The Neuronal Replicator Hypothesis: novel mechanisms for information transfer and search in brain ?

Eörs Szathmáry

Eötvös University, Budapest

Chrisantha Fernando

Queen Mary, University of London

PARMENIDES FOUNDATION

Munich

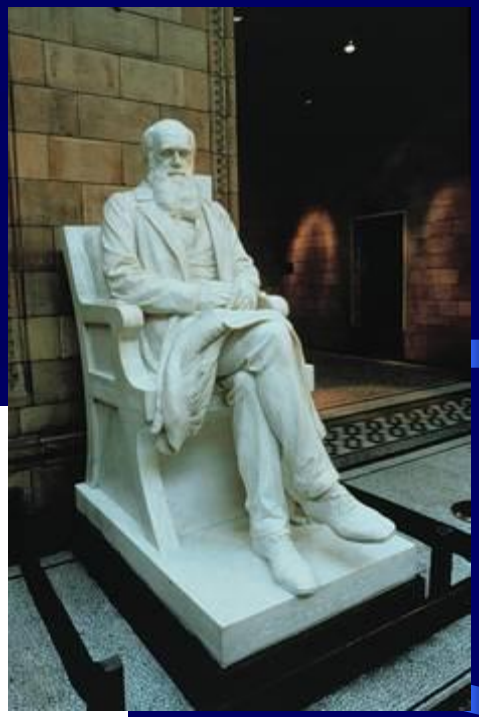
Exploring the idea

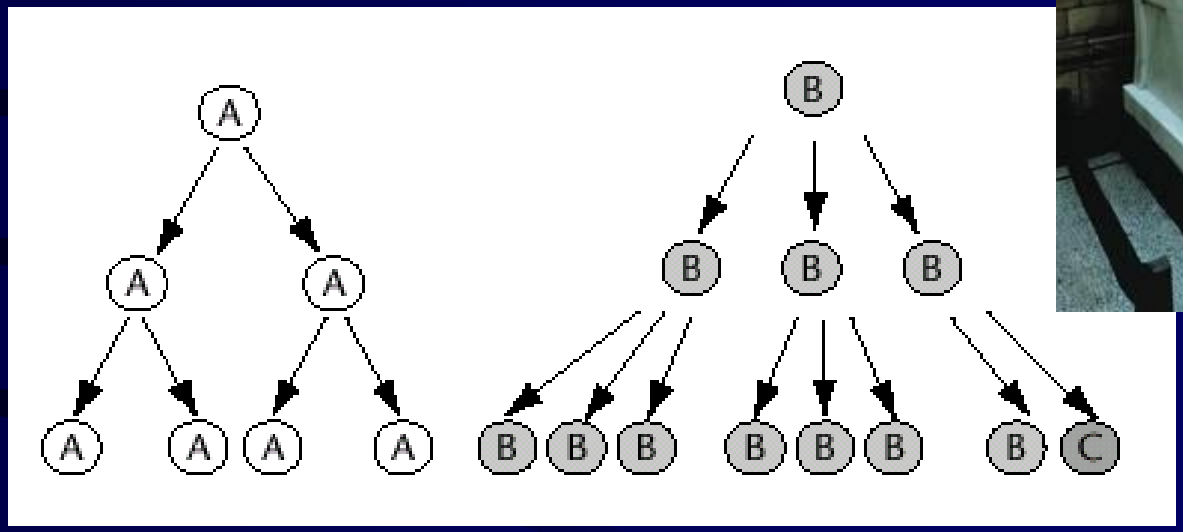
- Search for *a potentially missing piece* of the puzzle
- Not to claim that this could potentially explain everything
- Not to claim that other ideas do not explain anything
- Learning from biology AND evolutionary algorithms
- One of the main motivations: language, specifically Fluid Construction Grammar

Monod (1971)

- *For a biologist it is tempting to draw a parallel between the evolution of ideas and that of the biosphere. For while the abstract kingdom stands at a yet greater distance above the biosphere than the latter does above the nonliving universe,* ***ideas have retained some of the properties of organisms***
- *Like them, they tend to perpetuate their structure and to breed; they too can* ***fuse, recombine, segregate their content; indeed they too can evolve****, and in this evolution* ***selection*** *must surely play an important role.*
- *... one may at least try to define some of the principal factors involved in it. This* ***selection*** *must necessarily operate at* ***two levels: that of the mind itself and that of performance.***

Units of evolution: a tacit ‘algorithm’





1. multiplication
2. heredity
3. variability

Some hereditary traits affect survival and/or fertility

Units of selection and units of evolution!

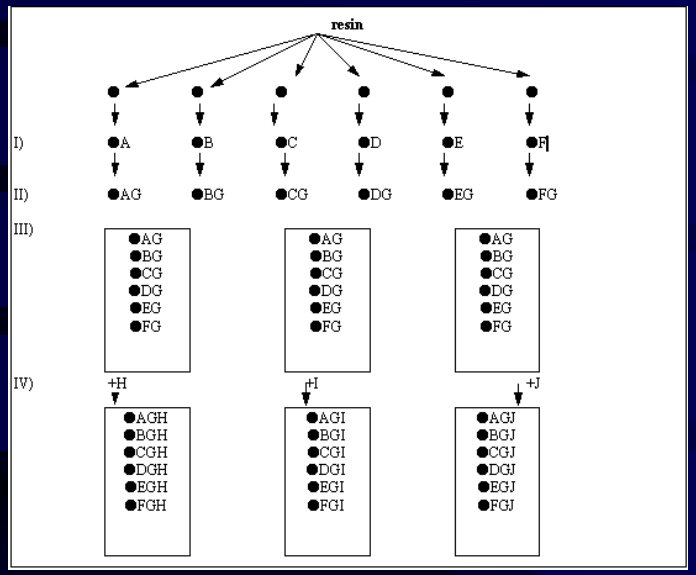
1. multiplication
2. heredity
3. variability

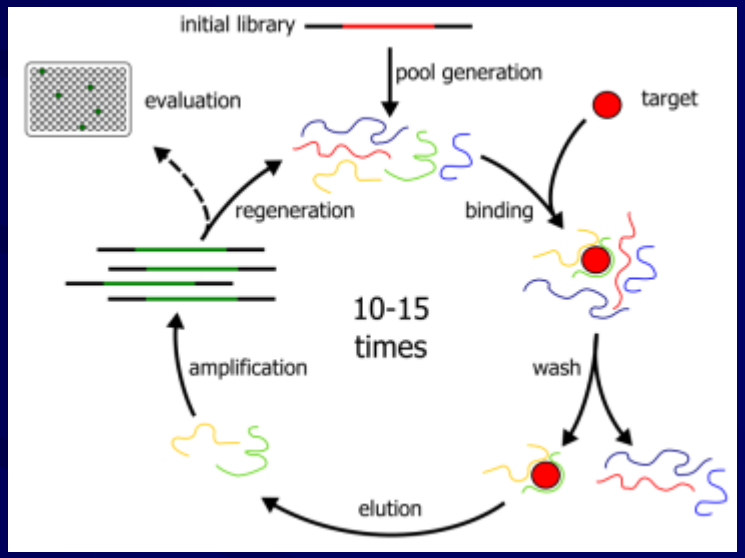
- Selection
- Evolution

Dawkins' replicator

- A an object that passes on its structure to a new one (the copy) largely intact
- Powerful replicators are essentially digital
- Microevolutionary fine-tuning
- Possibility of cumulative selection
- Does not have to be a linear structure!
- SPARSELY occupied sequence space

Combinatorial chemistry versus SELEX

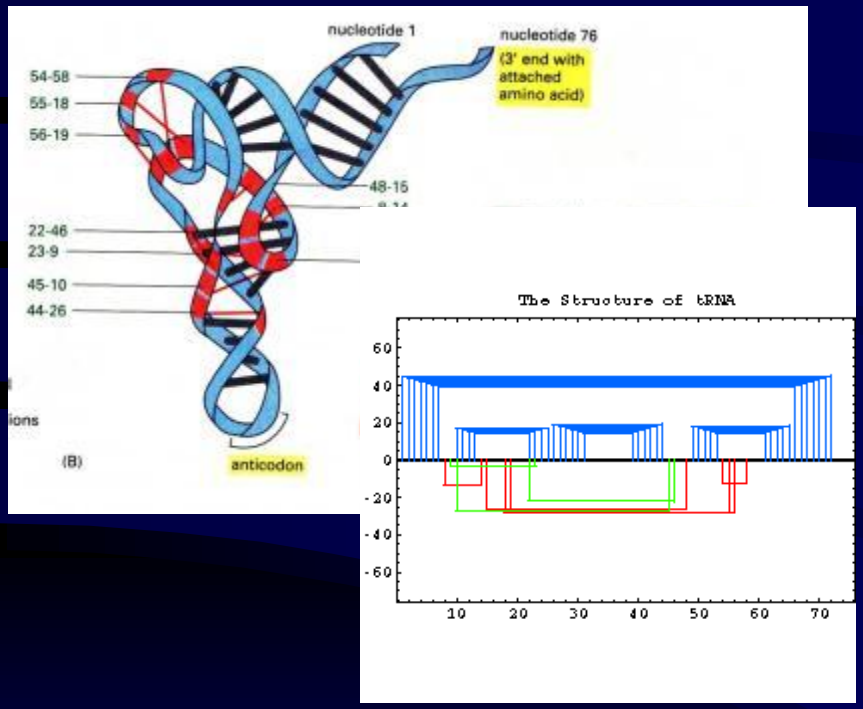




Short molecules

Complex molecules

Evolution of (more than simple phrase) tRNA-like structure



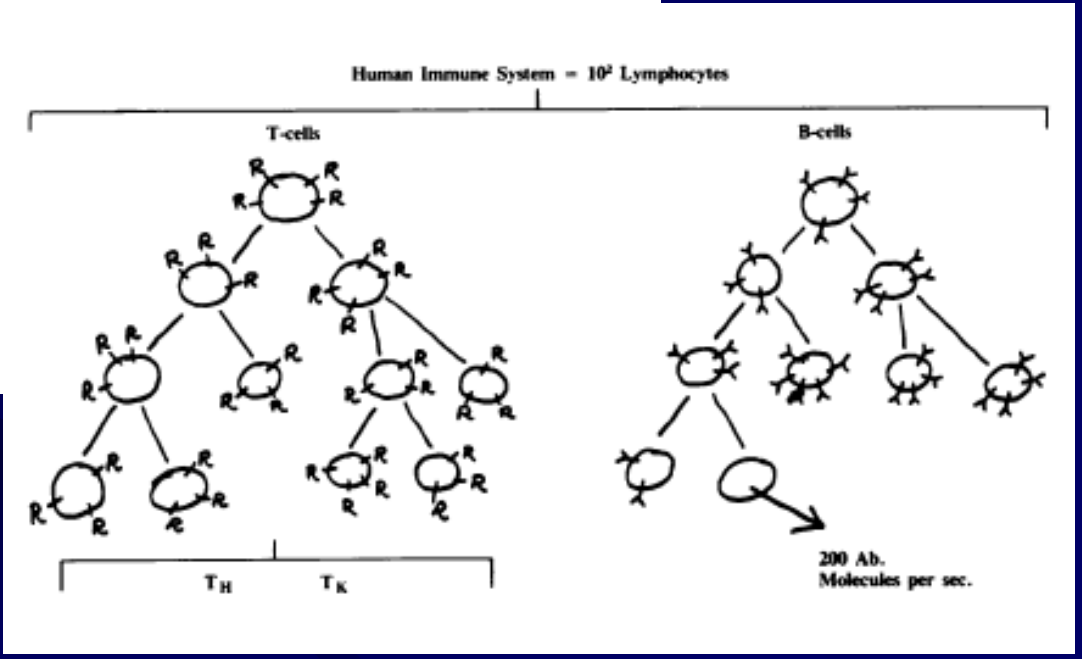
- One among 10^{11} random sequences
- Evolvable in a population of 500 in 500 generations, i.e. 2.5×10^5

When the Darwinian dynamic reinvented itself

The EMBO Journal vol.4 no.4 pp.847–852, 1985

The generative grammar of the immune system

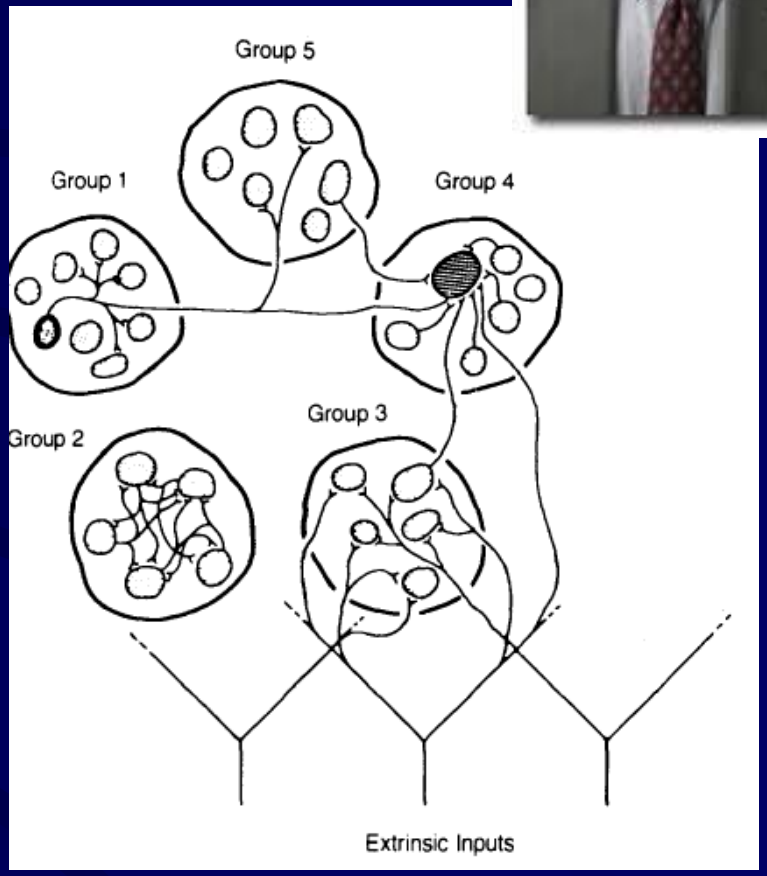
Niels K.Jerne M.D. FRS.



More like
articfial selection!

Neural ‘Darwinism’ (Edelman, 1987)

- The theory of neuronal group selection (NGS) proposes that a primary repertoire of neuronal groups within the brain compete with each other for stimulus and reward resources.
- This results in selection of a secondary repertoire of behaviourally proficient groups



Crick on Edelman (1989)

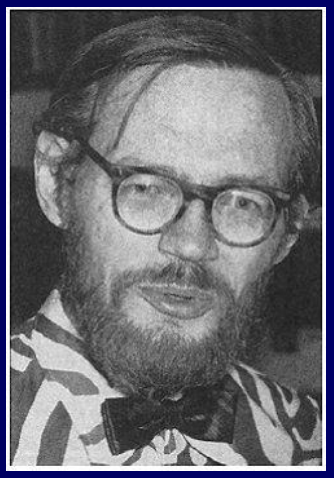
(1) I suggest that the term 'Neural Darwinism' be abandoned. I have not found it possible to make a worthwhile analogy between the theory of natural selection and what happens in the developing brain and indeed Edelman has not presented one. There is a lot of loose talk about 'populations' and 'selection', but very little of a concrete nature emerges from these comparisons. The same is true of the other possible analogy, that between the brain and the vertebrate immune system. There is, however, no need to pursue these topics further, since Edelman himself has written:

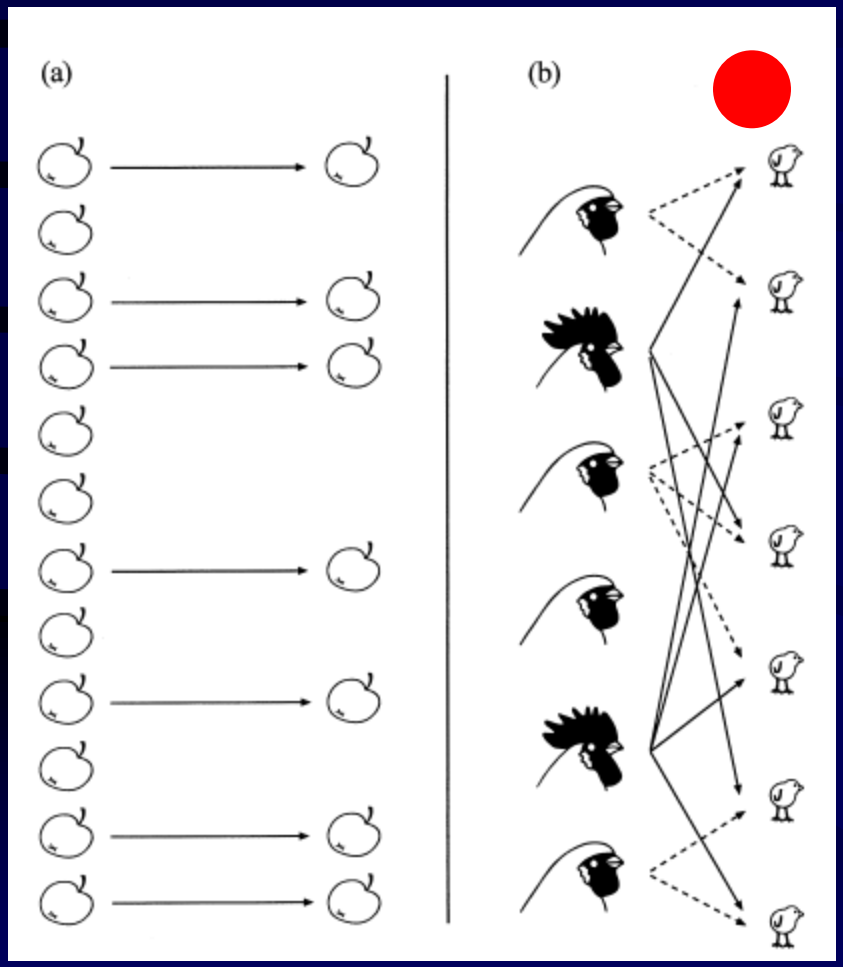
> 'there are enormous differences in detail and mechanism between natural selection, neuronal group selection and clonal selection in immunity' (ND, p. 9).

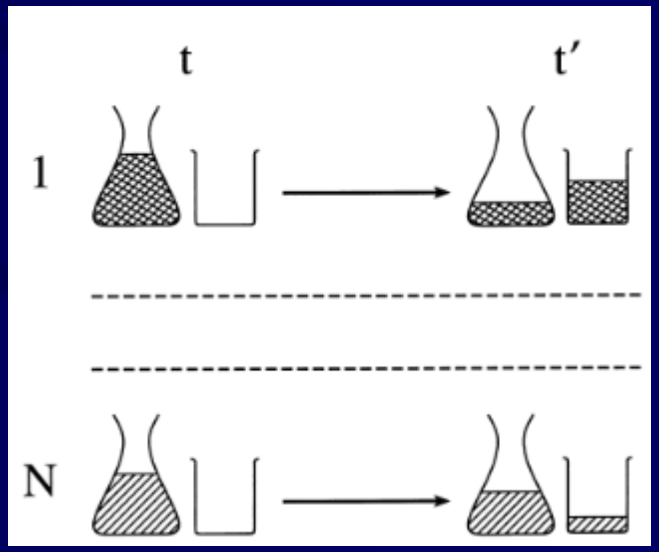
Michod on Edelman (1989)

- **Basic parallels between selection in genetic evolution and neuronal group selection:**
- In genetic evolution, differences in adaptedness to an environment lead to differences in reproductive success, which, when coupled with rules of genetic transmission, lead to a change in frequency of genotypes in a population.
- In neuronal group selection, differences in receptive fields and connectivity between neuronal groups lead to differences in initial response of groups to a stimulus, which, when coupled with rules of synaptic change, lead to a change in probabilities of further response to the stimulus.

The Price of selection (1970, 1995)

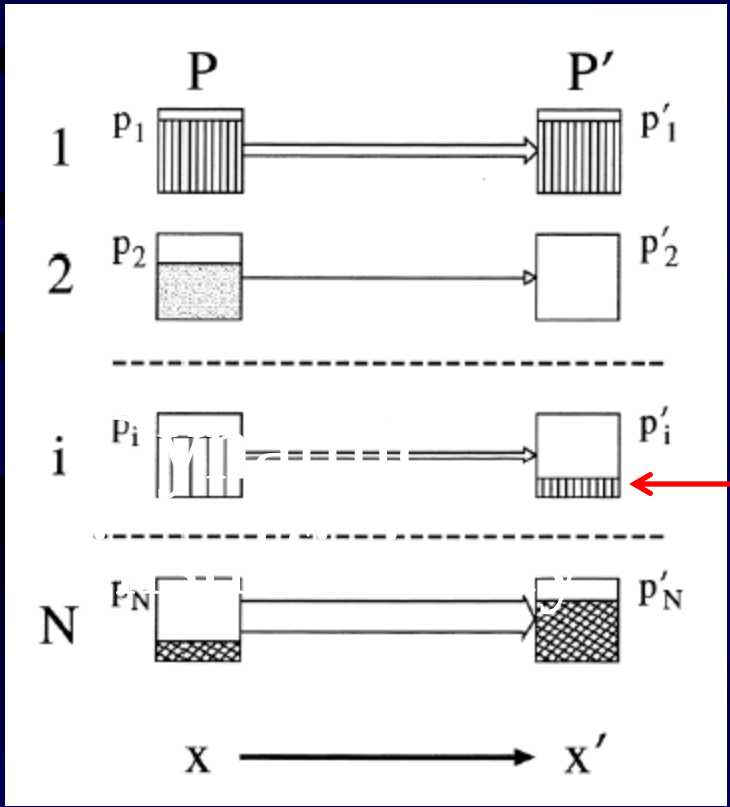






Subset sel. Darwinian sel. Chemical sel.

General selection model: NO Darwinian algorithm!



$$w\Delta z = \mathrm{cov}(w_i, z_i) + E(w_i \Delta z_i)$$

Mean fitness

transmission bias

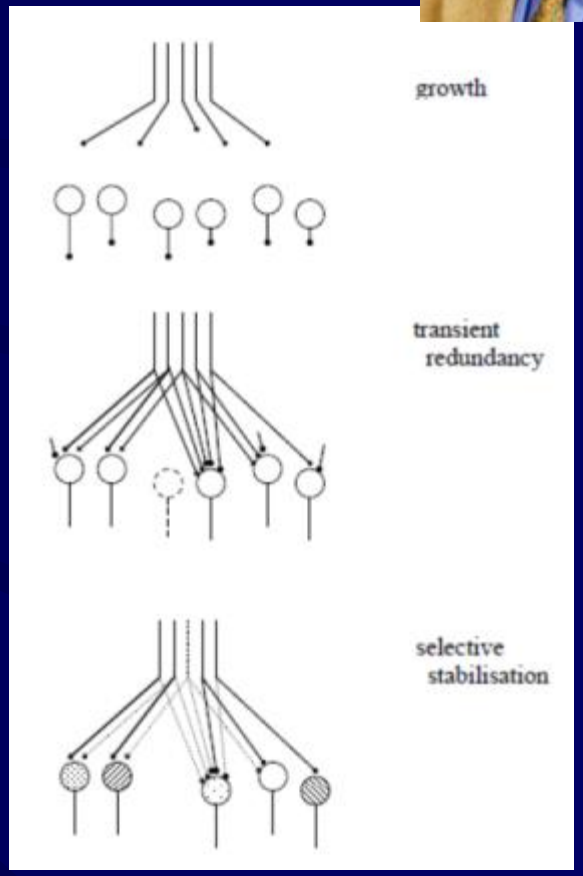
Dynamic insufficiency!!!

The false path...

- Price equation ALSO applicable to biological evolution
- Has been applied to interesting problems in evolution
- Neuronal groups are biological objects
- Price is ALSO applicable to them
- HENCE neuronal group selection is like biological evolution!

Changeux, 1973

- There is selection, in the sense of Price.
- This fundamentally important limitation is clearly stated: *"an organism can not learn more than is initially present in its pre-representations."*
- A **one-shot game!**



Synaptic selectionism (Fernando & Szathmáry, 2010)

$\mathbf{w}$ is the synaptic weight vector, v is the output rate, $\mathbf{u}$ is the input rate vector

$$\tau_w \frac{d\mathbf{w}}{dt} = v\mathbf{u} - \alpha v^2 \mathbf{w},$$

Oja rule

$$z_i = u_i / U.$$

$$\frac{dw_i}{dt} = \sum_{j=1}^{N_u} z_i z_j w_j - \alpha' w_i \sum_{j=1}^{N_u} \sum_{k=1}^{N_u} z_j z_k w_k,$$

$$\frac{dx_i}{dt} = A_i Q_i x_i + \sum_{j \neq i}^{N} m_{ij} x_j - \frac{x_i}{c} \sum_{j=1}^{N} \sum_{k=1}^{N} m_{ij} x_j,$$

Eigen equation

where x_i is the concentration of sequence i (of RNA for example), m_{ij} is the mutation rate from sequence j to i, A_i is the gross replication rate of sequence i and Q_i is its copying fidelity, N is the total number of different sequences, and formally $m_{ij} = A_i Q_i$

The ‘hedonistic’ synapse (Seung, 2003)

- “randomness is harnessed by the brain for learning, in analogy to the way genetic mutation is utilized by Darwinian evolution”
- (1) the probability of release is increased if reward follows release and is decreased if reward follows failure,
- (2) the probability of release is decreased if punishment follows release and is increased if punishment follows failure

Stochastic hill-climbing

- “The dynamics of learning executes a random walk in the parameter space, which is biased in a direction that increases reward.
- A picturesque term for such behaviour is “hill-climbing”, which comes from visualizing the average reward as the height of a landscape over the parameter space. The formal term is “stochastic gradient ascent”

Reinforcement and covariance (Loewenstein, 2010)

$$p_i(t) = p_i(\mathbf{W}(t)).$$

$$\Delta p_i(t+1) = p_i(\mathbf{W}(t+1)) - p_i(\mathbf{W}(t)),$$

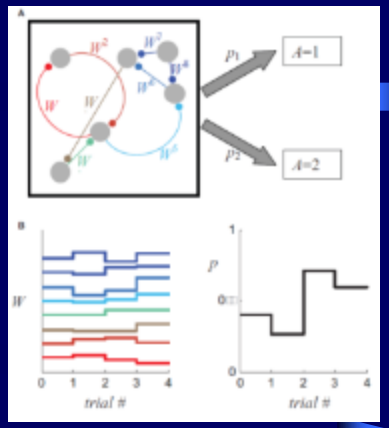
$$\Delta \mathbf{W} = \varphi R(N - \mathbf{E}[N])$$

$$\mathrm{E}[\Delta \mathbf{W}] = \varphi \mathrm{Cov}[R, N],$$

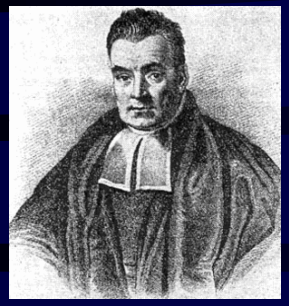
$$\frac{d\mathbf{W}}{dt} = \varphi \mathrm{Cov}[R, N].$$

$$\frac{dp_i}{dt} = \eta p_i(\mathrm{E}[R \mid A = i] - \mathrm{E}[R]),$$

Replicator equation



$p_i(t)$: probability of behaviour i
W: synaptic weight vector
N: any measure of neural activity
R: reward



Bayes and Darwin

Inductive inference is the only process known to us by which essentially new knowledge comes into the world. – R. A. Fisher, *The Design of Experiments* (1935) [8]

The theory of evolution by cumulative natural selection is the only theory we know of that is in principle capable of explaining the existence of organized complexity. – Richard Dawkins, *The Blind Watchmaker* (1987) [7]

Bayes and selection (e.g. Harper, 1010)

$$P(H_i \mid E) = \frac{P(E \mid H_i) P(H_i)}{P(E)}$$

$$x'_i = \frac{x_i f_i(\mathbf{x})}{f(\mathbf{x})},$$

Bayesian Inference	Discrete Replicator
Prior Distribution $(P(H_1), \ldots, P(H_n))$	Population state $x = (x_1, \ldots, x_n)$
New Evidence $P(E\|H_i)$	Fitness landscape $f_i(x)$
Normalization $P(E)$	Mean fitness $\bar{f}(x)$
Posterior distribution $P(H_1\|E), \ldots, P(H_n\|E)$	Population state $x' = (x'_1, \ldots, x'_n)$

Deeper than one might think

- For the continuous-time replicator dynamics

$$\dot{x}_i = x_i \left(f_i(x) - \bar{f}(x) \right).$$

- The Kullback-Liebler information divergence $D_{KL}(\hat{x}||x) = \sum_i \hat{x}_i \log \hat{x}_i - \sum_i \hat{x}_i \log x_i.$

  is a Lyapunov function at the ESS $\hat{x}$

- AND the solutions of the replicator equation can be realized as exponential families

…and even deeper!

- Kullback-Liebler divergence (KL)

↓

- Hessian of KL

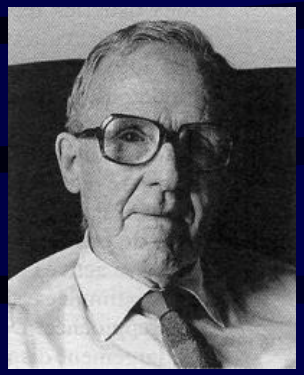
↓

- Fisher information metric (FIM):

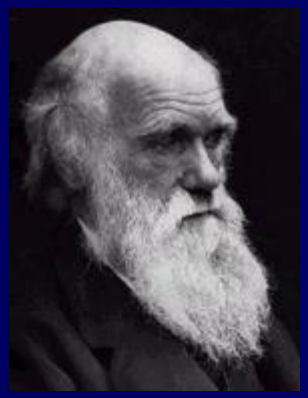
↓

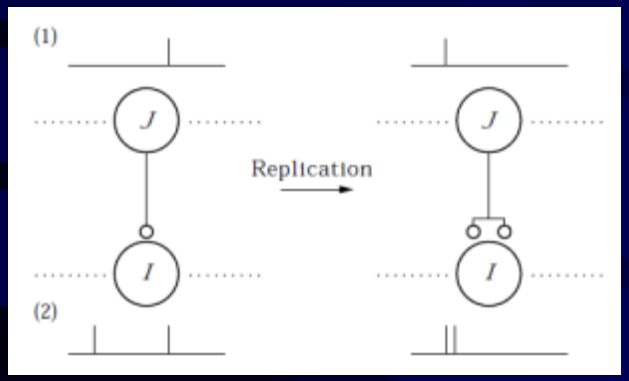
$$g_{ij}(x) = \mathbb{E}\left[\frac{\partial \log p}{\partial x^i}\frac{\partial \log p}{\partial x^j}\right].$$

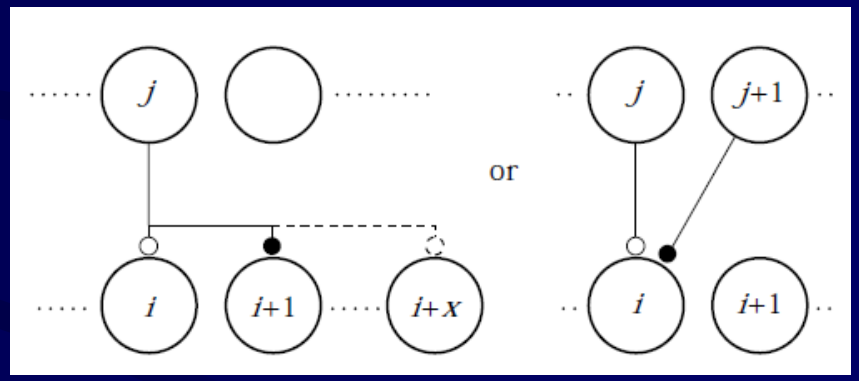
- Gradient flow of FIM: Replicator equation!



Hebb and Darwin (Adams, 1998)



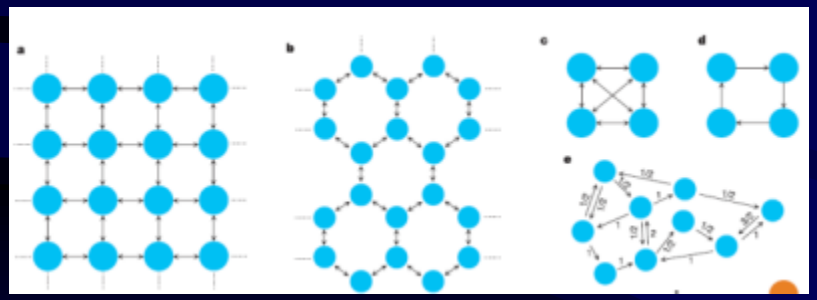




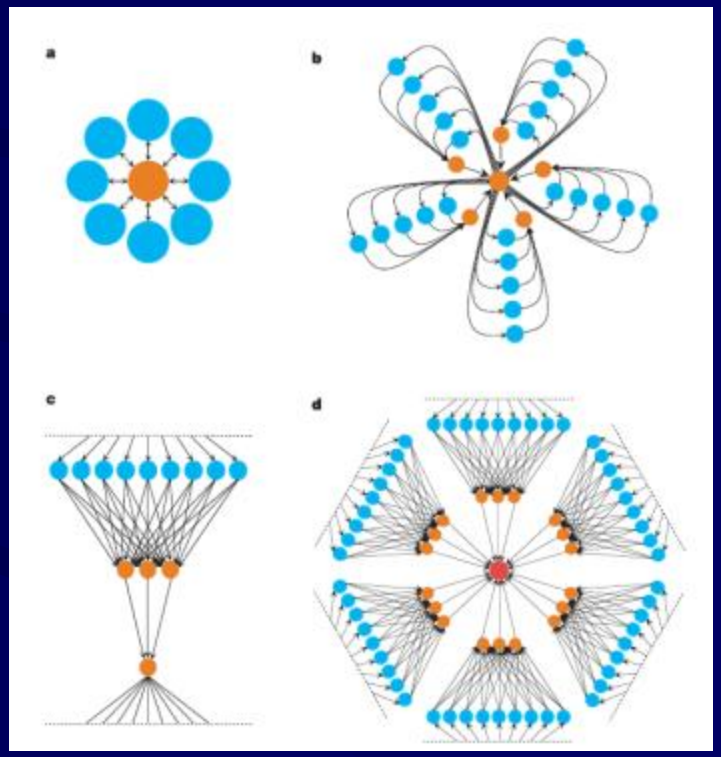
synaptic replication

synaptic mutation

Evolutionary graph theory (Liberman et al. 2005)



Like the Moran process without population structure



Selection amplifiers

Possible Neuronal Replication Mechanisms

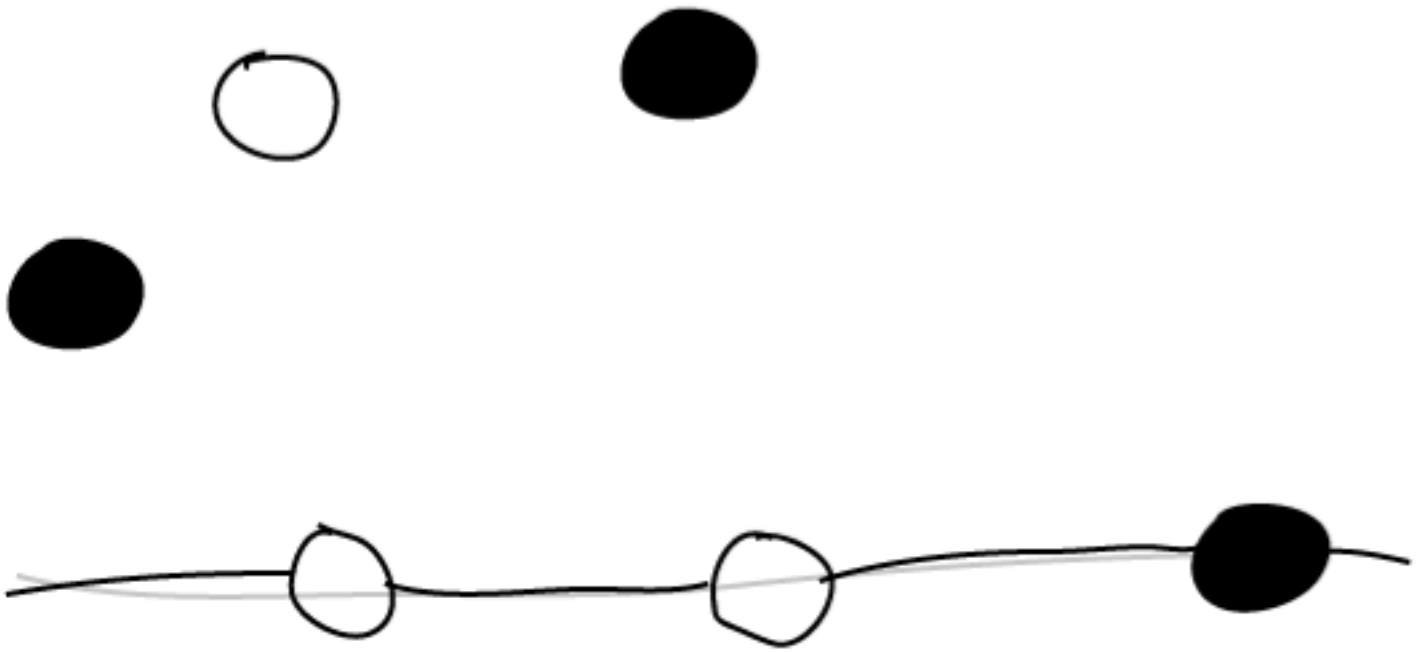
- Replication of Synaptic Connectivity Patterns
- Replication of Activity Patterns
- Evolvable Paths of Activity: Overlapping Units of Evolution

