Black holes as brains: Neural networks with area law entropy

Gia Dvali

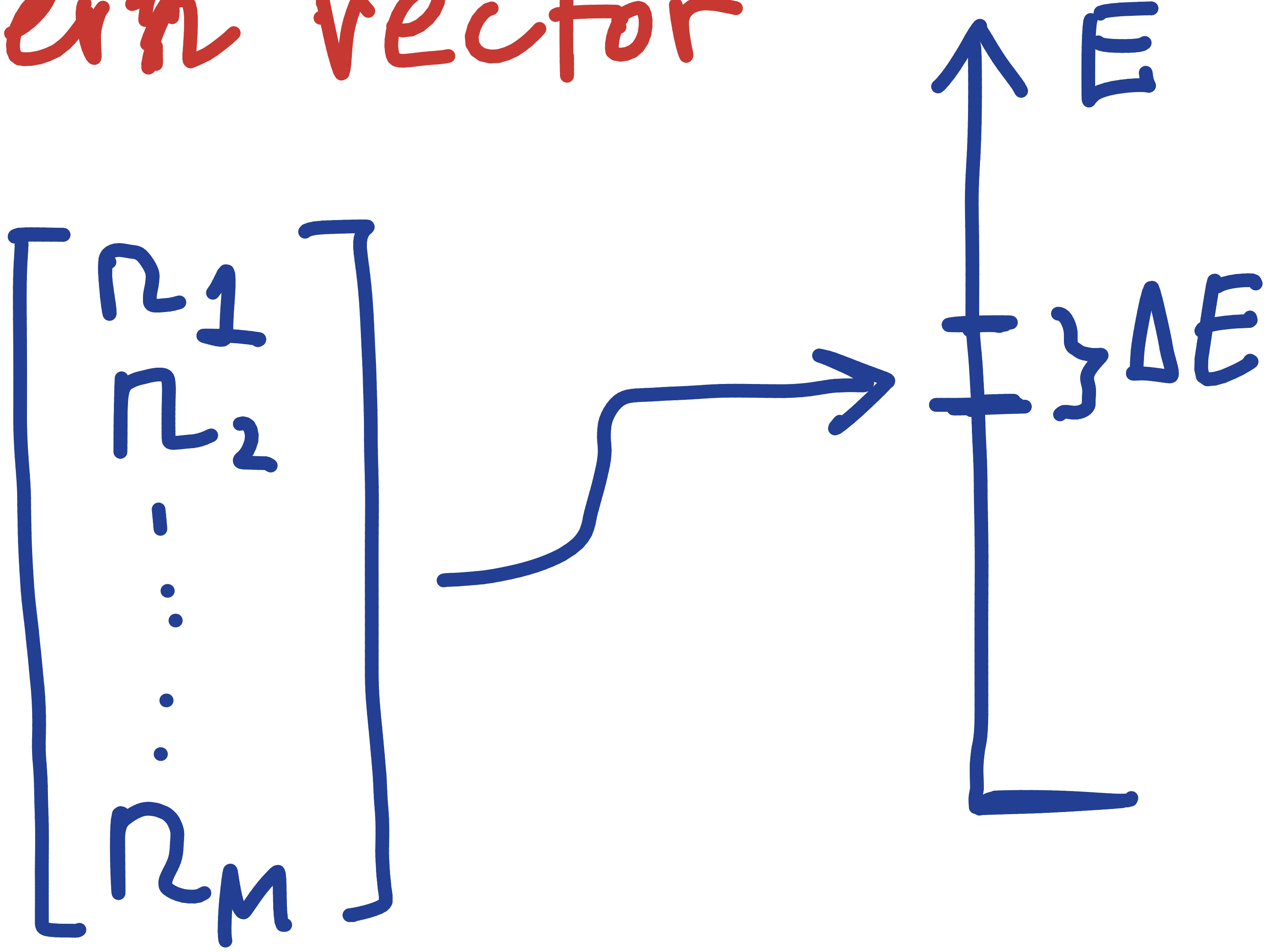
LMU - MPI & NYU

arXiv: 1801.03918
1712.02233
1711.09079

Enhanced capacity of memory storage:

Number of distinct patterns N_E that can be stored within energy gap ΔE.

