Storage and retrieval of ultrametric patterns in a network of CA1 neurons of the hippocampus

Daniela Bianchi Marco Piersanti *Brunello Tirozzi*

Department of Physics
–
Sapienza University of Rome

Contents

Contents

Associative memory and Hopfield model

- Associative memory (AM) is the ability of humans of remembering an information starting from a partial knowledge of it. *For example, remembering the entire title of a film knowing only a part of it.*
- The first question is to learn patterns and to conserve them, i.e. learning and storage.
- Once some patterns have been learned and stored the important ability is to retrieve them.
- In the past time a lot of work, both theoretical and experimental, has been done to construct models of associative memory.
- The important feature of these models is a measure of the retrieval and storage. The models are built with a number N of units, called **neurons**, and a number P of information called **patterns**. This measure is called **capacity** and is defined as the ratio

$$\alpha = \frac{P}{N} \tag{1}$$

Associative memory and Hopfield model

- The first attempt to simulate the behaviour of human long-term memory was the Hopfield model (Hopfield, 1982). The units were called neurons since their functioning was mimicking the evolution of real neurons, the set of N neurons was called **neural network**.
- The neurons were represented as an idealized two-state devices (McCulloch and Pitts, 1943) coupled through a symmetrical matrix J_{ij} that represents the synapses. Real neurons are coupled through very thin tubes called dendrites which end on the membrane of the neuronal cell in a structure called synapsis.
- The learning process is a particular evolution of the synaptic matrix J_{ij}, often a stochastic process, which makes the matrix to converge to some definite matrix which allows the retrieval of the information. The learning process is called *supervised* if the evolution is controlled by some external factor, *unsupervised* if there is no form of external control during the evolution.

Definitions

Here we summarize the basic assumptions and definitions of the Hopfield model.

- The **all-or-none** firing of a neuron is represented by a variable S_i taking two values: $S_i = +1$ (firing), $S_i = -1$ (rest). There are N of such variables. A configuration of the system is a collection of these variables $\underline{S} \equiv (S_1, S_2, \ldots, S_N)$.
- The dynamics of a neuron is a stochastic threshold dynamics:

$$S_i(t+1) = \text{sgn}\, h_i(t) \tag{2}$$

$$h_i(t) = \sum_{j, j \neq i}^{N} J_{ij} S_j(t) \tag{3}$$

- A pattern of activity, $\underline{\xi}^\mu$, of a network of N neurons is represented by a set of i.i.d.r.v. $\{\xi_i^\mu = \pm 1\}$, $i = 1, \ldots, N$, that lies at the corners of an N dimensional hypercube . There are P patterns $\underline{\xi}^\mu$ to store and retrieve, $\mu = 1, \ldots, P$.

Definitions

- Two patterns of activity, μ and ν, may be compared through their **overlap**:

$$\langle \mu | \nu \rangle = \frac{1}{N} \sum_{i=1}^{N} \xi_i^{\mu} \xi_i^{\nu} \tag{4}$$

- The overlap m^{μ} of a pattern $\underline{\xi}^{\mu}$ with a configuration $\underline{S}$ is a measure of the retrieval of the stored information in the network

$$m^{\mu} = \frac{1}{N} \sum_{i=1}^{N} \xi_i^{\mu} S_i \tag{5}$$

Definitions

- During *learning* the J_{ij} are modified by the system. A set of p patterns $\{\xi_i^\mu\}$, $i = 1, \ldots, N$, $\mu = 1, \ldots, P$, is embedded in the J_{ij}'s, via the **Hebbian** learning rule

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \tag{6}$$

The learning process of the Hopfield model is a supervised process.

- The patterns are memorized in the sense that each pattern $\underline{\xi}^\mu$ is a fixed point of the dynamics.

The result of the investigations made by Hopfield (1982); Amit et al. (1985) was that there is critical value of the capacity α_c such that all the overlap parameters are zero for $\alpha > \alpha_c$. The value found was $\alpha_c \sim 0.134$

Contents

Classes of patterns and ultrametricity

The memorization and retrieval of information is easier if one uses the classes.

Classes are like the atoms of the partition of a finite set and the single pattern is an element of the atom.

The hierarchical structure of the patterns is thus well described using the ultrametric distance in an ultrametric space.

Marc Krasner (1912–1985) invented this word in a note presented at the French Academy of Sciences on October 23, 1944, entitled “Nombres semi-réels et espaces ultramétriques”.