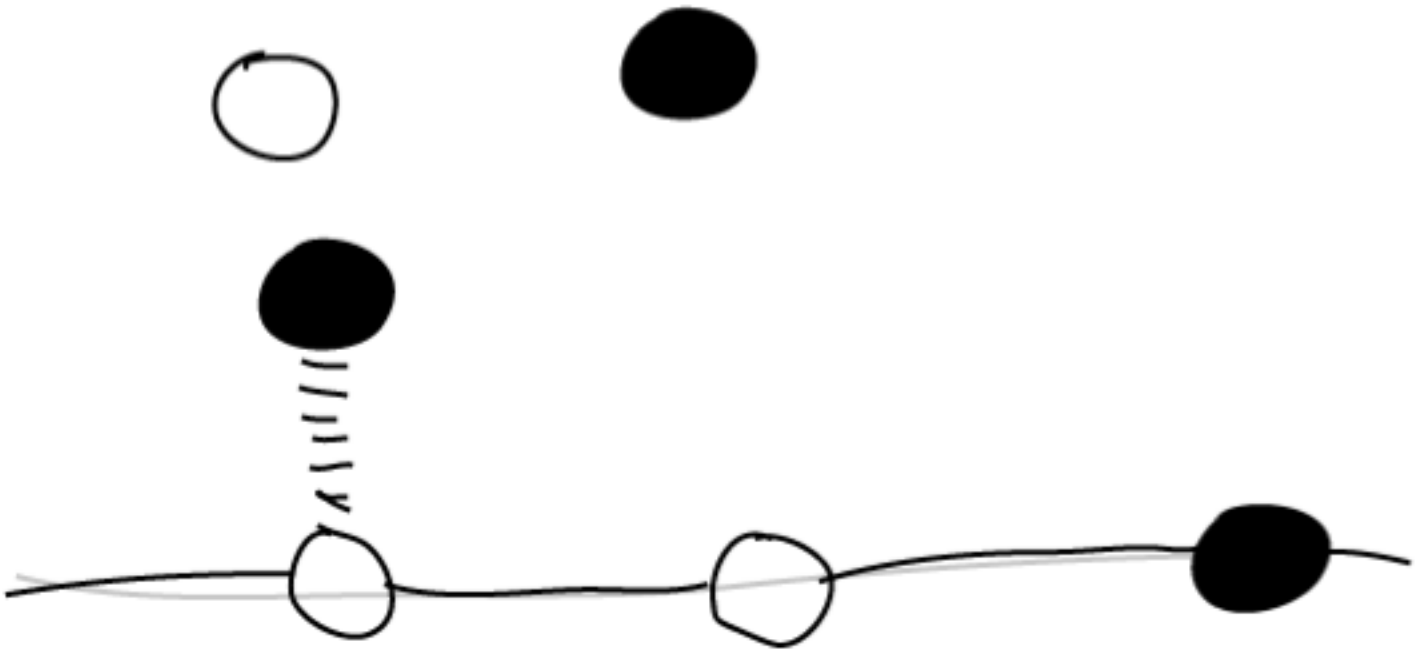
Possible Neuronal Replication Mechanisms

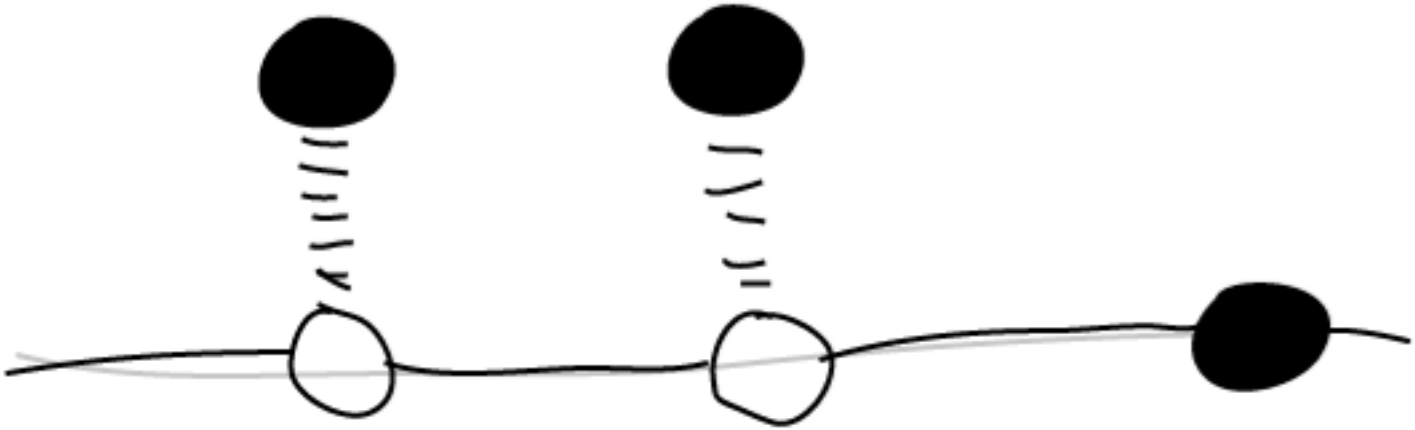
- Replication of Synaptic Connectivity Patterns
- Replication of Activity Patterns
- Evolvable Paths of Activity: Overlapping Units of Evolution

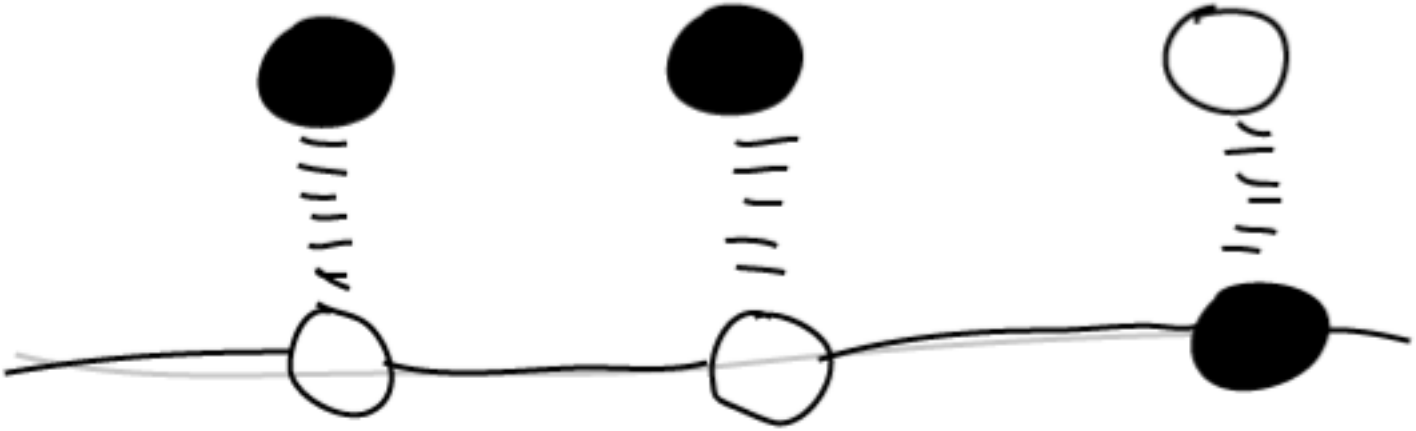
Possible Neuronal Replication Mechanisms?

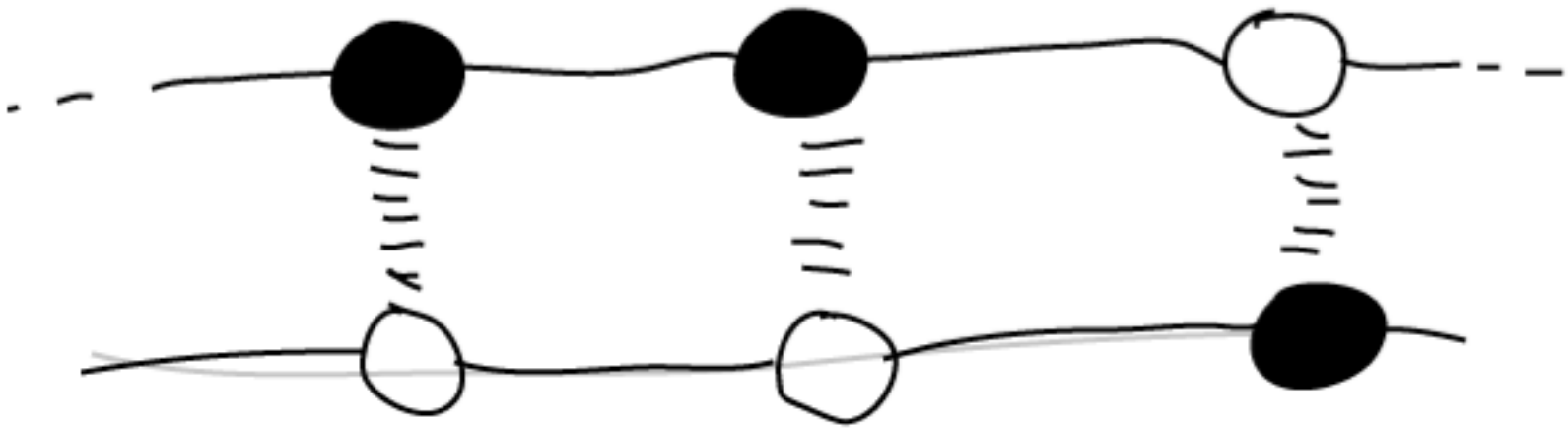
- Replication of Synaptic Connectivity Patterns
- Replication of Activity Patterns
- Evolvable Paths of Activity: Overlapping Units of Evolution

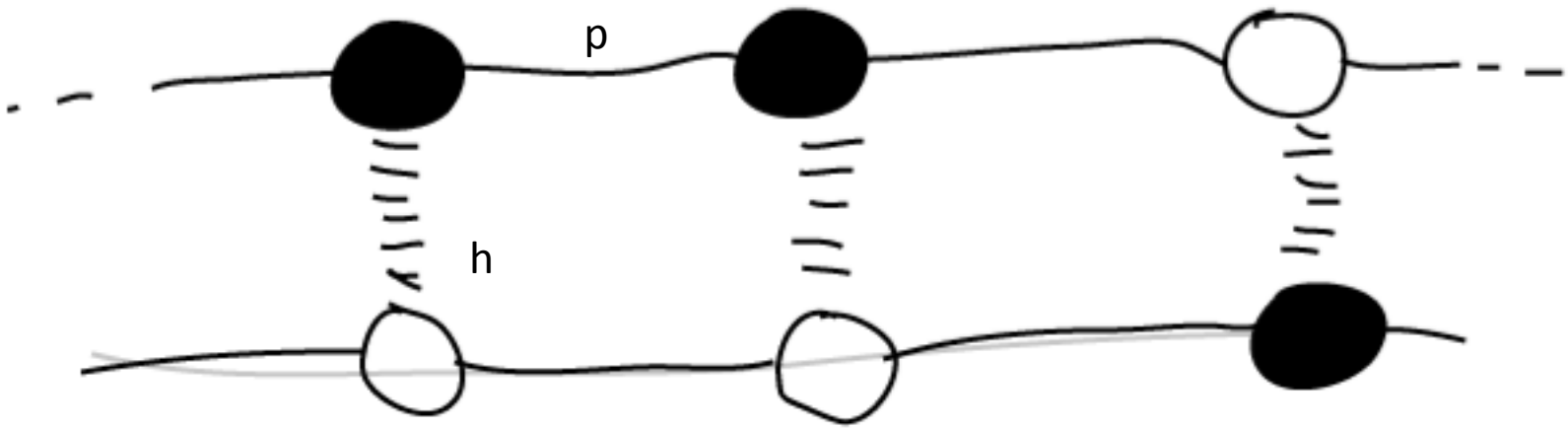


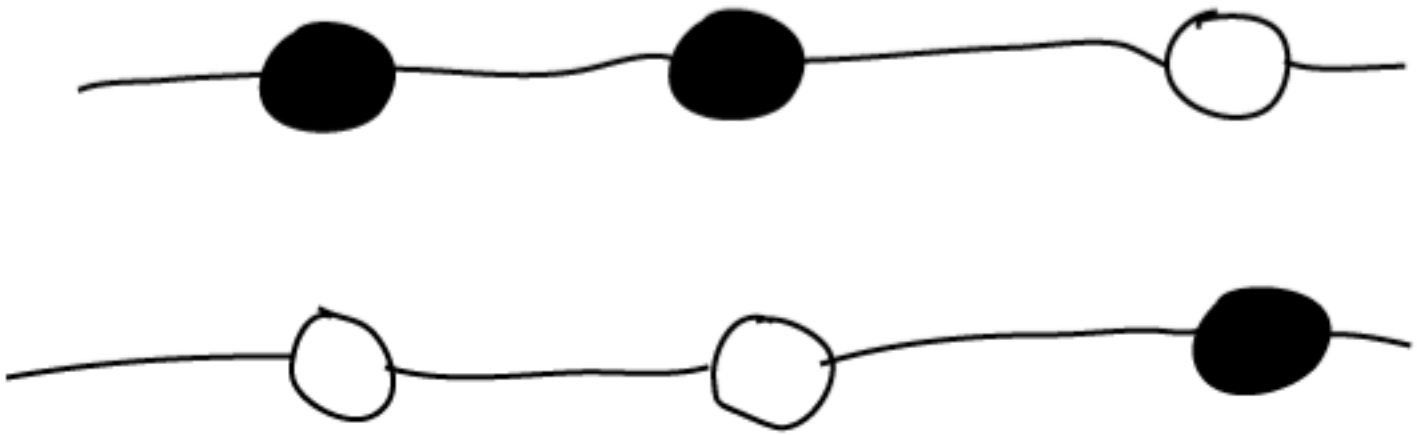




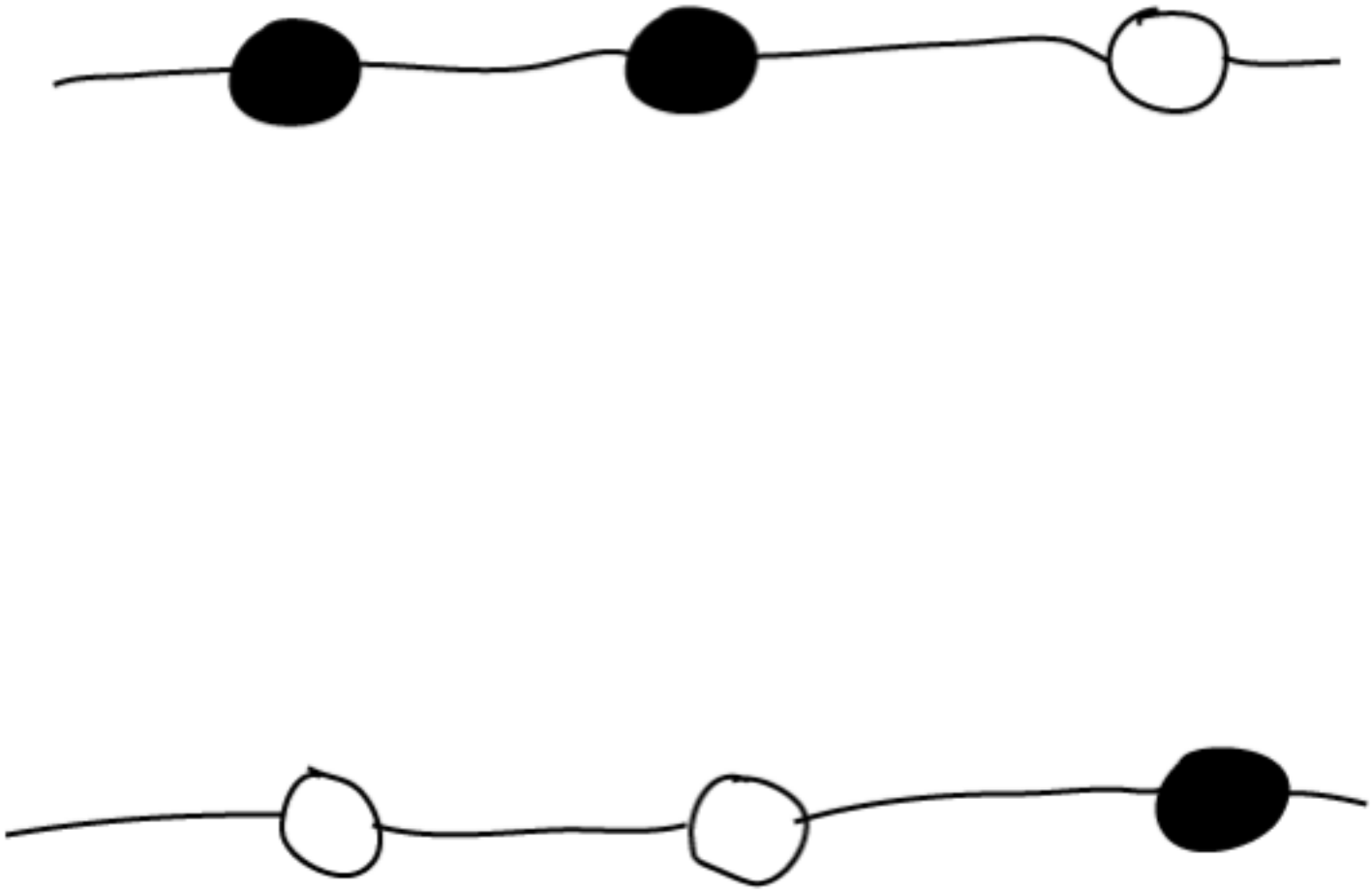


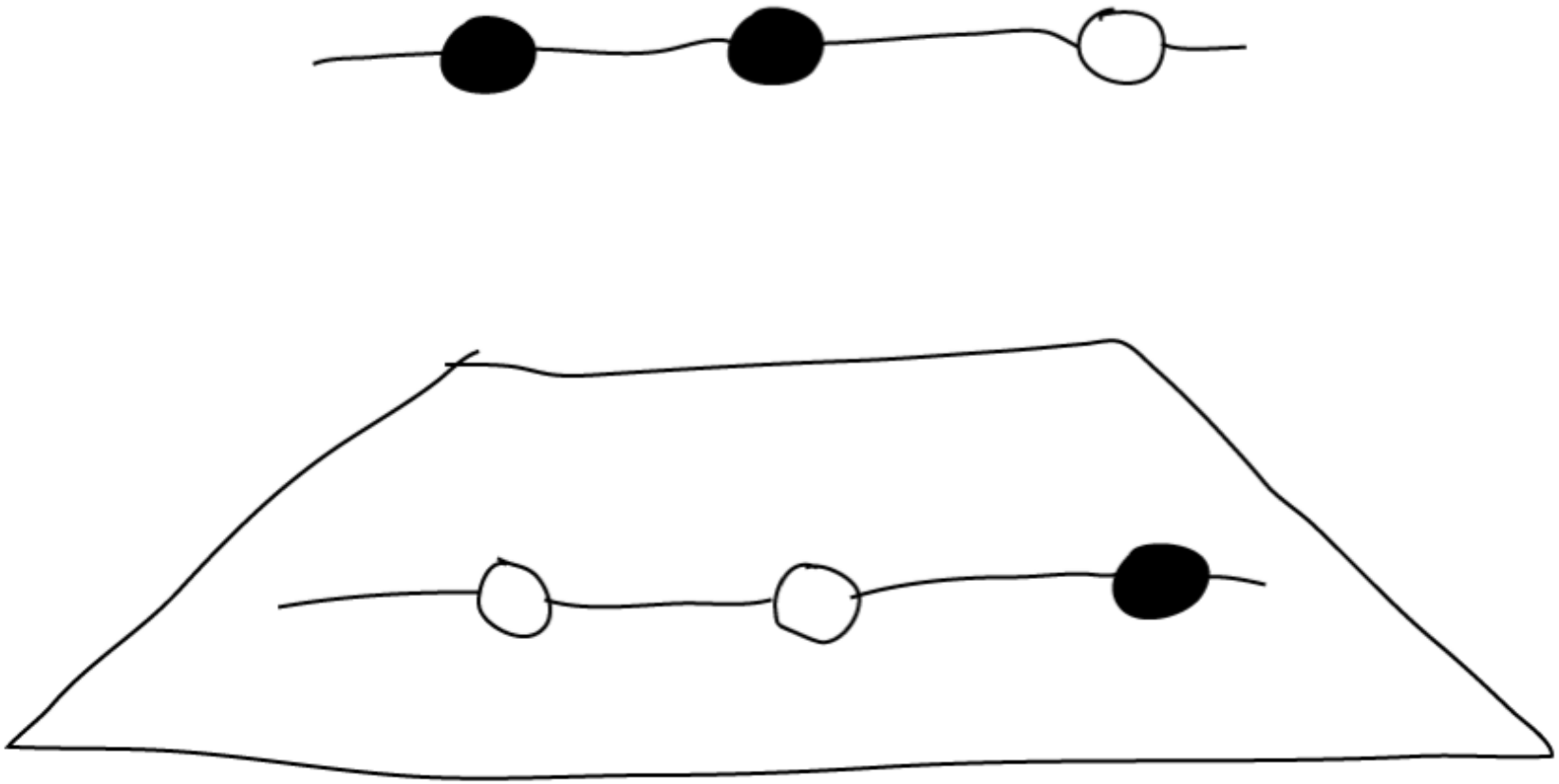


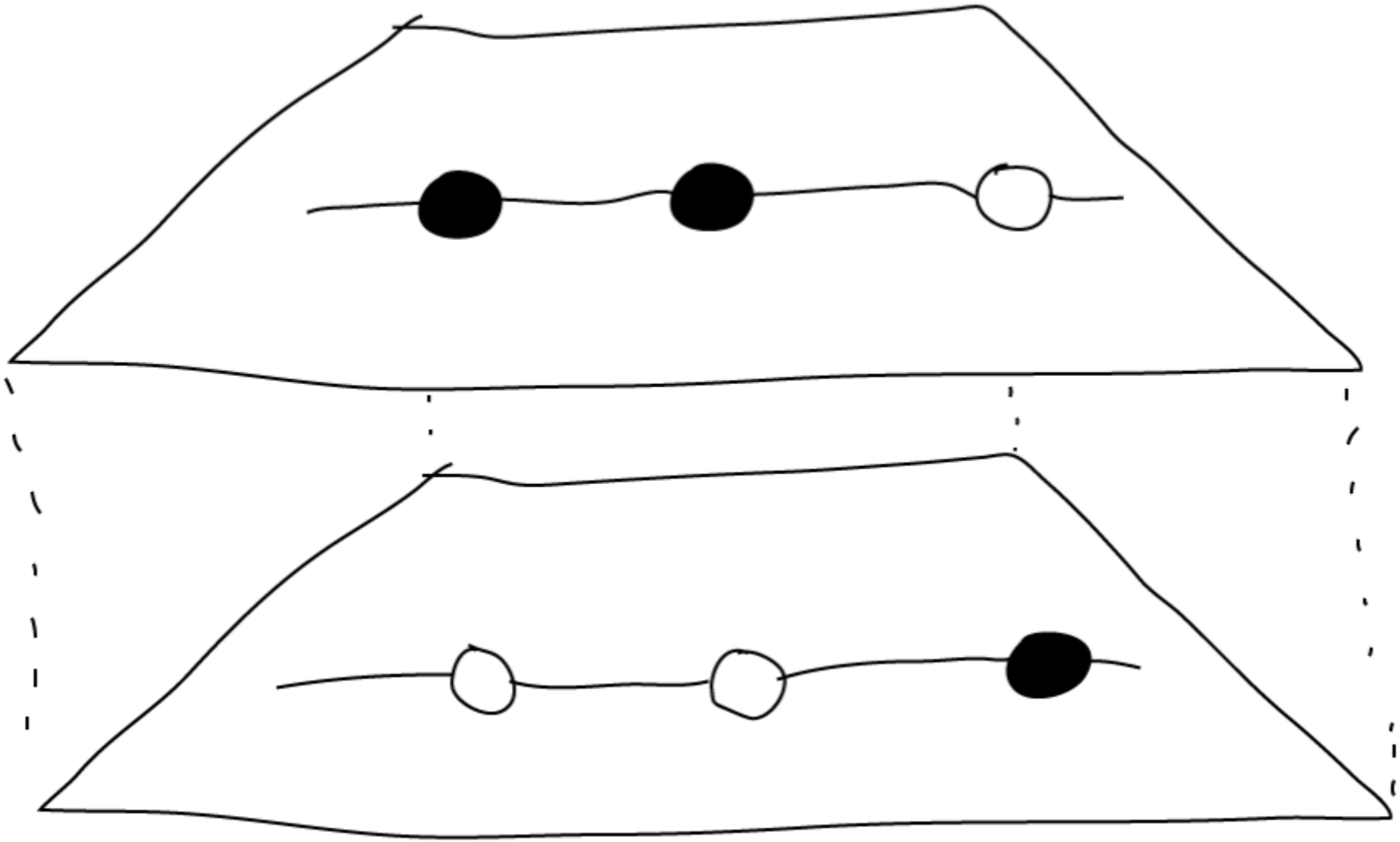

p
h

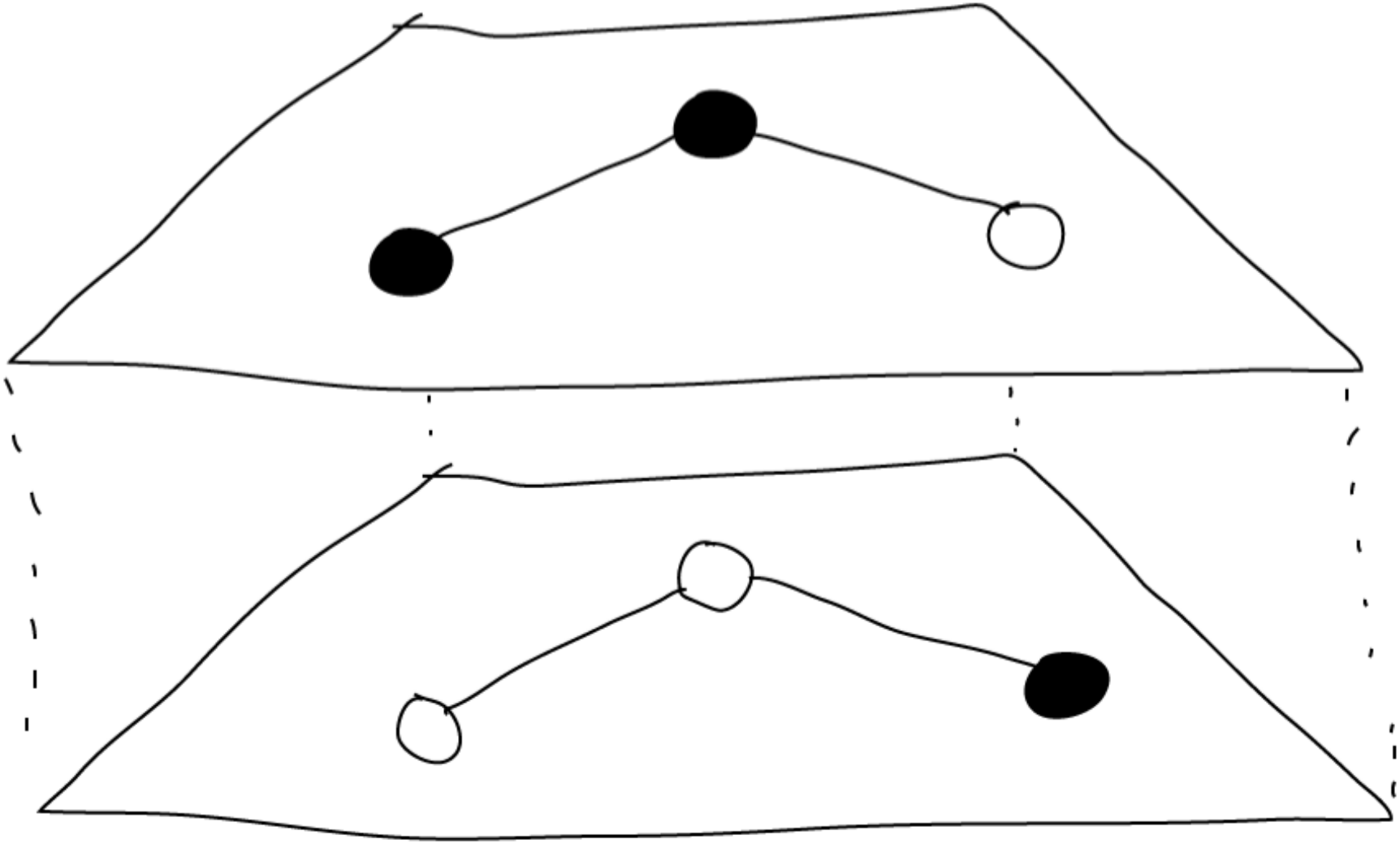


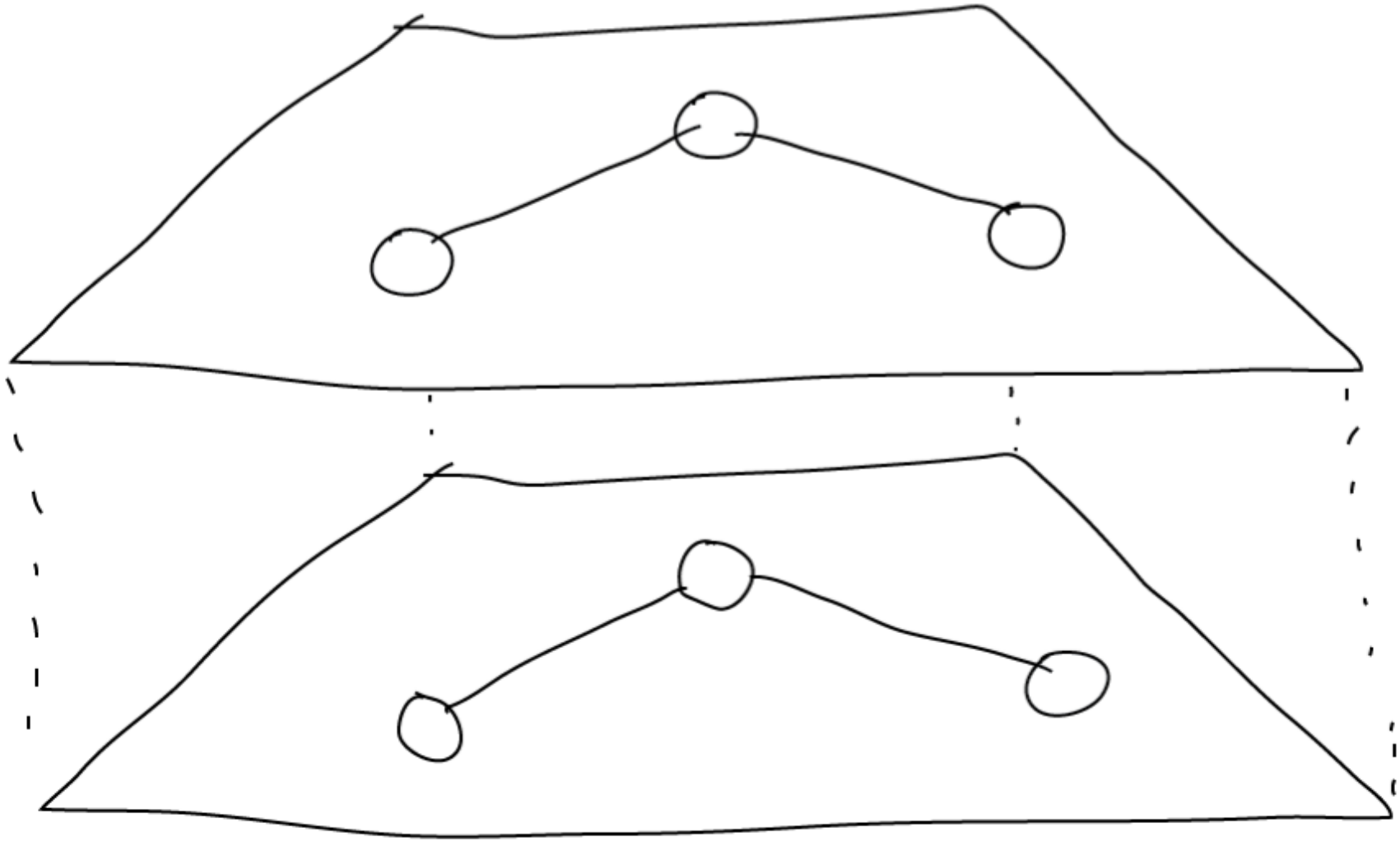
Synaptic connectivity replication is a little different

