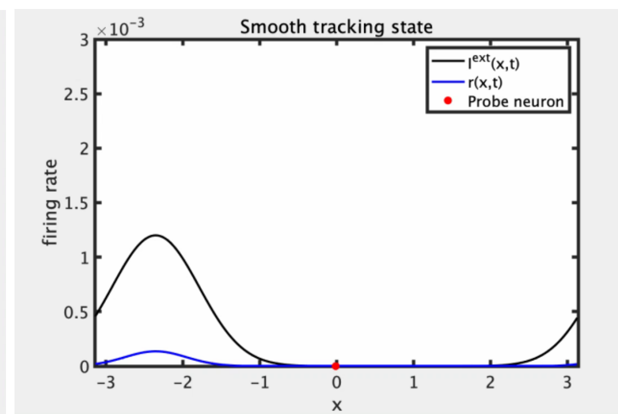
Travelling Wave



Oscillatory Tracking

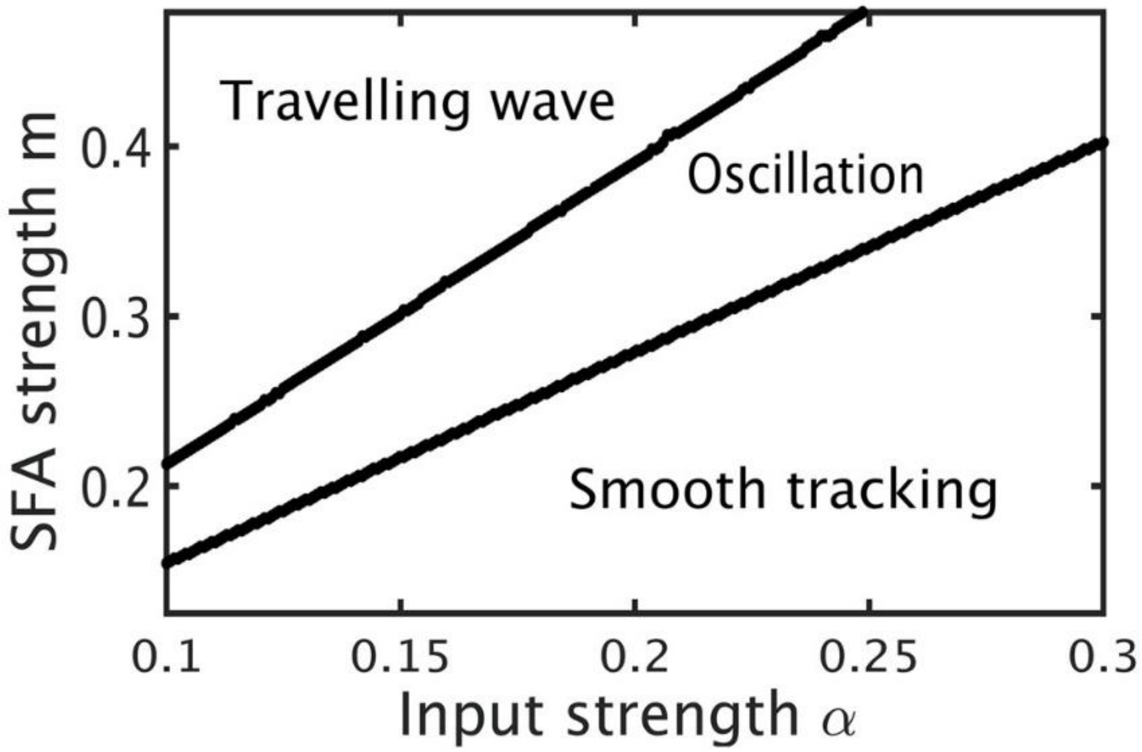


Smooth Tracking



Input strength α increase

Phase diagram of the network

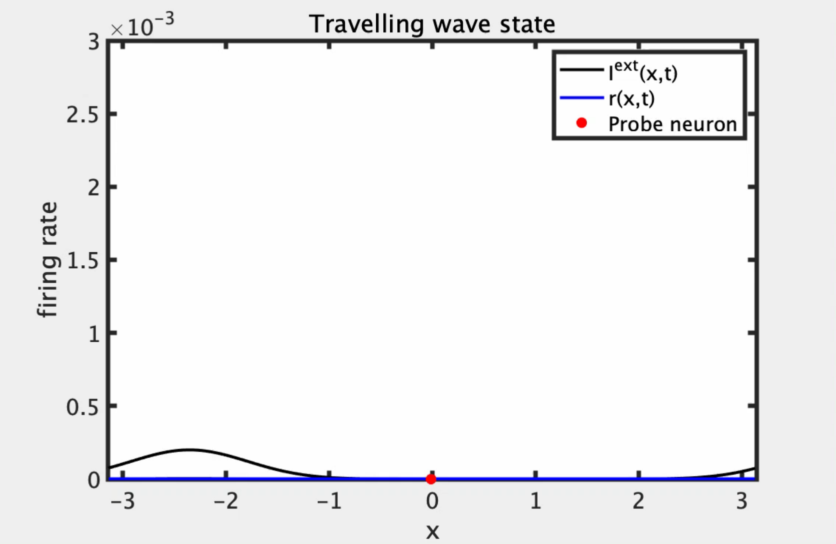


I. Computing with travelling wave

Traveling wave state

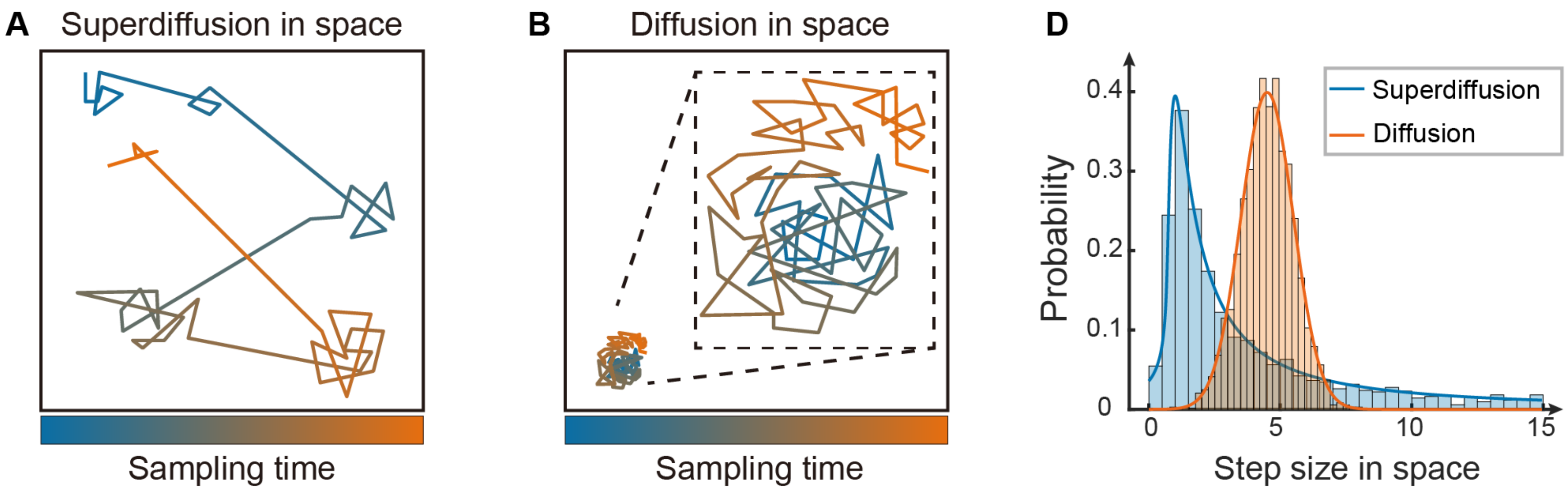
<u>Traveling Wave</u>: a moving bump in the network without relying on external drive.