Classes of patterns and ultrametricity

The ultrametric inequality is the inequality:

$$d(A, C) \leq \max\{d(A, B),\ d(B, C)\} \tag{7}$$

A distance that satisfies the ultrametric inequality is called an *ultrametric distance*. A space endowed with an ultrametric distance is called an *ultrametric space*.

The problem is then to organize the patterns in such a way that they are divided in classes and elements of the classes, and distinguish among them by means of a distance, which is the ultrametric distance.

The aim is to construct patterns which have an ultrametric structure and such that they can be used for storage and retrieval in a network of real neurons. There are many ways to organize patterns in an ultrametric space.

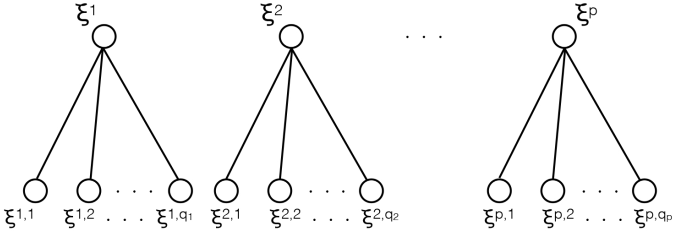
Definitions

Ancestors

p patterns $\{\xi^\mu = \pm 1\}$ with $\mu = 1, \ldots, p$, ξ^μ being IIDRV.

Descendants

Each ancestor has q descendants $\{\xi^{\mu,\nu}\}$.



Multi-ancestors and two levels

Pattern

$$\xi_i^{\mu,\nu} = \xi_i^{\mu} \eta_i^{\mu,\nu} \tag{8}$$

with

$$\Pr(\eta^{\mu,\nu} = \pm 1) = \frac{1}{2}(1 \pm a_\mu) \tag{9}$$

- Two patterns in the same bunch have distances < 1 ($a_\mu < 1$)

$$\langle \mu, \nu | \mu, \lambda \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} \xi_i^{\mu,\nu} \xi_i^{\mu,\lambda} = \frac{1}{N} \sum_{i=1}^{N} \eta_i^{\mu,\nu} \eta_i^{\mu,\lambda} = a_\mu^2 \tag{10}$$

- Two patterns in different bunches have distance equal to 1

$$\langle \mu, \nu | \rho, \lambda \rangle = 0 \tag{11}$$

Properties

The synaptic matrix that stabilises these patterns is

$$T_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^{\mu}\xi_j^{\mu}\left(1 + \frac{1}{\Delta}\sum_{\nu=1}^{q_\mu}(\eta_i^{\mu,\nu} - a_\mu)(\eta_j^{\mu,\nu} - a_\mu)\right) \tag{12}$$

- When $\Delta = 1 - a^2$, the degeneracy between parents and descendants sets in. The storage capacity is the familiar $\alpha \approx 0.15$, where α refers to the total number of memorized pattern, i.e.

$$\alpha N = p + \sum_{\mu=1}^{p} q_\mu \tag{13}$$

  All these states become attractors.
- If $\Delta > 1 - a^2$ the degeneracy is lifted and the parents become lower in energy than the descendants.
- The total storage capacity remains the same, but the ancestors appear first, at higher loading levels, and then the detailed descendants become retrieval states at lower loading levels.

Contents

Indexed hierarchies

Let E be a finite set. A *hierarchy* $\hat{H}$ on E is a special set of partitions of E, $\hat{H}(E)$, such that

i) $E \in \hat{H}(E)$;

ii) each single element $a \in E$ belongs to $\hat{H}(E)$, i.e. the atoms of the partitions can be also single elements, (singleton);

iii) for each pair of partitions $r, r' \in \hat{H}$, such that $r \cap r' \neq 0 \Longrightarrow r \subset r'$ or $r' \subset r$

An *indexed hierarchy* on E is a pair $\{\hat{H}, f\}$ where $\hat{H}$ is a given hierarchy on E and f is a positive function satisfying the following conditions

i) $f(a) = 0$ if and only if a is a single element of E (a singleton);

ii) if $a \subset a'$ then $f(a) < f(a')$.

f corresponds to the index of the levels of the hierarchies introduced in the Markov chain approach.

Definition of the distance

The distance among two subsets of E is

$$\delta(a,b) = \min\{d(x,y) \mid x \in a,\ y \in b\} \tag{14}$$

Example: trivial ultrametric

If $E: E = \cup_i E_i$, then $d(x,x) = 0$, $d(x,y) = 1$ if $x \in E_i$, $y \in E_j$ $(i \neq j)$, and $d(x,y) = a$ if $i = j$, $0 < a < 1$.