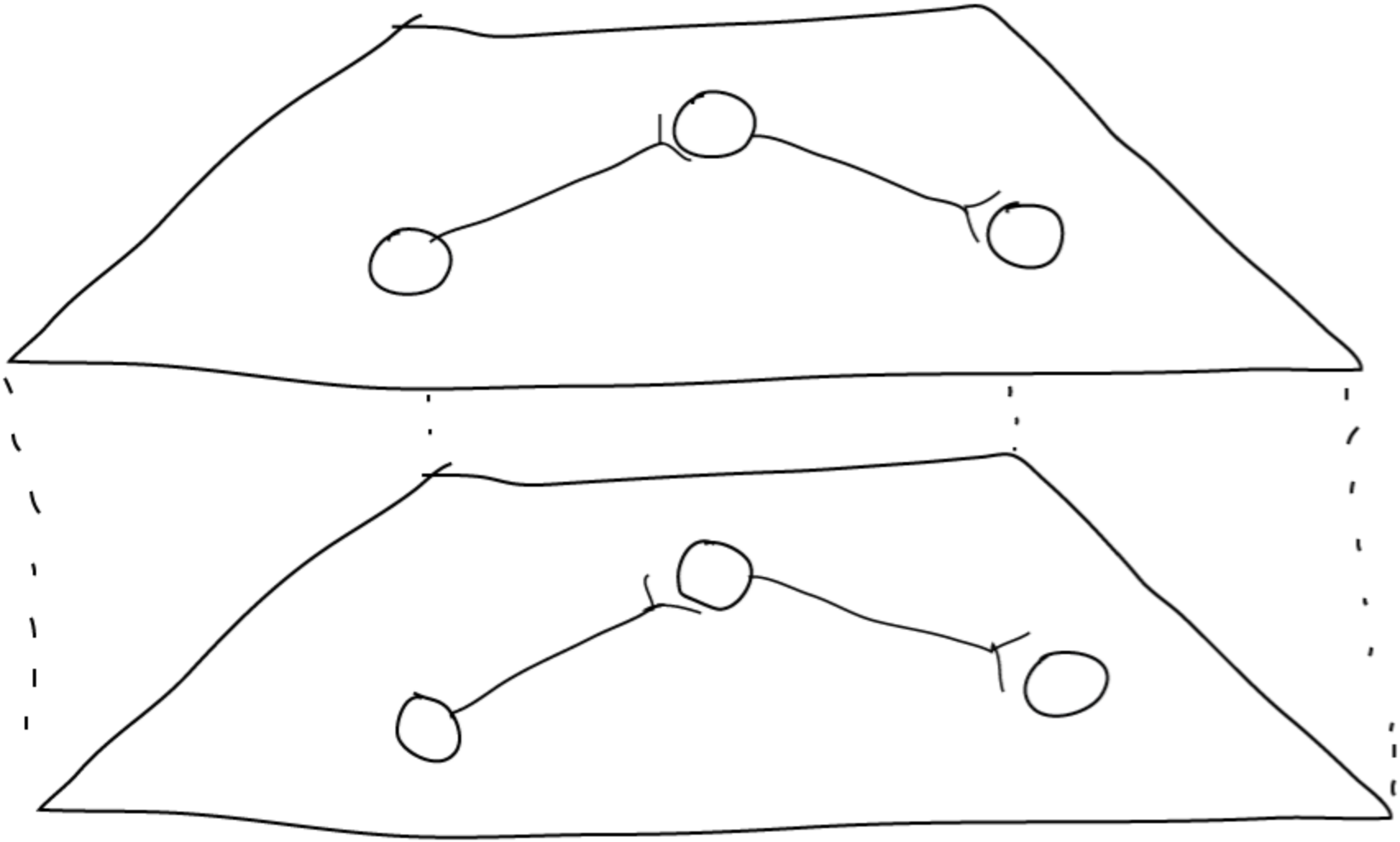


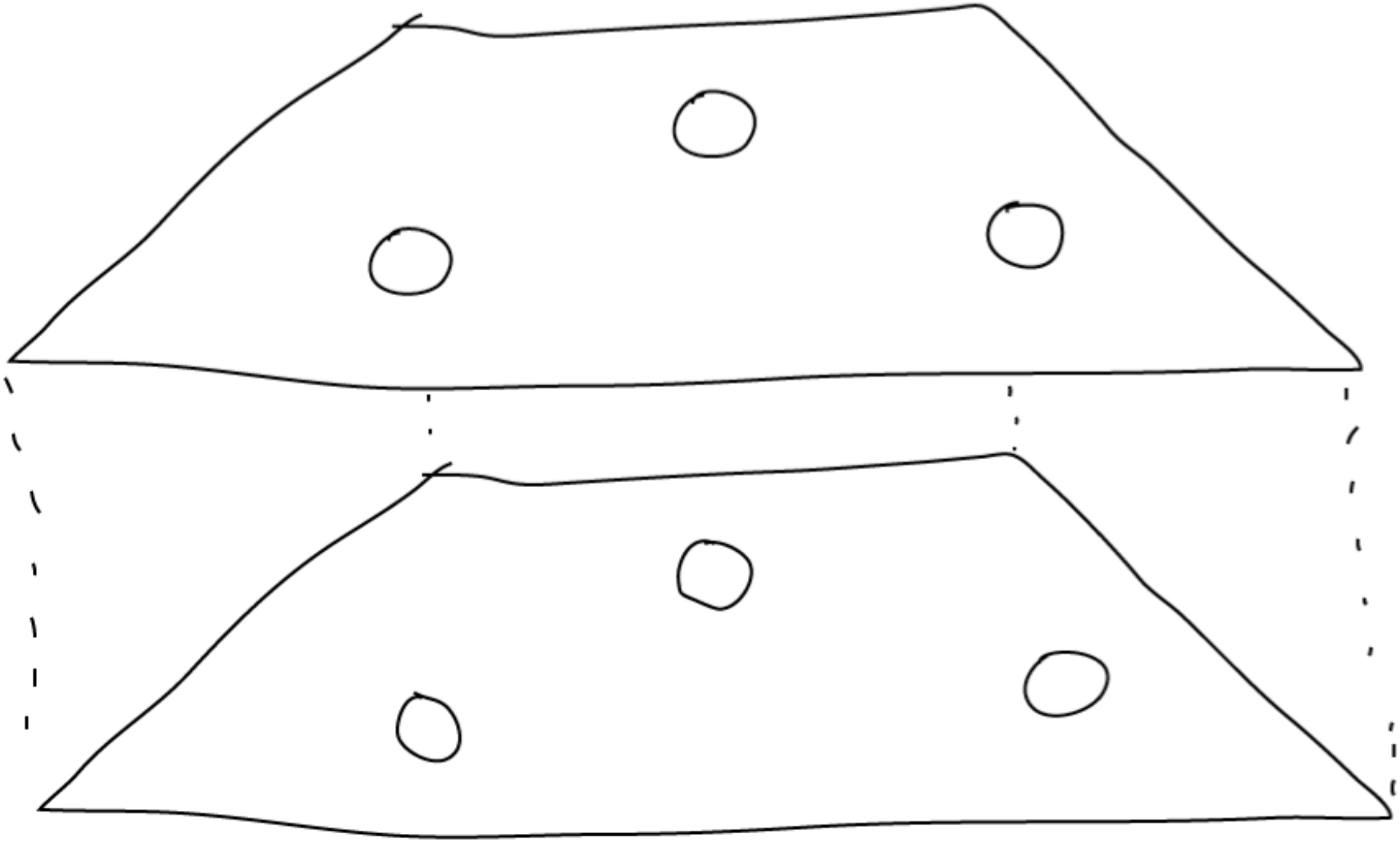


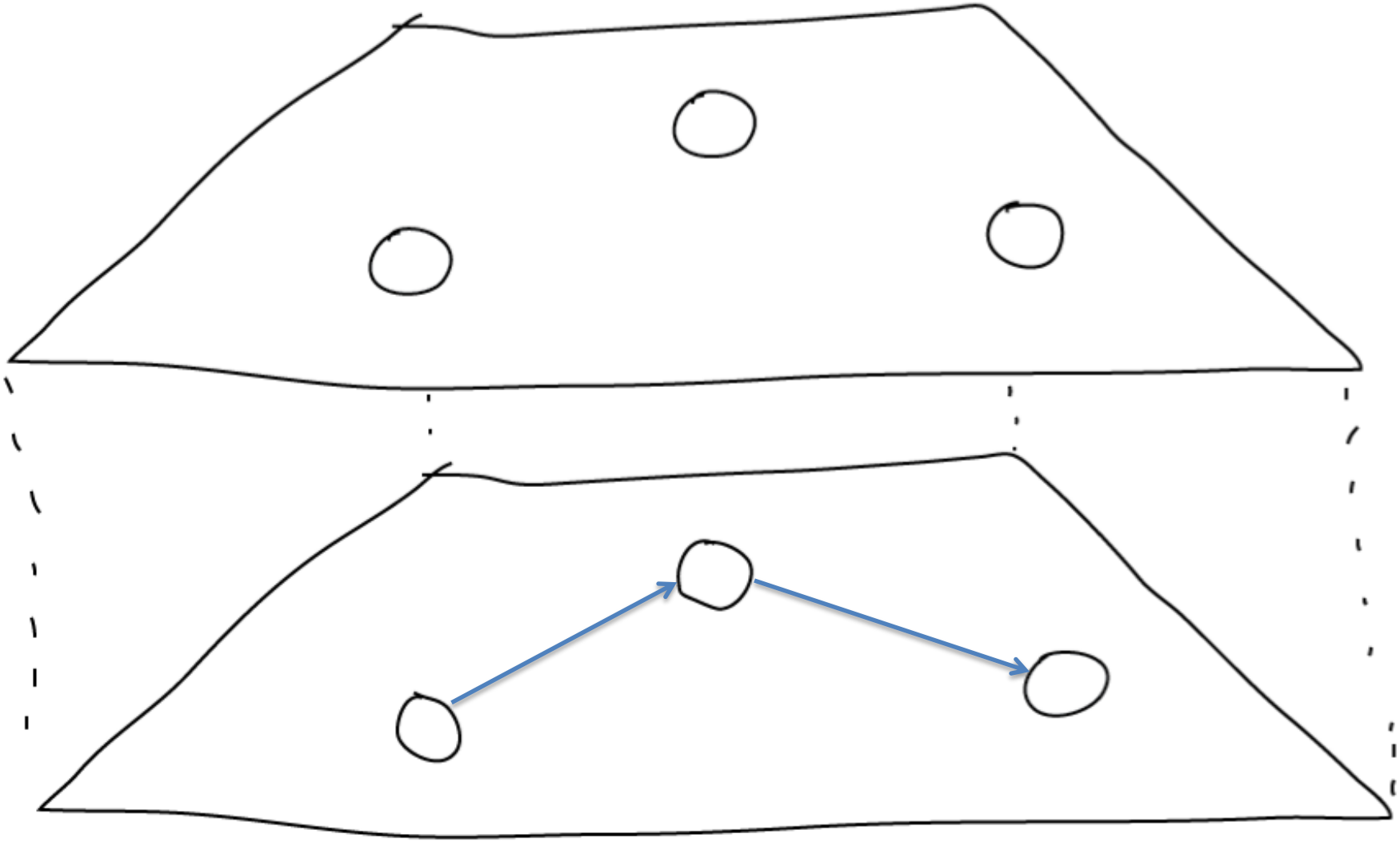


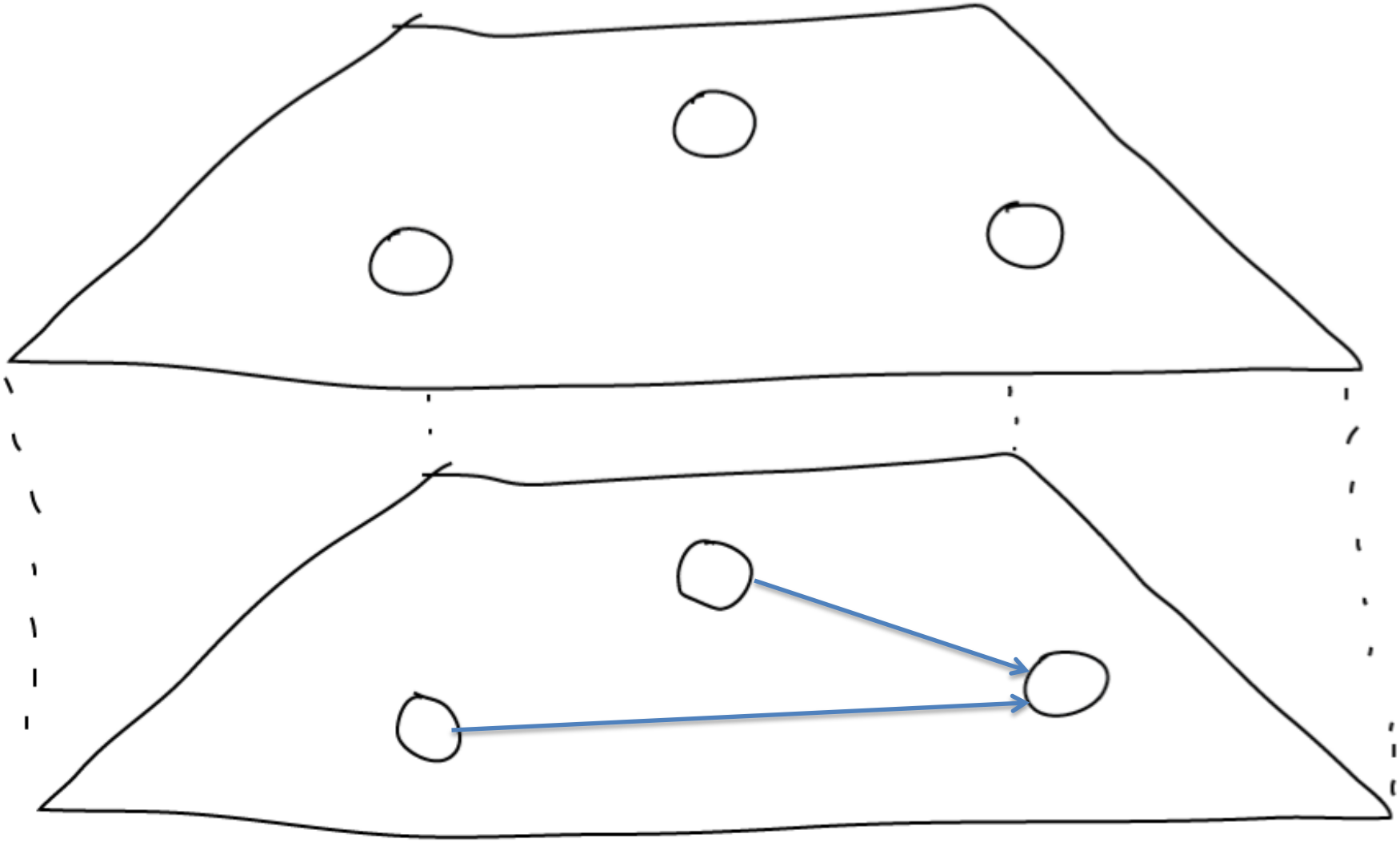


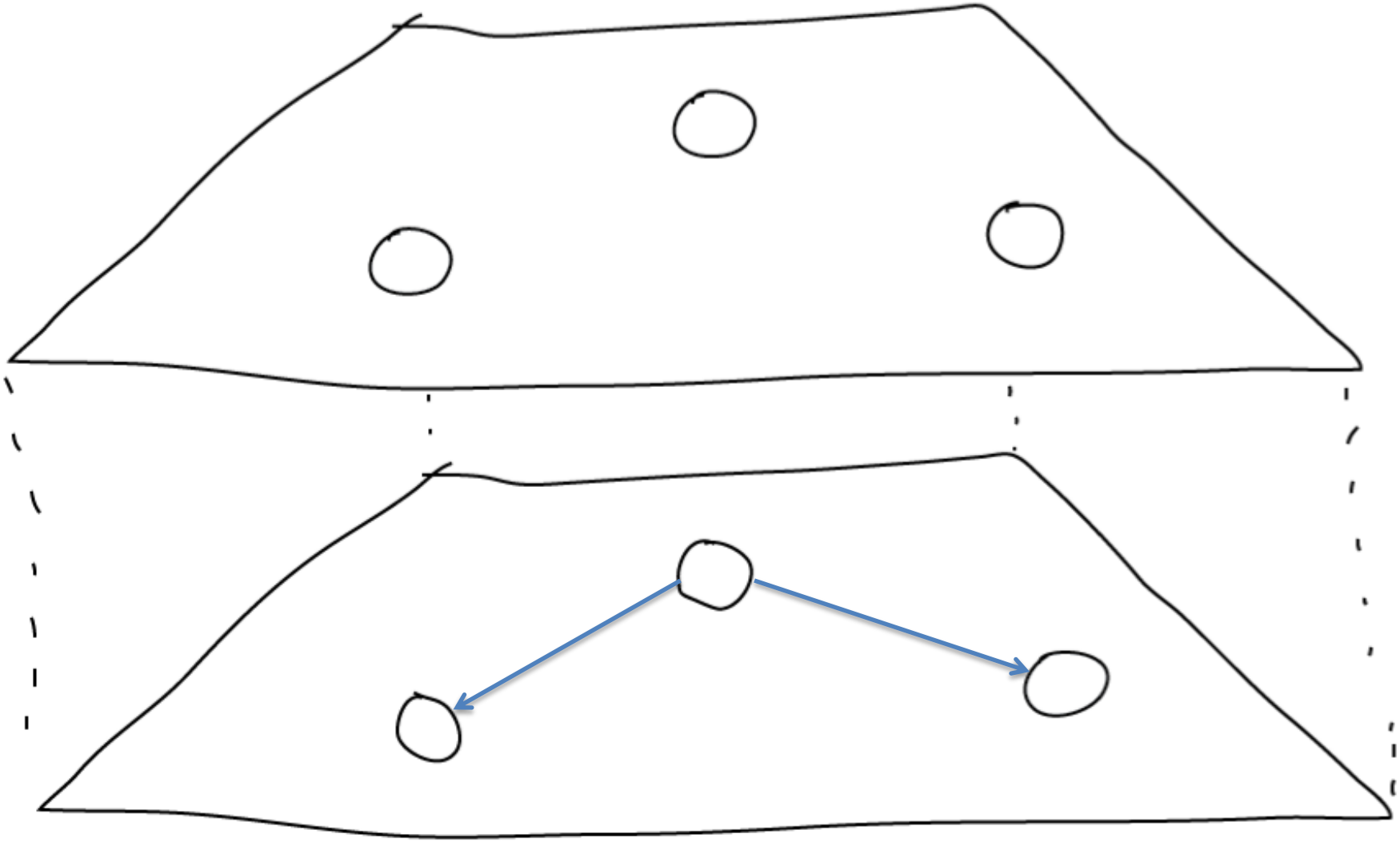


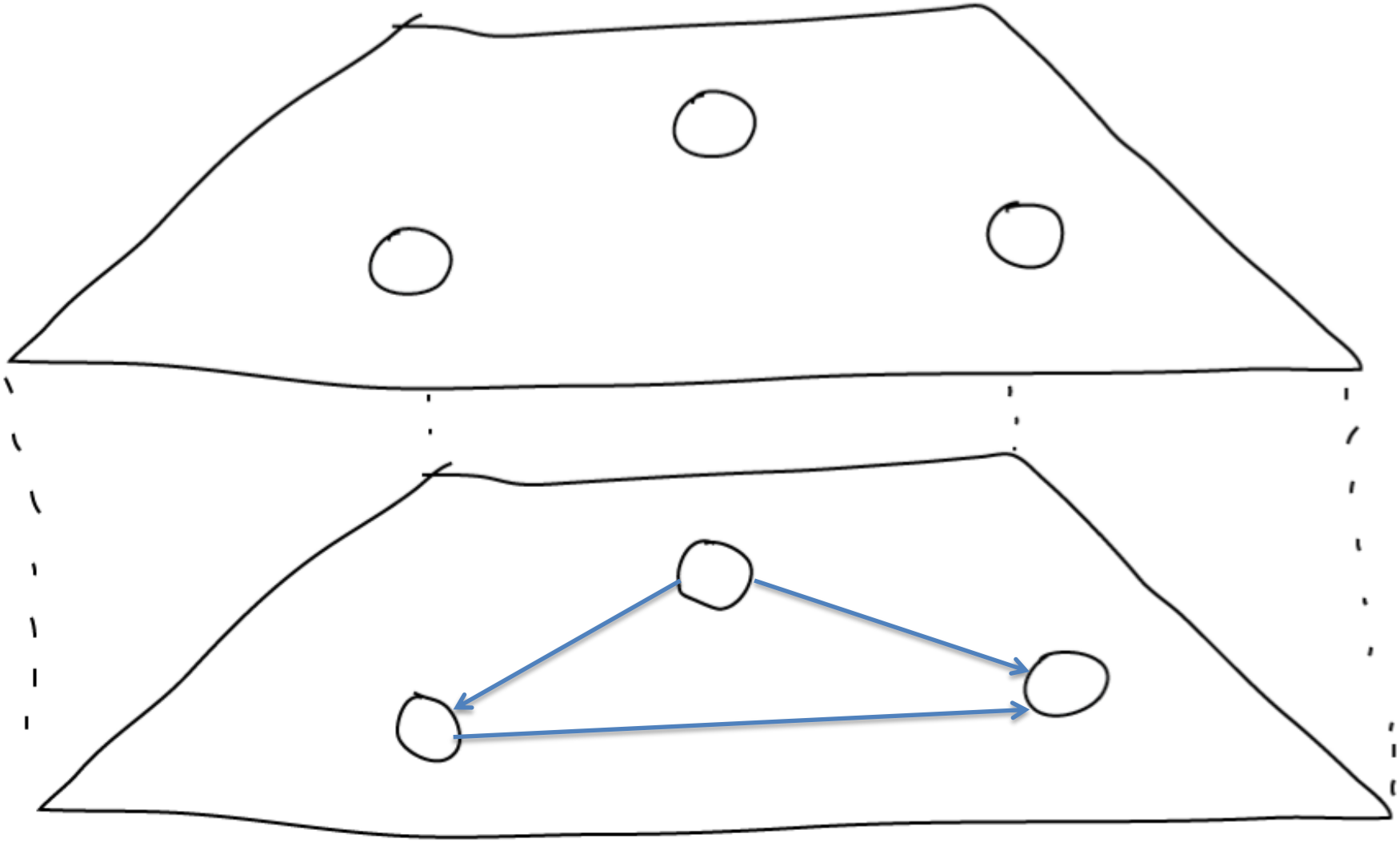


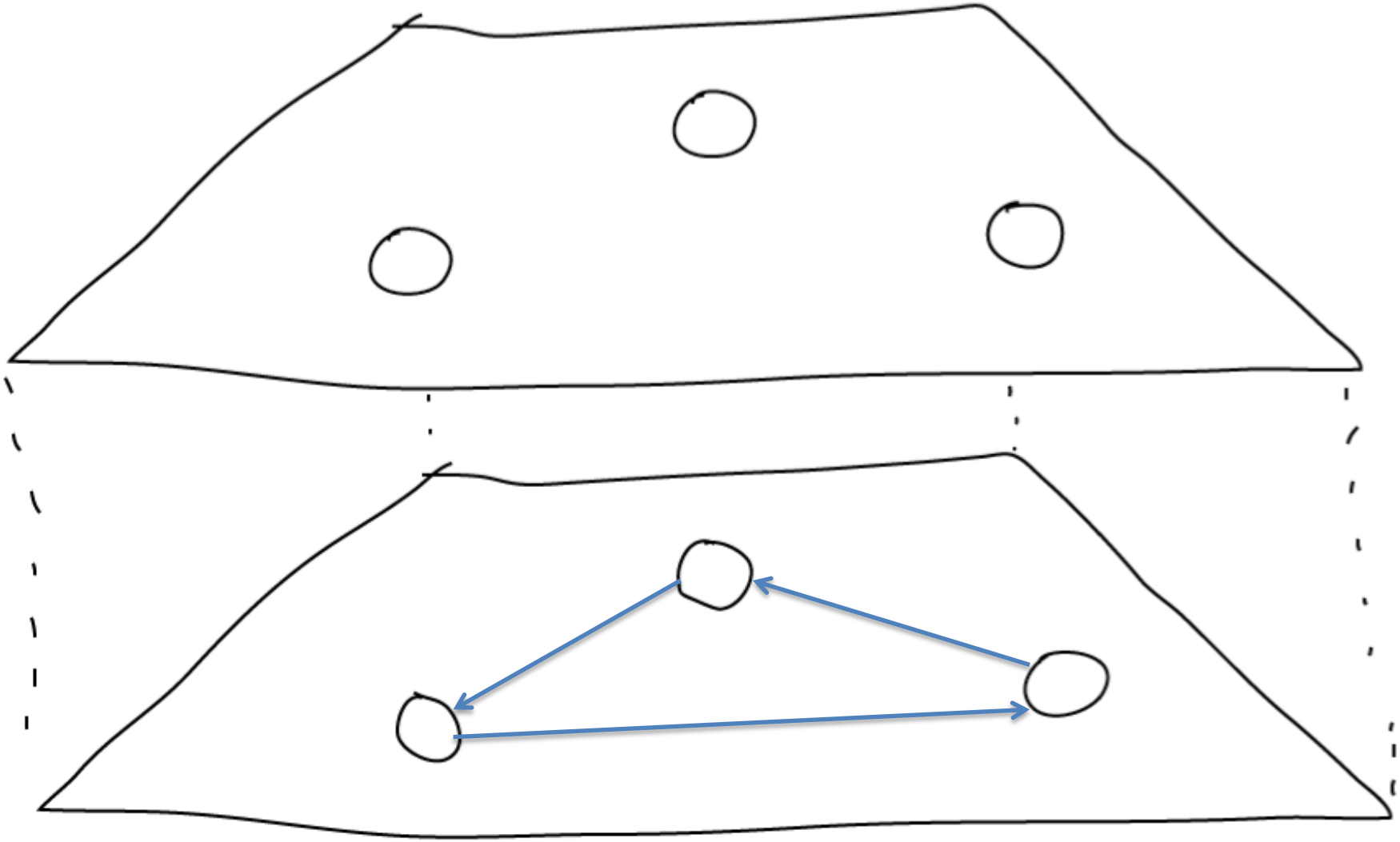


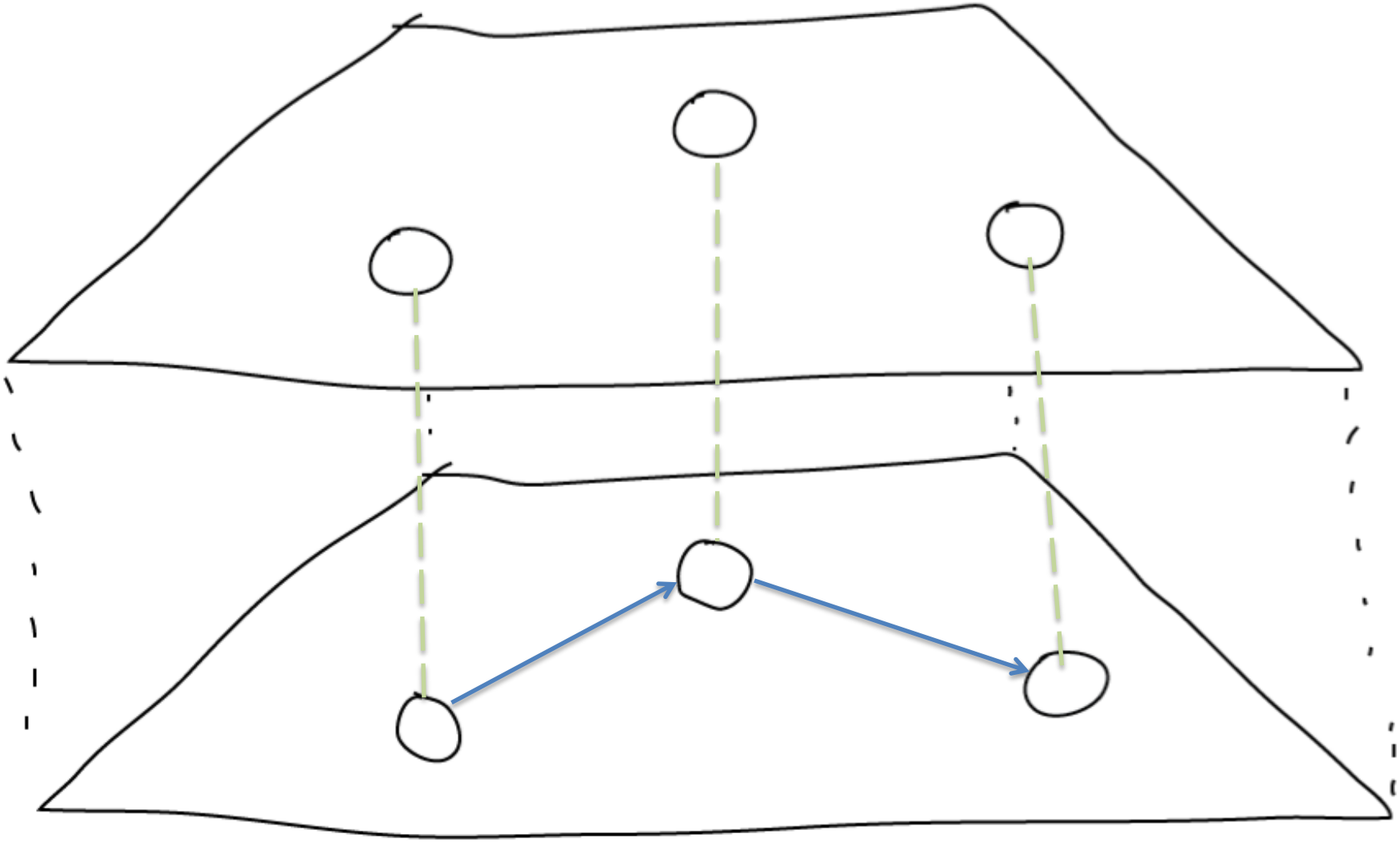


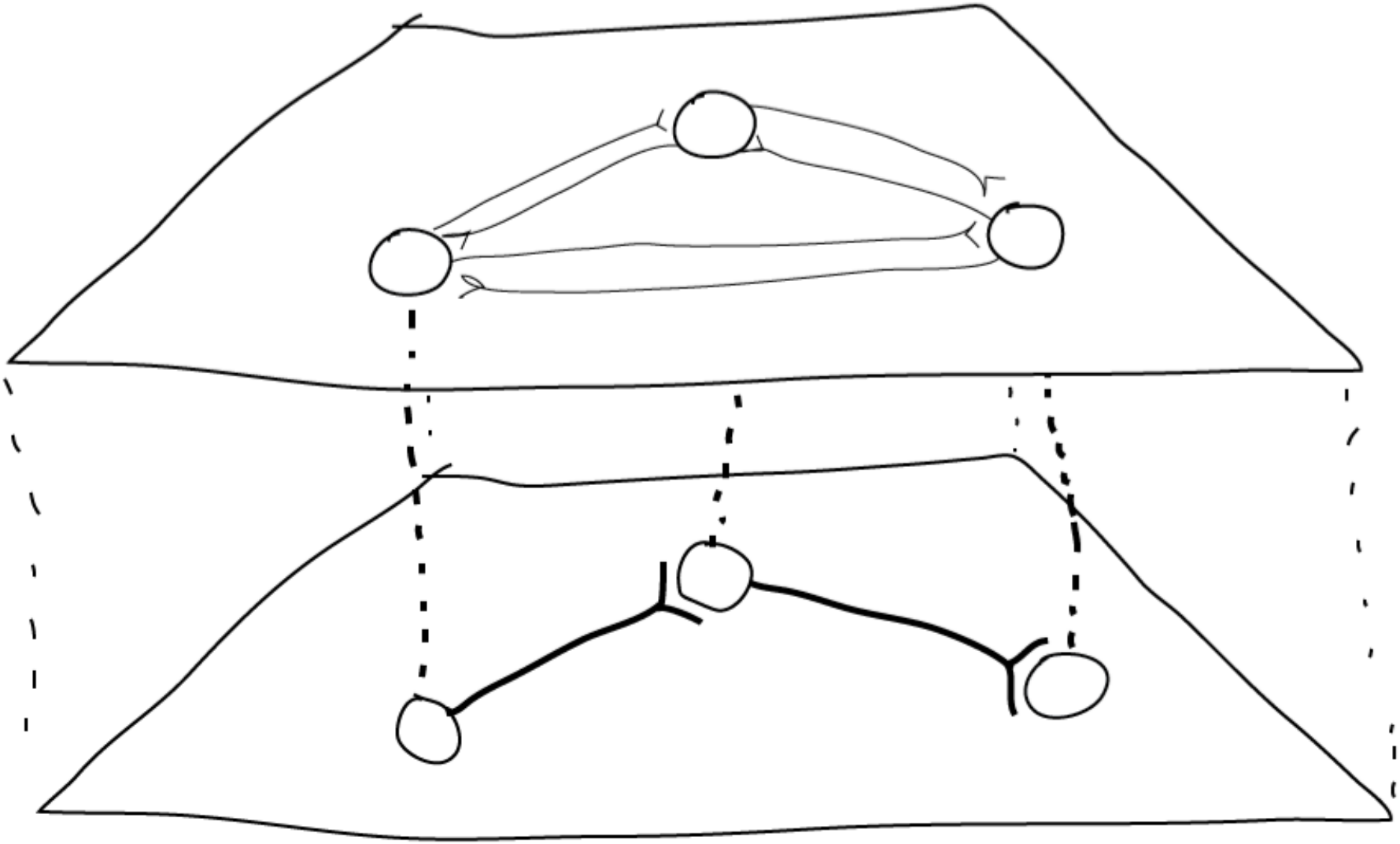


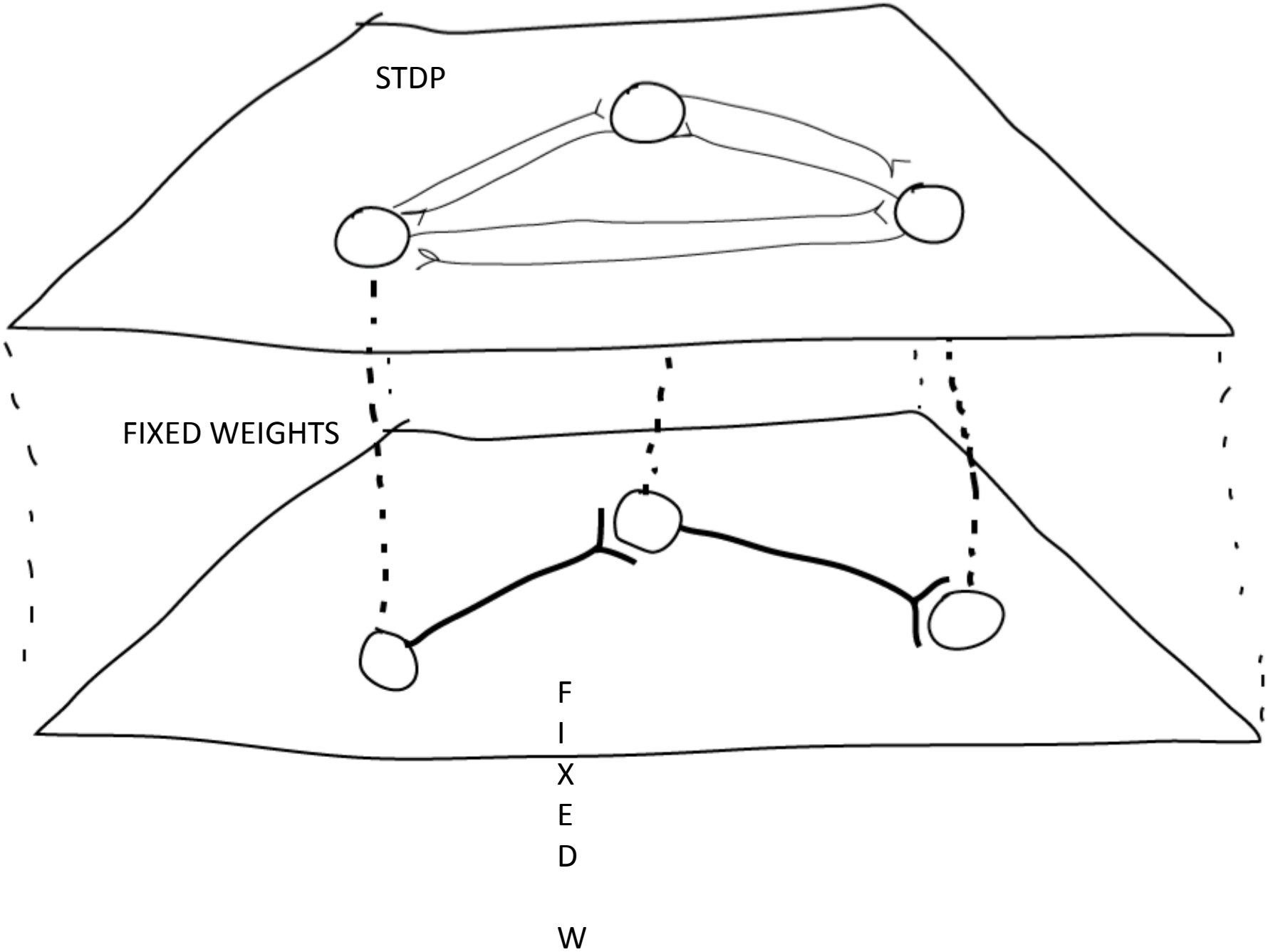


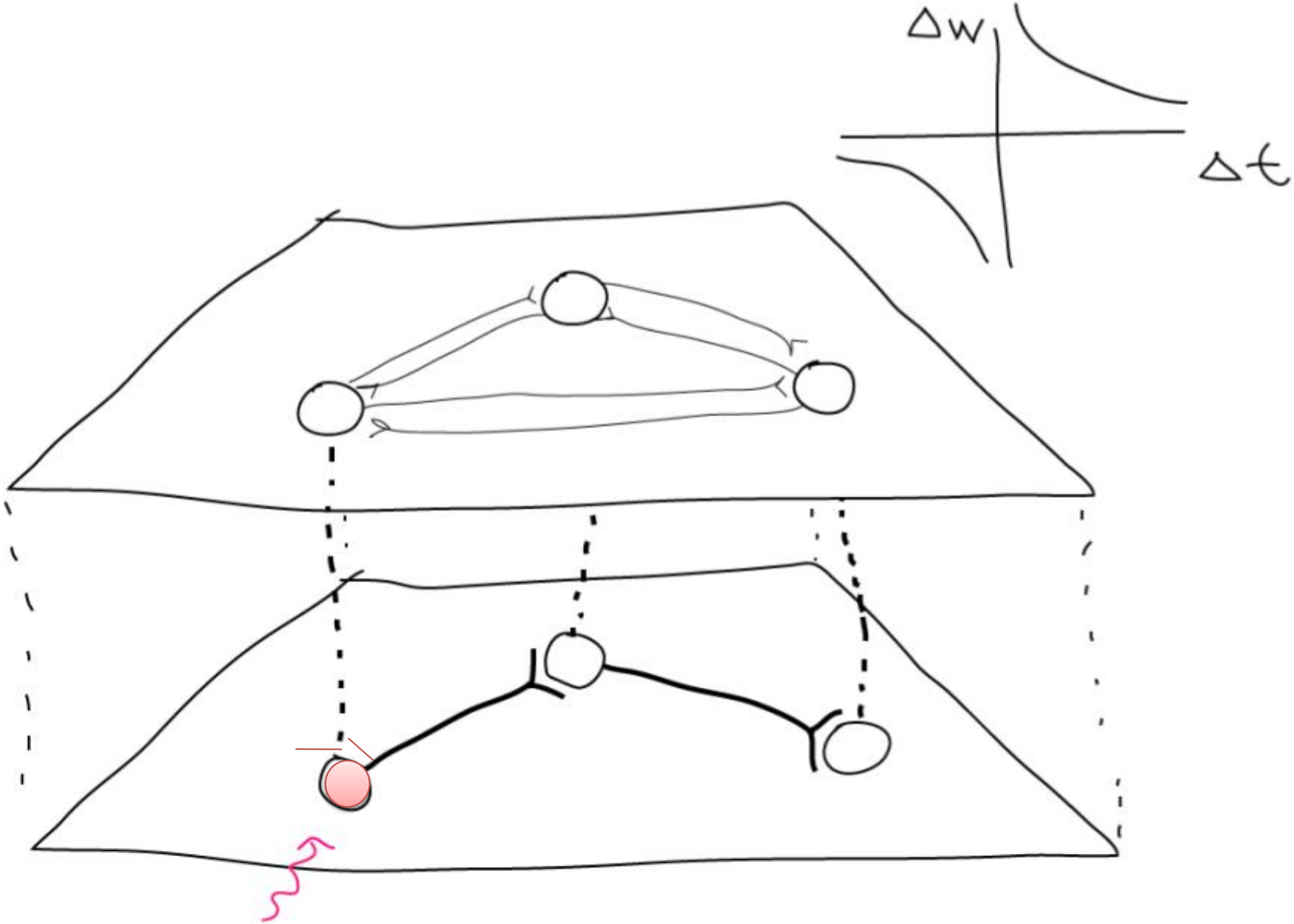


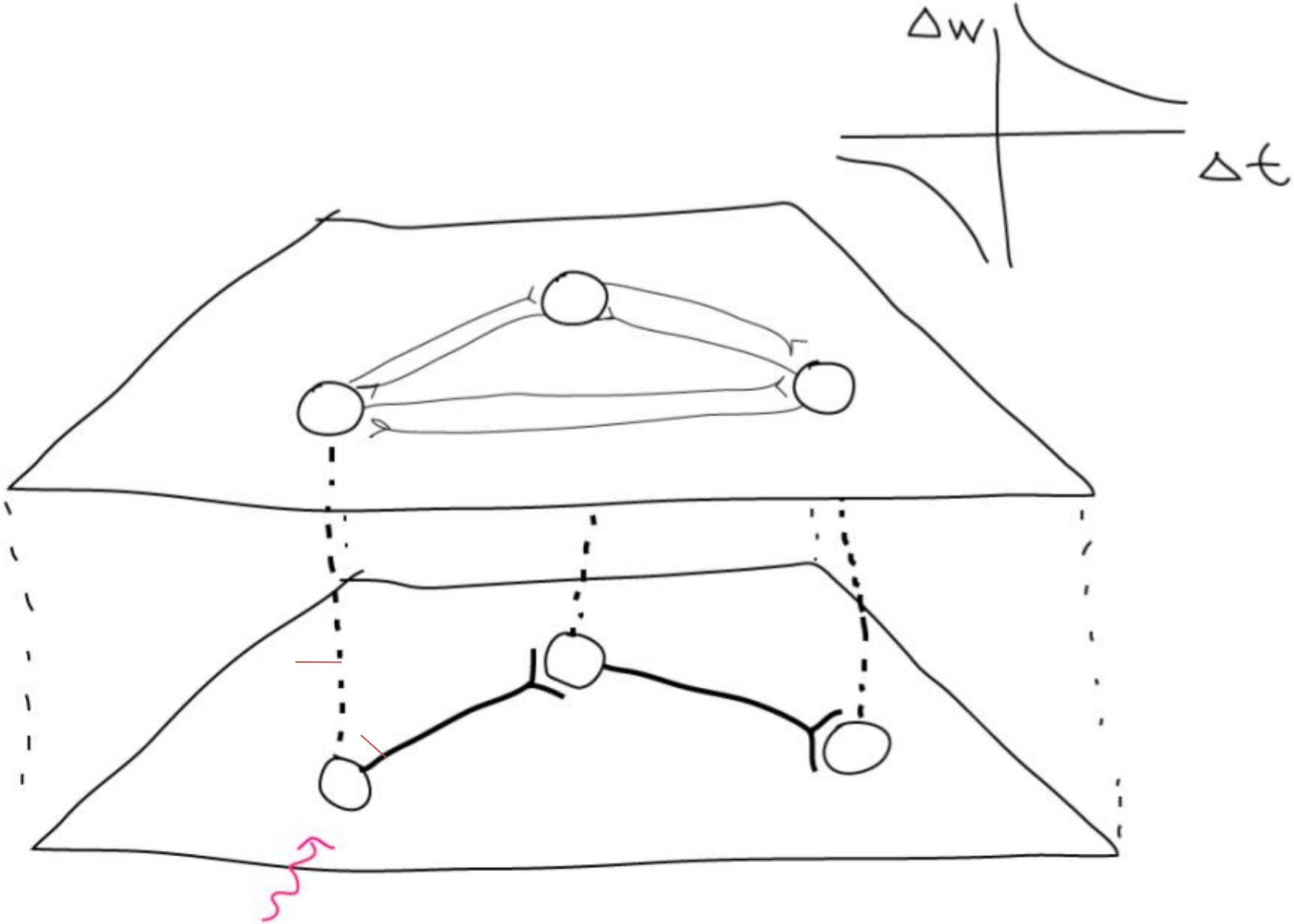


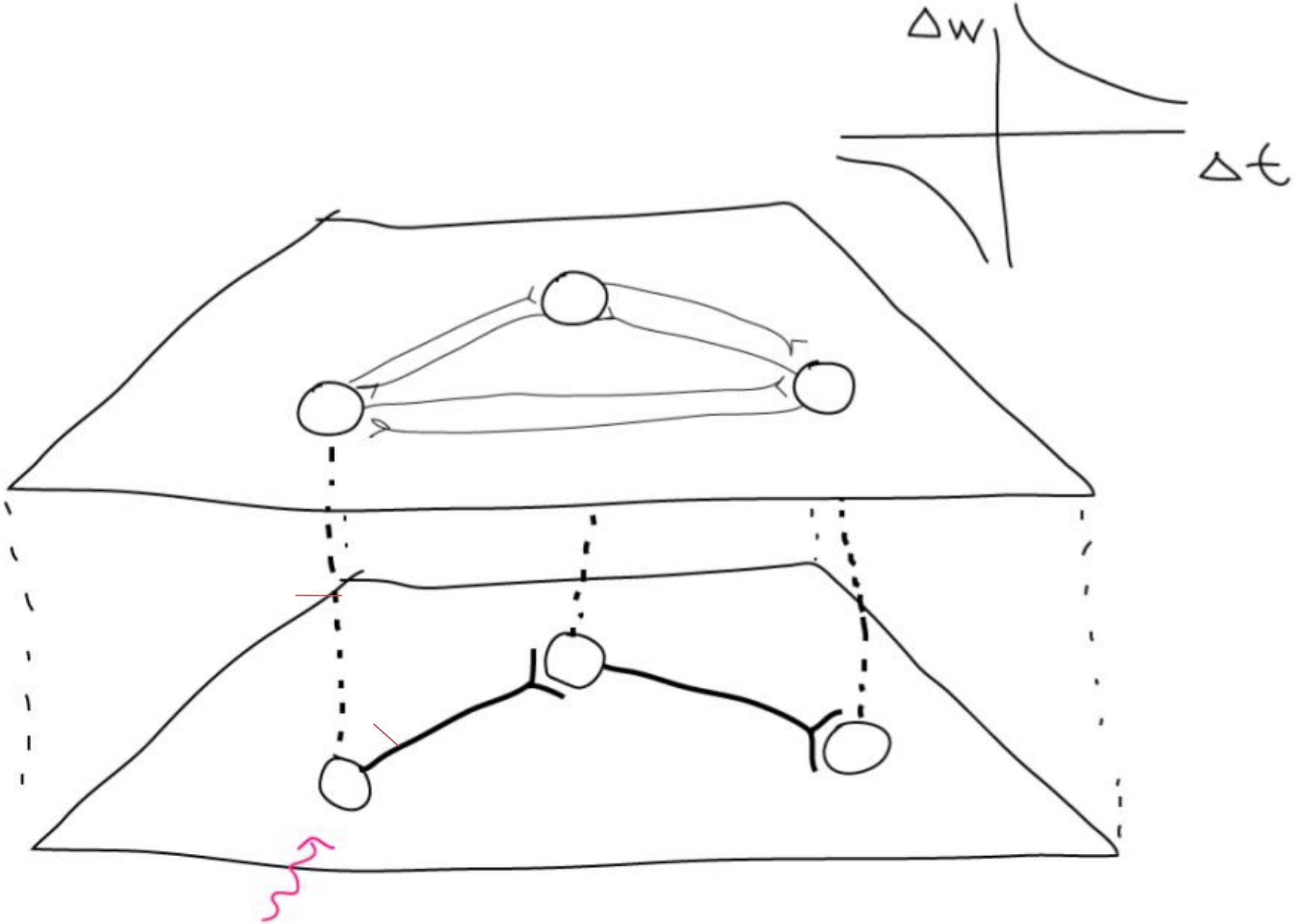


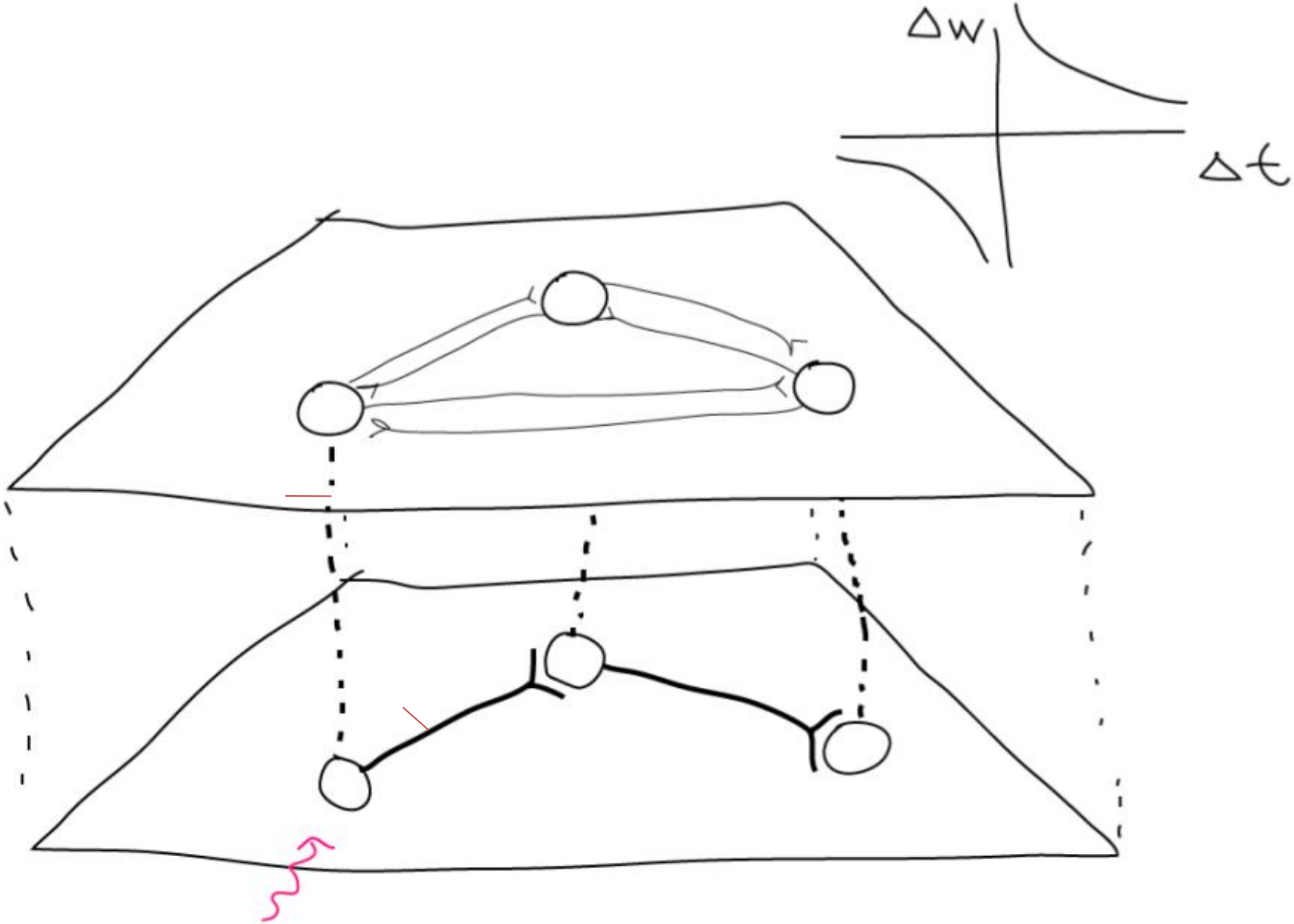


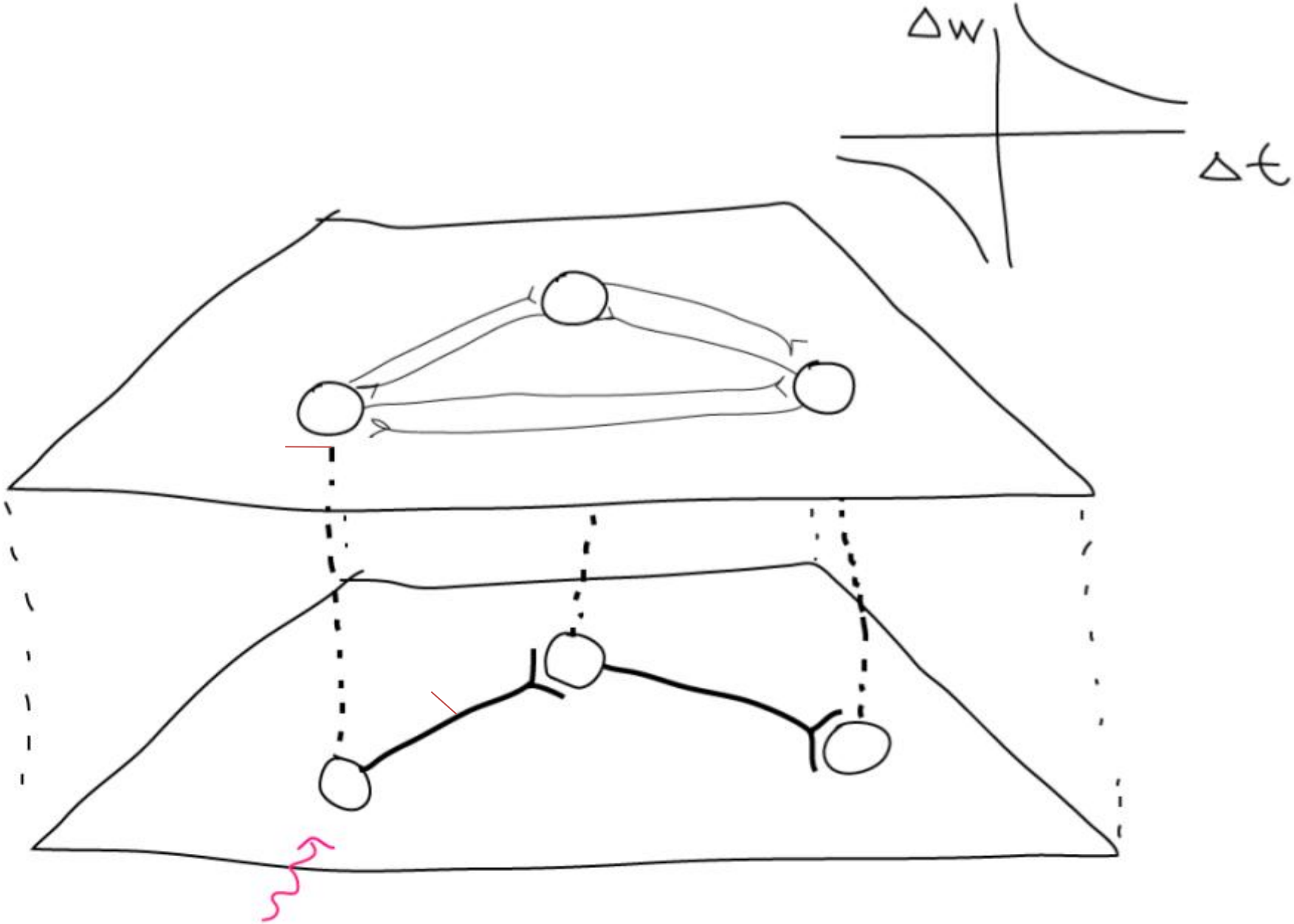

STDP
FIXED WEIGHTS
FIXED W

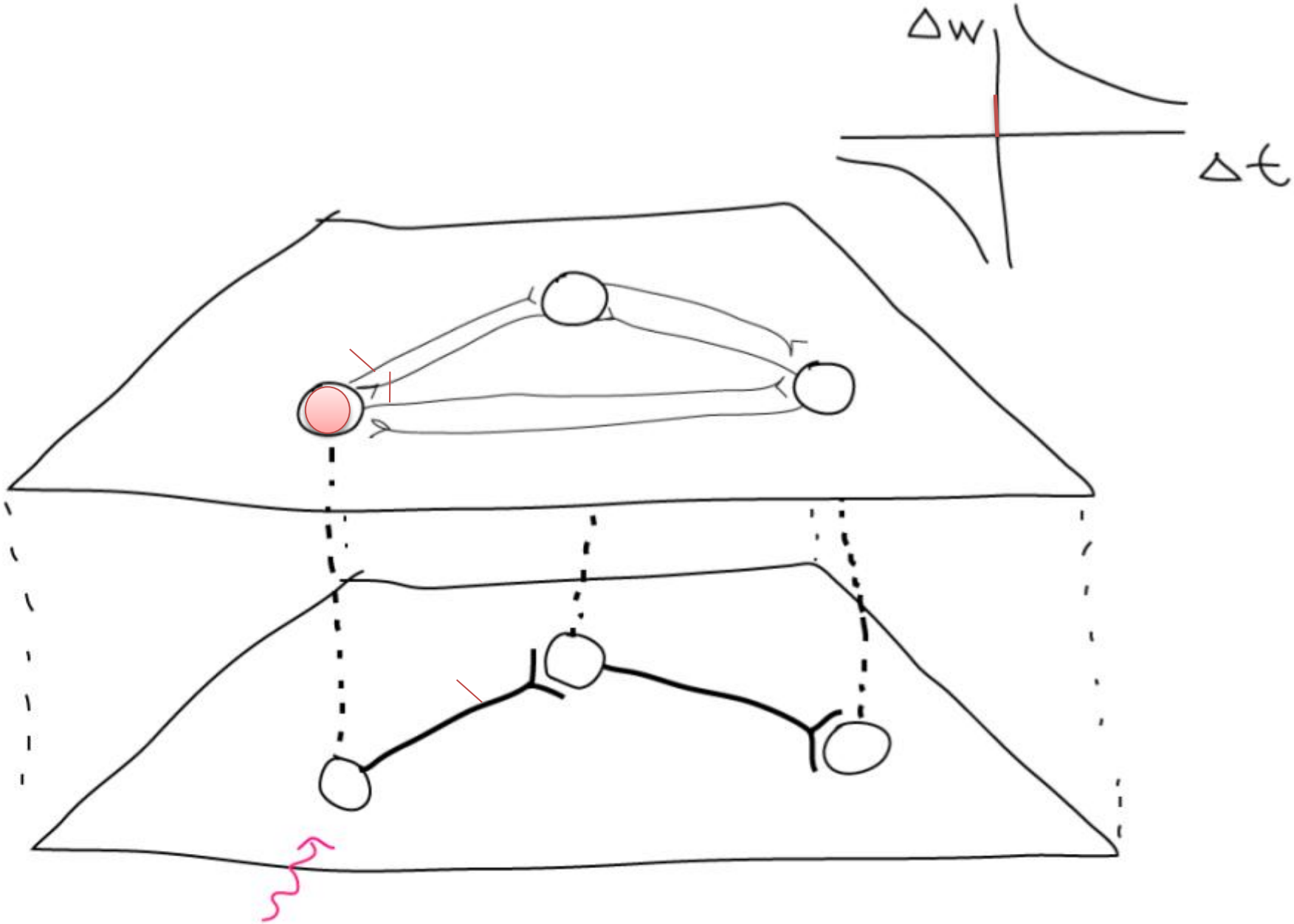

Δw
