Brain Memory Working

Franco Cardin\textsuperscript{1}, Alberto Lovison\textsuperscript{2}, Amos Maritan\textsuperscript{3}, Aram Megighian\textsuperscript{4}

PINN-PAD: Physics Informed Neural Networks in Padova - Dipartimento di Ingegneria Civile, Edile e Ambientale

\textsuperscript{1}Dipartimento di Matematica “Tullio Levi-Civita” - Università di Padova
\textsuperscript{2}Dipartimento di Matematica e Fisica “Ennio De Giorgi” - Università del Salento
\textsuperscript{3}Dipartimento di Fisica e Astronomia “Galileo Galilei” - Università di Padova
\textsuperscript{4}Dipartimento di Scienze Biomediche and Padova Neuroscience Center - Università di Padova
Original discrete Hopfield model with $N$ neurons. At the $n$-th time step activation potential at neuron $i$:

$$V^{(n)}_i, \quad i = 1, \ldots, N.$$ 

$T_{ij} =$ conductance between neurons $i$ and $j$. Potentials updating rule:

$$V^{(n+1)}_i = g \left( \sum_{j=1}^{N} T_{ij} V^{(n)}_j \right), \quad g(u) := \begin{cases} +1 & u \geq a, \\ -1 & u < a. \end{cases}$$

A neuronal circuit.
The associated continuous dynamics

\[
\dot{u}_i = \sum_{j=1}^{N} T_{ij} g(u_j) - u_i
\]

where \( g(x) \) is a sigmoidal activation function, is described by the vector field

\[
X_i(u) = \sum_{j=1}^{N} T_{ij} g(u_j) - u_i.
\]
Symmetry: $T_{ij} = T_{ji}$

Symmetry + Constancy: Energy landscape

$$E(V) := -\frac{1}{2} T_{ij} V_i V_j + \sum_{i=1}^{N} \int_0^V g^{-1}(x) \, dx.$$ 

Gradient dynamics

$$X = -\nabla E$$

Dynamics drives the potential pattern $(V_1, \ldots, V_N)$ towards the local energy minimum.

Contour plot of the energy landscape. Gradient-type dynamics. Minima = red dots. Maxima = blue dots.
Hebbian updates are discontinuous and can only add new patterns until saturation.

\[ T_{ij}^{new} = T_{ij}^{old} + \frac{1}{N} \hat{V}_i \hat{V}_j \]

Excessively rigid updating scheme: the network is forced to learn a pattern.
the interaction matrix $T_{ij}$ varies with electric potential $V_i$ according to this “unusual” rule:

$$T_{ij} \rightarrow T_{ij}(V) := \frac{\partial^2 \Phi}{\partial V_i \partial V_j}(V), \quad \text{(still symmetric)} \quad (\spadesuit)$$

$T_{ij}(V)$ is the Hessian of the Lagrangian $\Phi(V)$.

via Legendre transform we obtain the Hamiltonian energy:

$$E(V) := -\left( \nabla \Phi(V) \cdot V - \Phi(V) - \sum_{i=1}^{N} \int_{0}^{V_i} g^{-1}(x) dx \right).$$

new gradient vector field:

$$\widehat{X}_i(V) := -\nabla_i E(V) = \nabla_{ij}^2 \Phi(V) \cdot V_j - g^{-1}(V_i).$$
We have decrypted condition (♠):

**Theorem**

*In a simply connected domain, the closure condition:*

\[(T_{kj,i} - T_{ki,j})V_k = 0, \quad (\star)\]

*is equivalent to the gradient structure for \(\hat{X}\) and to the existence of a Lyapunov-like energy function.*

**Remark**

*Under the stronger condition:

\[T_{kj,i} - T_{ki,j} = 0, \quad (\diamond)\]

*we gain the Krotov hypothesis (♠): \(T = \nabla^2 \Phi\), but not to (\(\star\)).*
In the more general condition (⋆), we define now the corresponding Energy function. Let:

\[ W(x) := \int_0^1 T_{ij}(\lambda x)\lambda x_i x_j d\lambda, \]

we set

\[ E(V) := -W(V) + \sum_{i=1}^N \int_0^V g^{-1}(\lambda) d\lambda, \]

and obtain

\[ \hat{X}_i = -\nabla E(V) = T_{ij}(V)V_j - g^{-1}(V_i). \]

**gradient-like dynamics**
Physiology states that $T_{ij}$ is asymmetric: connections are directed, i.e., specific structures are dedicated to outgoing (axons) and incoming (dendrites) connections.