<u>The mechanism</u>: SFA suppresses localized neural activity and triggers the bump to move.



$$v_{\text{int}} = \frac{2a}{\tau_v}\sqrt{\ln \frac{m\tau_v}{\tau}}$$

Lévy-flight vs. Brownian motion



$$p(x) \sim x^{-u}$$

or

$$p(x) \sim x^{-1-\alpha}$$

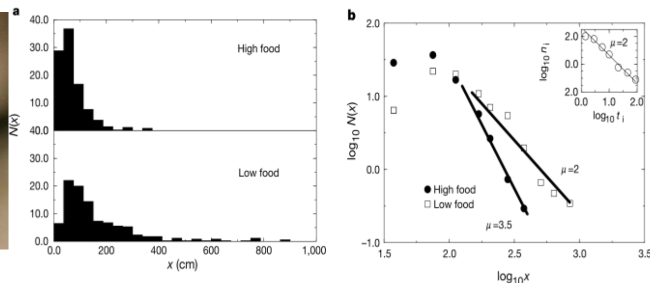
$0 < \alpha < 2$

Superdiffusion / Lévy motion

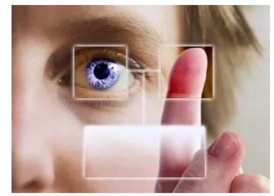
$\alpha \geq 2$

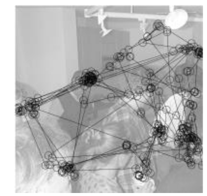
Diffusion / Brownian motion

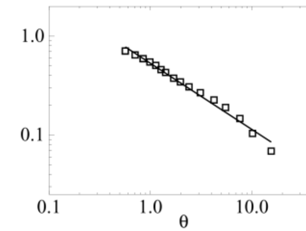
Lévy flights in ecology and human cognitive behaviors



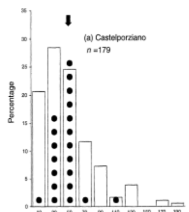
Heinrich, B. (1979). *Oecologia*

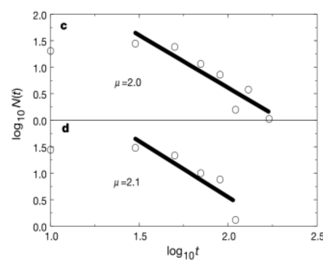




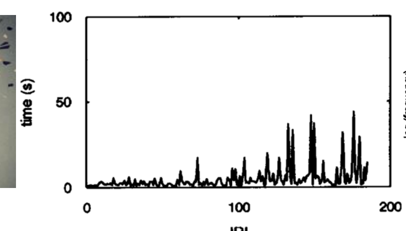


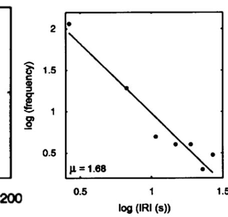
Brockmann, D., & Geisel, T. (2000). *Neurocomputing*





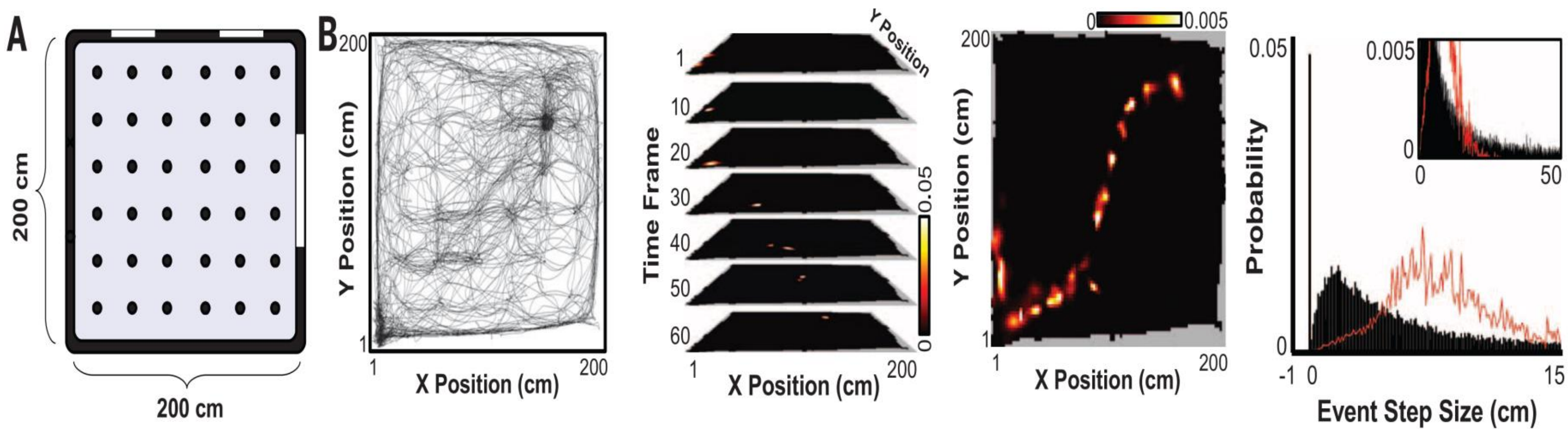
Focardi et al. (1996) *Journal of Animal Ecology*





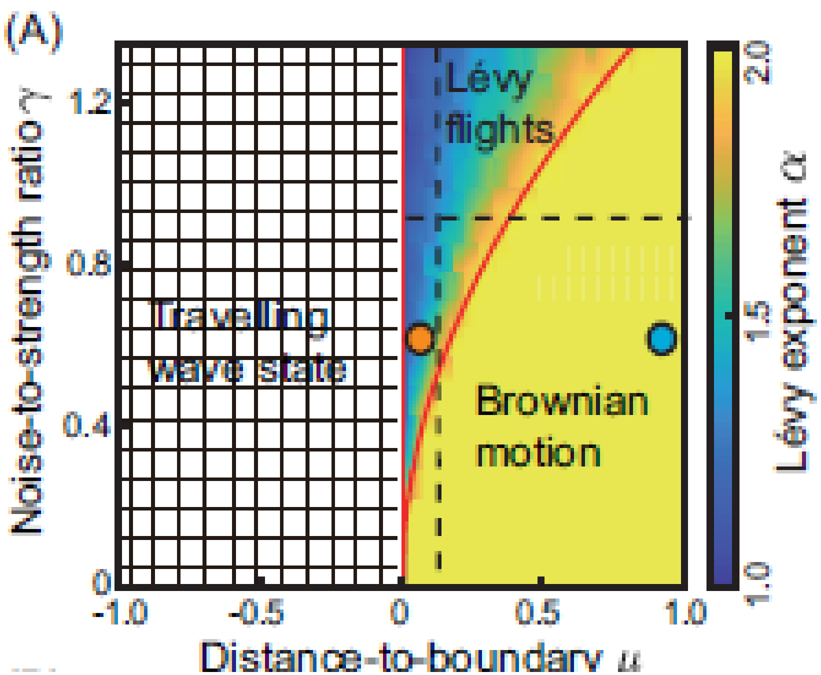
Rhodes, T., & Turvey, M. T. (2007). *Physica A*

Lévy flights in brain dynamics

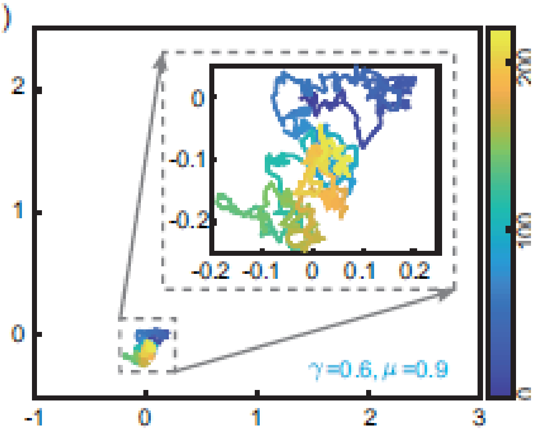


Pfeiffer, B. E., & Foster, D. J. (2015). *Science*.