Hierarchies and ultrametrics

Associated with each indexed hierarchy $\{\hat{H}(E), f\}$ on E is the following ultrametric:

$$\sigma(x, y) = \min_{a \in \hat{H}(E)} \{f(a) \,|\, x \in a,\, y \in a\} \tag{15}$$

This means that the distance $\sigma(x, y)$ between two elements x and y in E is given by the *index of the* smallest element in $\hat{H}(E)$, which contains both x and y (rule of the closest common ancestor).

The measure of approximation of the measure d (**proximity index**) is

$$\Delta_0(d, \delta) = \max_{x,y \in E} |d(x, y) - \delta(x, y)| \tag{16}$$

Our goal is then to find the best approximating $\delta(x, y)$ of $d(x, y)$.

Subdominant ultrametric

Limit the search on a subset of $\mathcal{U}$ (the set of all the ultrametrics on E).

$$\mathcal{U}^s = \{\delta \in \mathcal{U} \mid \delta \leq d\} \tag{17}$$

Definition

The *subdominant ultrametric* d^s is defined as the upper limit of $\mathcal{U}^s$. This is the maximal element in $\mathcal{U}^s$, and by definition

$$d^s(x,y) = \max\{\delta(x,y) \mid \delta \in \mathcal{U}, \delta \leq d\} \tag{18}$$

$$\Delta(d, d^s) = \min\{\Delta(\delta, d) \mid \delta \in \mathcal{U}, \delta \leq d\} \tag{19}$$

We use the **Minimum-Spanning-Tree** (MST) construction method (Murtagh, 1983; Prim, 1957).
Note that the although the MST is not uniquely defined, d^s is unique.

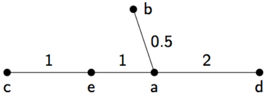
Subdominant ultrametric

If A is a MST on E, the distance $d^s(x, y)$ between two elements x and y in E is given by

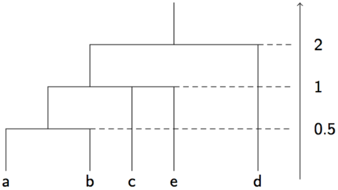
$$d^s(x, y) = \max\{d(w_i, w_{i+1}), i = 1, \ldots, n-1\} \tag{20}$$

where $\{(w_1, w_2), (w_2, w_3), \ldots, (w_{n-1}, w_n)\}$ denotes the unique chain in A, between x and y ($w_1 = x$, $w_n = y$).

	a	b	c	d	e
a	0	0.5	3	2	1
b		0	4	3	4
c			0	5	1
d				0	3
e					0



	a	b	c	d	e
a	0	0.5	1	2	1
b		0	1	2	1
c			0	2	1
d				0	2
e					0



Examples

Distortion index

$$\mathcal{D} = 1 - \frac{\sum_{x,y \in E} d^s(x,y)}{\sum_{x.y \in E} d(x,y)} \tag{21}$$

d is the input metric on E, d^s is the associated subdominant ultrametric.

In general, $0 \leq \mathcal{D} \leq 1$, vanishes if d is already an ultrametric (i.e., $d^s = d$) and provides a quantitative measure of ultrametricity.

Take $E = \{x_1, \ldots, x_n\}$, $x_i \in R$ and $d(x_i, x_j) = |x_i - x_j|$ is the usual Euclidean metric. If $x_i = i$, then the MST is the set of edges going from x_i to $x_{i+1} \rightarrow d^s(x_i, x_j) = 1$: all triangles are equilateral. It can be shown that for large n

$$\mathcal{D} \simeq 1 - \frac{3}{n+1} \sim 1 \tag{22}$$

Euclidean spaces are far from being ultrametric spaces.

Examples

Let E be a set P binary words of N bits each, taken randomly from among the 2^N possible words. The distance among two words $\xi^1 = (\xi_1^1, \dots, \xi_N^1)$, $\xi^2 = (\xi_1^2, \dots, \xi_N^2)$ is the Hamming distance

$$d(\xi^1, \xi^2) = \sum_{i=1}^{N} | \xi_i^1 - \xi_i^2 | \tag{23}$$

- for $P = 2^N$, d^s reduces to the trivial ultrametric and $\mathcal{D}_{\mathcal{N}}(x = 1) = 1 - 2/N \sim 1$ at large N where $x = P/2^N$ is the filling factor of the hypercube of all the configurations $\{0, 1\}^N$.
- For fixed but large N numerical calculations show that $\mathcal{D}_{\mathcal{N}}$ approaches zero as x goes to zero. This means that if the number of patterns is small (10 or 20) and the dimension of the vector is 100 we have ultrametricity. Ultrametricity holds in the case of large spaces (sparse coding).
- This is our case with the patterns of the CA1 neural network.