Δt

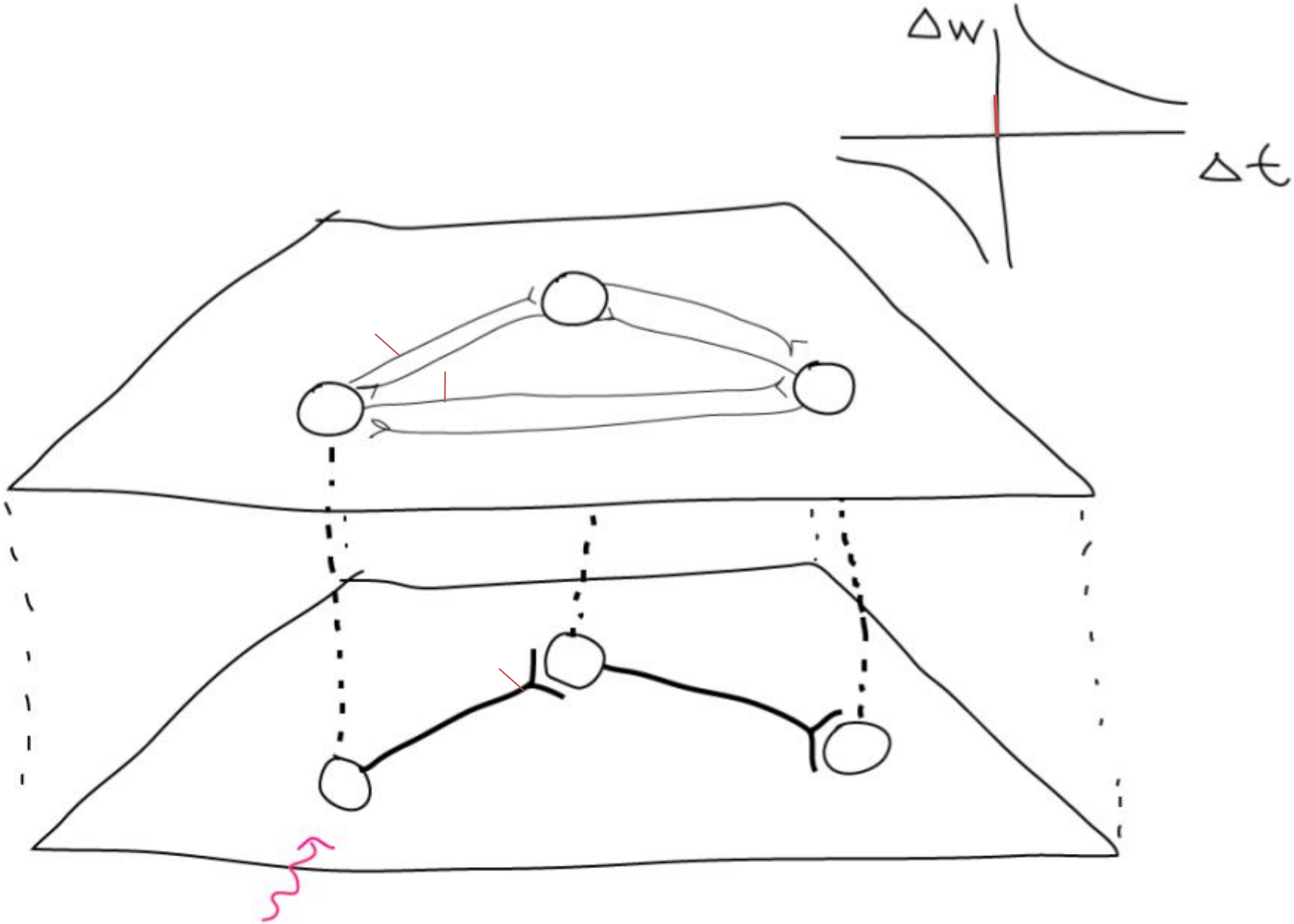

Δw
Δt

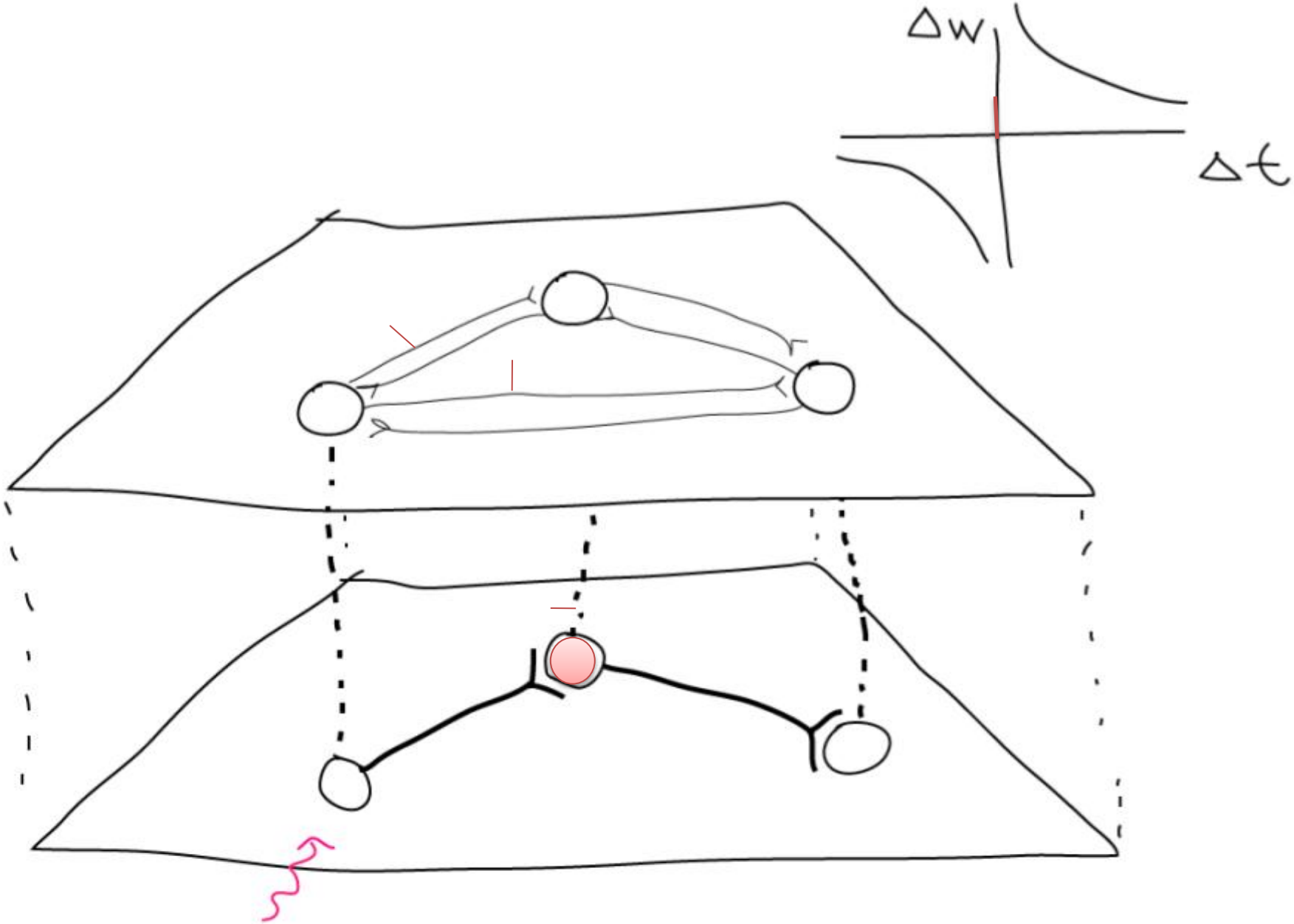

Δw
Δt

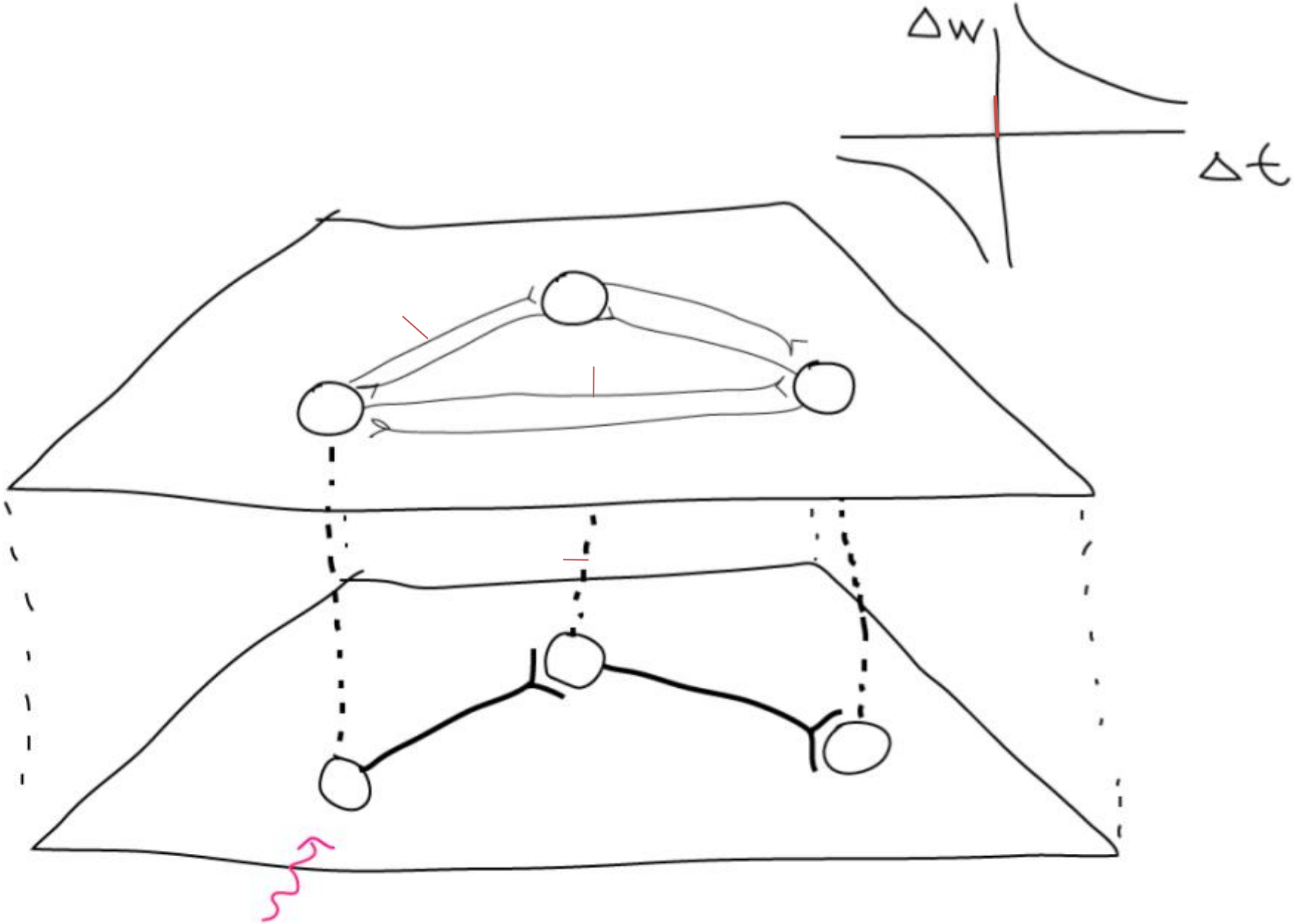

Δw
Δt

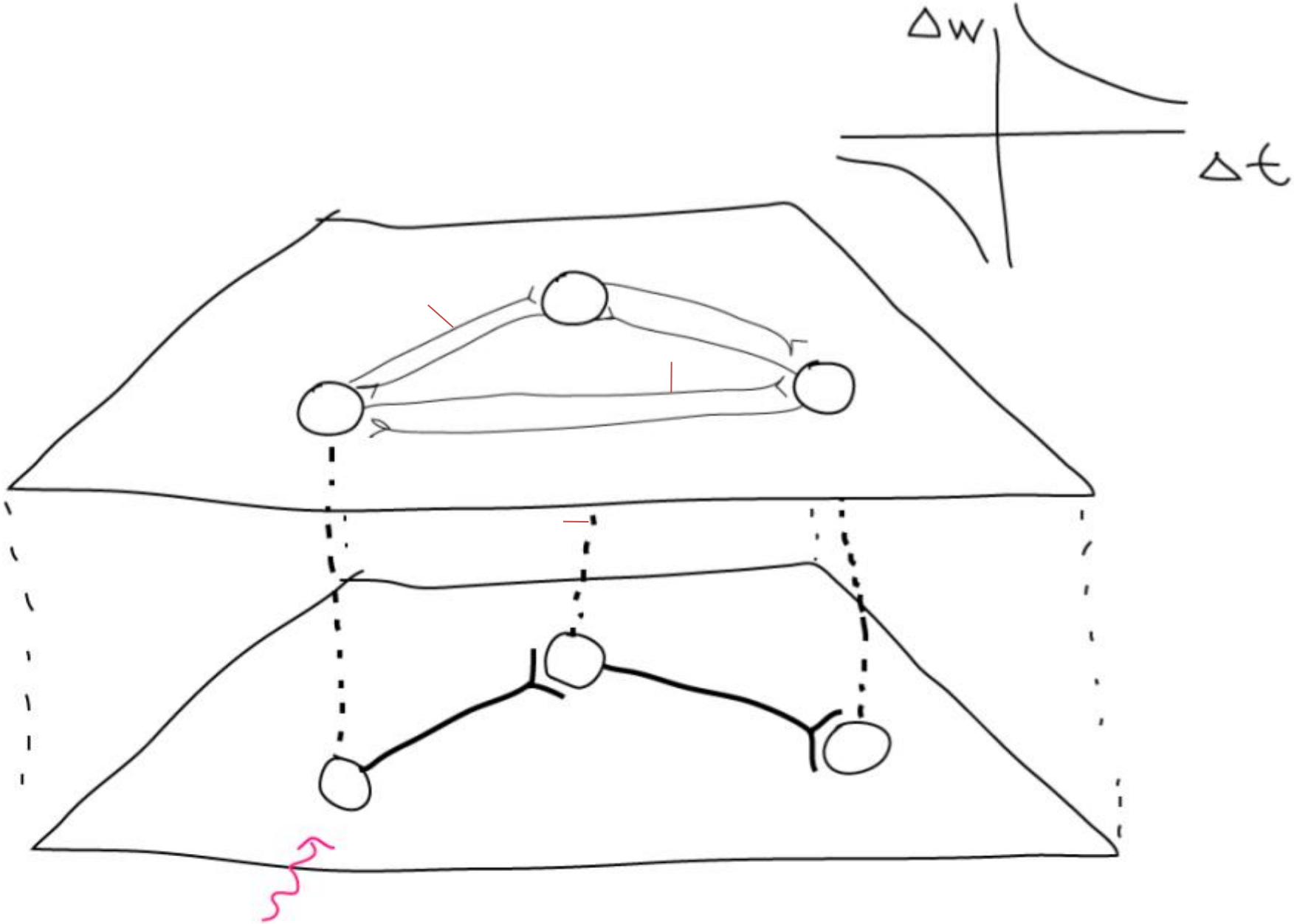

Δw
Δt

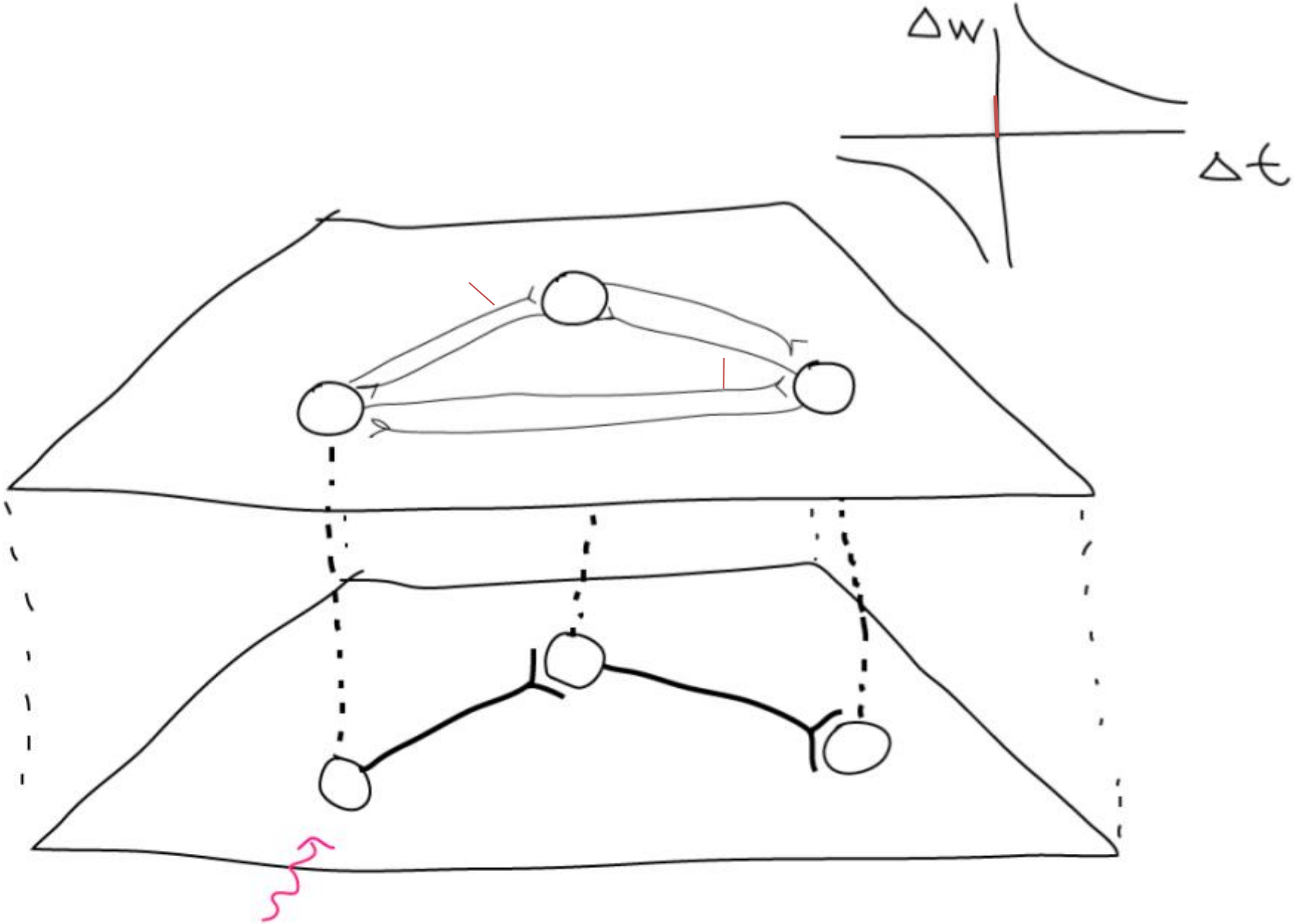

Δw
Δt

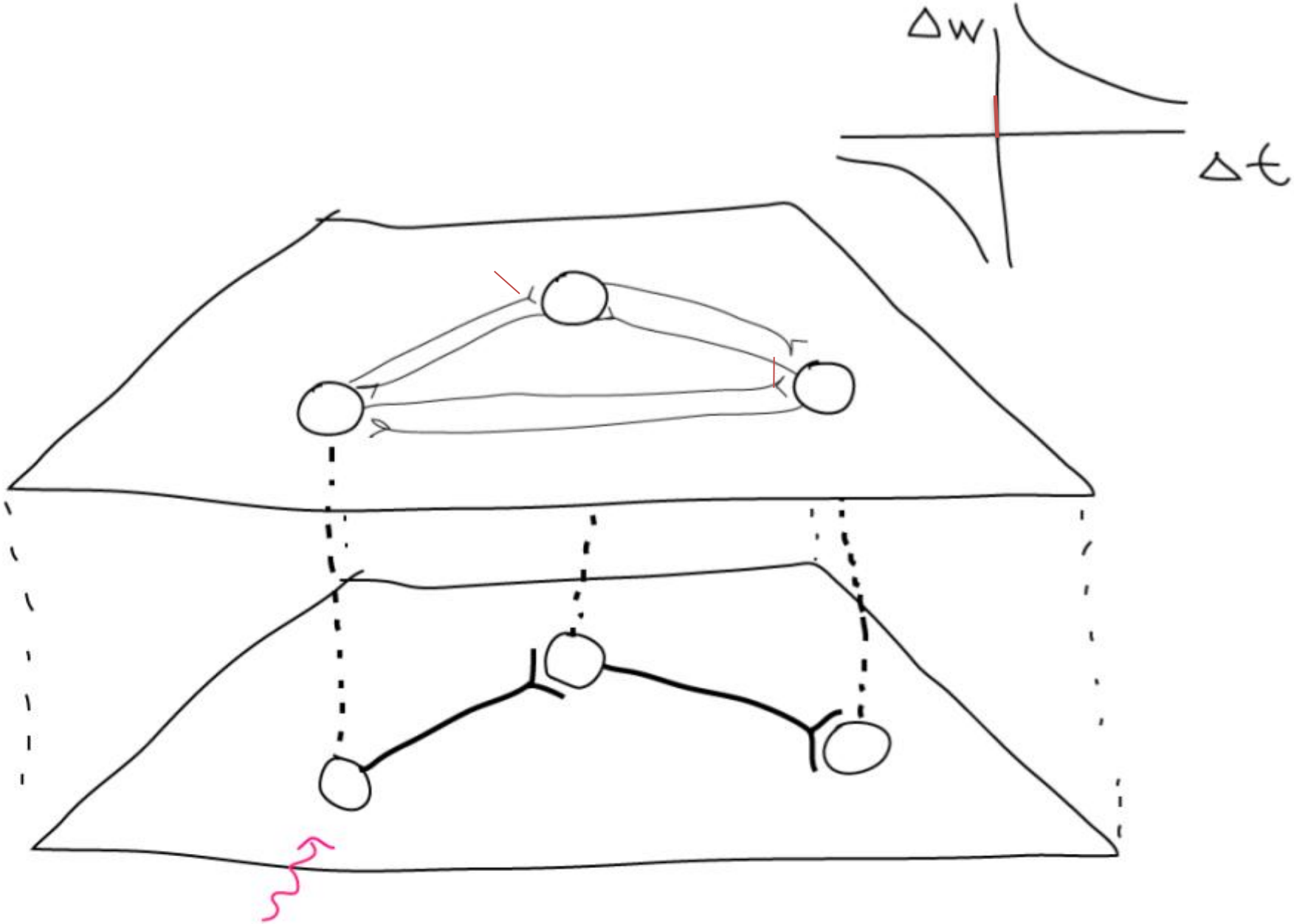

Δw
Δt

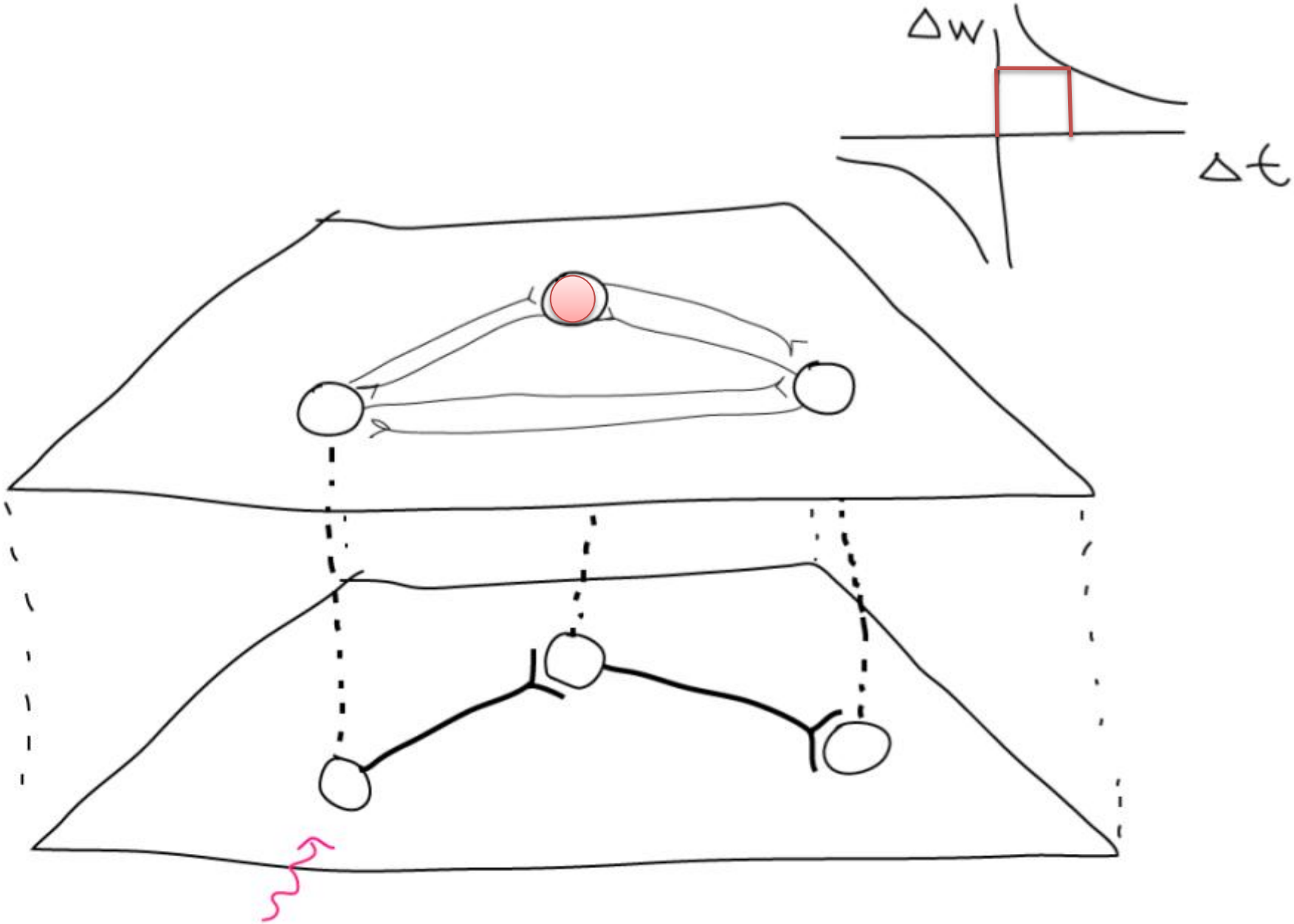

Δw
Δt

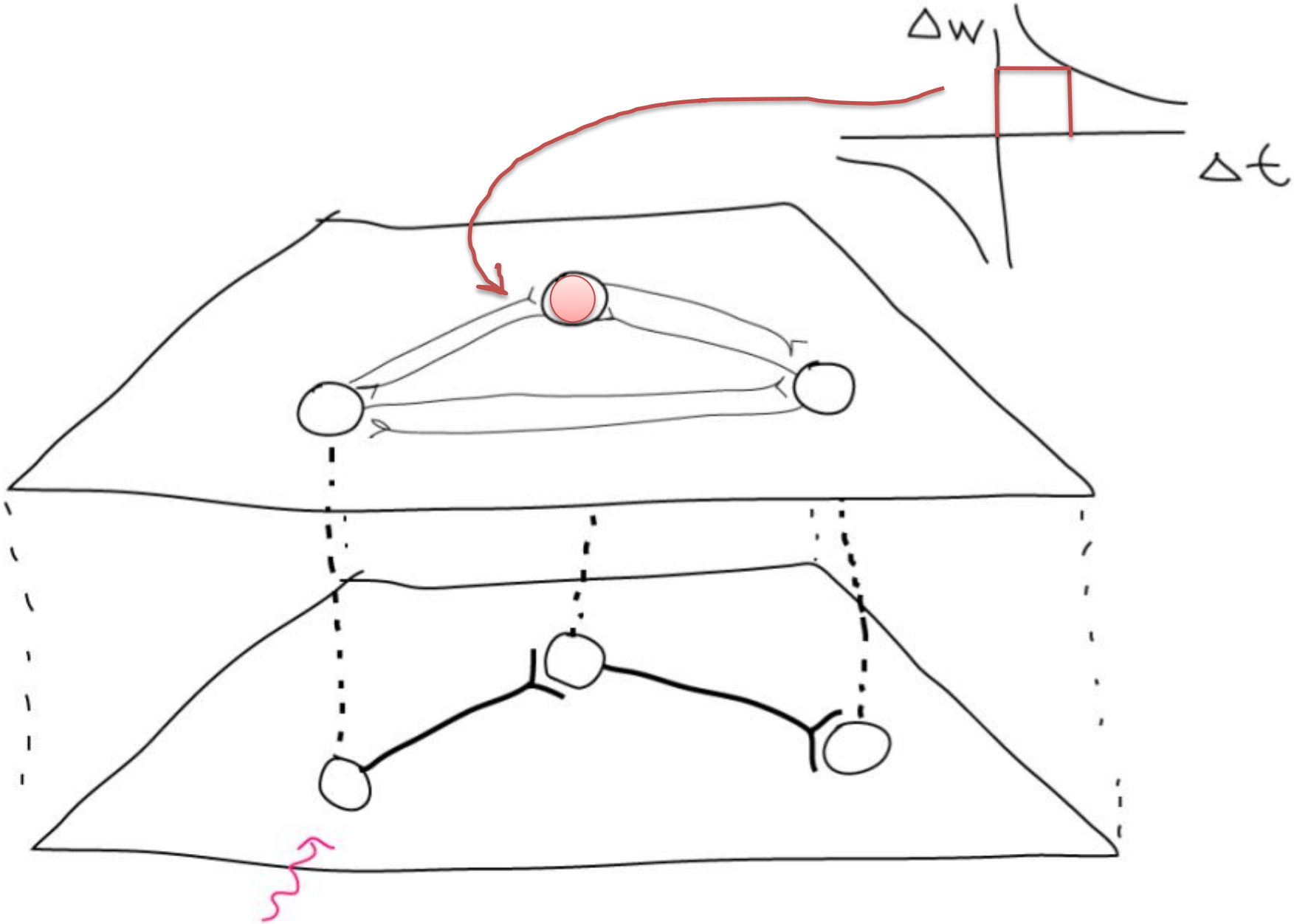

Δw
Δt

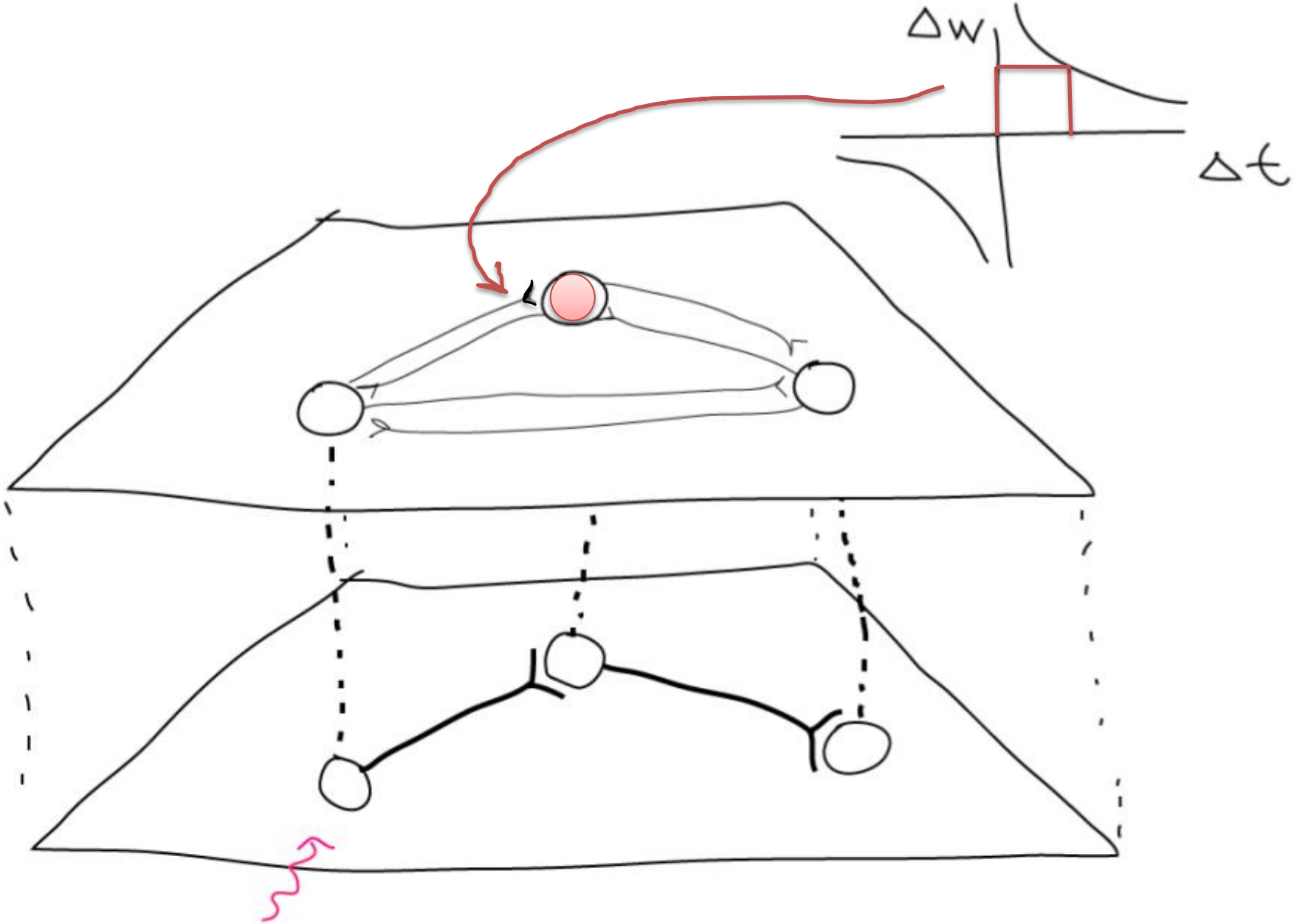

Δw
Δt

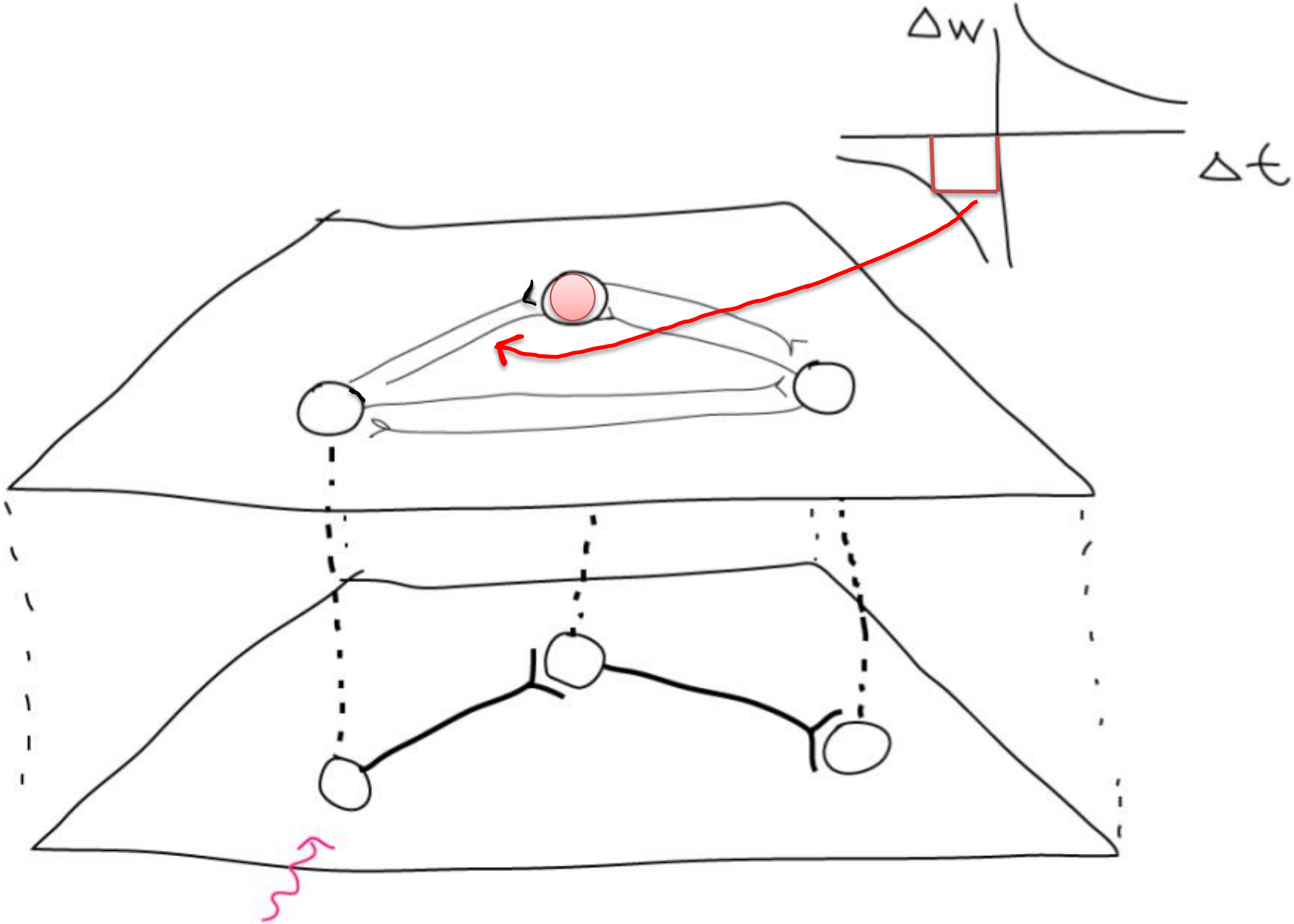

Δw
Δt

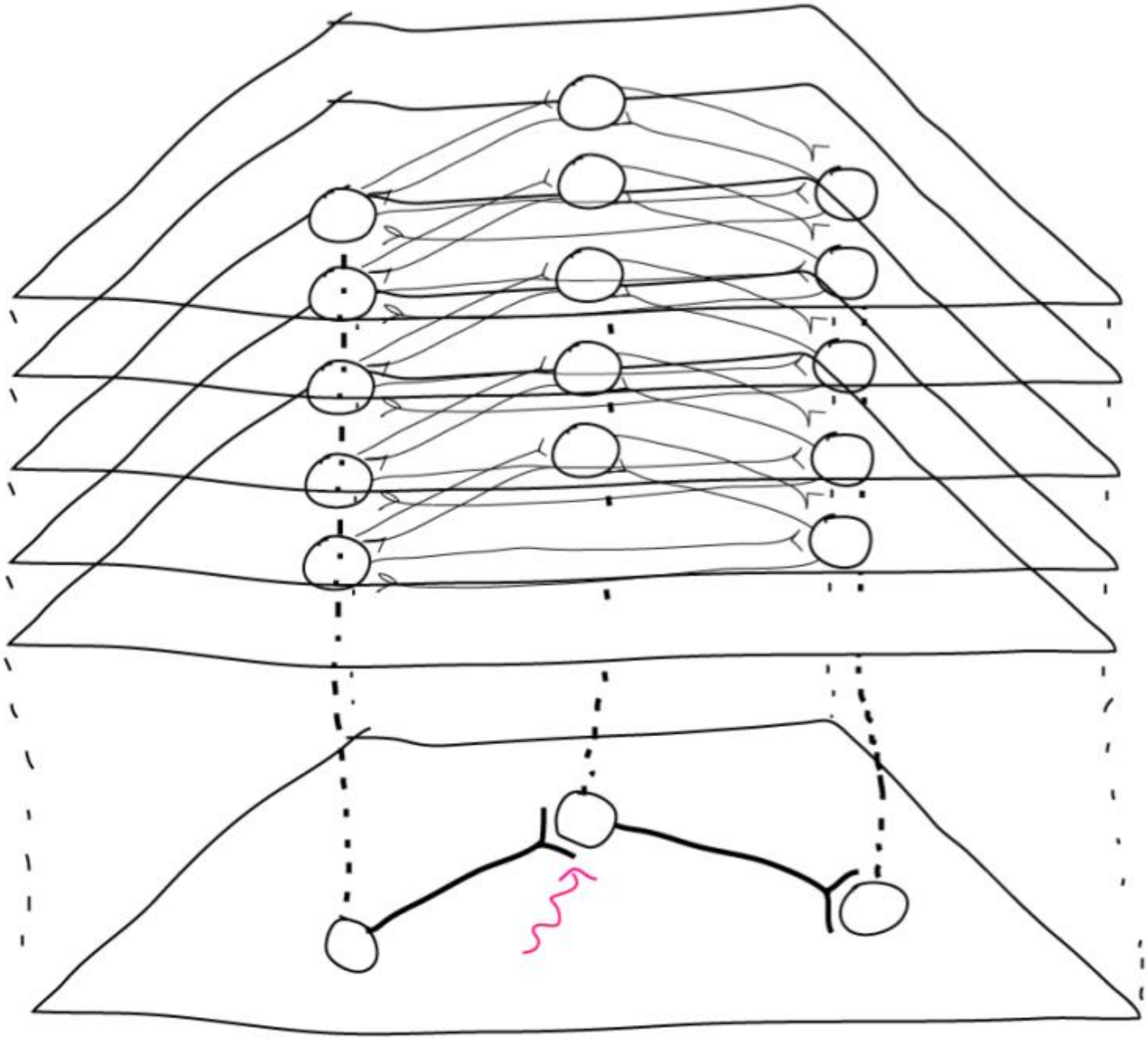

Δw
Δt

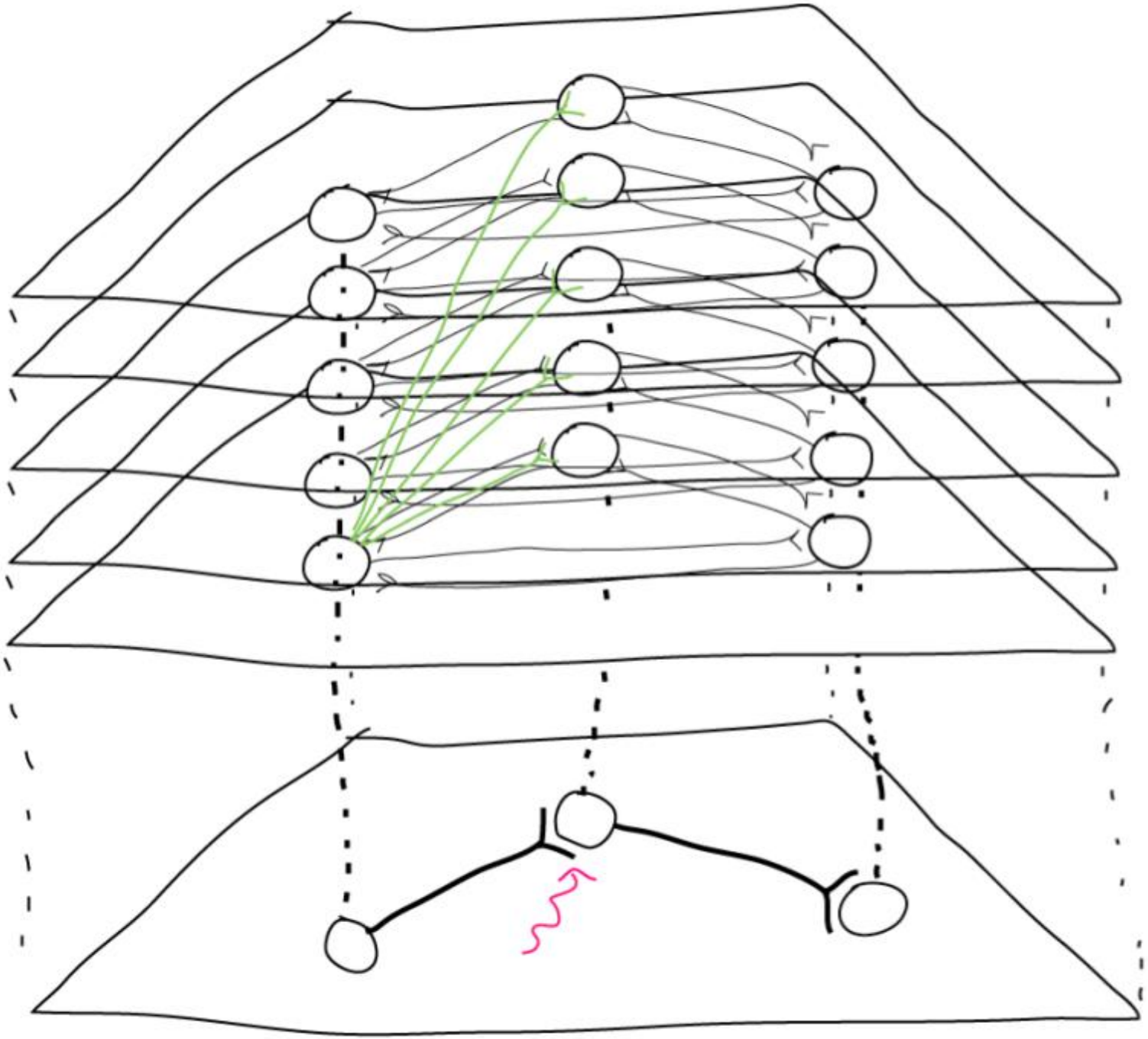

Δw
Δt

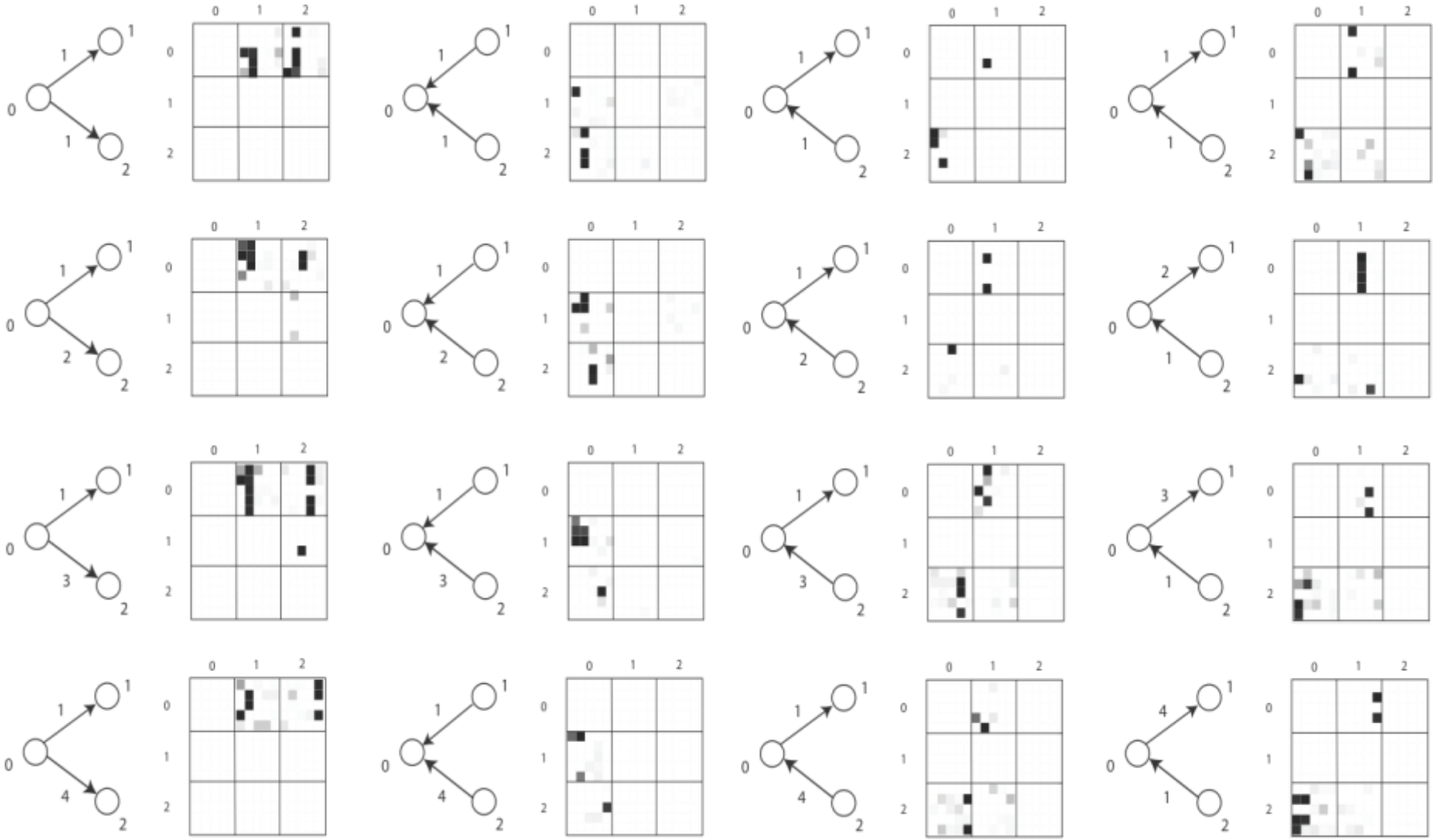

Δw
Δt


Δw
Δt


Δw
Δt





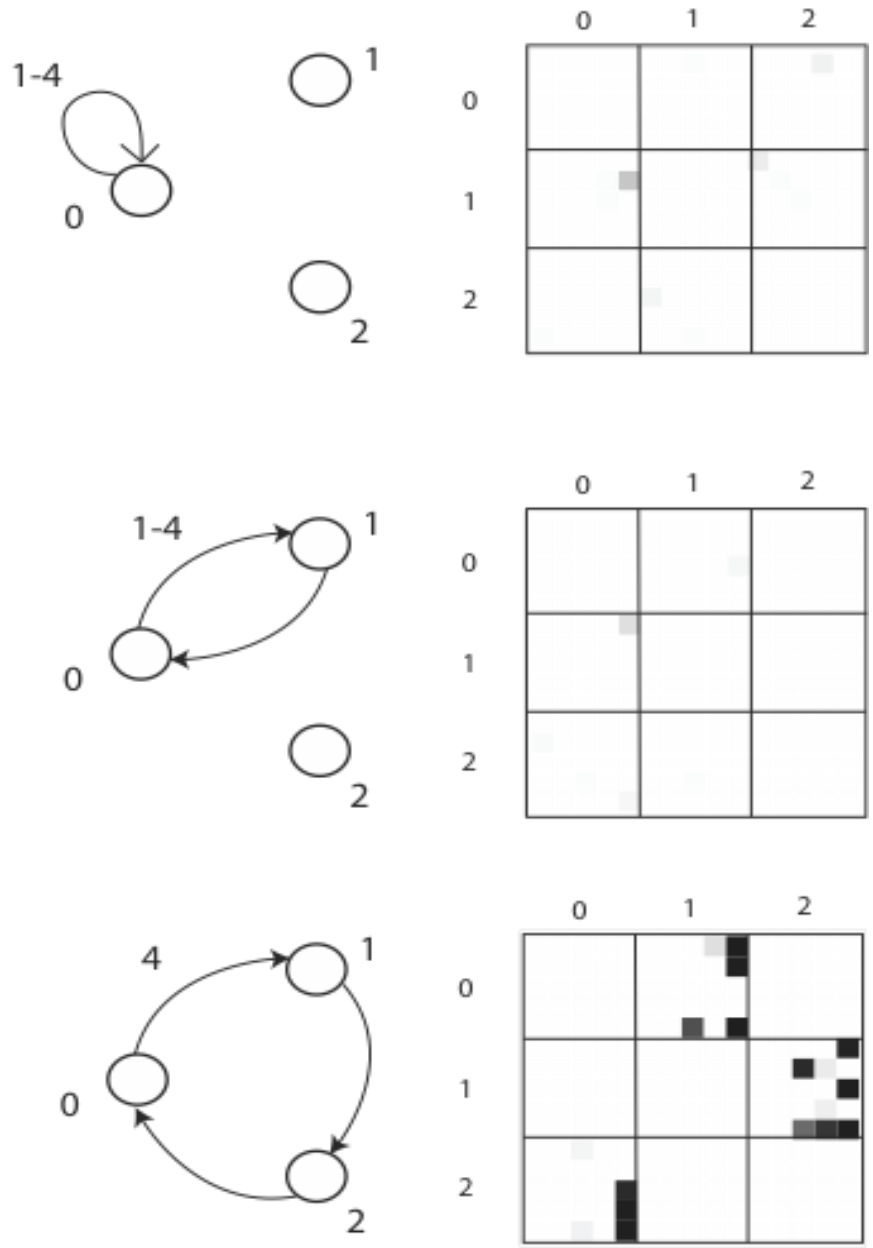


Common Cause