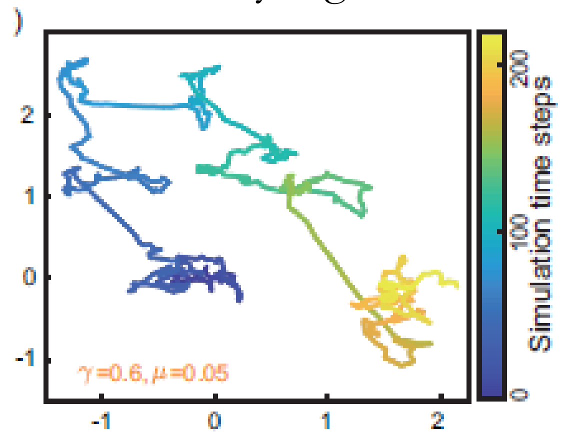
Noisy adaptation induces Levy flight



Brownian motion

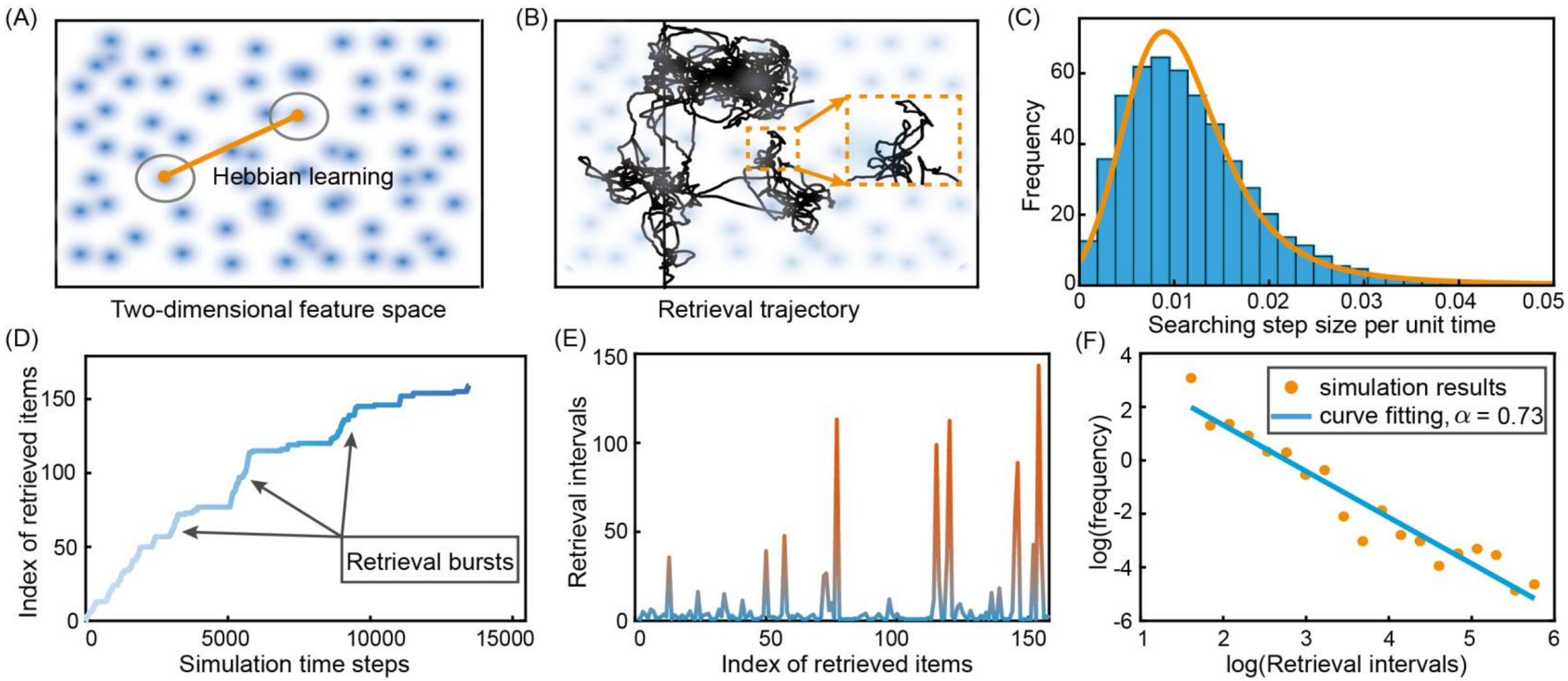


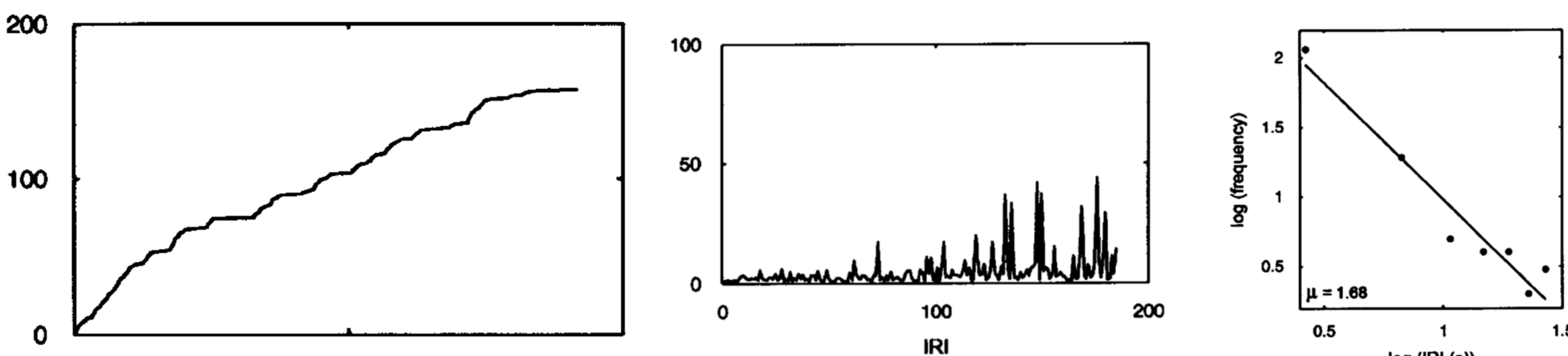
Levy flight



Dong et al. (submitted)

Modeling human free memory recall





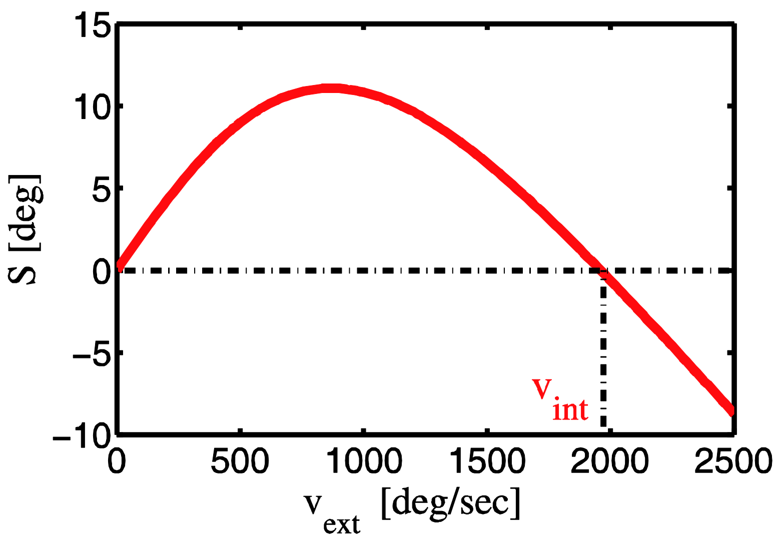
Rhodes, T., & Turvey, M. T. (2007). *Physica A*

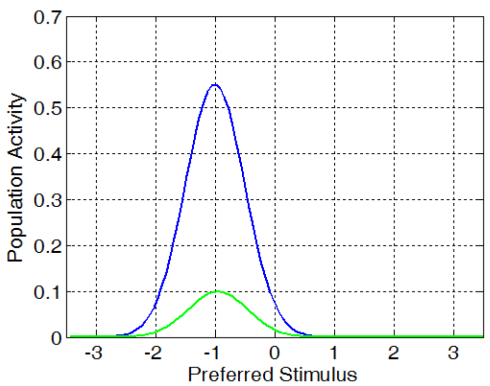
II. Computing with anticipative tracking

Smooth tracking state:
Interplay between Intrinsic Mobility & External Drive

When $v_{\text{int}} > v_{ext}$, bump leading external drive

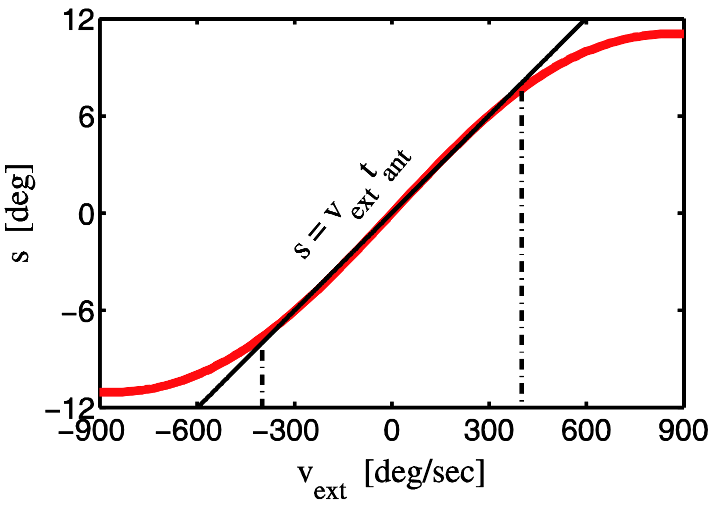
When $v_{\text{int}} < v_{ext}$, bump lags behind