Pattern vector



Motivation: Black holes, human brains, AI, ...

understanding capacity of memory storage in black holes in language of neural networks

and

understanding enhanced capacity of memory storage in neural networks in the language of Quantum Field.

Describing neural network as quantum field:

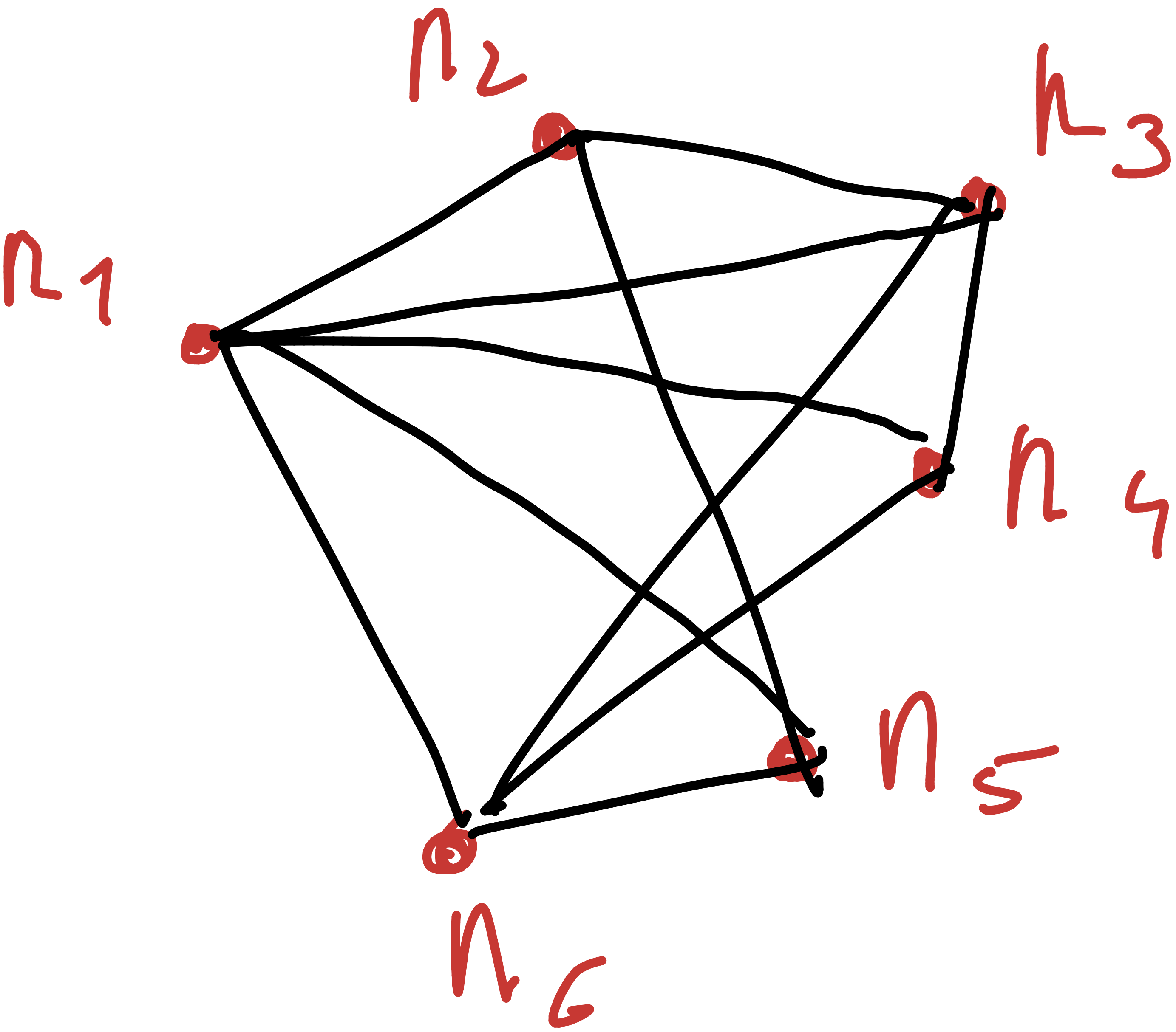
Enhanced memory state of a neural network

$\downarrow$

Critical state of large micro-state entropy of a quantum field.