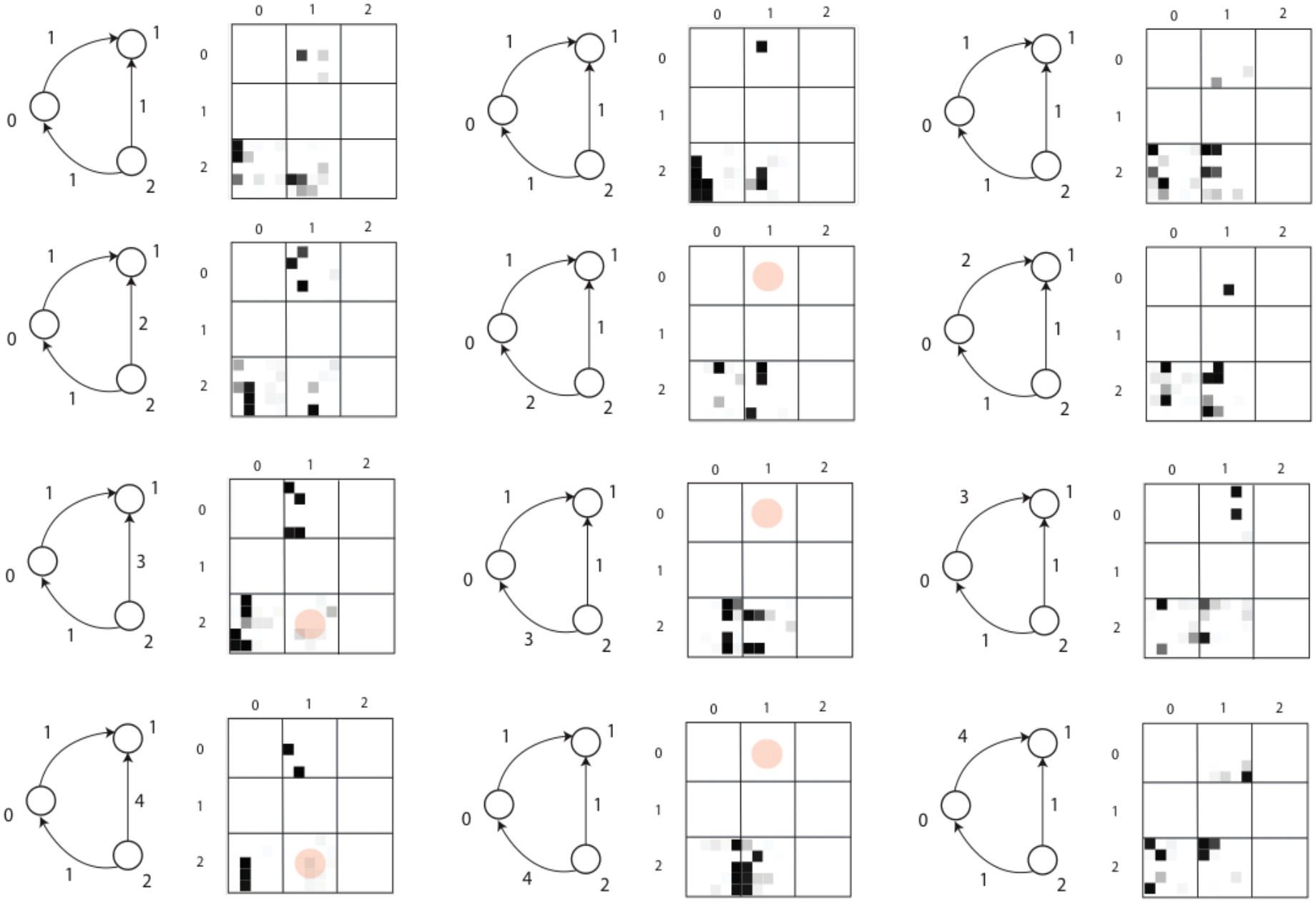
Common Effect

Causal Chain (Slow/Fast)

Causal Chain (Fast/Slow)



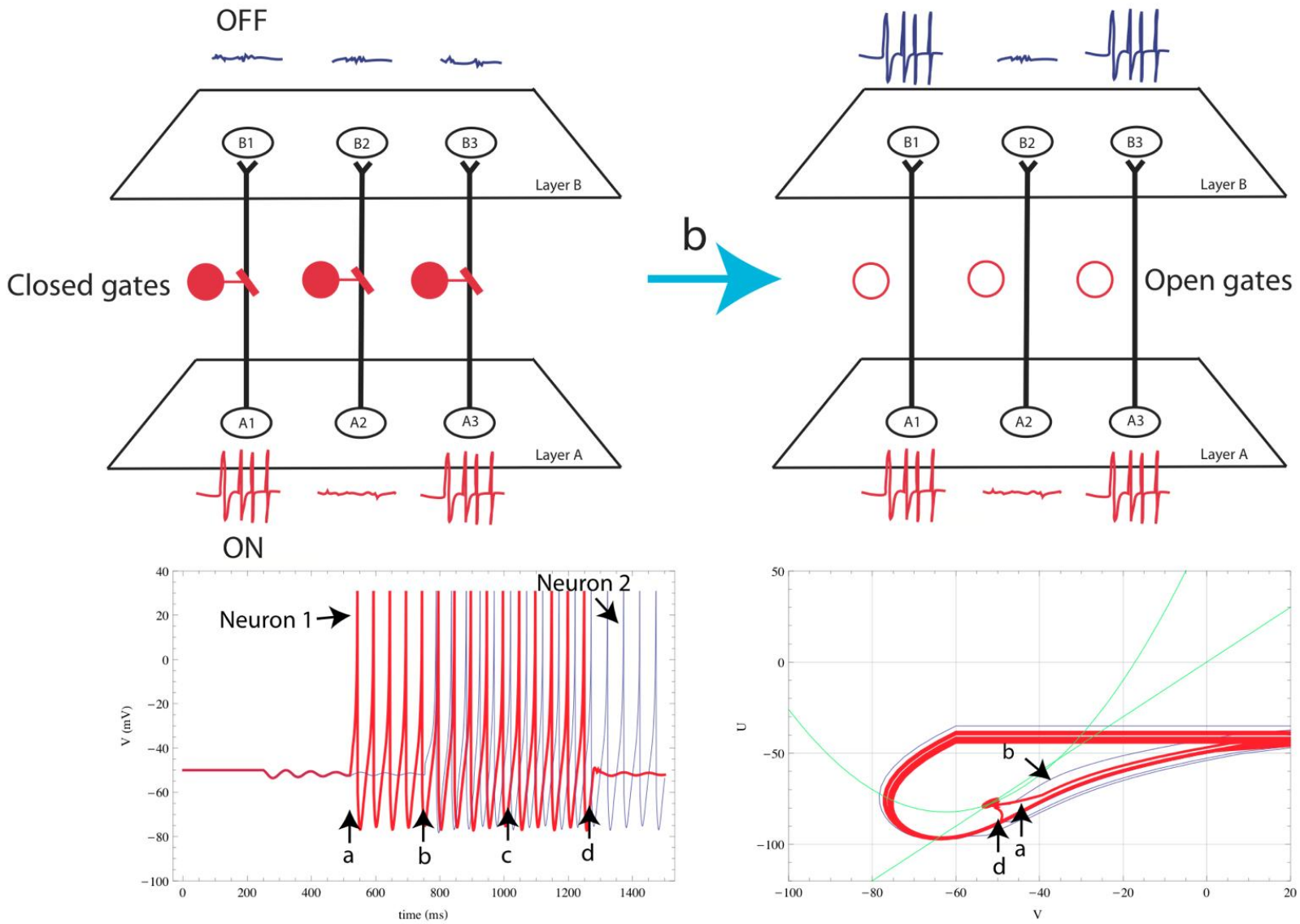
1-,2- and 3-cycles



Feedforward Loops

Possible Neuronal Replication Mechanisms

- Replication of Synaptic Connectivity Patterns
- Replication of Activity Patterns
- Evolvable Paths of Activity: Overlapping Units of Evolution


OFF
B1
B2
B3
Layer B
Closed gates
b
Open gates
A1
A2
A3
Layer A
ON
Neuron 1
Neuron 2
V (mV)
time (ms)
a
b
c
d
U
V

Hebbian Learning can Structure Exploration Distributions

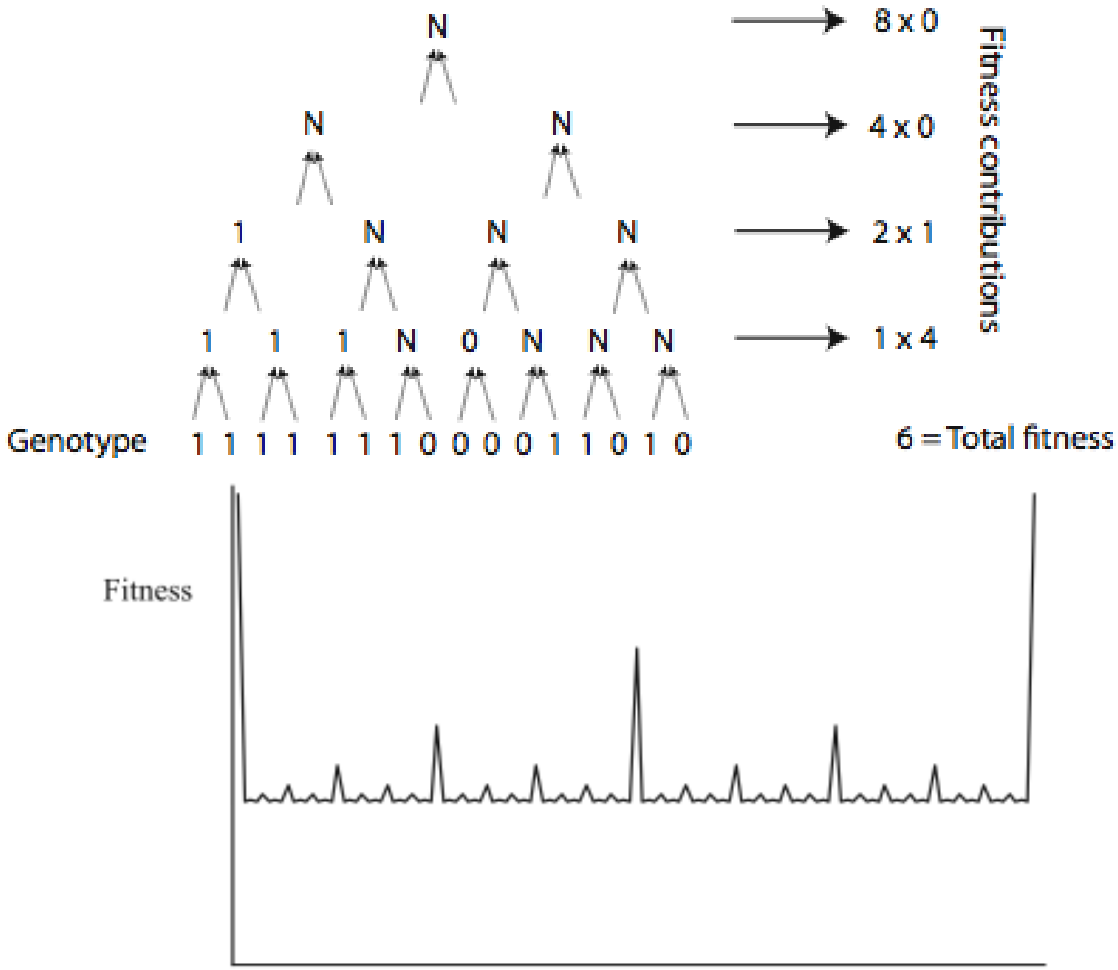
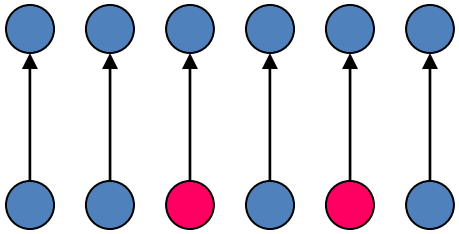


Figure 4.8
A particular cross section through the HIFF fitness landscape.