Features non comprised by symmetric interactions:

- oscillations / memory association,
- wandering (instability),
- forgetting and recovering memories.

Oscillations and instability need for asymmetry in $T_{ij}$

Oscillations or limit cycles are only possible with asymmetry.
A NEW PROPOSAL: ASYMMETRIC OPTIMAL CONTROL

- Starting constant matrix $A_{ij}$ non-symmetric.
- $\xi$-controlled adjustments:
  
  \[
  T_{ij}(\xi) := A_{ij} + \xi_{ij}, \quad |\xi_{ij}| \leq K
  \]
- $\xi$-controlled Hopfield dynamics:
  
  \[
  \dot{u}_i(t) = X_i(u(t), \xi(t)) = \sum_{j=1}^{N} \left( A_{ij} + \xi_{ij}(t) \right) g(u_j(t)) - u_i(t).
  \]

Trajectories dynamically evolving during motion. Dynamical evolution of the energy landscape (symmetric).

Ideas already appeared for instance in

- D. Vardalaki et al, Filopodia are a structural substrate for silent synapses in adult neocortex. Nature 2022
A NEW PROPOSAL: ASYMMETRIC OPTIMAL CONTROL

Remark on sparsity:

- If \( N \) is large, \( A_{ij} \) is sparse.
- \( \xi_{ij} \) may update only \( A_{ij} \neq 0 \), or...
- \( \xi_{ij} \) may also act on \( A_{ij} = 0 \), lighting up
  - existing but silent synapses
  - build brand new synapses not existing before

Proposal: fix \( 0 < k \ll K \):

- if \( A_{ij} \neq 0 \) \( \implies |\xi_{ij}(t)| \leq K \),
  i.e., if a connection \( A_{ij} \) between neurons \( i \) and \( j \) already exists, then the corresponding update may be “strong”: \( \xi_{ij} \leq K \).

- if \( A_{ij} = 0 \) \( \implies |\xi_{ij}(t)| \leq k \ll K \),
  i.e., if \( A_{ij} \) is silent, then only smaller updates are possible \( \xi_{ij} < k \ll K \).

Resuming the updating scheme we write: \( |\xi_{ij}(t)| \leq (k, K) \).

- D. Vardalaki et al, Filopodia are a structural substrate for silent synapses in adult neocortex. Nature 2022
A NEW PROPOSAL: ASYMMETRIC OPTIMAL CONTROL

Surprisingly: there exist a perfectly fit powerful mathematical framework:

Infinite Horizon Optimal Control Problem.

- Differential Constraint:

\[ \dot{u}_i(t) = X_i(u(t), \xi(t)) = \sum_{j=1}^{N} (A_{ij} + \xi_{ij}(t)) g(u_j(t)) - u_i(t). \]  

- \( e^{-\lambda t} \)-discounted variational principle:

\[
\min_{\xi(\cdot)} J\left(u^{(0)}, \xi(\cdot)\right) = \min_{\xi(\cdot)} \int_0^{+\infty} \left\{ \left| X(u(t, u^{(0)}, \xi(\cdot)), \xi(t)) \right|^2 + |\xi(t)|^2 \right\} e^{-\lambda t} dt
\]

Lagrangian: \( \ell(u, \xi) \)

where \( u(u^{(0)}, \xi(\cdot)) \) solves the differential constraint (\( \dagger \)).
A NEW PROPOSAL: ASYMMETRIC OPTIMAL CONTROL

Infinite Horizon Optimal Control Problem.