Basic definitions

Here we follow the ideas of Lerman (1981).

Partition

Let E be a finite set. A *partition* of E is a set of disjoint subsets of E such that their union is E, the classes of the partition being the subsets.

Let us consider an example: $E = \{a, b, c, d, e, f, g\}$
A partition of E is

$$\Big\{ \{a, b, c, d\} \ , \ \{e, f\} \ , \ \{g\} \Big\}$$

while the classes of the partitions are

$$\{a, b, c, d\} \ , \ \{e, f\} \ , \ \{g\}$$

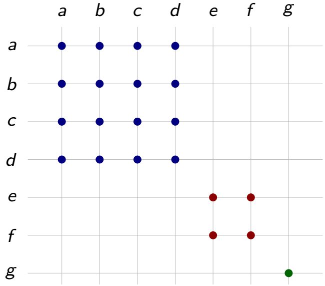
We will indicate with $\hat{P}(E)$ the set of the partitions of E, P being a generic partition.

Basic definitions

Two elements of E are **equivalent** if they belong to the same class of the partition P.
The graph of an equivalence relation induced by P is indicated with $\mathrm{Gr}(P)$,

$$\mathrm{Gr}(P) = \{(x, y) \,|\, x \in E,\ y \in E \text{ and } xPy\} \tag{24}$$

$\mathrm{Gr}(P)$ is a subset of $E \times E$. This inclusion allow us to define an ordering in $\hat{P}$.



Ordering

$P < P'$ if $\mathrm{Gr}(P) < \mathrm{Gr}(P')$ or if, $\forall x, y \in E$, $xPy \Longrightarrow xP'y$.

For example the partition

$$\Big\{\{a, b, c, d\}, \{e, f\}, \{g\}\Big\}$$

is smaller than the partition

$$\Big\{\{a, b, c, d\}, \{e, f, g\}\Big\}$$

$\hat{P}(E)$ is an ordered set with the structure of a lattice in the sense that for any pair of partitions there is a "greatest smaller" partition $P \wedge P'$ and a "smallest greater" partition $P \vee P'$. The partition $P \wedge P'$ is defined by its graph

$$\mathrm{Gr}(P \wedge P') = \mathrm{Gr}(P) \wedge \mathrm{Gr}(P') \tag{25}$$

$x(P \wedge P')y$ if and only if xPy and $xP'y$. On the other hand, $\mathrm{Gr}(P \vee P')$ is the smallest graph containing the set $\mathrm{Gr}(P)$ or $\mathrm{Gr}(P')$.

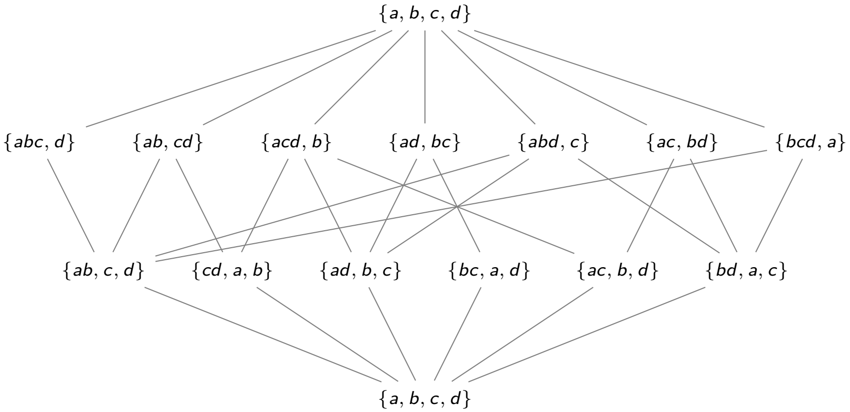
Example

Consider the same partition P introduced before and a new one P':

$$P = \Big\{ \{a, b, c, d\}, \{e, f\}, \{g\} \Big\}$$
$$P' = \Big\{ \{a, b\}, \{c, d\}, \{e, f, g\} \Big\}$$
$$P \wedge P' = \Big\{ \{a, b\}, \{c, d\}, \{e, f\}, \{g\} \Big\}$$
$$P \vee P' = \Big\{ \{a, b, c, d\}, \{e, f, g\} \Big\}$$

The classes of the smallest partition of $\hat{P}$ are the single elements of E, the largest is the set $\{a, b, c, d\}$. We show the example of lattice constructed starting from the set of four elements.