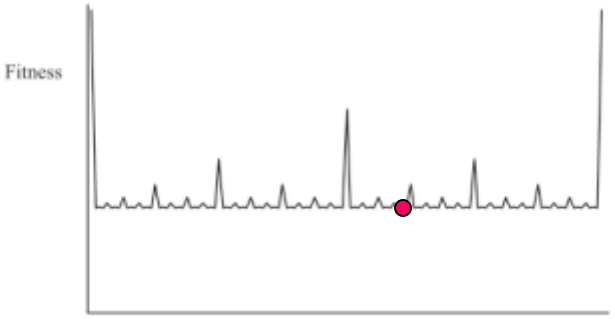
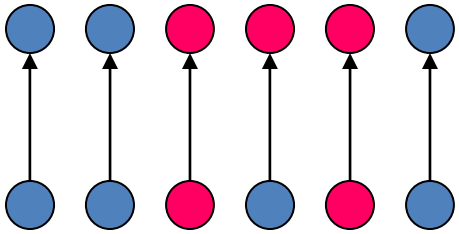
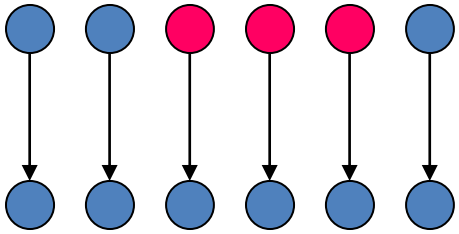
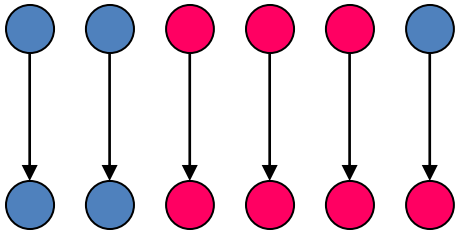
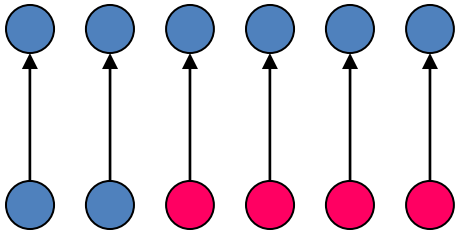


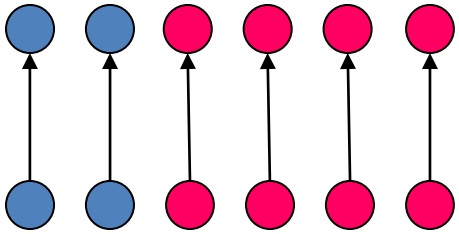
Figure 4.8
A particular cross section through the HIFF fitness landscape.











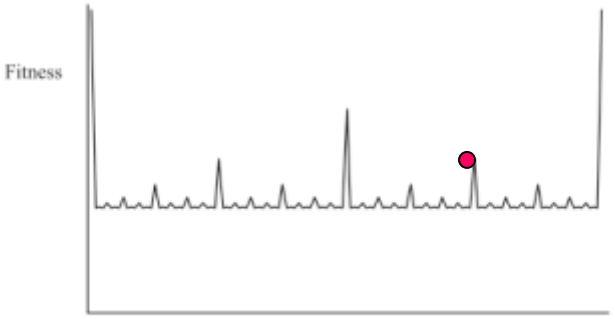
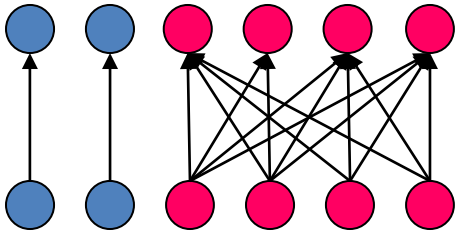
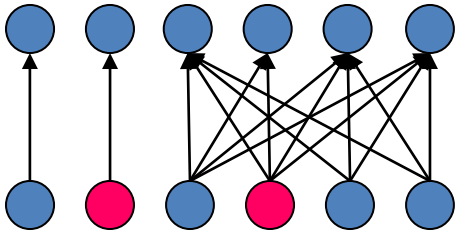
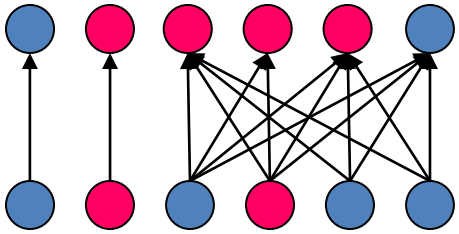


Figure 4.8
A particular cross section through the HIFF fitness landscape.

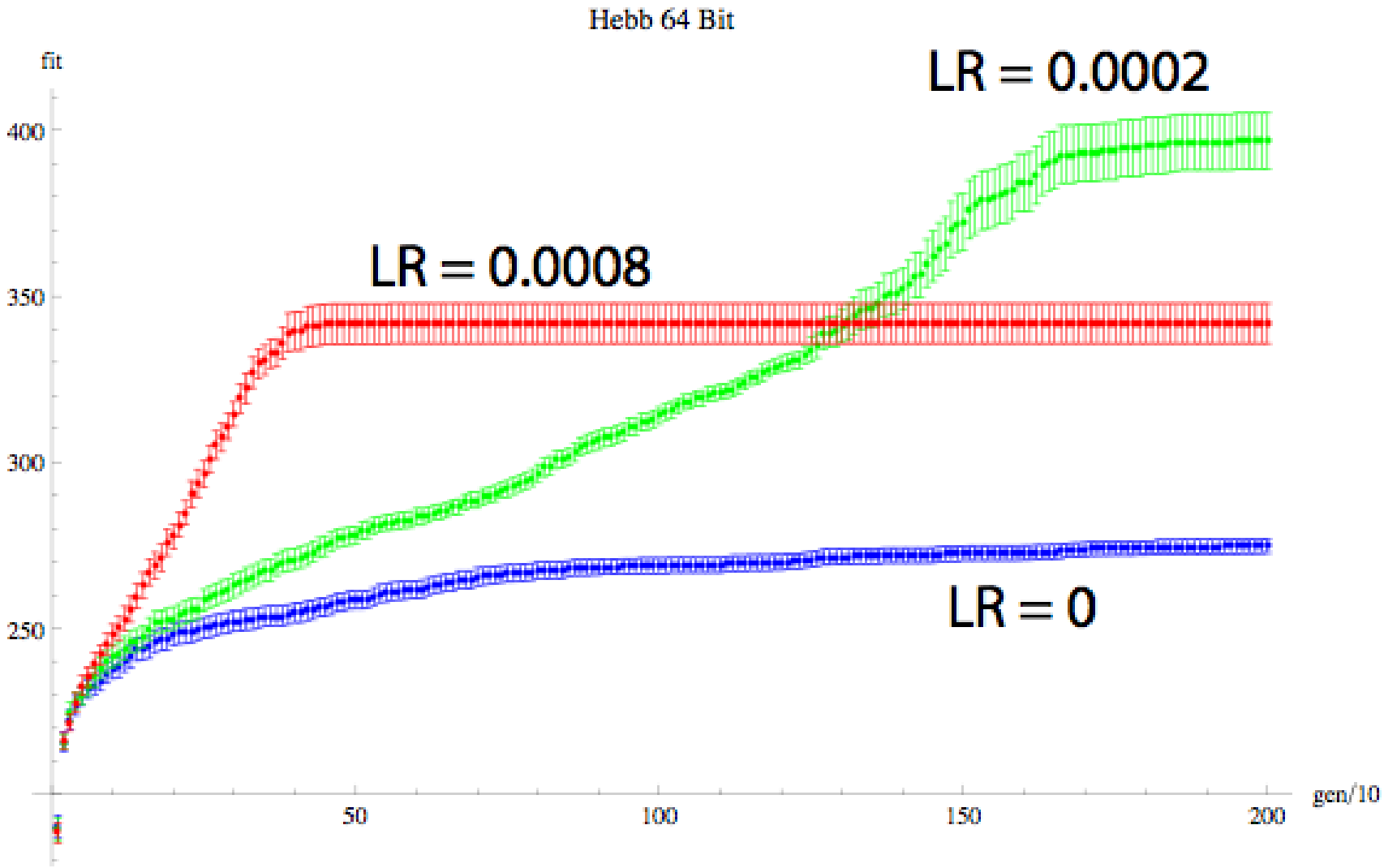




Synergy between classical Darwinian dynamics and other neurobiological component mechanisms

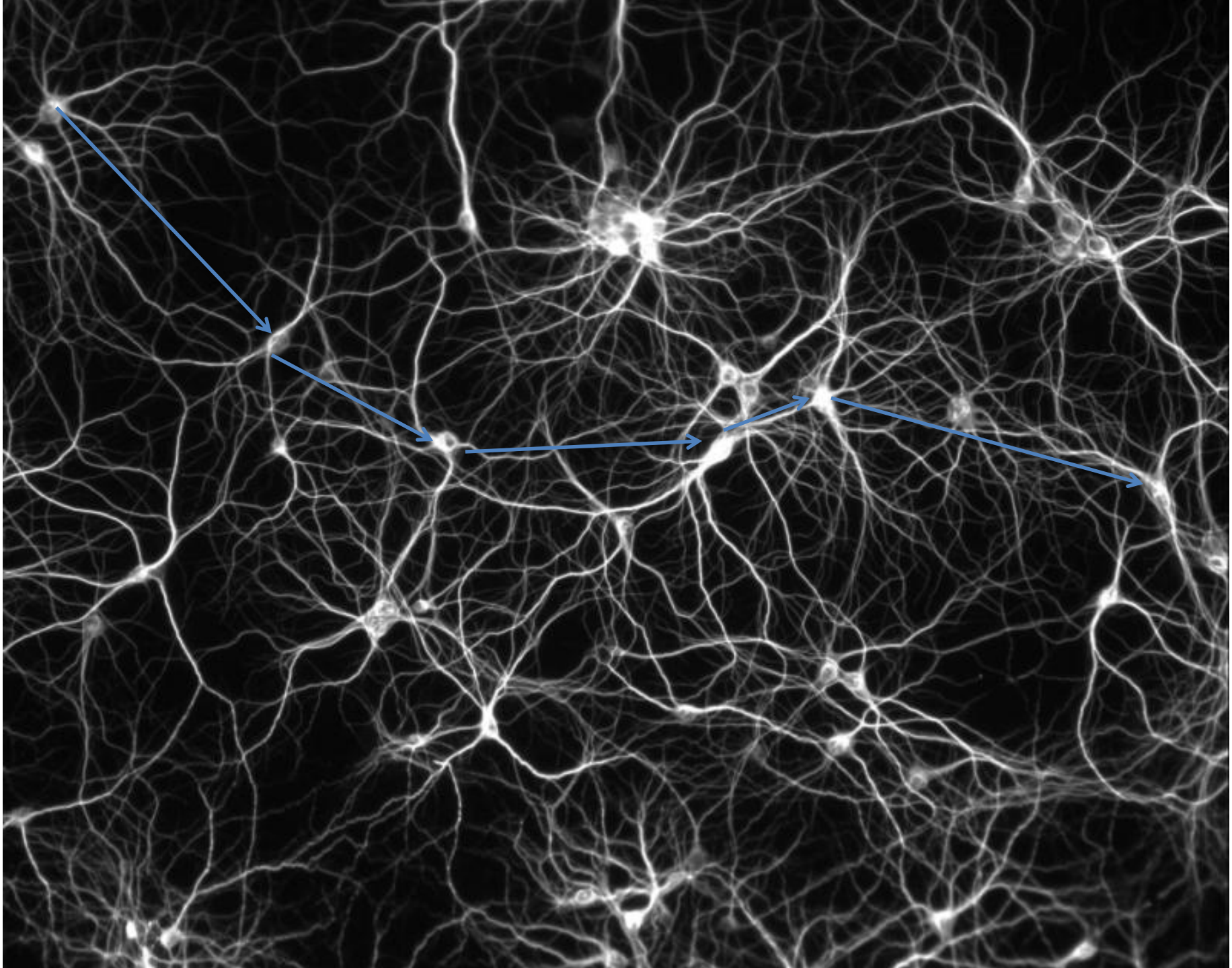


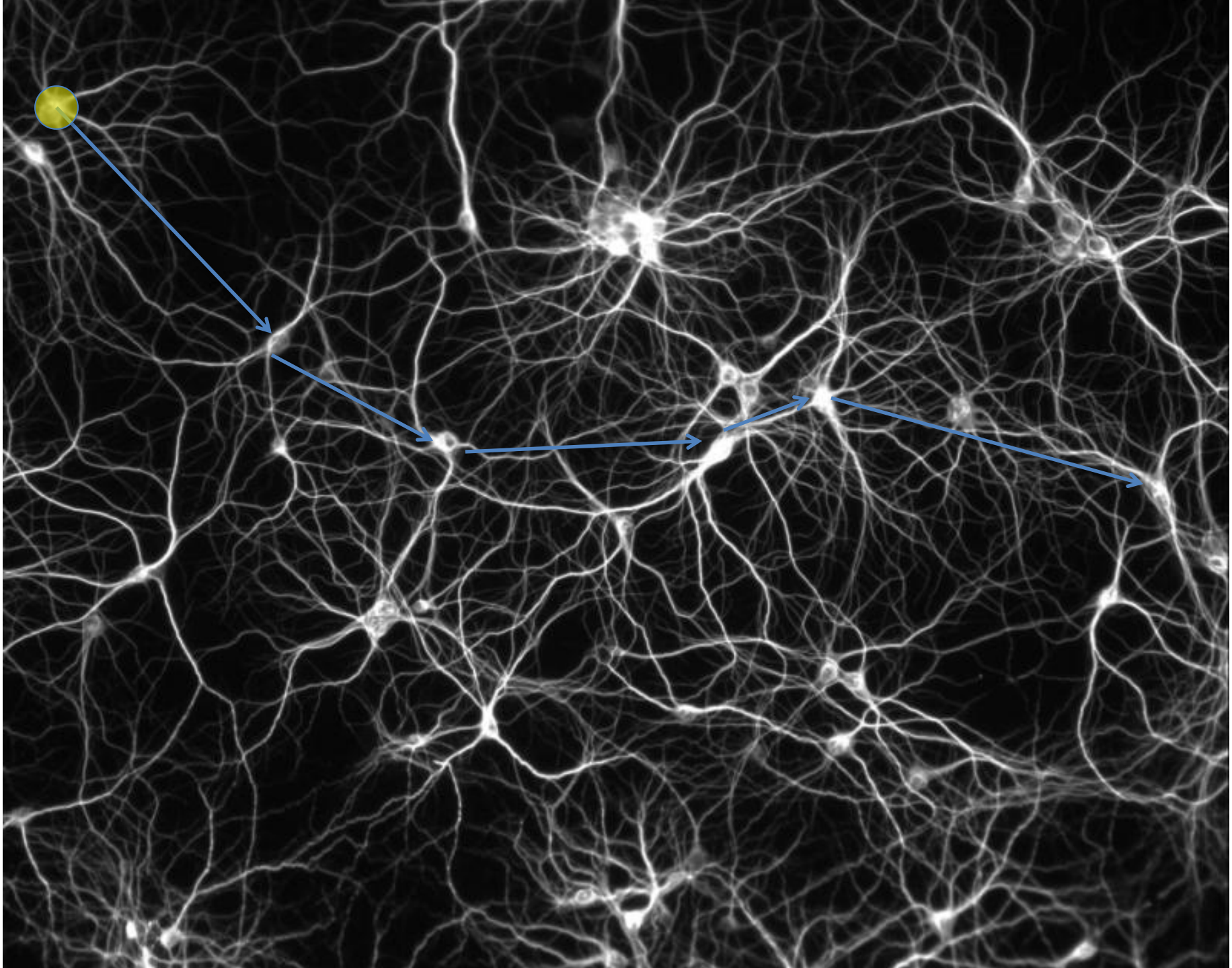
- Search is biased towards previous local optima

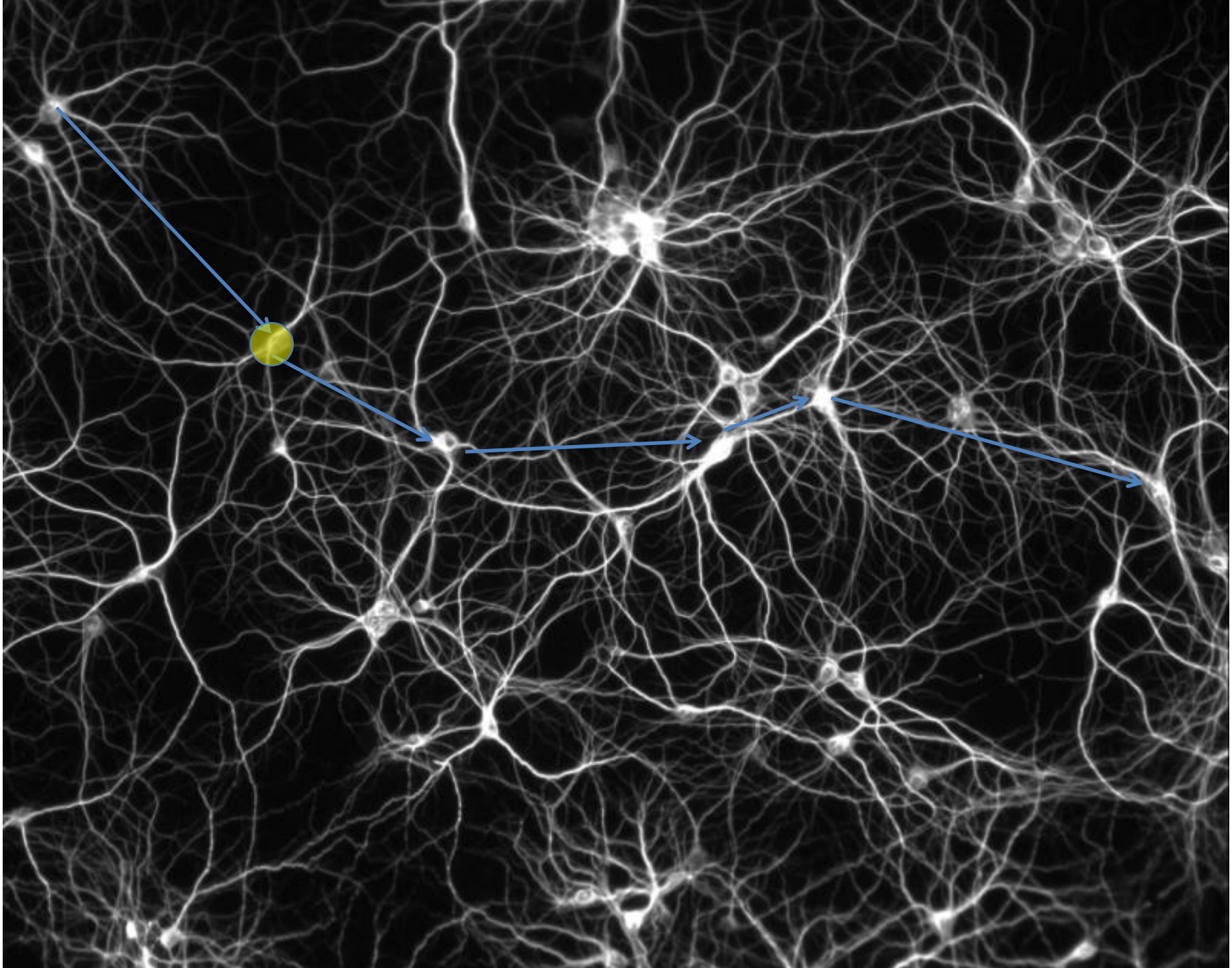

Hebb 64 Bit
fit
400
350
300
250
LR = 0.0002
LR = 0.0008
LR = 0
50
100
150
200
gen/10

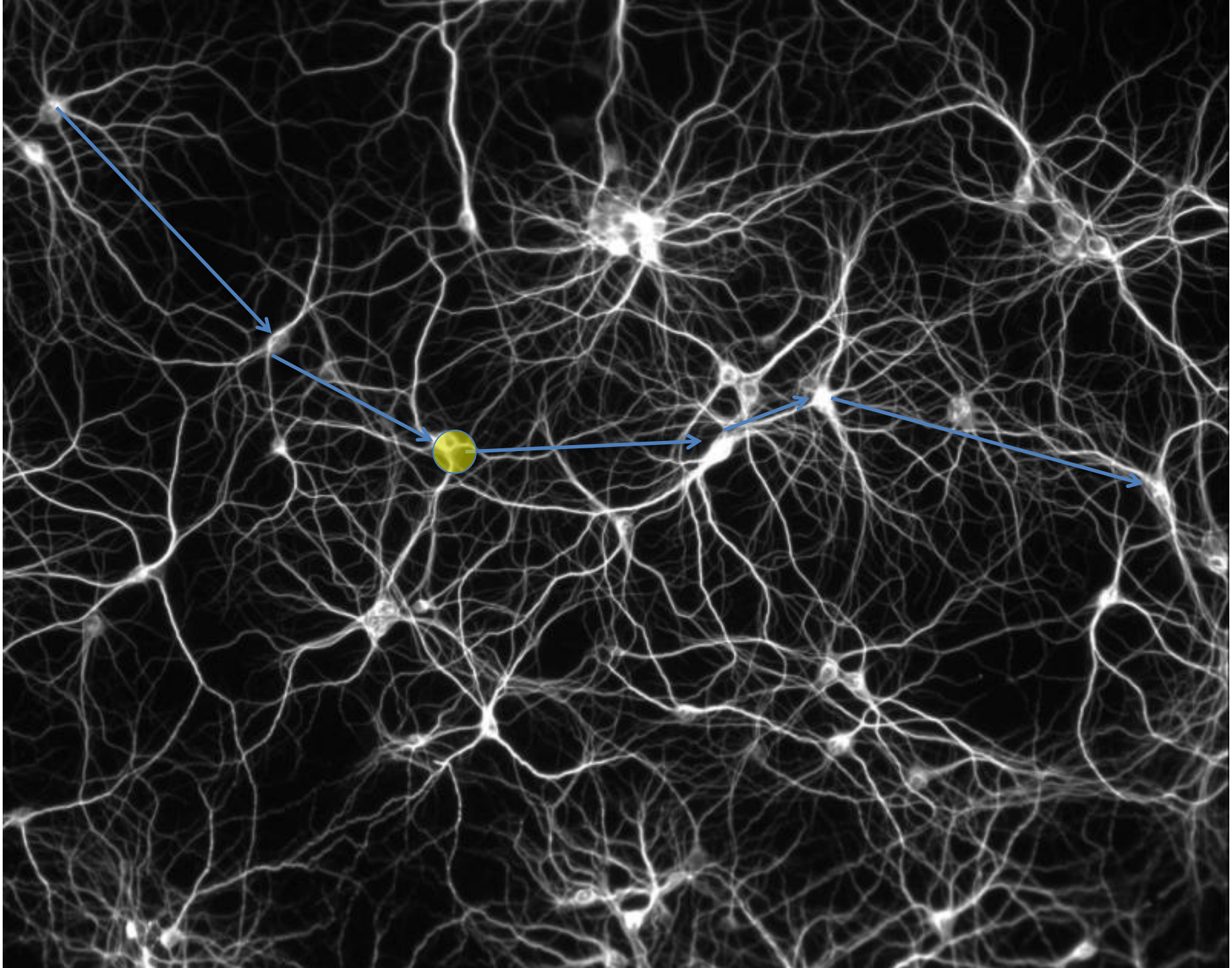
Possible Neuronal Replication Mechanisms

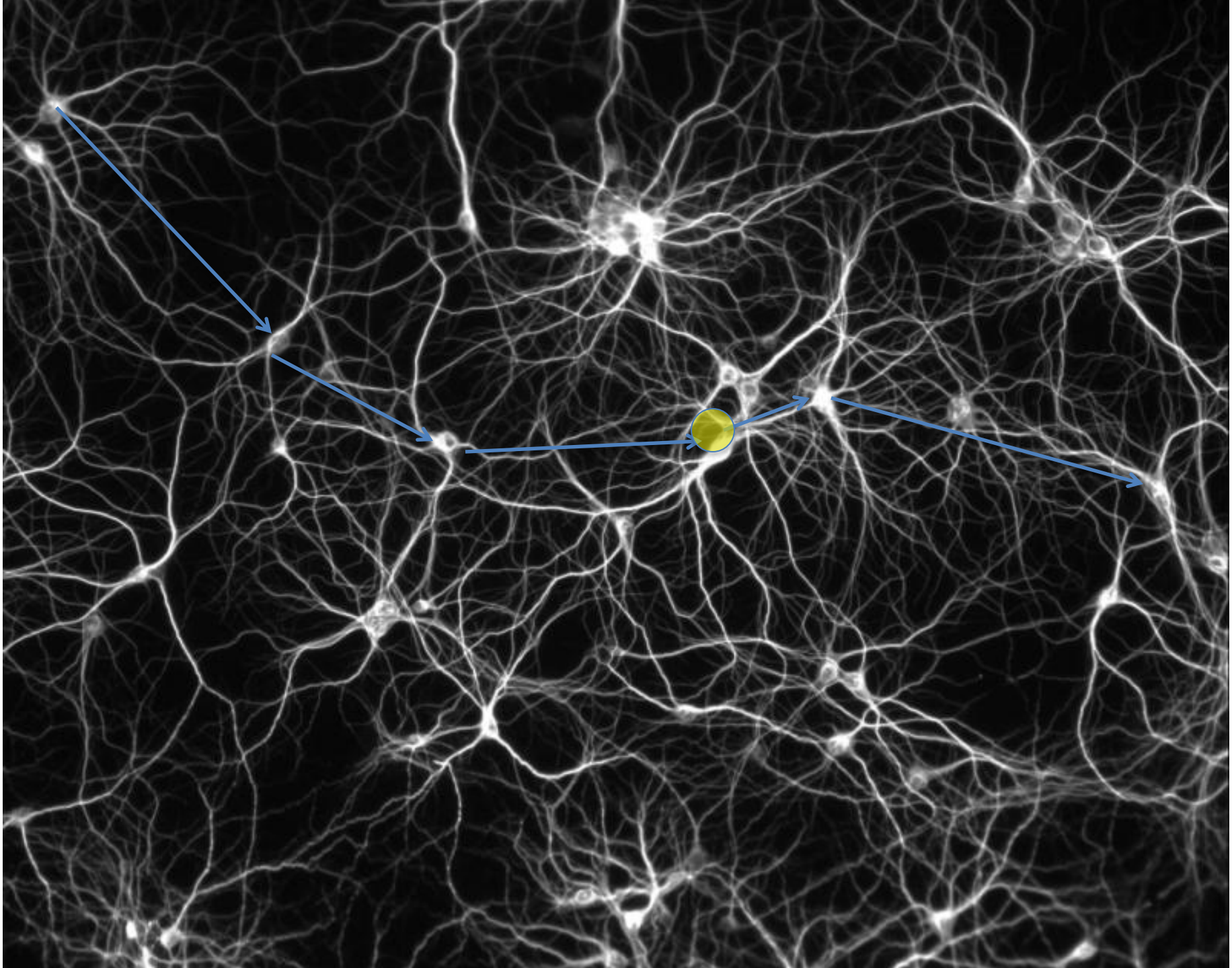
- Replication of Synaptic Connectivity Patterns
- Replication of Activity Patterns
- Evolvable Paths of Activity: Overlapping Units of Evolution