No SFA, delayed tracking

Anticipative Tracking

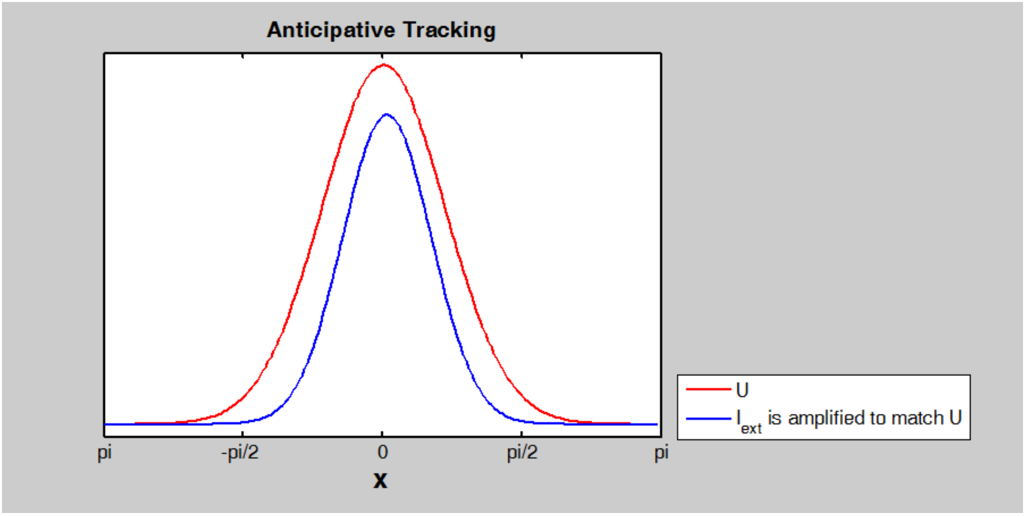


The parameter conditions:

$$m > \frac{\tau}{\tau_v}; \quad v_{ext} \square \frac{a}{\tau_v}$$

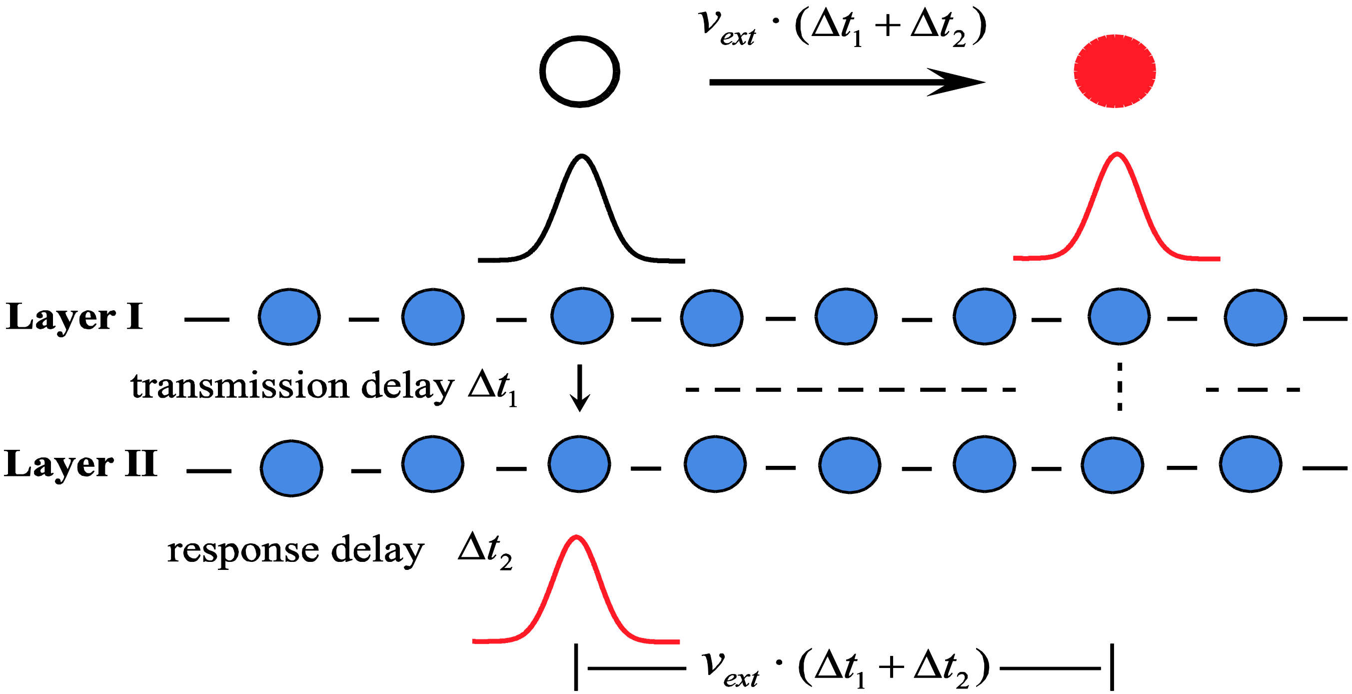
Observed in ADN, LMN

(Blair et al., J. Neurosci. 1995; Neuron 1998)

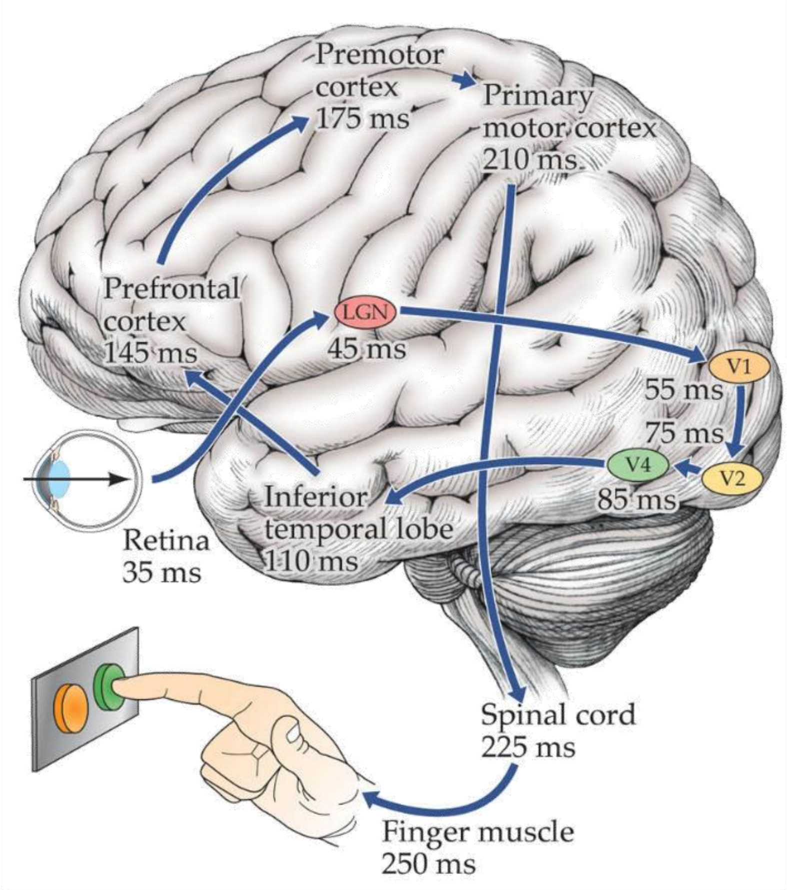


Mi et al. NIPS 2014

Time delay is inevitable in hierarchical neural information processing



Time delay is significant in neural signal transmission



(e.g. Maunsell and Gibson 1992; Raiguel et al. 1989; Nowak et al. 1995; Schmolesky et al. 1998; Thorpe, Fize, & Marlot 1996)

Prediction is required to compensate for time delays

From retina to V1 ~50ms

- Federer's serve speed: ~200km/h
- 50 ms delay implies

  displacement ~ 3 m!