Framework: Effective description

System is described by effective degrees of freedom, $\hat{n}_K(\hat{a}^+_K, \hat{a}_K)$, and their interactions



Dictionary:

Quantum field $\longleftrightarrow$ Neural network

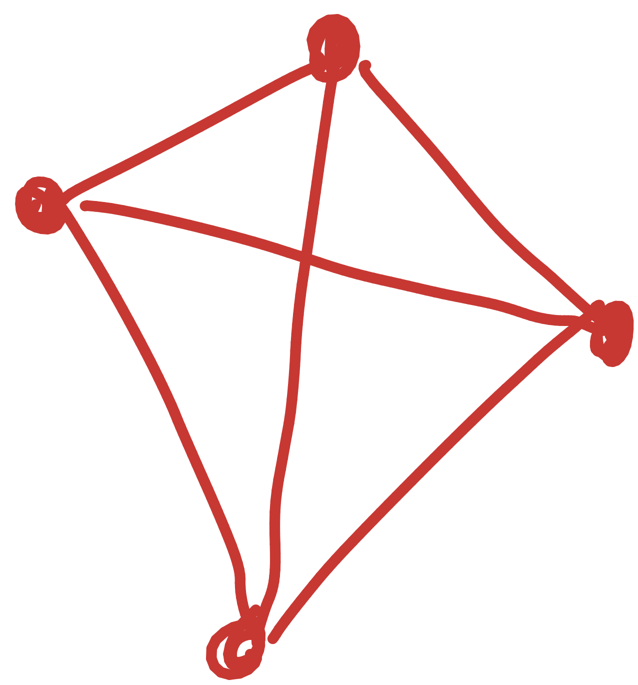
Momentum mode $\longleftrightarrow$ Neuron

$\hat{n}_k \equiv \hat{a}^+_k a_k$

$\hat{a}^+_k |0\rangle = |1_k\rangle$

Hamiltonian interactions between modes $\longleftrightarrow$ Synaptic connections

$$\hat{H}_{int} = d_{kj} \hat{n}_k \hat{n}_j$$



Exitation level of neuron n_k

Occupation number of k-th mode

$$n_k = \langle \hat{n}_k \rangle$$

$$\hat{\psi}(x) = \sum_{k} Y_k(x)\,\hat{a}_k$$

Complete set of harmonics.

This introduces a sense of geometry and locality in neural network

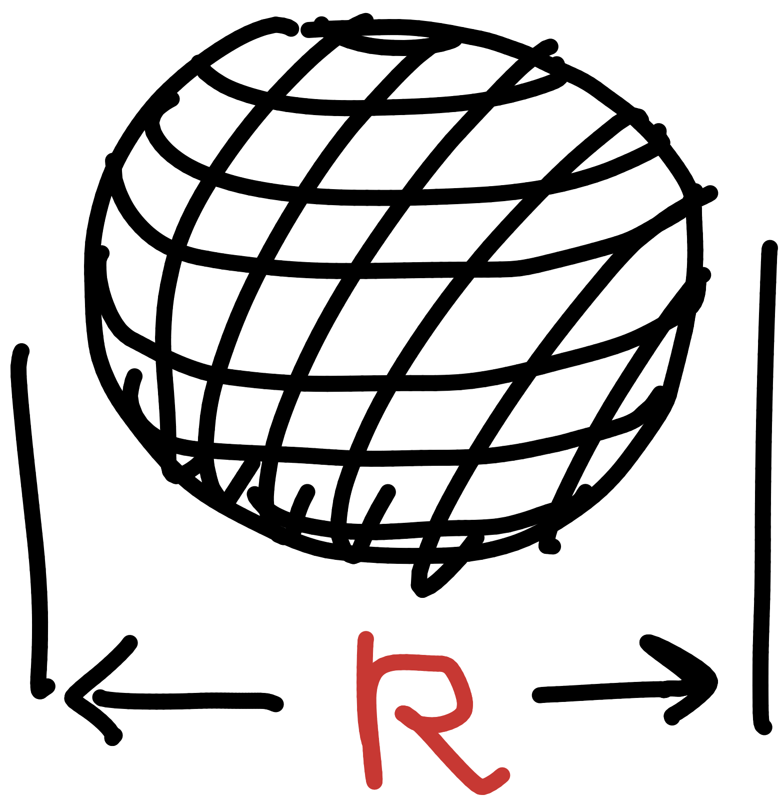
$$\hat{H}_{int} = \int \hat{\psi}^{+} \hat{\psi}^{+} \psi \psi$$

local in position space

Lesson from black holes:

Bekenstein entropy

$$S = \frac{R^2}{L_p^2}$$



This implies number of patterns

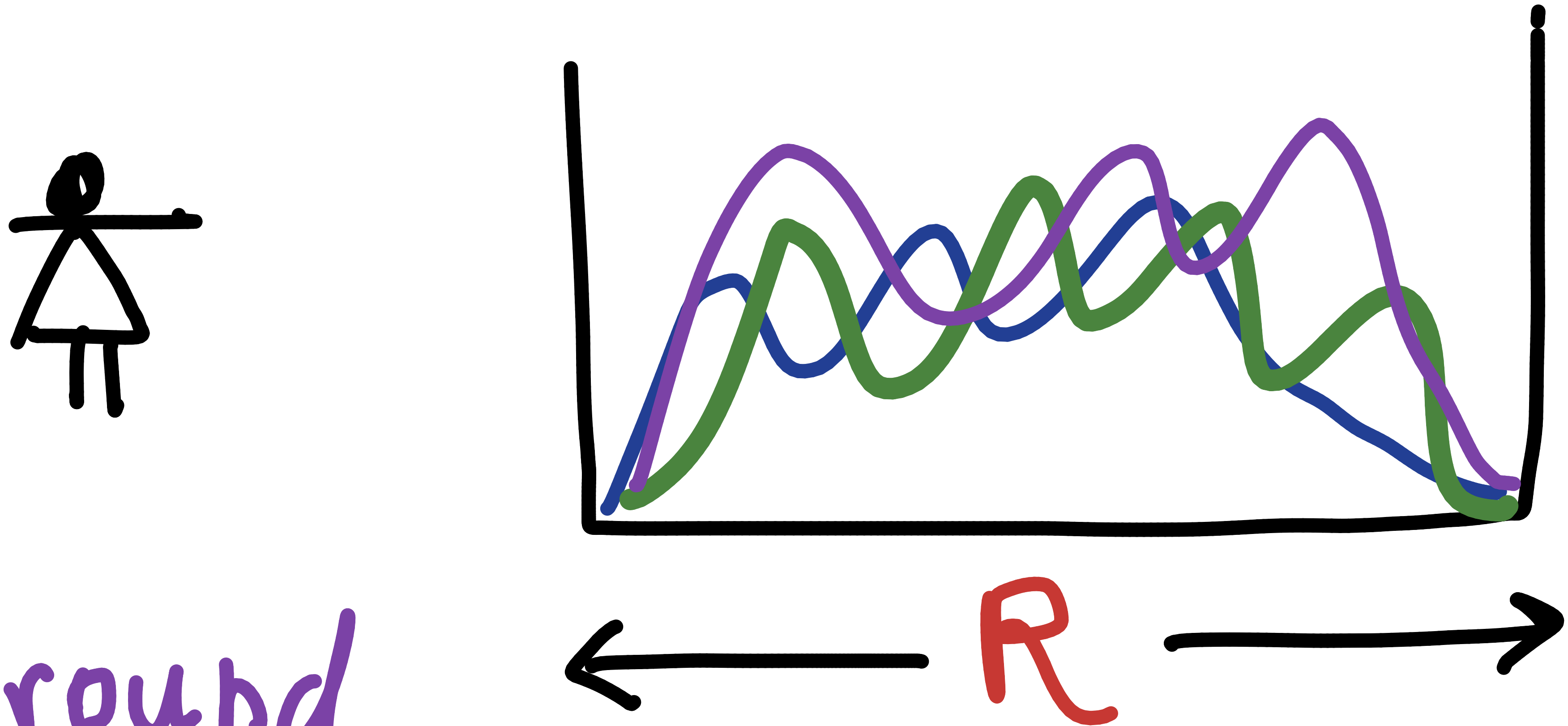
$$\Delta N_E = e^S$$ per energy gap $$\Delta E = \frac{\hbar}{R}$$!

We shall not speculate about how black hole manages this.

Our task:

Using well-known properties of black hole, design a neural network (equivalently a quantum field) with similar properties.