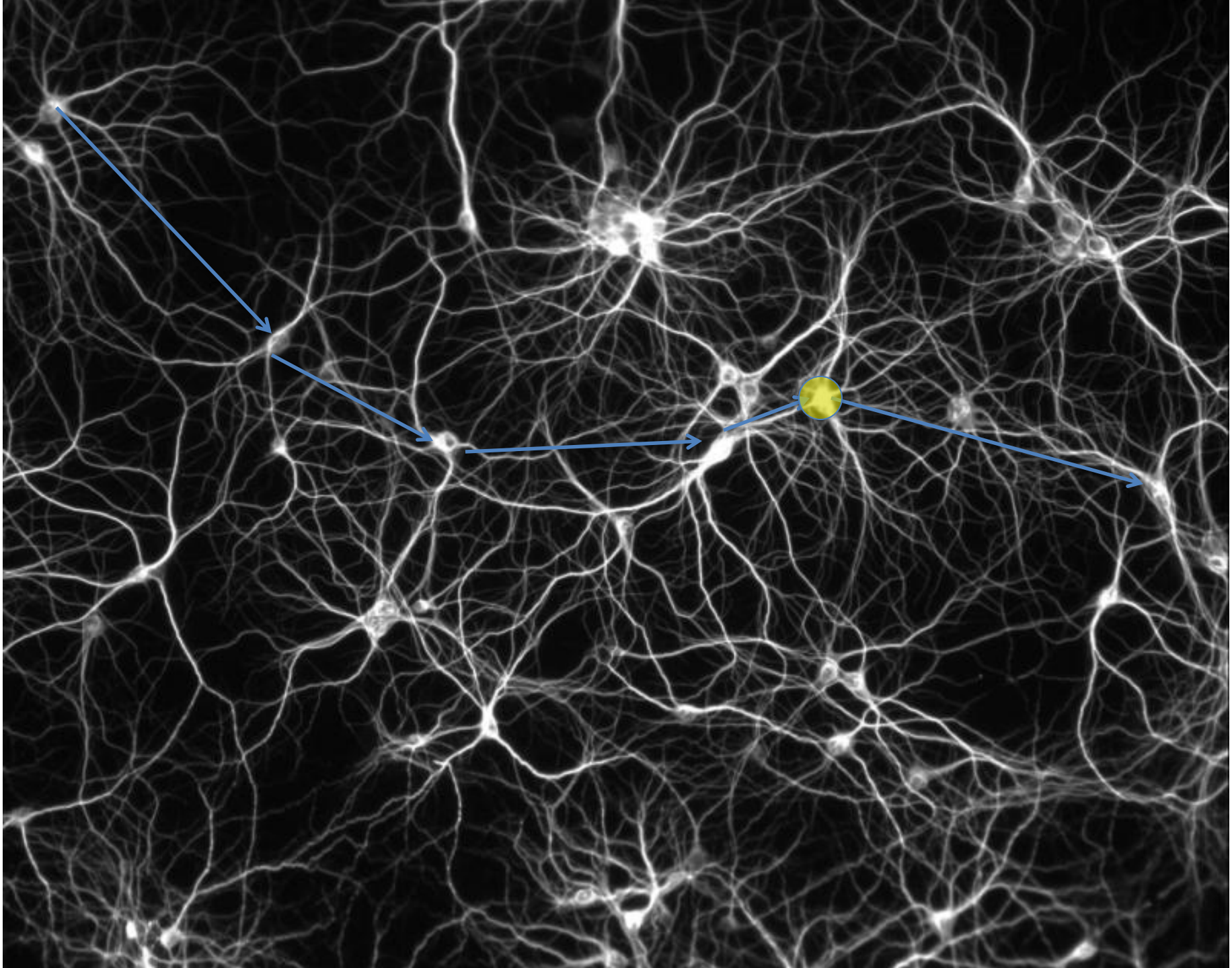


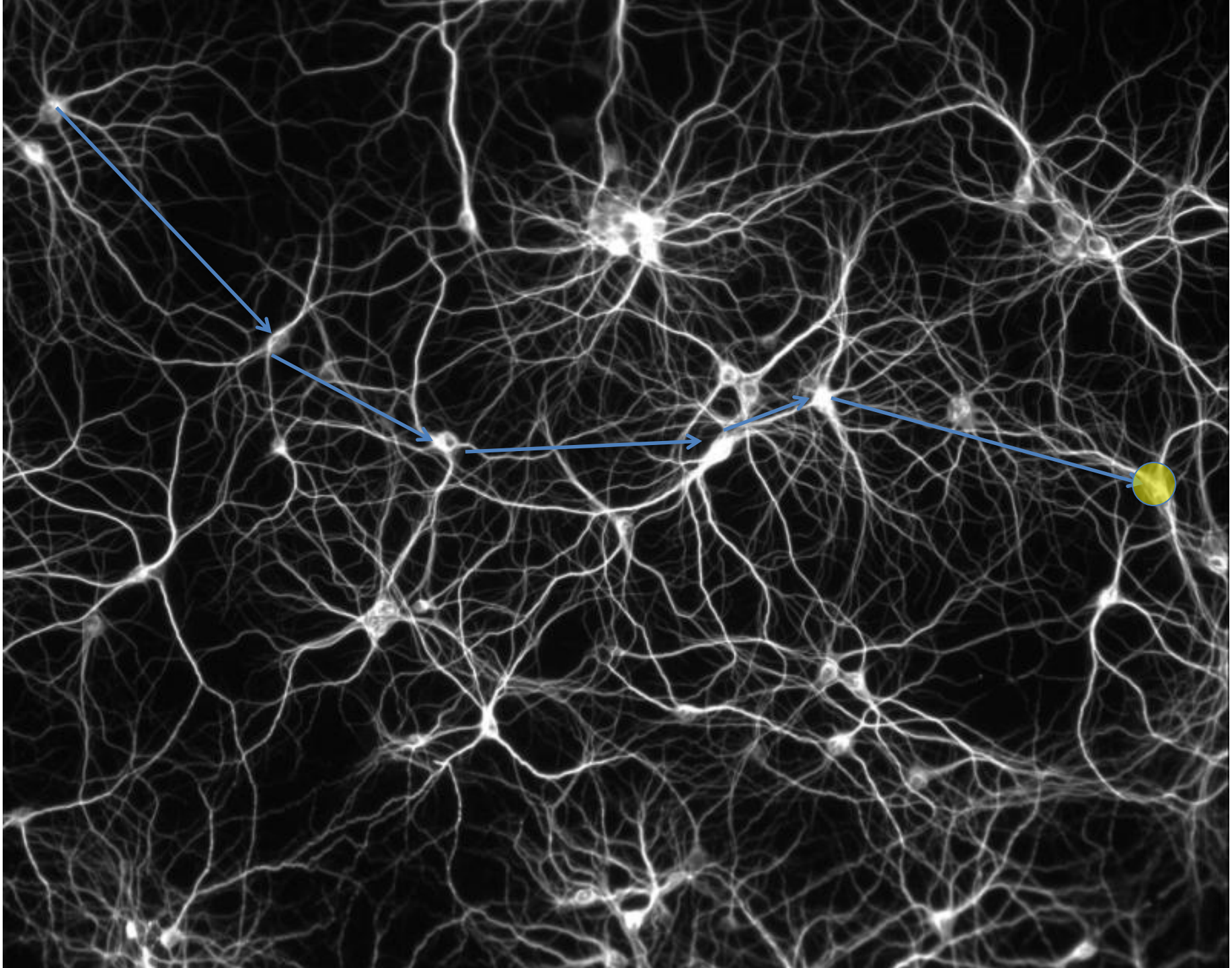


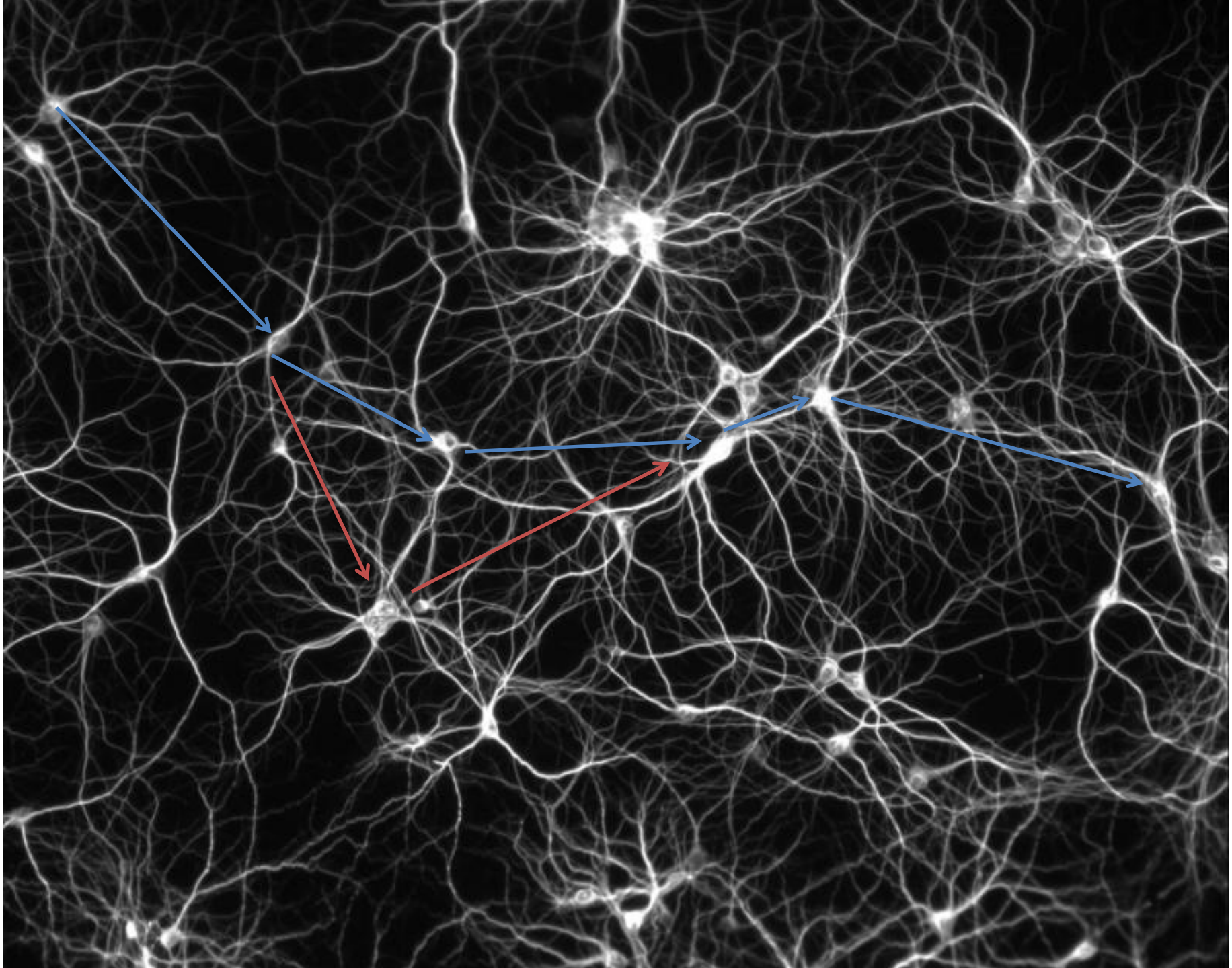


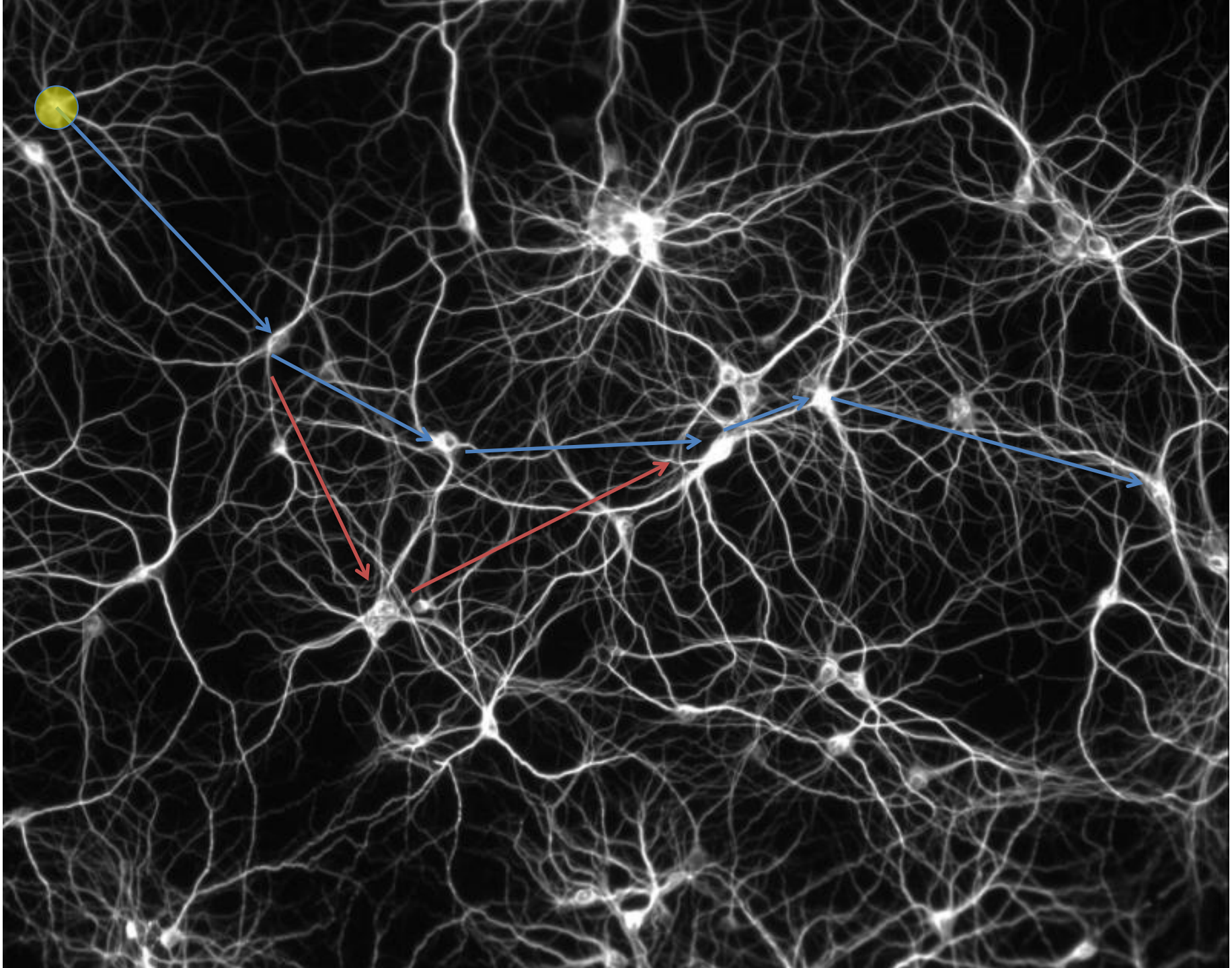


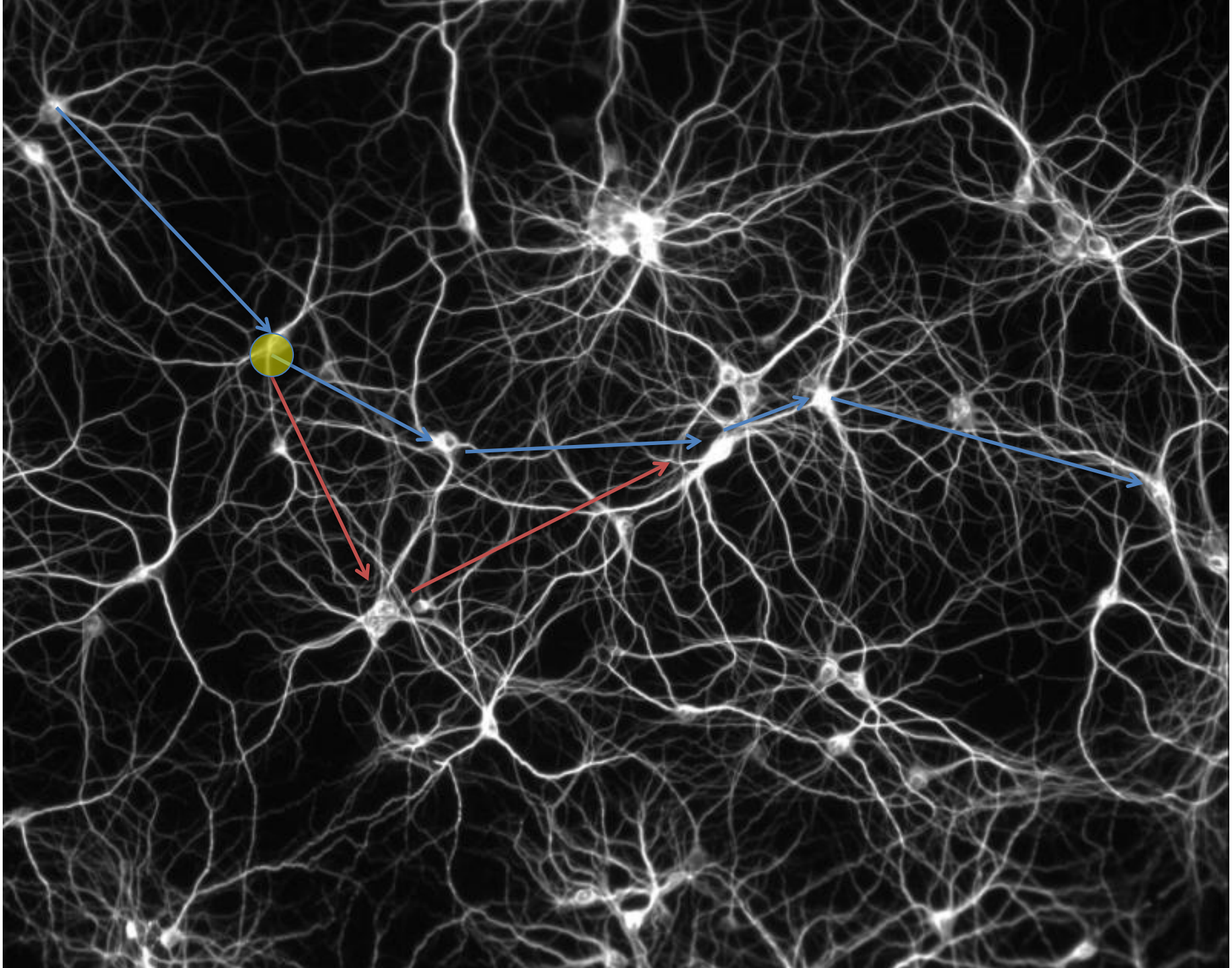


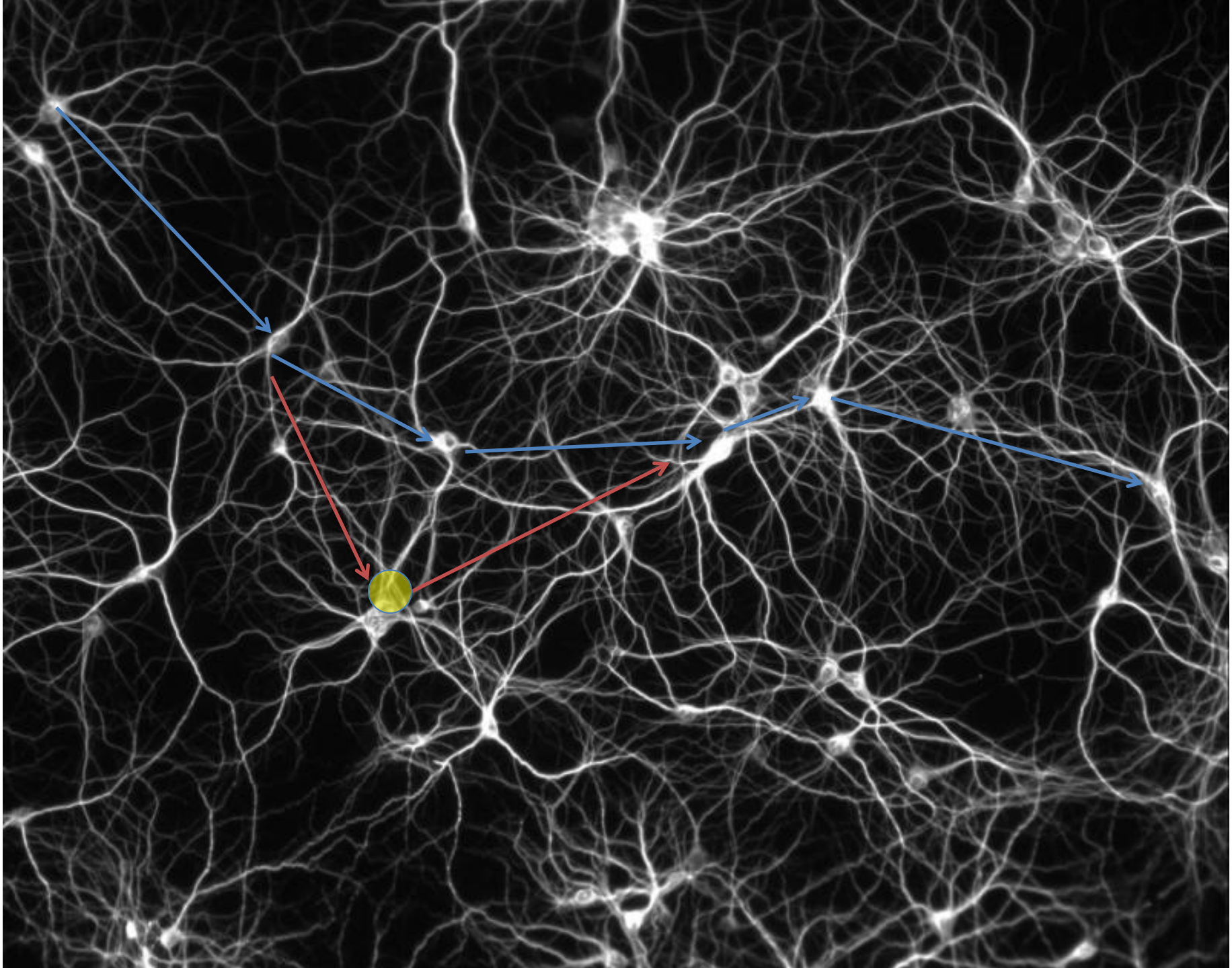


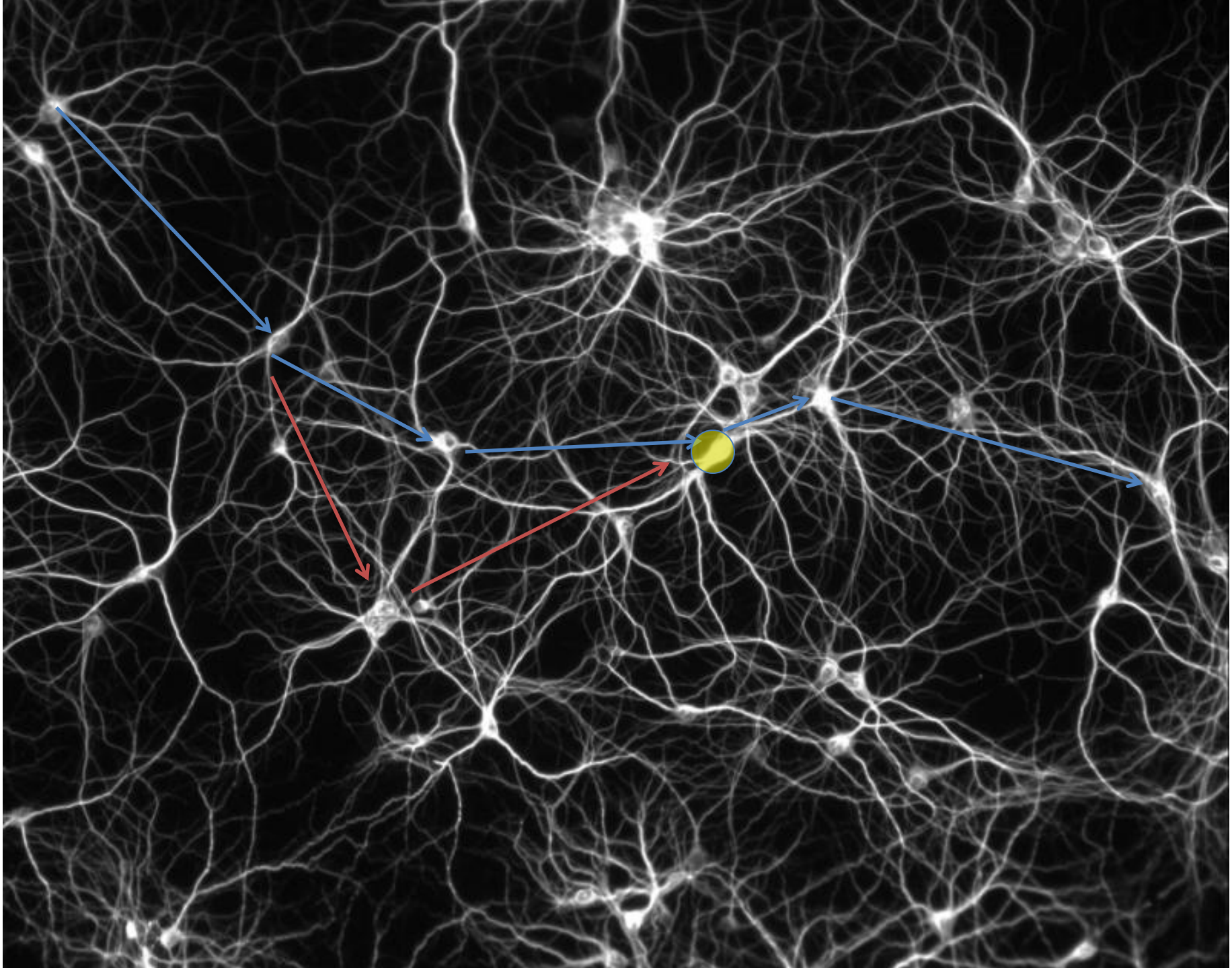


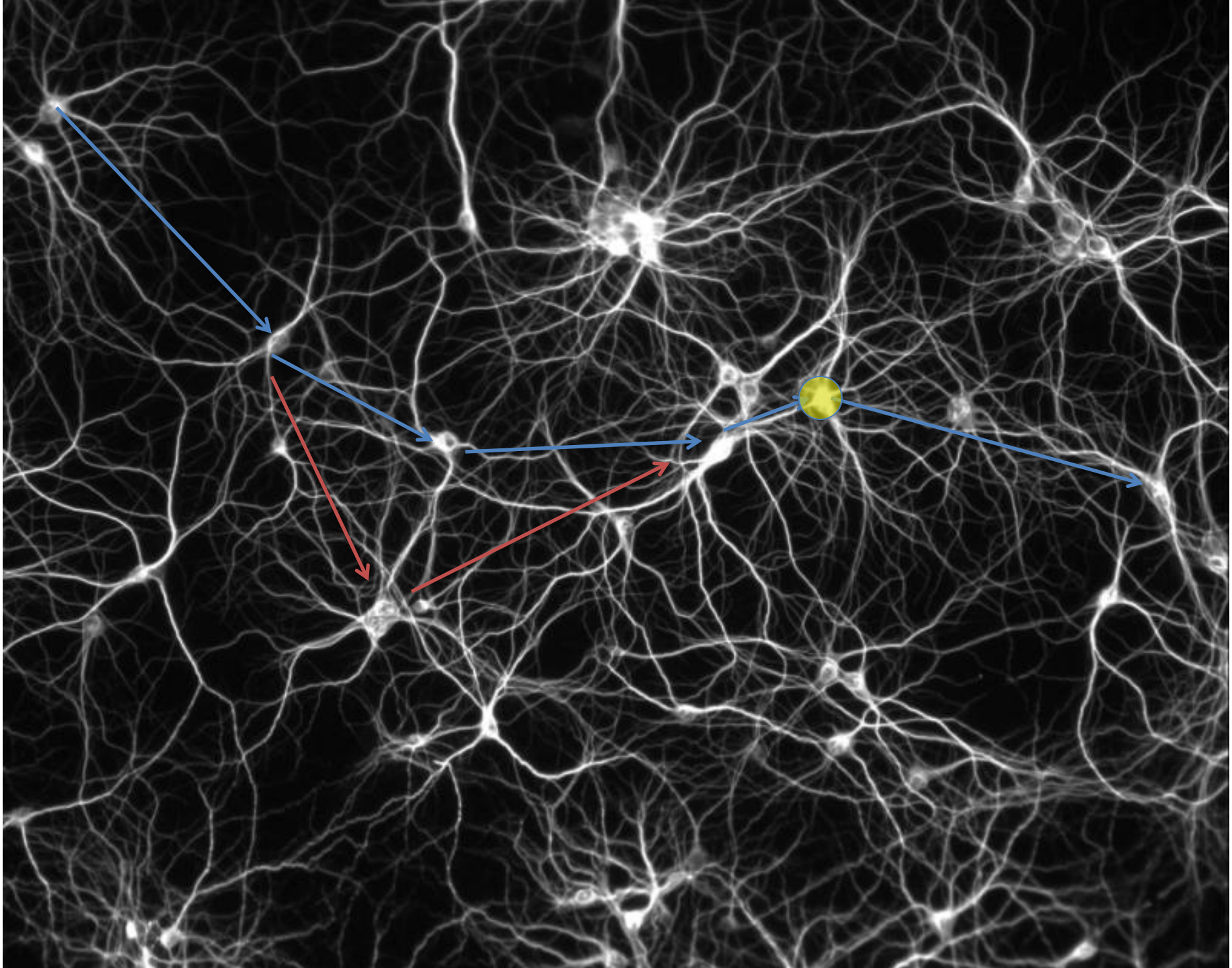


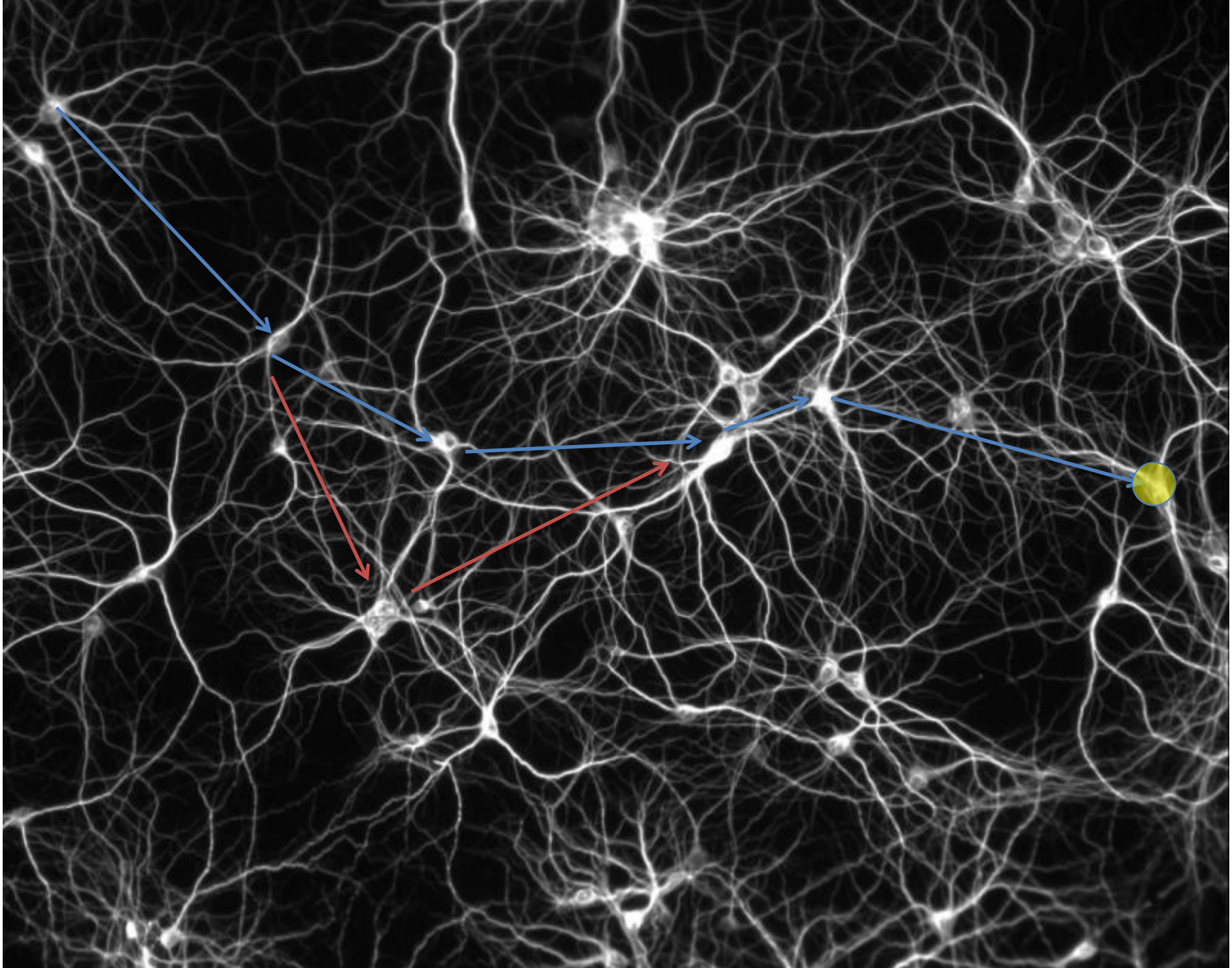


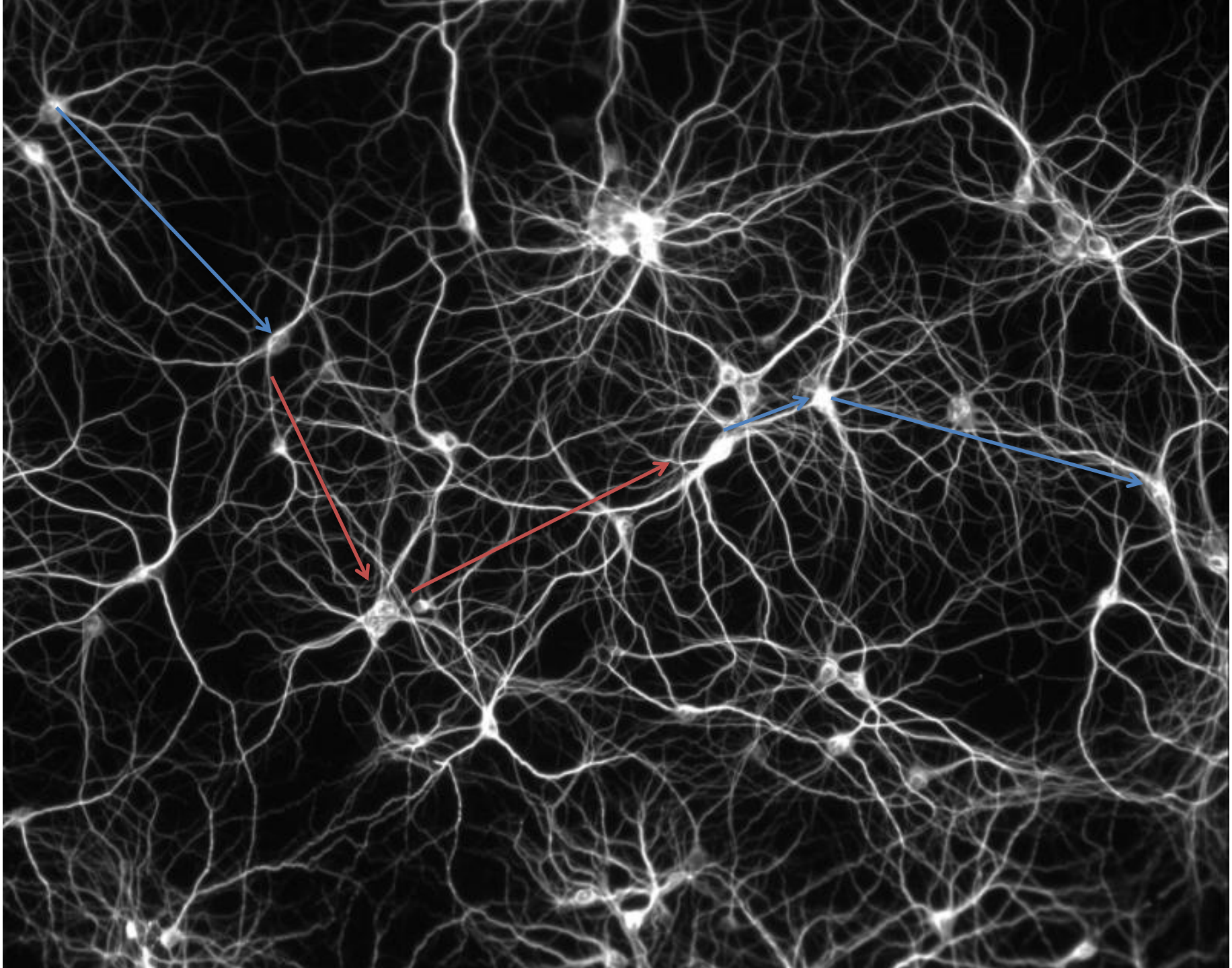


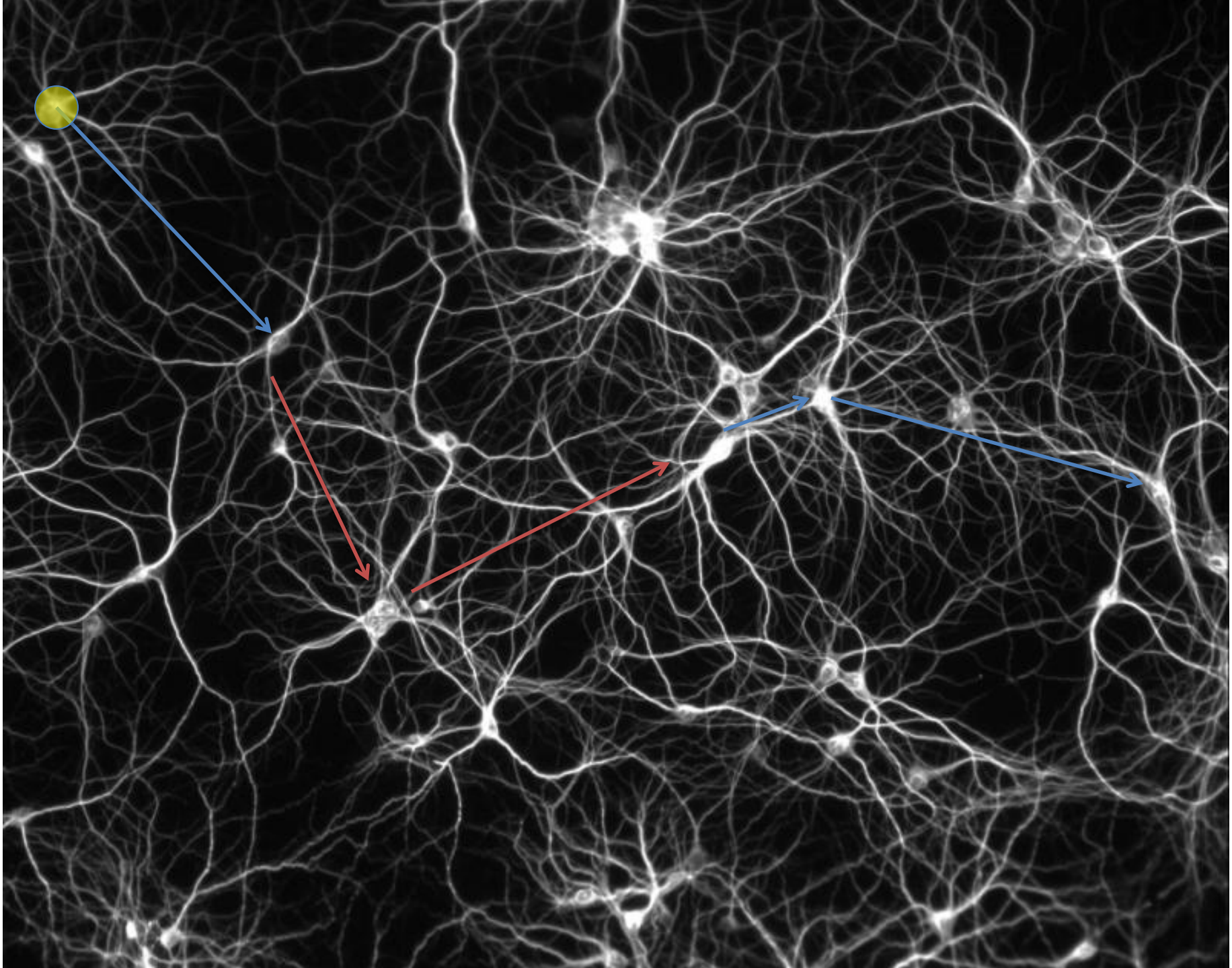


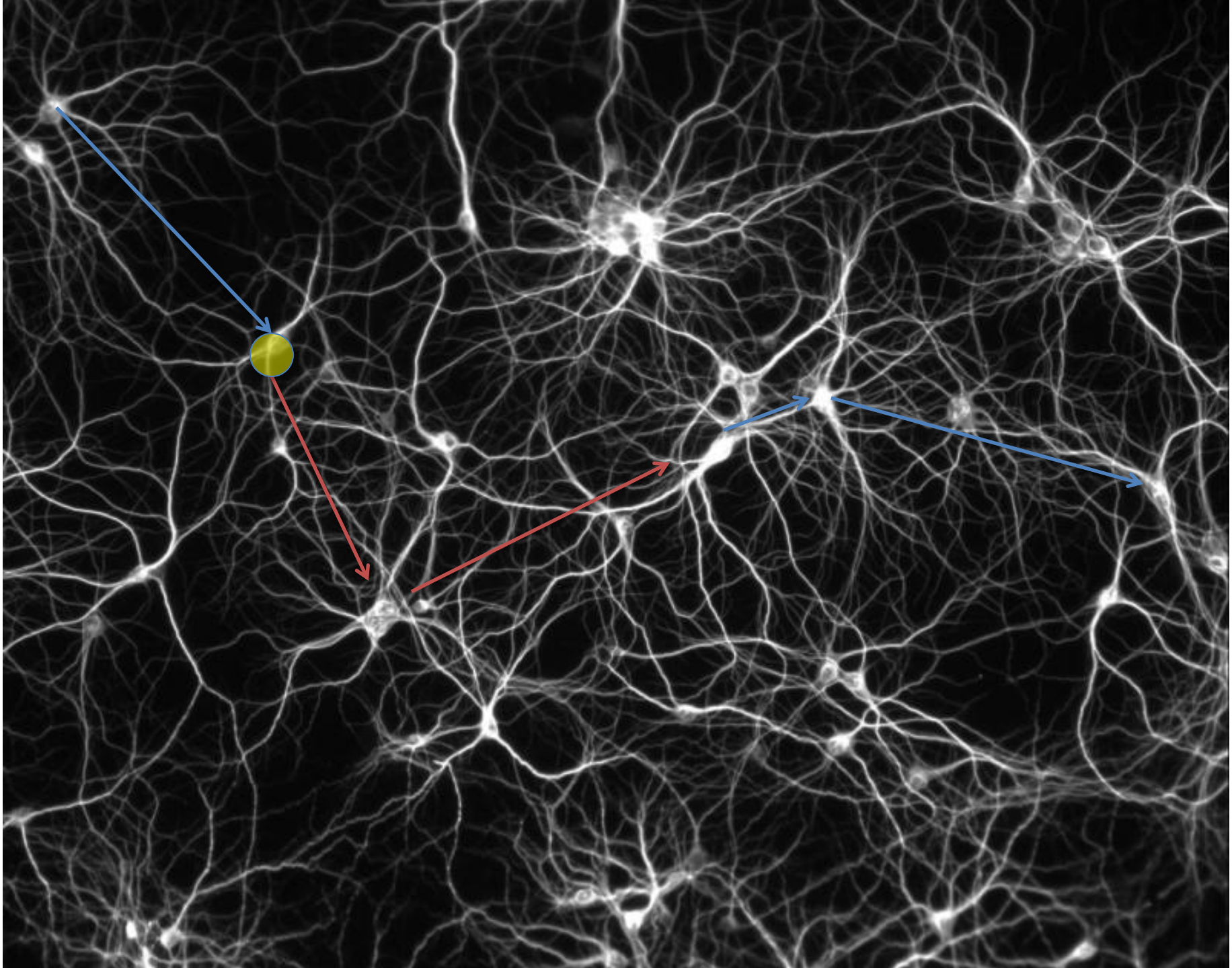


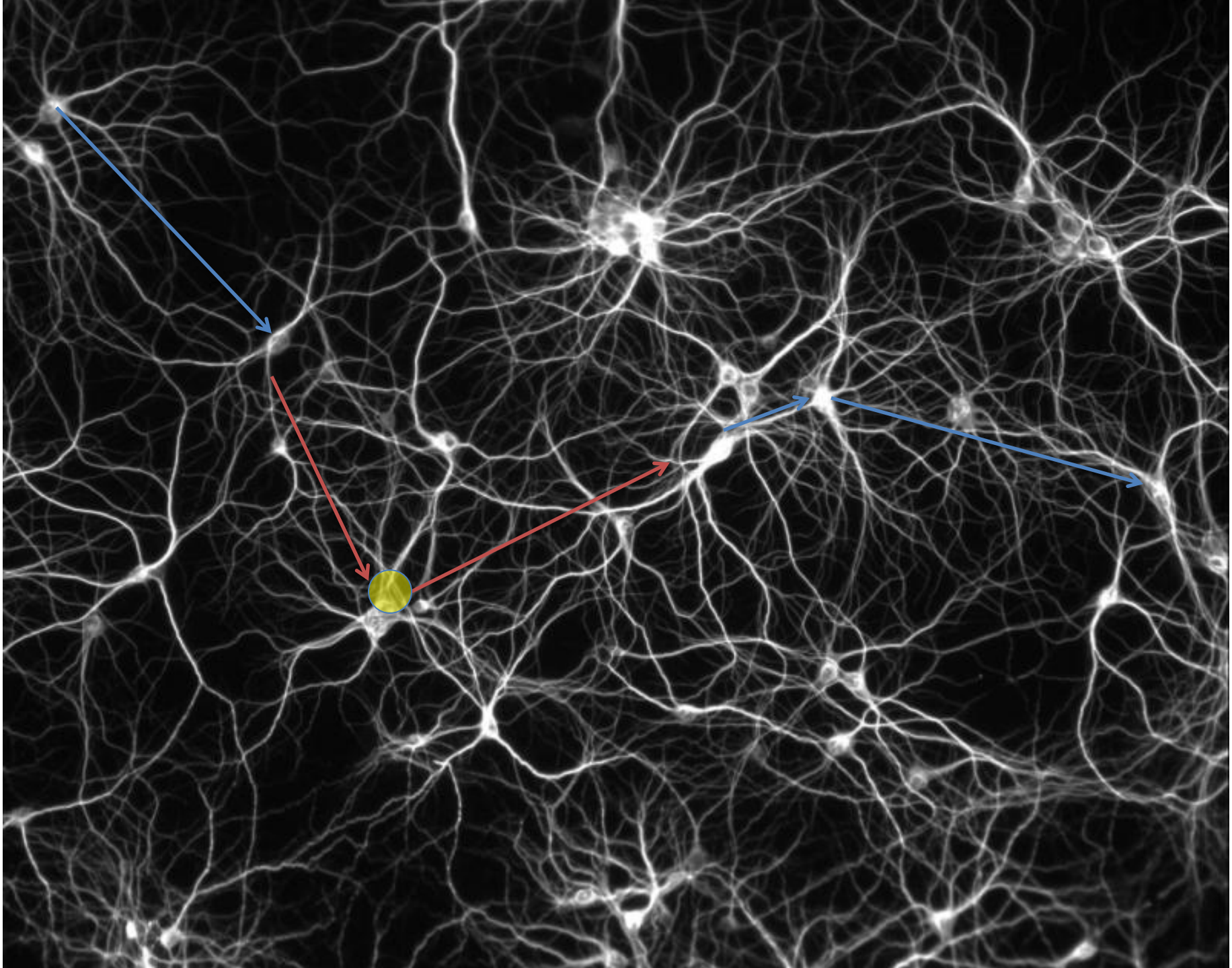


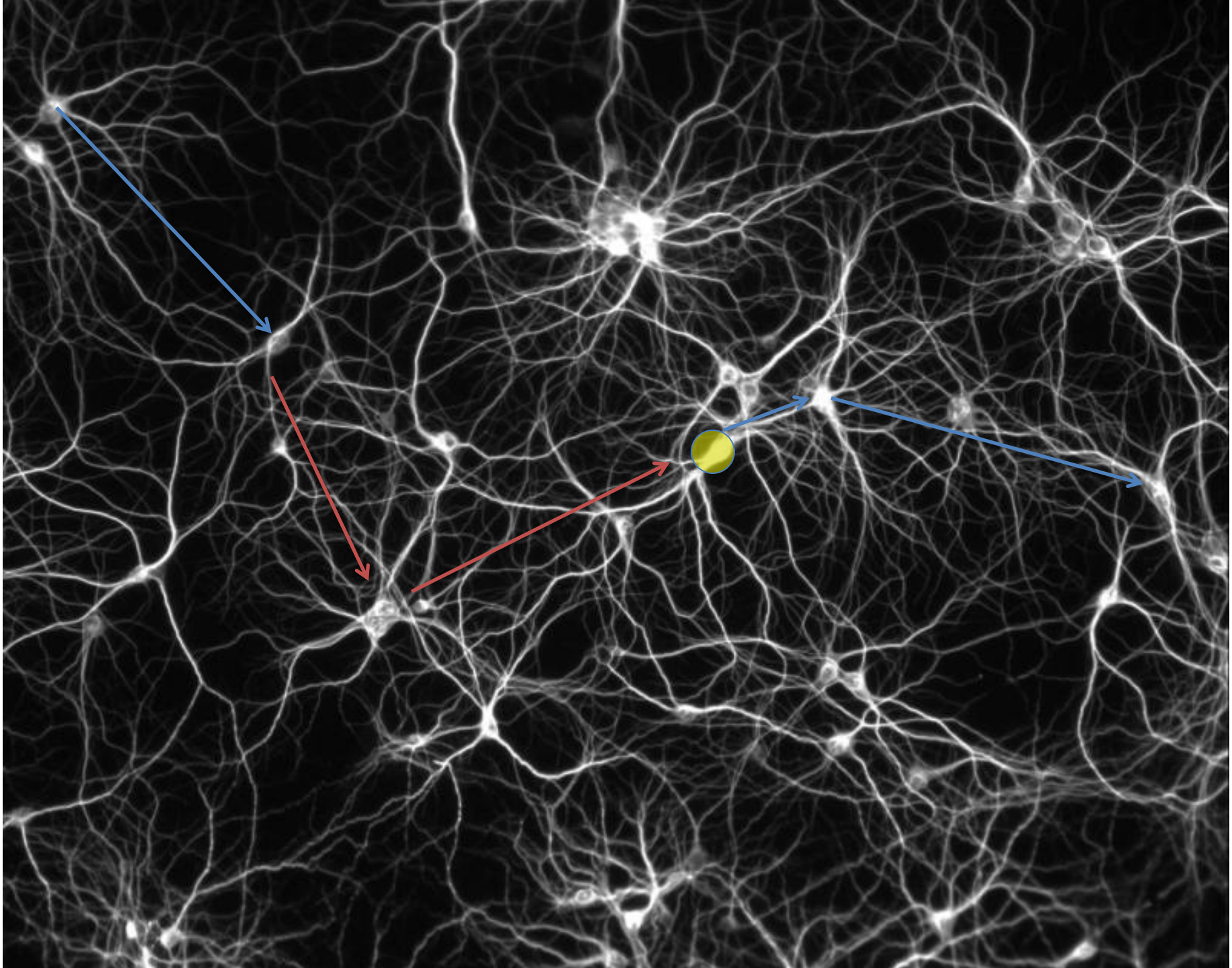


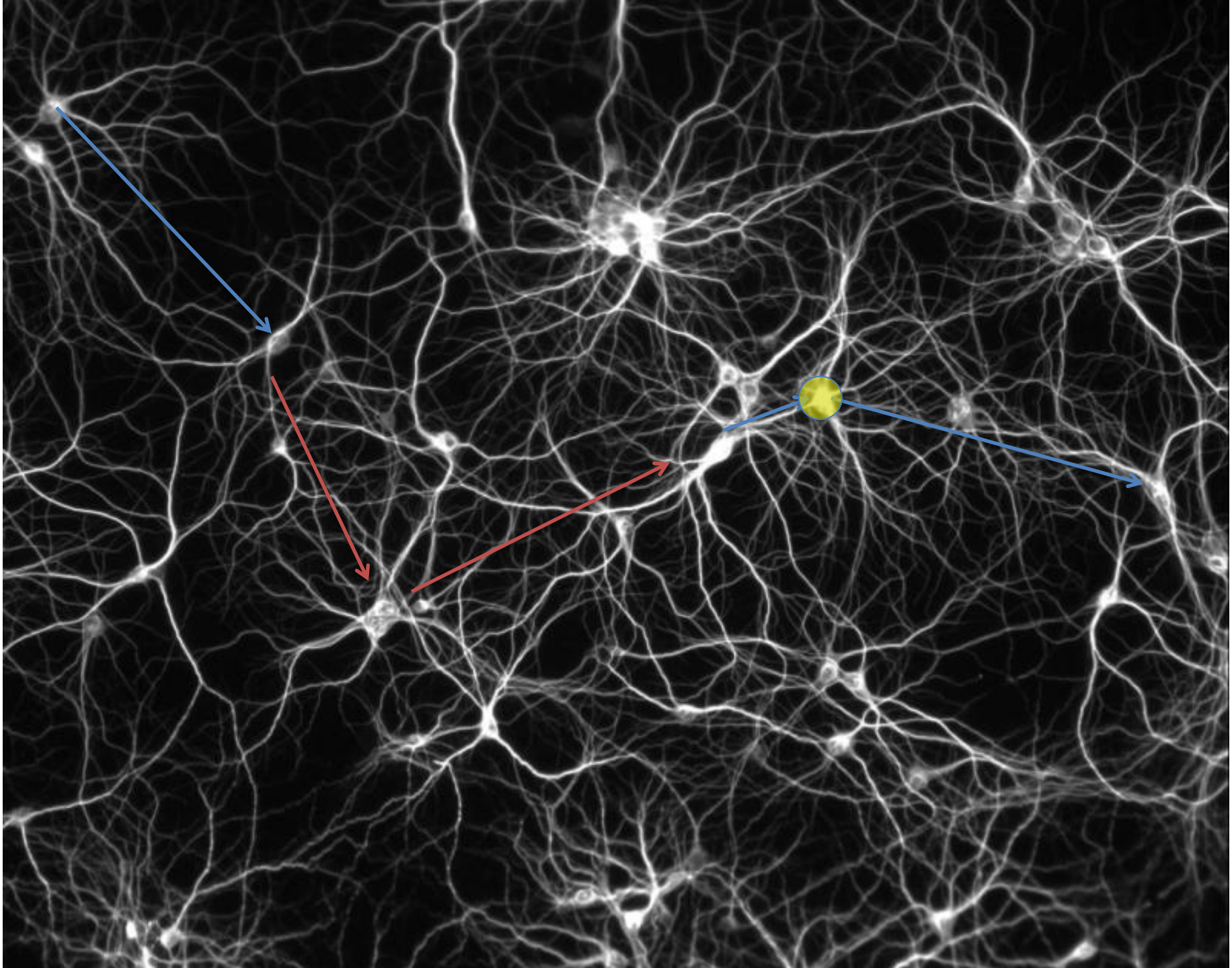


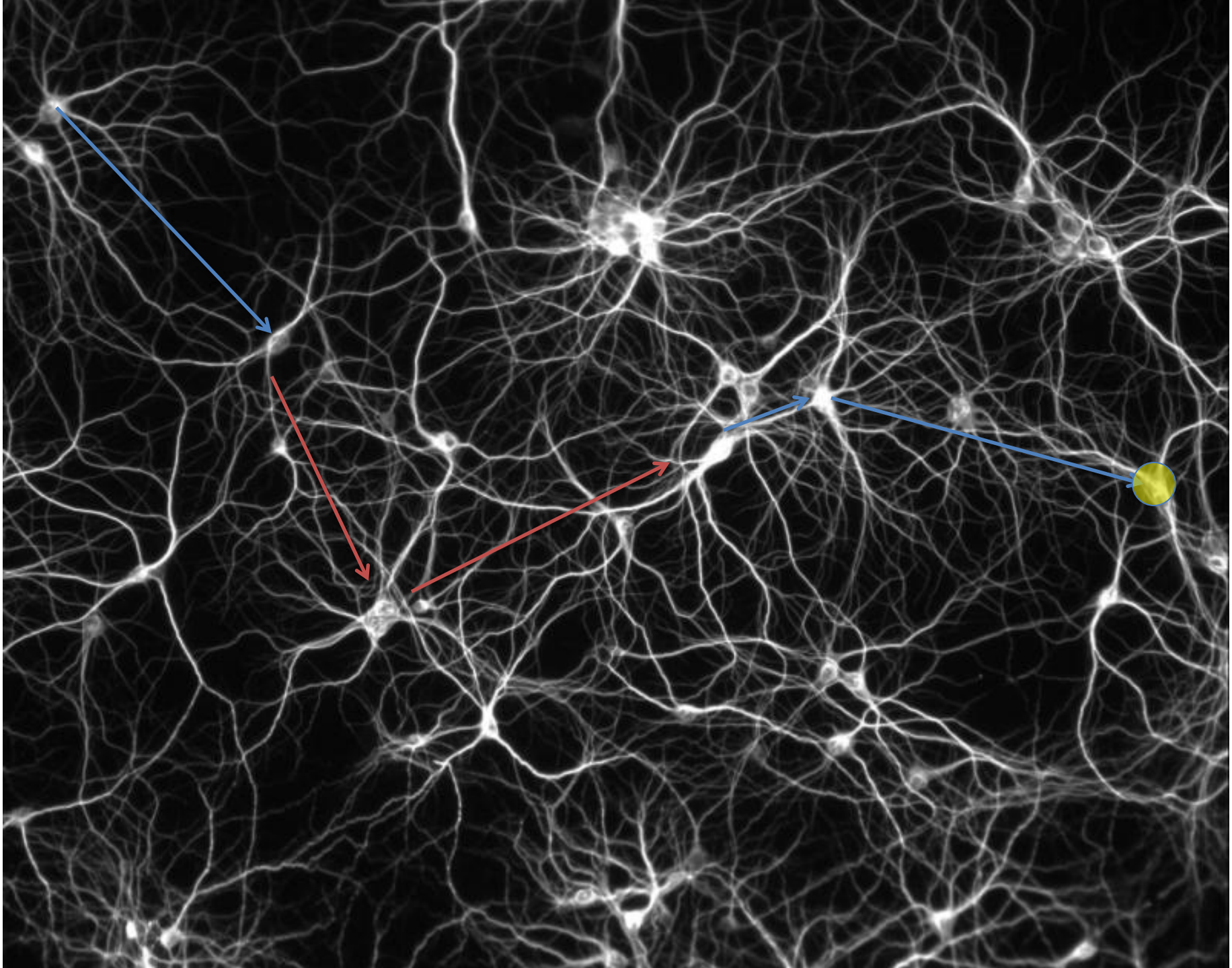




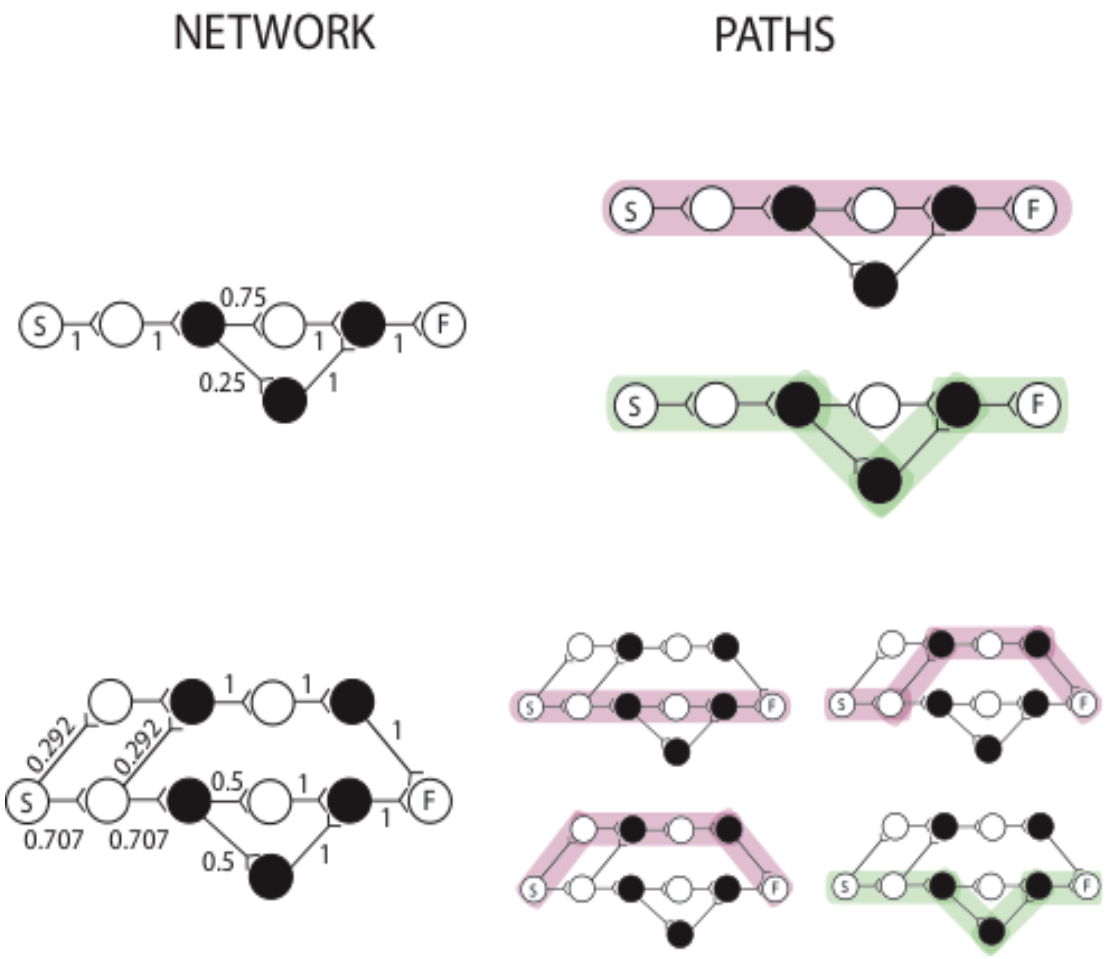




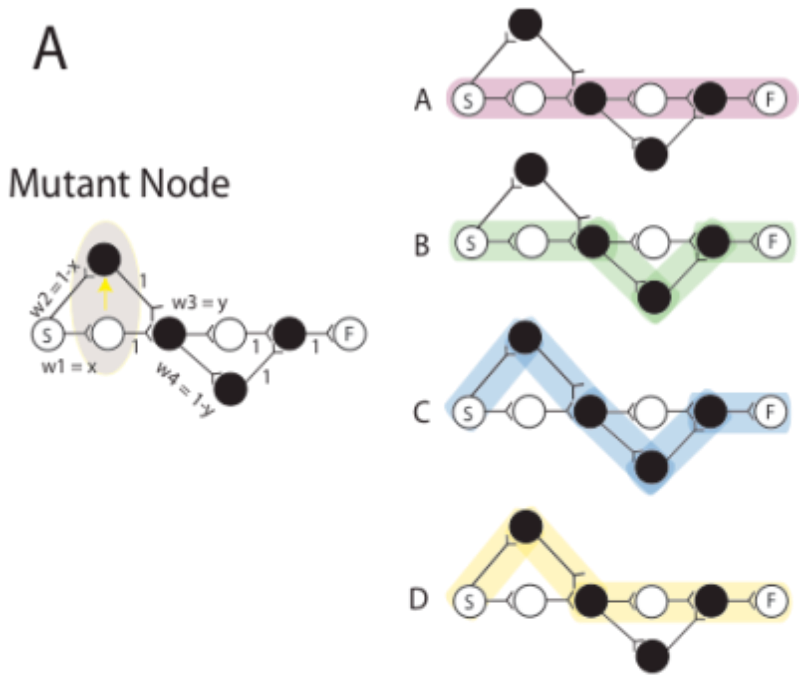


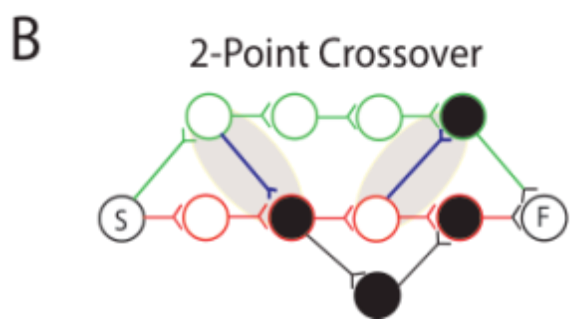


Paths as Units of Evolution

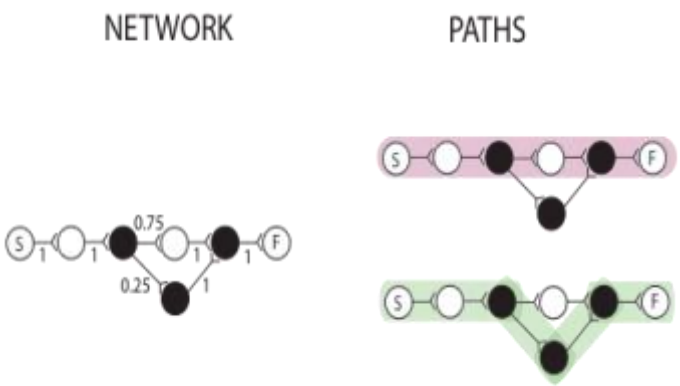


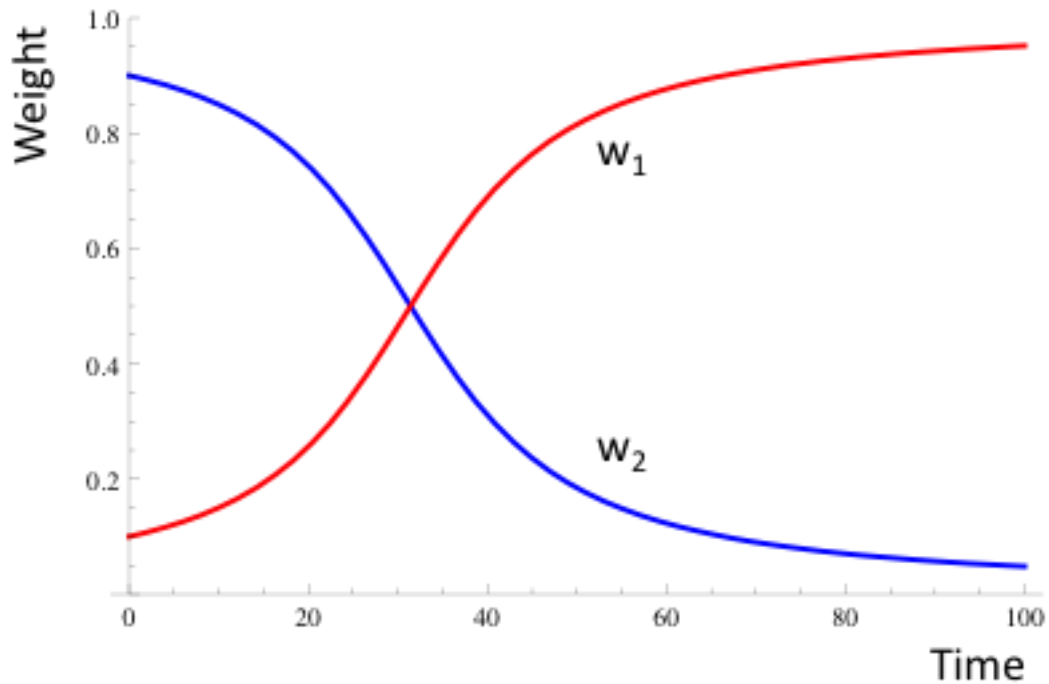
Mutation and Crossover of Paths



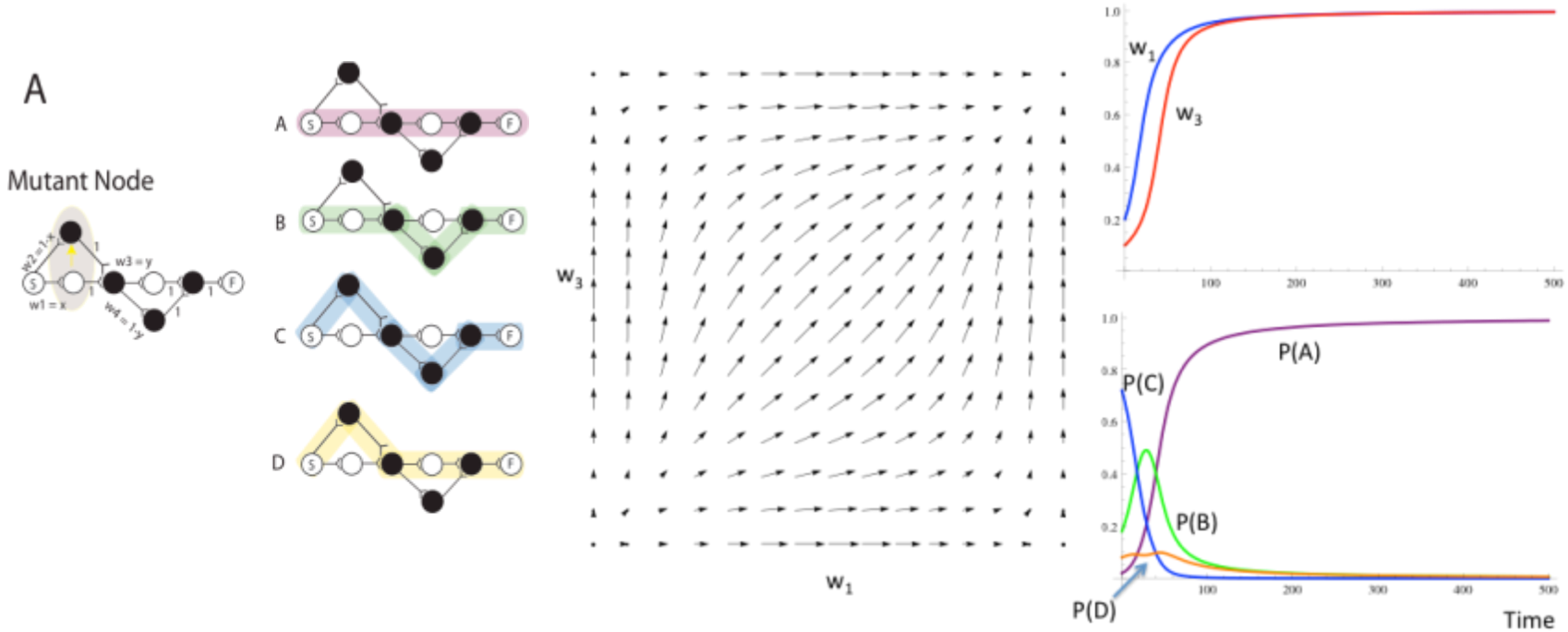