A Life or Death

- Escaping from a predator

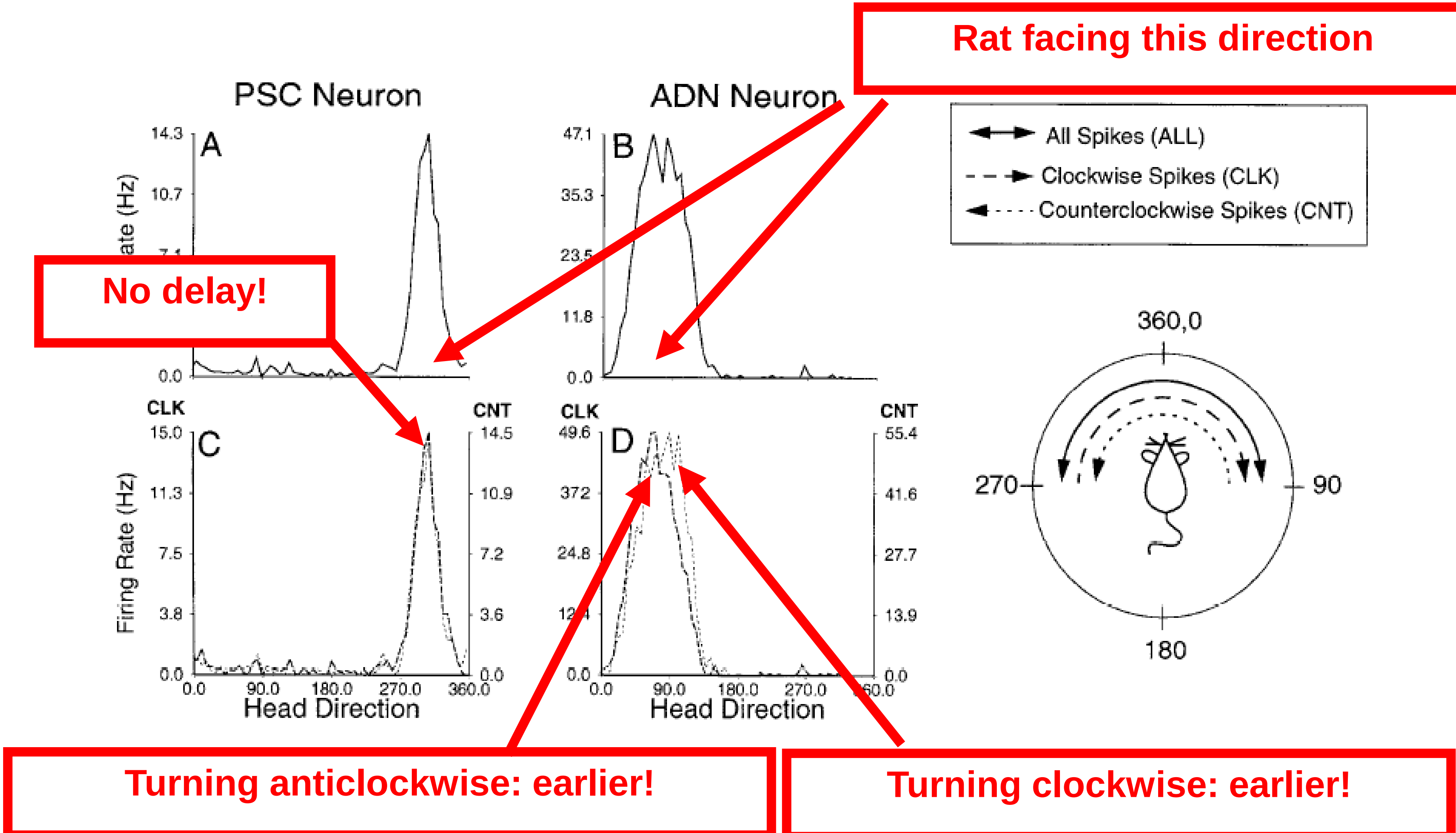
- Catching a prey

We are good at prediction!

躲闪帝(escaping king)

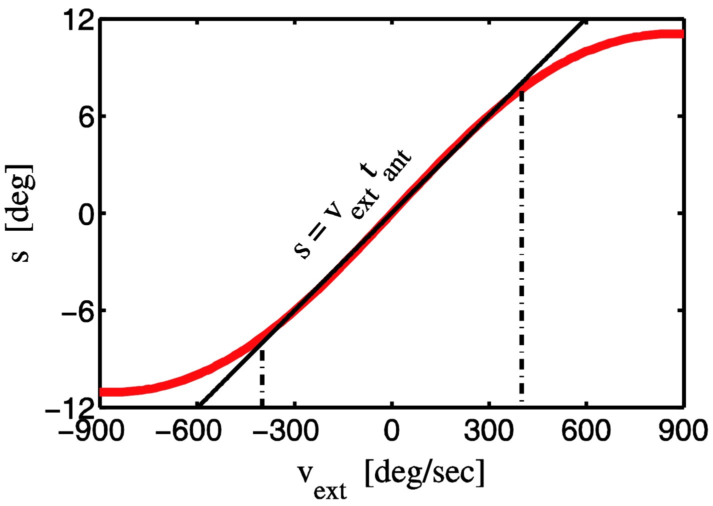
奔跑哥 (running brother)

Prediction of head-direction



Blair and Sharp, J. Neurosci. (1995)

Anticipative Tracking

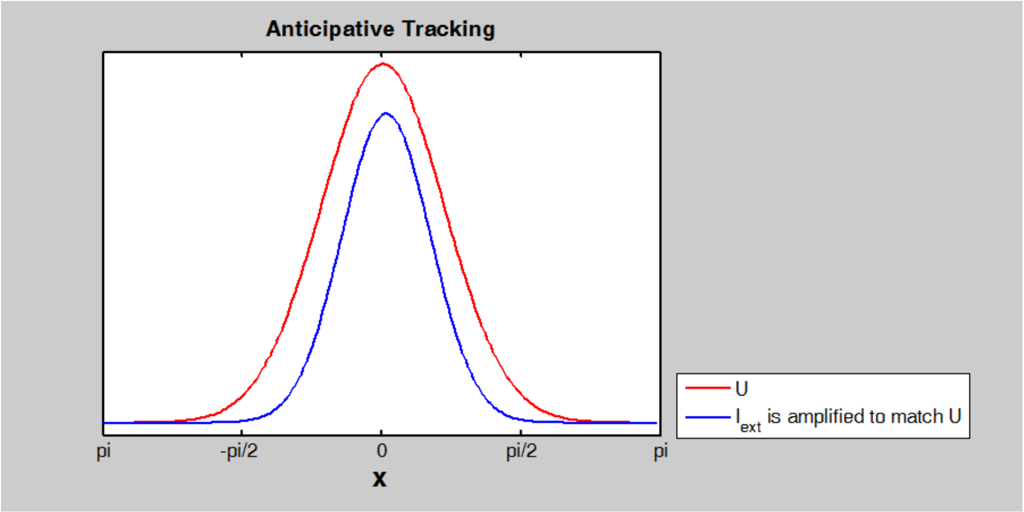


The parameter conditions:

$$m > \frac{\tau}{\tau_v}; \quad v_{ext} \square \frac{a}{\tau_v}$$

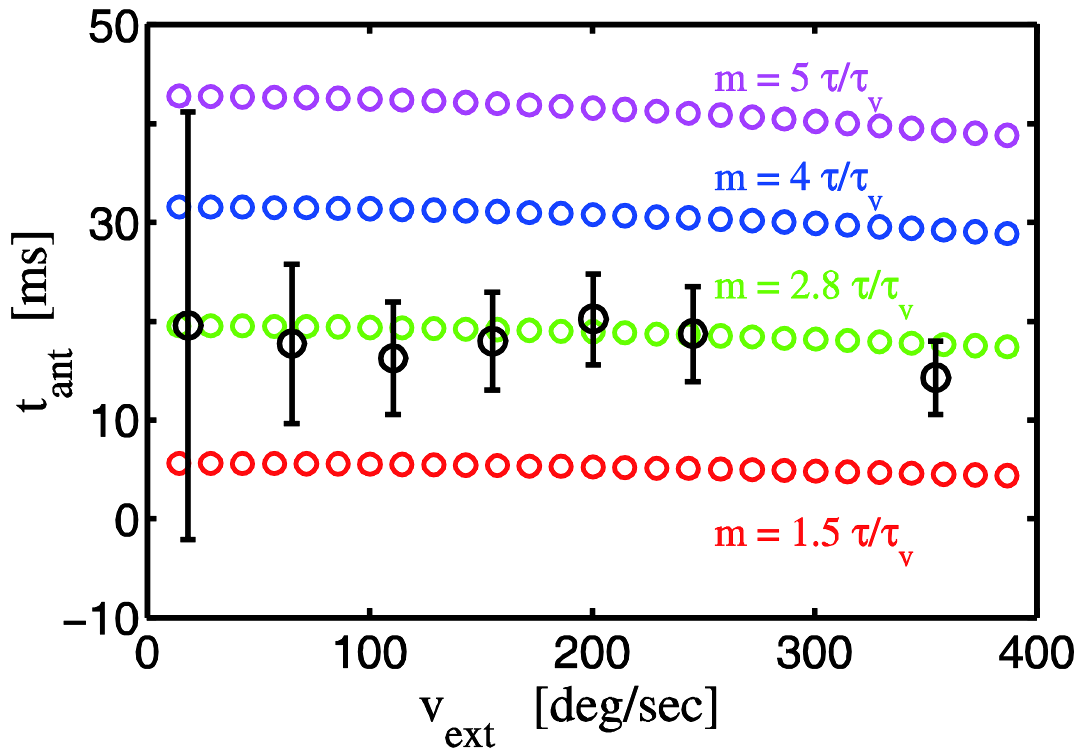
Observed in ADN, LMN

(Blair et al., J. Neurosci. 1995; Neuron 1998)



Mi et al. NIPS 2014

Adaptive Tracking



The anticipation time t_{ant} is insenstive to input speed

t_{ant} is controllable by adjusting the SFA effect m

In ADN, t_{ant} □ $25ms$, irrespective to input speed

Why Neural Delays?