- The Lagrangian of the Control Problem:
  \[ \ell(u, \xi) = |X(u, \xi)|^2 + |\xi|^2, \]

- \(|X|^2\) small ⇒ towards equilibra,
- \(|\xi|^2\) small ⇒ cheap solutions in terms of matrix modification.
- The discount \(e^{-\lambda t}\) ensures convergence.
- Control problem: for fixed \(u^{(0)}\) find the minimizing controls \(\xi(\cdot)\):
  \[ \inf_{|\xi(t)| \leq (K,k)} J\left(u^{(0)}, \xi(\cdot)\right), \]
A controlled trajectory starting from the input pattern $u^{(0)}$ may fall in one of the following classes:

- reach existing equilibrium without activating the controls $\xi = 0$:

\[
\lim_{t \to \infty} X(u(t, u^{(0)}, 0), 0) = X(u^*, 0) = 0,
\]

i.e., the initial pattern $u^{(0)}$ has been recognized.

- The controls $\xi(t) \neq 0$ operate to minimize $J(u, \xi)$ and asymptotically drive to a new equilibrium:

\[
\lim_{t \to \infty} X(u(t, u^{(0)}, \xi(t)), \xi(t)) = X(u^{**}, \xi^{**}) = 0,
\]

i.e., the initial pattern $u^{(0)}$ has been recorded in the network $T_{ij} \to T_{ij} + \xi^{**}$ and a new equilibrium $u^{**}$ has been created.
Assume that \( \bar{u} \) is an equilibrium for the synaptic matrix \( T_{ij} \). A sequence of alterations to \( T_{ij} \) are operated:

\[
T_{ij} \rightarrow T_{ij} + \xi^\alpha + \cdots + \xi^\omega, \quad |\xi^\alpha + \cdots + \xi^\omega| > K.
\]

In the new configuration the pattern \( \bar{u} \) cannot be recognized, i.e., the pattern \( \bar{u} \) has been forgot.

(continued) Successive alteration \( \xi^n_\infty \) may bring back the synaptic network closer to the starting configuration:

\[
|\xi^\alpha + \cdots + \xi^\omega + \xi^n_\infty| \leq K,
\]

allowing to recover the old equilibrium \( \bar{u} \), i.e., a memory has been restored.
Discussion of the controlled model

Given the asymmetry of $T_{ij}$, limit cycles are possibly approached ($\xi = 0$) or created ($\xi_\infty \neq 0$) during the controlled motion:

$$\lim_{t \to \infty} \text{dist} \left( u(t, u(0), \xi(t)), \mathcal{U} \right) = 0, \quad \mathcal{U} \subseteq \mathbb{R}^N \quad \text{(limit cycle)},$$

Instability with oscillations: this situation can be interpreted as memory association.

Controls are activated during the motion ($\xi(t) \neq 0$) but they are not able to reach or create any equilibrium:

$$\lim_{t \to \infty} u(t, u(0), \xi(t)) \quad \text{does not exist.}$$

Instability with wandering: pattern not found nor created.
FURTHER DISCUSSION / CONCLUSIONS

- **Final Value Theorem**
- **Hamilton-Jacobi-Bellman Equation**
- **Dynamic Programming Principle**
- **Pareto optimization: conservative/innovative attitudes:**

\[
J_\mu \left( u^{(0)}, \xi(\cdot) \right) := \int_0^{+\infty} \left( (1 - \mu) \left| X(u(t, u^{(0)}, \xi(\cdot)), \xi(t)) \right|^2 + \mu \left| \xi(t) \right|^2 \right) e^{-\lambda t} dt
\]

Lagrangian: \( \ell_\mu(u, \xi) \)

- if \( 0 < \mu \ll 1 \) large values of \( \xi \) are allowed, letting the network explore innovative configurations,
- if \( 0 \ll \mu < 1 \) large values of \( \xi \) are penalized and the network is more prone towards existing minima: conservative attitude.
Thanks for your attention!

Brain memory working. Optimal control behavior for improved Hopfield-like models
https://arxiv.org/abs/2305.14360

alberto.lovison@unipd.it