One Locus Two Allele Model of Path Evolution

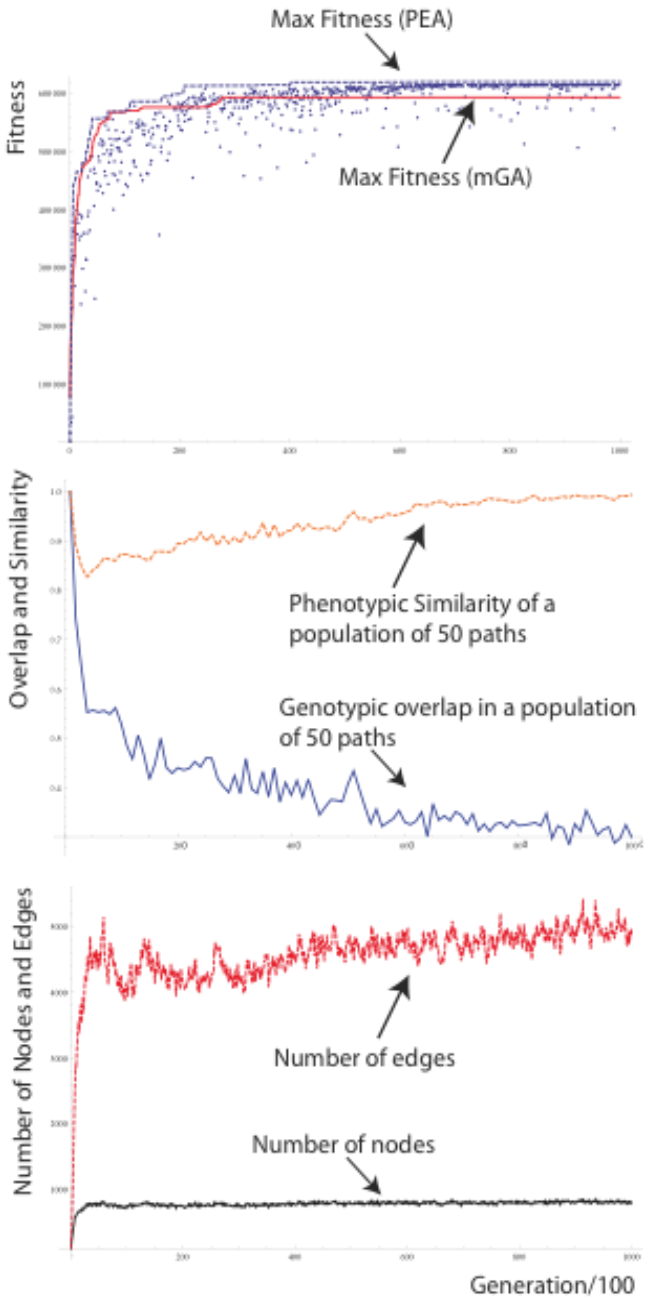


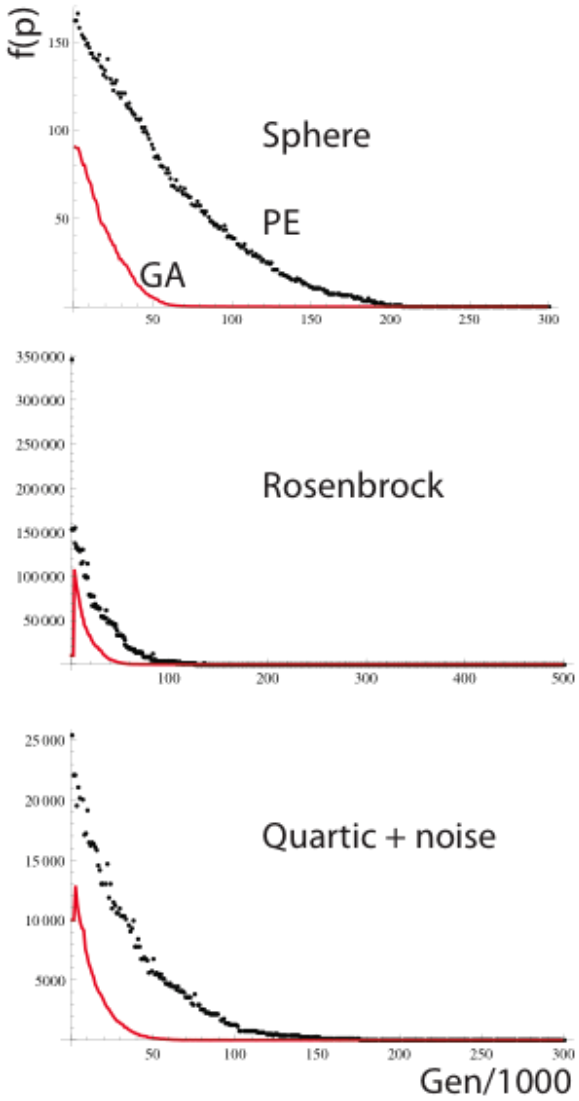


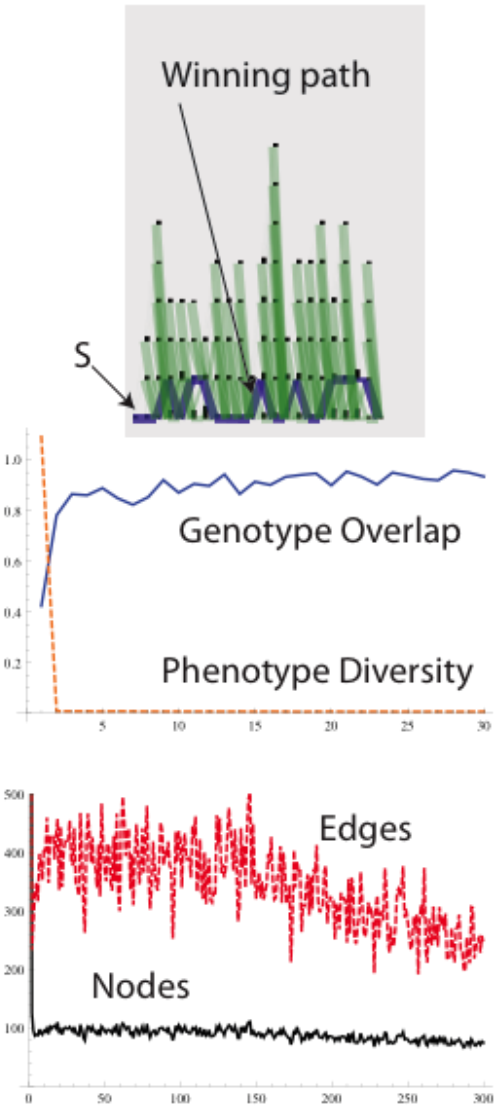
Two Locus Two Allele Model of Path Evolution

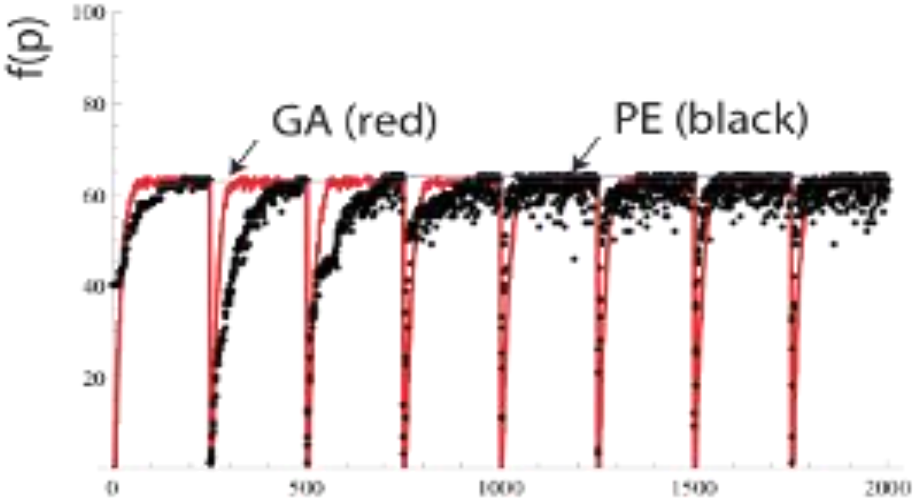


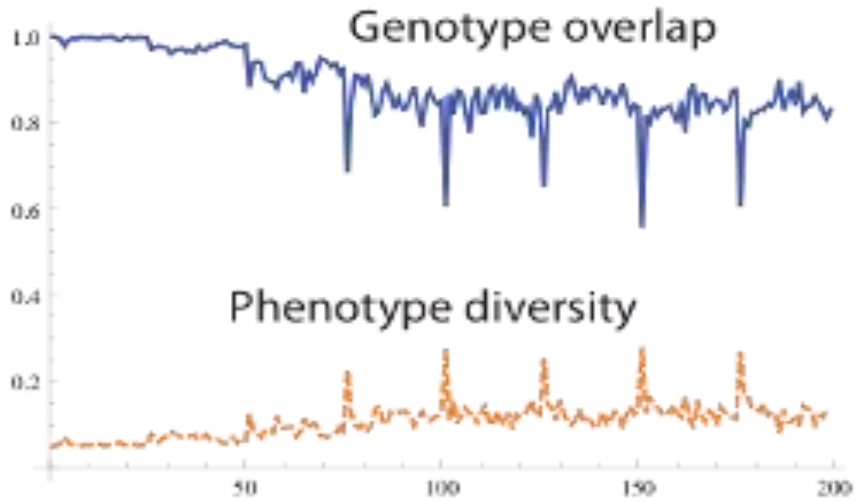
Weing8-105 Knapsack Problem

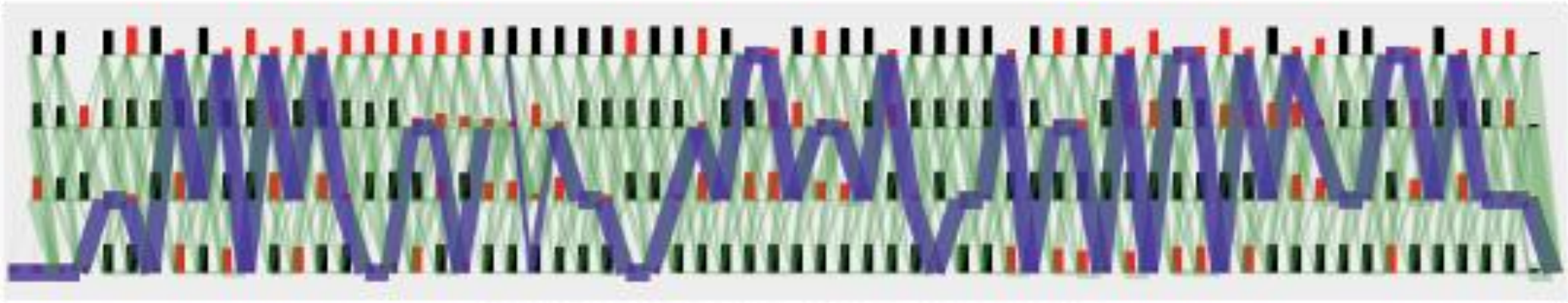




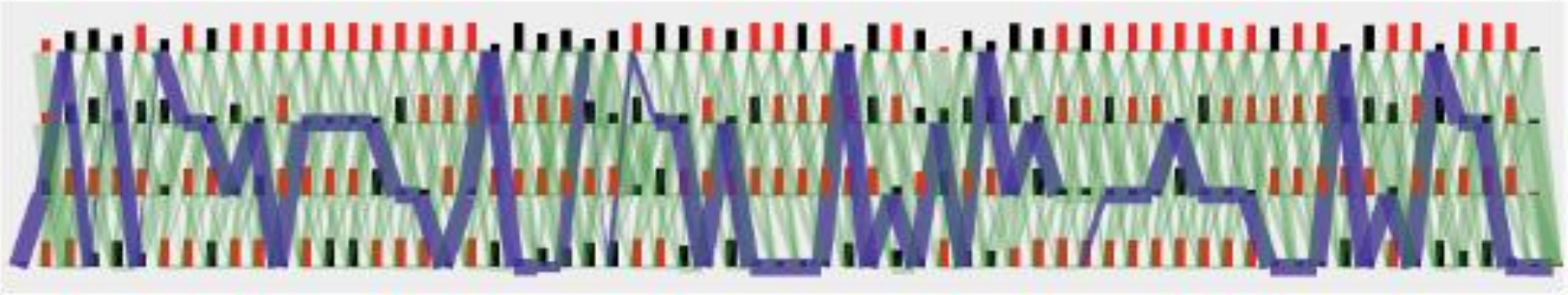








A winning pathway for counting 0's



A winning pathway for counting 1's

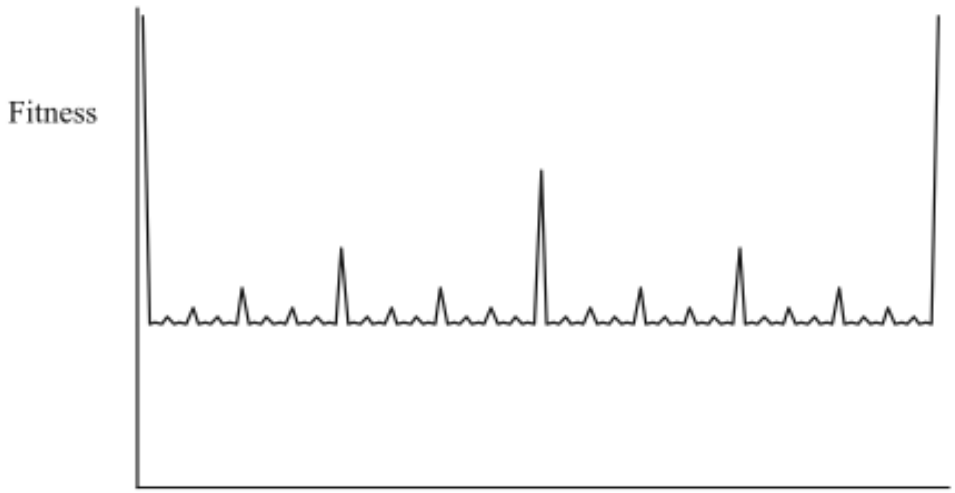
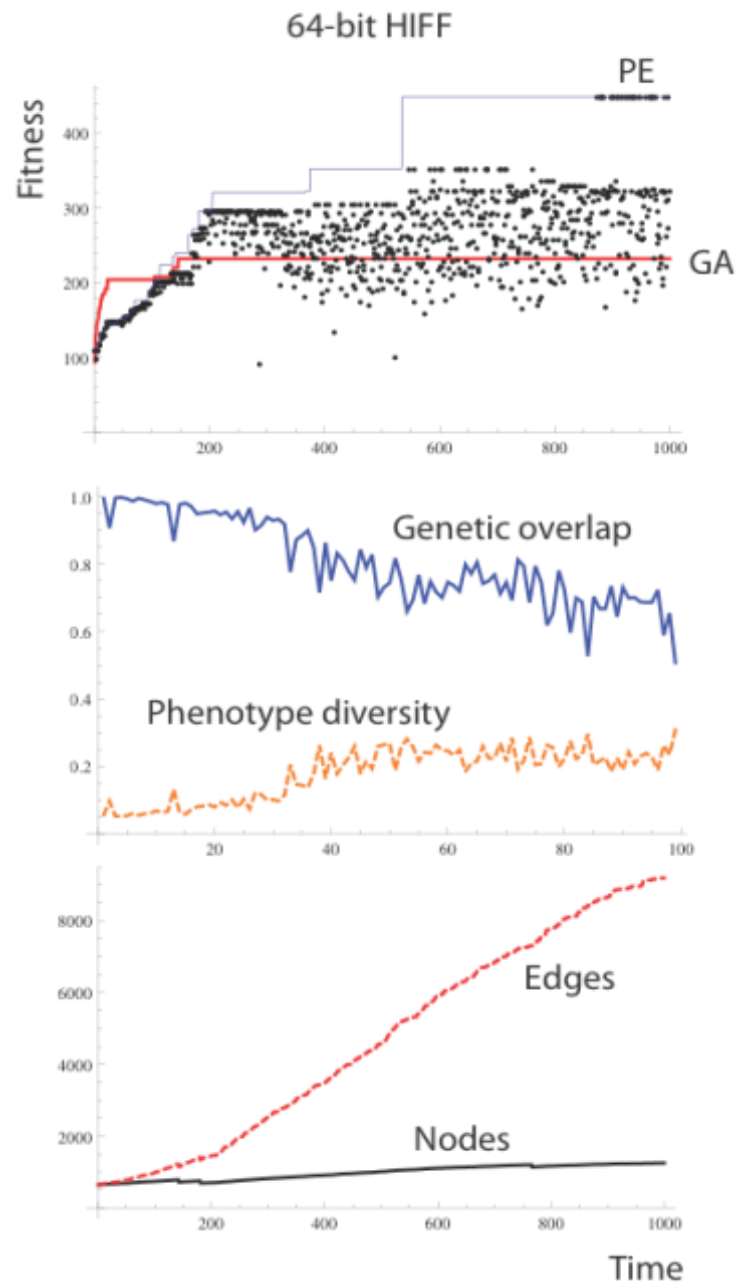


Figure 4.8
A particular cross section through the HIFF fitness landscape.



Are paths units of evolution?

- Parent and offspring are not independent units, they overlap.
- Path phenotypes can be ‘expressed’ only in series, not in parallel.
- But there is information transfer between solutions so its more than independent SHC, and even more than SHC with competition.
- How do true Darwinian replicators differ from neuronal paths in their evolvability?