- **Advantages**
  - To integrate temporal information over time for reliable responding
  - To integrate multiple sensory cues
  - To implement temporal code
  - And many others
- **Disadvantages**
  - Delayed response to fast moving objects or varying temporal information

The art of being slow

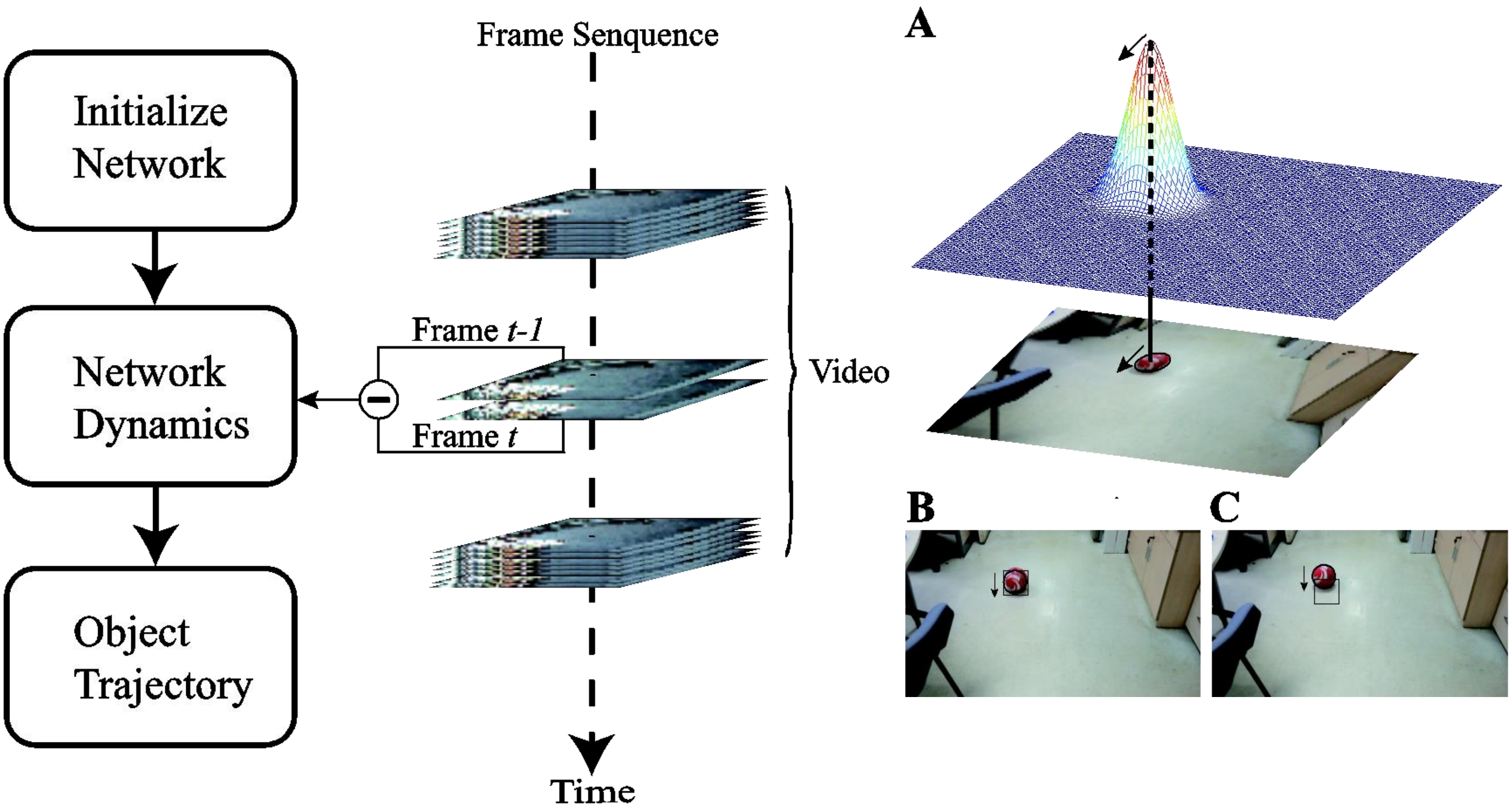
Two sides of the same coin

- Every animal adopts to its own optimal time scale suitable for its own survival in the natural environment
- The brain co-evolves strategies to compensate delays

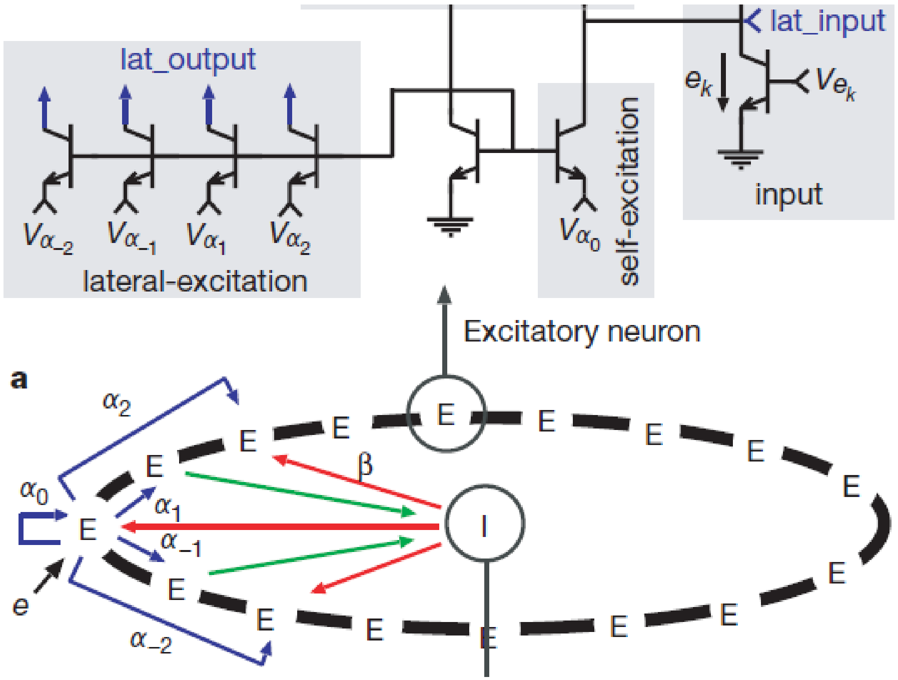
Appealing properties of CANN tracking

- Network computation, no extra computational effort (e.g., feature extraction)
- Mobility, implementable by neuromorphic devices
- Adaptive to speed
- Guided by formal theory
- Robust to noises