Alice and Bob



around vacuum $|0\rangle$ = empty box

Pattern storage is very costly in energy.

e.g. $|0\rangle \to |1\rangle$

$(0, 0, \dots 0) \longrightarrow (1, 0, \dots \dots 0)$

$$\Delta E \sim \hbar / R \,!$$

After the particle number becomes

$$N \sim \frac{1}{\alpha_{gr}} = \frac{R^2}{L_P^2}$$

a black hole is formed

$$M_{BH} = \frac{R^2}{L_P^2} \frac{\hbar}{R}$$

and memory capacity becomes huge

Elementary gap:

$$\Delta E \sim \frac{\hbar}{R} \frac{1}{N} \,!$$

Bob wants to achive the same in neural network:

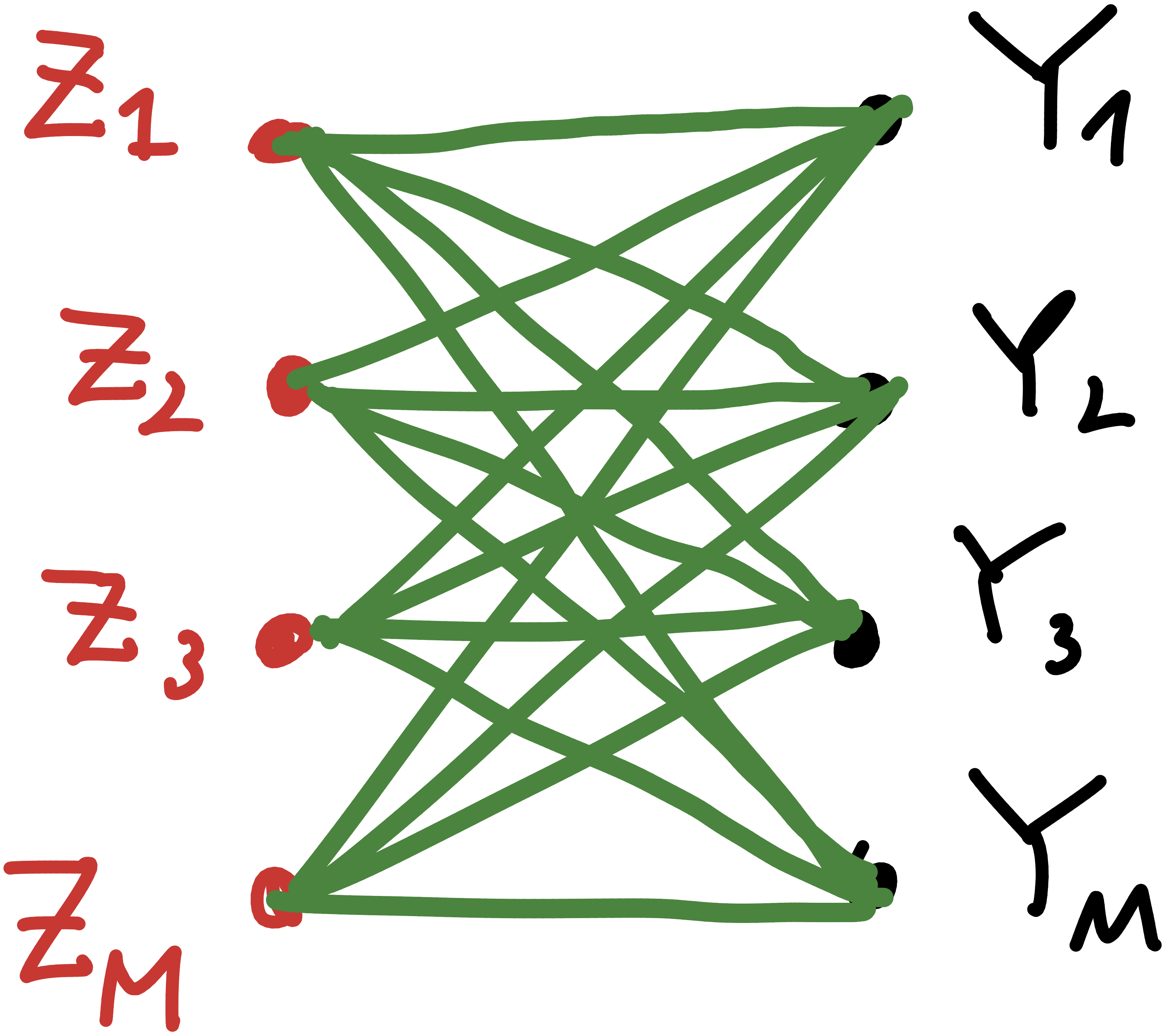
1. Design a network with gravity-like synaptic connections;