Example


{a, b, c, d}
{abc, d}
{ab, cd}
{acd, b}
{ad, bc}
{abd, c}
{ac, bd}
{bcd, a}
{ab, c, d}
{cd, a, b}
{ad, b, c}
{bc, a, d}
{ac, b, d}
{bd, a, c}
{a, b, c, d}

Chain of partitions and ultrametric spaces

Let (E, d) be a metric space with a finite number of elements. A *divisor* of E is an equivalence relation D in E such that

$$\forall a, b, x, y \in E \quad aDb \quad \text{and} \quad d(x, y) \leq d(a, b) \Longrightarrow xDy \tag{26}$$

We can associate to each sequence of increasing lattices of partitions of E, $\hat{P}(E)$, an ultrametric space. That is, if we consider a finite sequence of partitions of E, P_i, with $P_i < P_{i+1}$, these partitions decrease their fineness as i increase.

We define the distance function $d(x, y) : E \times E \to I \subset \mathbb{N}$ as the smallest i such that x, y belong to the same class P_i. I is a finite set of integers. Thus defined, d has the following property.

d is an ultrametric distance on E such that the divisors are the P_i. The inverse also holds.

Definitions

The relations among the various classes of the partitions and the elements of these classes are described in general terms which we have to define for the sake of clarity. For any two objects in the finite space E we have defined a distance among them. Based on this definition, we introduce some general *binary relations* among the pairs.

i) a binary relation on E is a **preorder** if it is reflexive and transitive;
ii) a binary relation on E is an **equivalence** relation if it is reflexive, transitive and symmetric;
iii) a binary relation on E is an **order** if it is reflexive, transitive and antisymmetric;
iv) a binary relation is called *total* if it holds for all the pairs $(i, j) \in E \times E$, or *partial* otherwise.

Ultrametric preorder

A preorder can be associated to a set of partitions organized in a lattice $\hat{P}(E)$. Let F be the set of all pairs of elements in E. The distance d defines a *total preorder* in F:

$$\forall\{(x,y),(z,t)\} \in F : (x,y) \leq (z,t) \iff d(x,y) \leq d(z,t) \qquad (27)$$

The preorder is indicated with ω. Two distances on a given set E are equivalent if the preorderings associated with each of them on E are identical. A total preorder is equivalent to a partition which defines an equivalence relation on F, and to a total order on the set of classes.

A preorder ω is called ultrametric if

$$\forall x,y,z \in E : \rho(x,y) \leq r\,,\ \rho(y,z) \leq r \Longrightarrow \rho(x,z) \leq r$$

$\rho(x,y)$ is the rank of the pair, for ω defined by the non-decreasing values of the distance d in E.

A necessary and sufficient condition for a distance d to be ultrametric is that the associated preordering is ultrametric.

Difference between preorder and ultrametric preorder

It is possible to introduce a quantity which measures the degree of ultrametricity starting from these definitions. Let J be the set of all the triplets (x, y, z) of elements of E. Consider the application τ of J in F that, given (x, y, z) and the preorder ω, associates to them the open interval $]M(x, y, z), S(x, y, z)[$, which are respectively the median and the maximum among the three couples (x, y), (y, z) and (z, x).

We just say that a triplet (x, y, z) for which $(x, y) \leq (y, z) \leq (x, z)$, given the preordering ω, is such that the interval $](y, z), (x, z)[$ is empty if ω is ultrametric. Considering such a triplet, the preorder ω is less and less ultrametric as the cardinality of $](y, z), (x, z)[$ become bigger. To take into account the set J of all the triplets, we may adopt as a measure of the discrepancy between ω and an ultrametric preorder:

$$H(\omega) = \frac{1}{|J|} \sum_{J} \frac{\left|]M(x, y, z), S(x, y, z)[\right|}{|F|} \tag{28}$$

where we have normalized with number of the triples $|J|$ and with the number of the pairs $|F|$.