A CANN for Object Tracking



CANN implementation on Chip

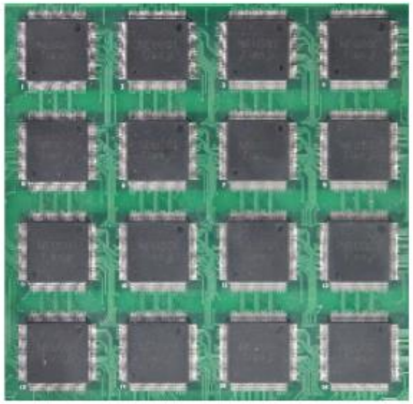


Nature 2000

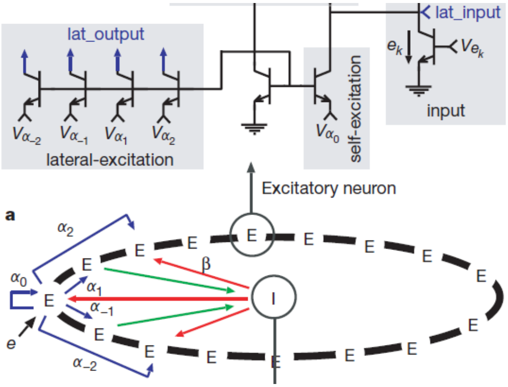
IBM True North

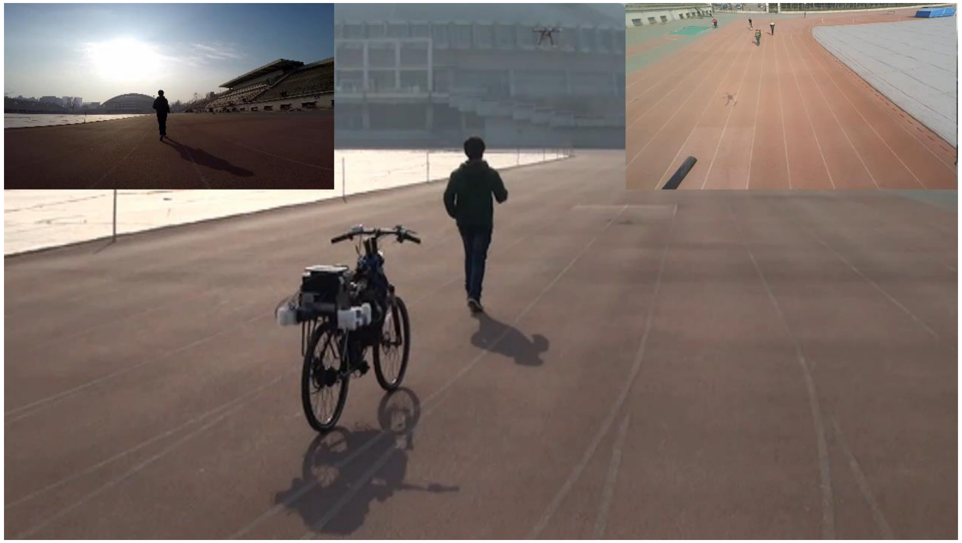
Object Tracking on Chip

“Tianji” Chip



CANN on Chip

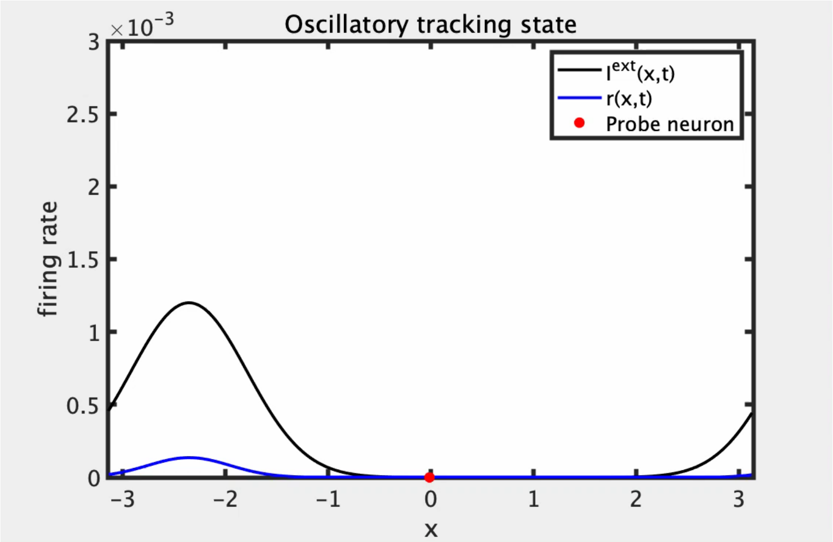


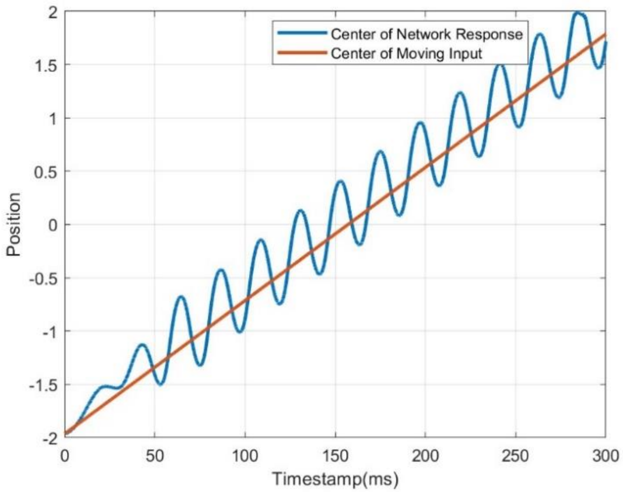


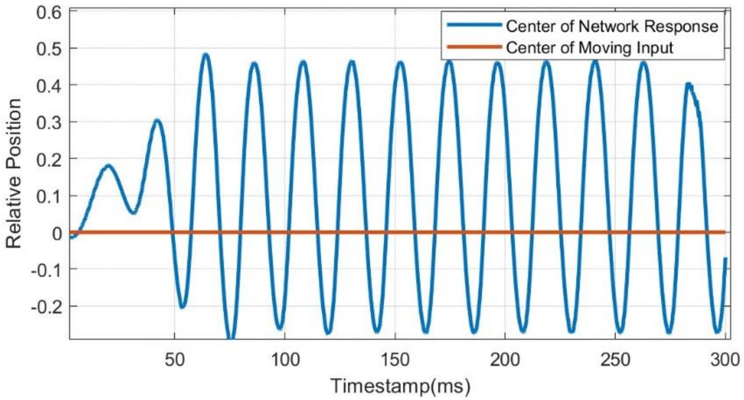
In collaboration with
Tsinghua Univ.
Pei et al., Nature 2019

III. Computing with oscillatory tracking

Oscillatory tracking state







$$z(t) = vt + s(t) = vt + c_0 \sin(\omega t) + d_0$$

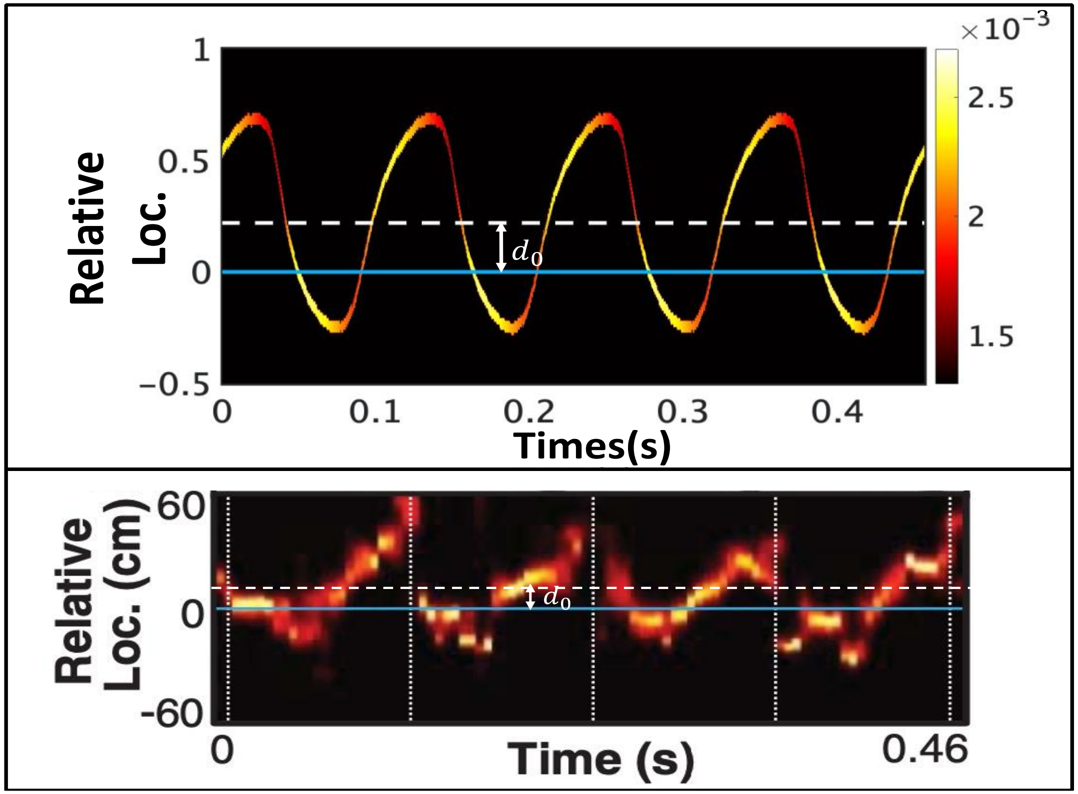
$$\omega = \sqrt{\frac{2\sqrt{\pi}ak(1+m)\alpha}{\tau\tau_v(J_0 + 2\sqrt{\pi}ak\alpha)}}$$

Chu et al. Cosyne 2021

Oscillatory tracking generates alternating forward- and reverse- ordered sequences

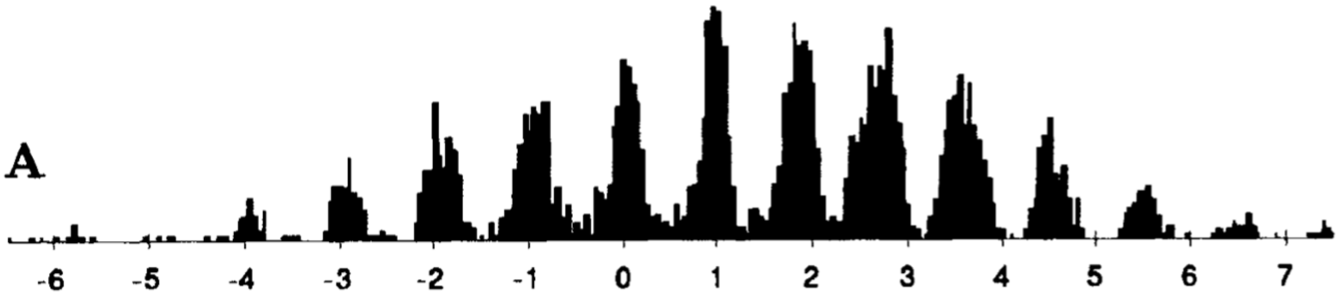
Model:

Data:

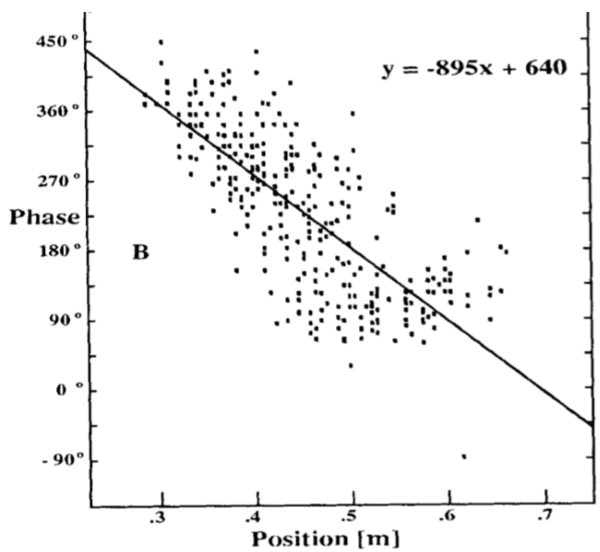


Wang et al. Science 2020.