2. Look for the states with lowest momentum modes occupied to

$$N \sim \frac{1}{(\text{coupling strength})}$$

First the essence:

$$\hat{H} = \epsilon_\alpha \hat{Z}_\alpha + \epsilon_j \hat{Y}_j + - g_{\alpha j} \hat{Z}_\alpha \hat{Y}_j$$



Effective threshold for Y-neurons:

$$\mathcal{E}_j = \epsilon_j - g_{j\alpha} Z_\alpha$$

Thus, around the state

$$Z_\alpha = (g^{-1})_{\alpha j} \epsilon_j$$

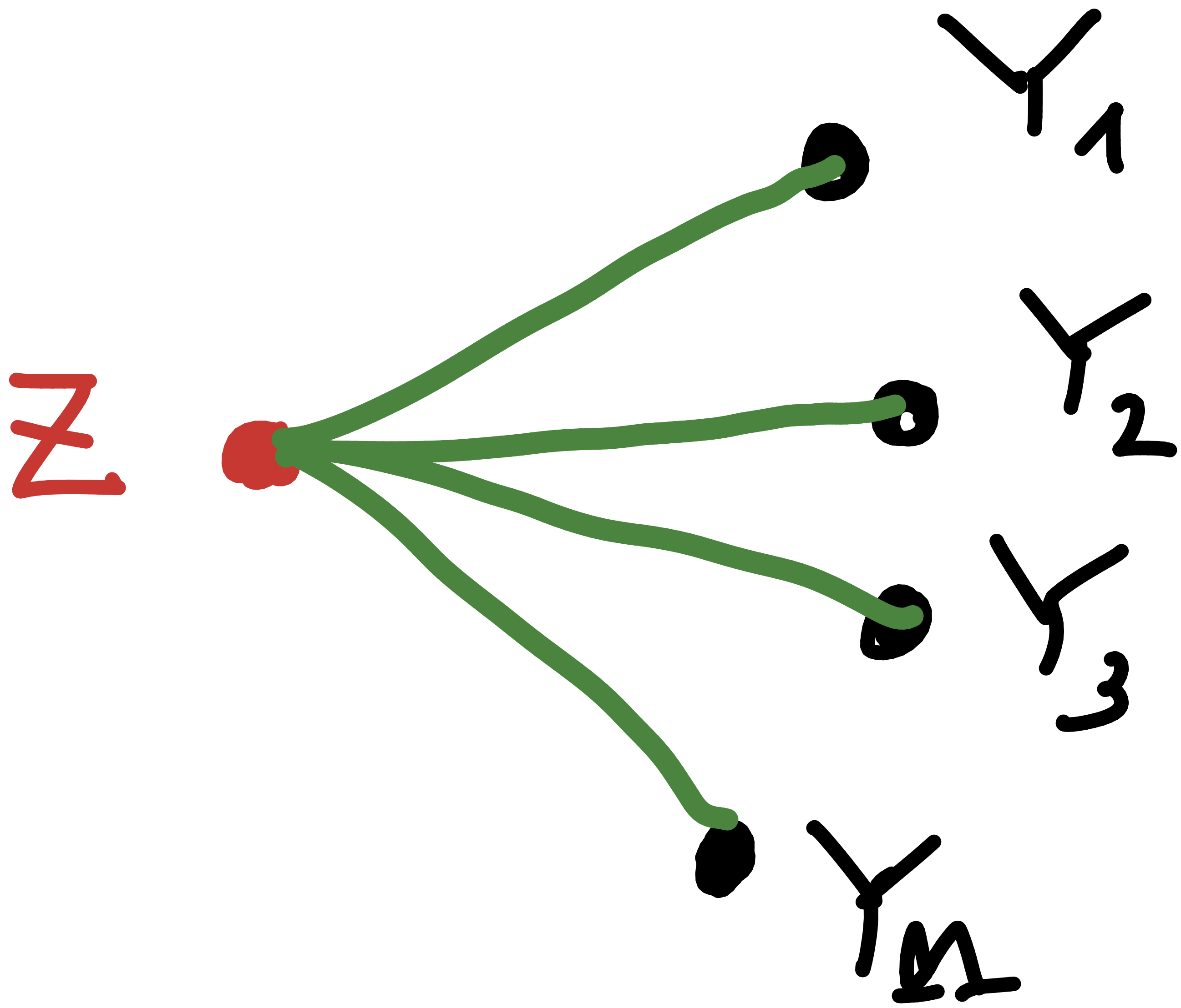
the Y-neurons become gapless and memory storage capacity is exponentially large

$$N_P \sim Y_{max}^M$$

In the presence of a symmetry, extra degeneracy appears.

e.g.,

$$\hat{H} = (\epsilon - \hat{Z} g)\hat{Y}$$



Gravity-like network with area low entropy

$$\hat{H} = \epsilon_k \hat{a}^\dagger_k \hat{a}_k + \hat{a}^\dagger_k \hat{W}_{kr} \hat{a}_r$$

where

$$\hat{W}_{kr} = \frac{\epsilon_k \epsilon_r}{2\Lambda} \cdot \left[3 - \frac{\epsilon_k \epsilon_r}{\epsilon_*}\right] \cdot \sum_{s,q} C_{skqr} \hat{a}^\dagger_s \hat{a}_q$$

$$C_{skqr} = \int d^d\Omega \, Y_s^* Y_k^* Y_q Y_r$$

Harmonics on S_d

$$Y_k \equiv Y_{k_1 \cdots k_d}$$

$$|k_1| \leq k_2 \leq \ldots \leq k_d = 0, 1, \ldots \infty$$

$$\Delta Y_k = - \frac{k_d(k_d + d - 1)}{R^2} Y_k$$

Laplace on S_d.

Eigenvalue degeneracy

$$N_k \sim (k_d)^{d-1}$$

Quantum field description of network

$$\hat{\psi} = \sum_{k} \gamma_k \hat{a}_k$$

Hamiltonian:

$$\hat{H} = \int_{S_d} -\hat{\psi}^{+} \Delta \hat{\psi} - \frac{3}{2\Lambda} (\hat{\psi}^{+} \Delta \hat{\psi}^{+}) \cdot (\hat{\psi} \Delta \hat{\psi}) + \frac{1}{2\Lambda \epsilon_*^2} (\hat{\psi}^{+} \Delta^2 \hat{\psi}^{+}) (\hat{\psi} \Delta^2 \hat{\psi})$$

$\Lambda, \epsilon_* \leftarrow$ parameters

The state of enhanced memory capacity:

$$\langle \hat{a}_0^+ \hat{a}_0 \rangle = N_0 = \frac{\Lambda}{\epsilon_*} \gg 1$$