Bursting Responses and Phase Precession of a Place Cell

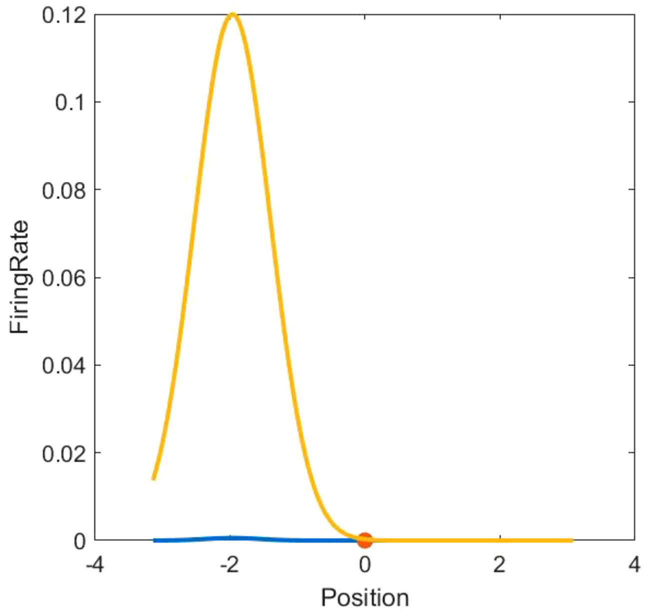


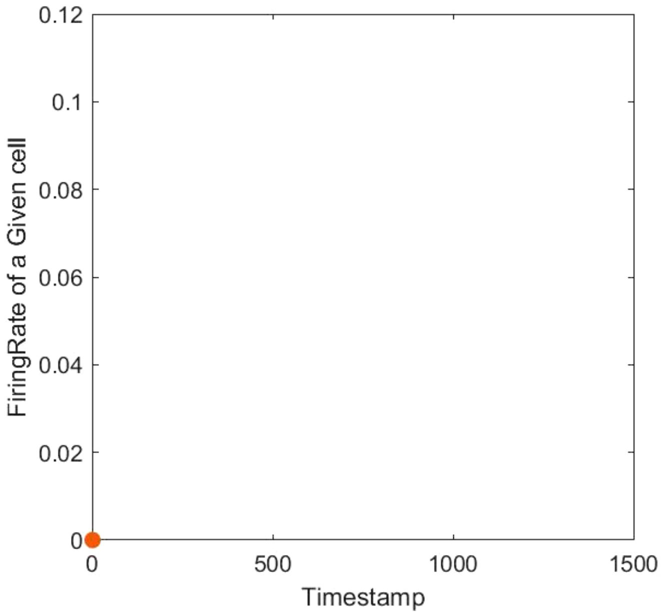
Skaggs et al. Hippocampus, 1996

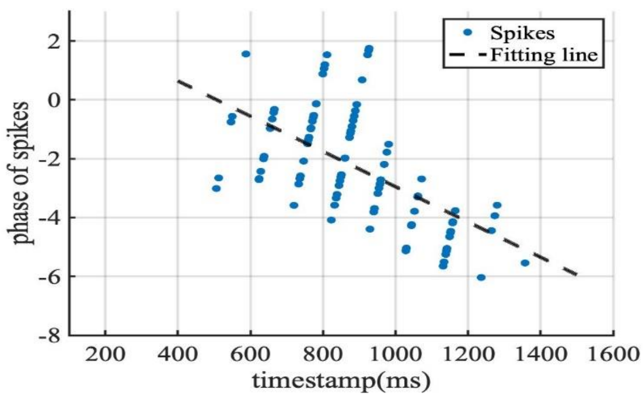


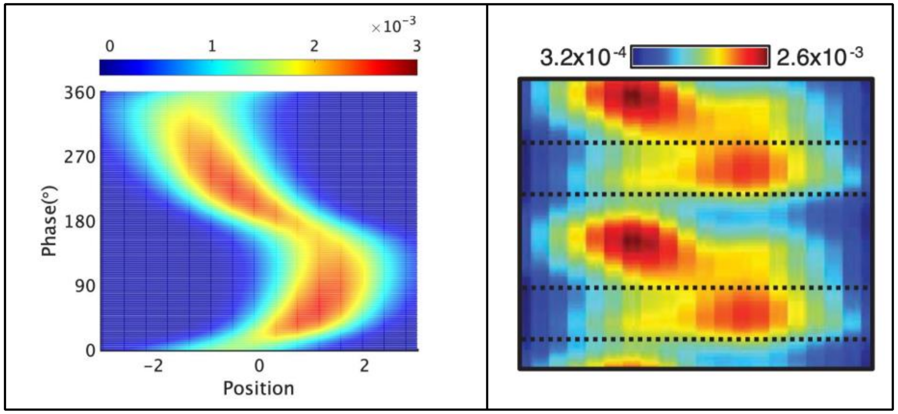
O'keefe and Recce, Hippocampus, 1993

Oscillatory tracking generates phase pre- and pro-cession



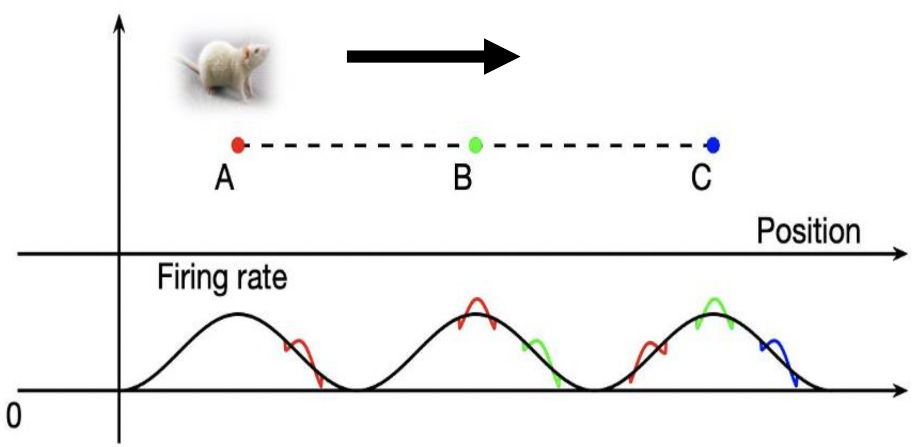




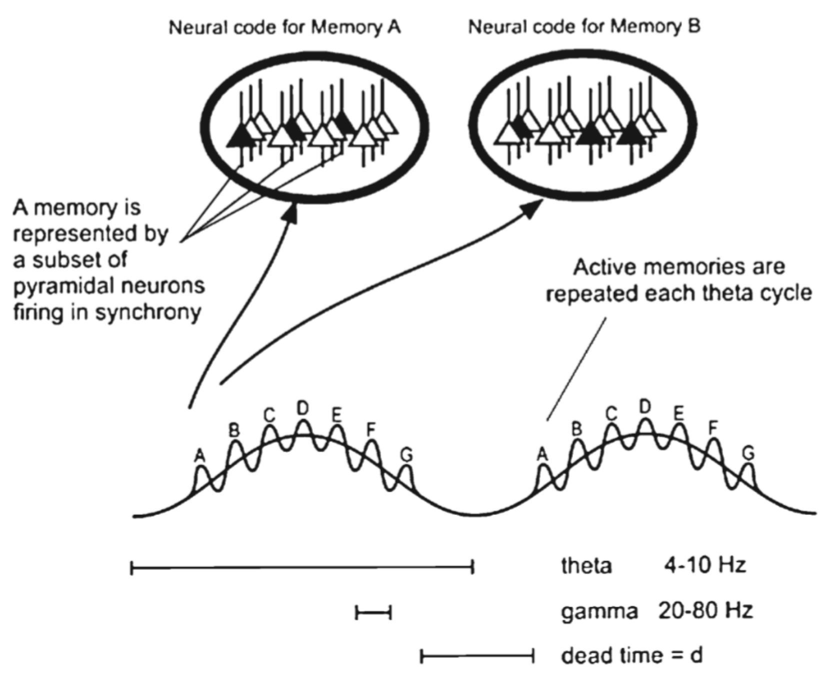


Functions of phase precession?

Encoding temporal order



Working memory



Summary

- Brain performs computation via network dynamics.
- Information is encoded as attractors of network dynamics.
- Hopfield model is appealing for associative memory.
- CANN is appealing for processing continuous information.
- A generic attractor network encodes not only memory items, but also their relationships.
- Adaptation destabilizes attractors, realizing anticipative tracking and efficient information search.
- CANN can be applied in AI.