Double-scaling limit

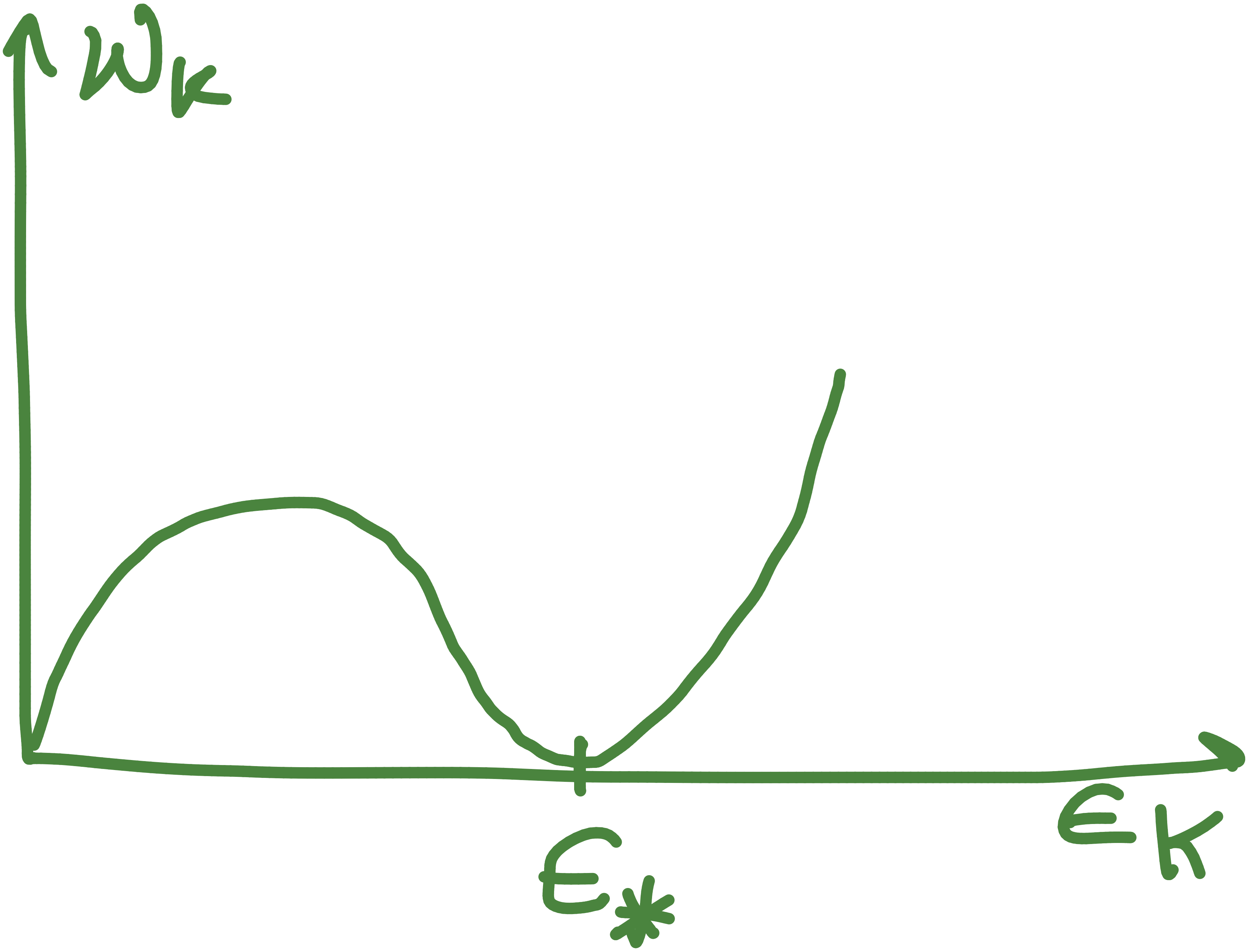
$$N_0 \to \infty, \quad \frac{N_0}{\Lambda} = \text{finite}$$

modes $\epsilon_k = \epsilon_*$

are <u>exactly</u> gapless!

$$\hat{H} = \sum_k \omega_k \hat{a}_k^\dagger \hat{a}_k + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$$

$$\omega_k = \frac{\epsilon_k}{2}\left(\frac{\epsilon_k}{\epsilon_*} - 1\right)^2\left(\frac{\epsilon_k}{\epsilon_*} + 2\right)$$



Number of gapless modes

$$N_{gapless} = k_*^{d-1} =$$

$$= \left(\frac{R}{L_*}\right)^{d-1}$$

$$L_* \equiv \sqrt{\epsilon_*^{-1}}$$

$N_{gapless} \propto$ Area of S_{d-1}!

Patterns can be stored in micro-states obtained by exciting gapless modes, say up to

$$\langle \hat{a}^{\dagger}_{k_*} a_{k_*} \rangle < n$$

$$N_p = n^{N_{gapless}}$$

$$\downarrow$$

$$\text{Entropy} \sim \left(\frac{R}{L_*}\right)^{d-1} \ln(n)$$

Summary:

- A neural network with most naive gravity-like synaptic connections exhibits a state of sharply enhanced memory storage capacity.
- In such a state some neurons become effectively gapless and can store exponentially large number of patterns within narrow energy gap.

⊛ The phenomenon has a smooth classical limit.

⊛ Neural network can be mapped on a quantum field. The state of high memory capacity is then translated as a critical state of enhanced micro-state entropy of some gapless modes.

⊛ For spherical symmetry the systems exhibits area law entropy.

* This gives an explicit microscopic realization of "holography":

$$N_{\text{gaplen}} \propto \text{Area of } S_{d-1}$$

* There are many open interesting questions