Example

Let $E = \{a, b, c, d, e\}$, and ω the preorder on E

$$\{a, d\} = \{a, c\} < \{a, e\} < \{c, e\} < \{b, d\} = \{c, d\} < \{b, c\} \\ < \{d, e\} < \{a, b\} < \{b, e\}$$

J \ F	(a,d)	(a,c)	(a,e)	(c,e)	(b,d)	(c,d)	(b,c)	(d,e)	(a,b)	(b,e)
(a,b,c)		●					●	×	●	
(a,b,d)	●				●		×	×	●	
(a,b,e)			●						●	●
(a,c,d)	●	●	×	×		●				
(a,c,e)		●	●	●						
(a,d,e)	●		●	×	×	×	×	●		
(b,c,d)					●	●	●			
(b,c,e)				●			●	×	×	●
(b,d,e)					●			●	×	●
(c,d,e)				●		●	×	●		

Example

In the table we reported on the rows the set of the triples J and on the columns the set of the pairs F. On each row, a "$\bullet$" indicates the pairs contained in the triple and a "$\times$" the pairs which are strictly between the median and the maximum. If there are no crosses the median and the maximum are in the same class and the preordering is ultrametric, hence $H(\omega) = 0$. Summing the number of crosses for the pairs which are strictly included between some median and maximum one obtains a quantitative measure of the deviation of (E, ω) from the ultrametric preordering.

$H(\omega)$ is a more reliable measure of the deviation from ultrametricity than the the distortion index of the subdominant metric introduced before because the subdominant metric can be very different from the metric d.

Comments

The structure of $H(\omega)$ suggests to define a measure on the space F introducing the number of pairs in F which are strictly included in the interval $]M(x,y,z), S(x,y,z)[$. Given any pair $p \in F$, we define the subset J_p of J such that, for any triple $\{x,y,z\} \in J_p$, p is strictly included in the interval $]M(x,y,z), S(x,y,z)[$. It is possible then to define a measure m_p on the space of pairs F such that for any $p \in F$

$$m_p = \frac{|J_p|}{|J|}.$$

For any preorder ω we can then define the vector $D(\omega)$ as the set of $m_p, p \in F$. If the preorder is ultrametric this vector has all the components equal to 0. Thus the number of components of $D(\omega)$ which are different from zero and also the values of these components are a measure of the deviation from ultrametricity of the preorder ω.

Comments

In this example we have that

$$H(\omega) = \frac{13}{10 \times 10}$$

$$D(\omega) = (0.3, 0.3, 0.2, 0.2, 0.1, 0.1, 0.1, 0, 0, \ldots)$$

For large n, the number of elements of E, and for a large sample Q of sets of triples J obtained by generating the triples with uniform probability. We have that the $H(\omega)$ has a gaussian distribution since is the sum of indipendent uniformly distributed random variables:

$$H'(\omega) = \frac{1}{|Q|} \sum_{\{x,y,z\} \in Q} \Lambda(x, y, z) \tag{29}$$

where $\Lambda(x, y, z)$ is the cardinality of $]M(x, y, z), S(x, y, z)[$.

A neuronal example

Consider a set of 6 patterns, organized in a hierarchy as described in the Hopfield model section, with 2 ancestors and 2 descendants for each ancestor. Then we can calculate the quantities so far introduced.

```
Minimum-spanning-tree

2 -> 0 : 12
0 -> 1 : 14
1 -> 5 : 47
5 -> 3 : 17
3 -> 4 : 14

Deviation from ultrametricity (Rammal index): -0.00788955
```

A neuronal example

```
Preorder \omega

(0,1): 14
(0,2): 12
(0,3): 47
(0,4): 47
(0,5): 47
(1,2): 14
(1,3): 47
(1,4): 47
(1,5): 47
(2,3): 47
(2,4): 47
(2,5): 47
(3,4): 14
(4,5): 17

--> (0,2)<(0,1)=(1,2)=(3,4)<(4,5)=(0,3)=(0,4)=(0,5)=(1,3)=(1,4)=(1,5)=(2,3)
     =(2,4)=(2,5)=(3,5)
```

J \ F	(0,2)	(0,1)	(1,2)	(3,4)	(4,5)	(0,3)	(0,4)	(0,5)	(1,3)	(1,4)	(1,5)	(2,3)	(2,4)	(2,5)	(3,5)
(0,1,2)	•	•	•												
(0,1,3)		•				•	–	–	•						
(0,1,4)		•					•	–	–	•					
(0,1,5)		•						•	–	–	•				
(0,2,3)	•					•	–	–	–	–	–	•			
(0,2,4)	•						•	–	–	–	–	–	•		
(0,2,5)	•							•	–	–	–	–	–	•	
(0,3,4)				•		•	•								
(0,3,5)						•		•	–	–	–	–	–	–	•
(0,4,5)					•		•	•							
(1,2,3)			•						•	–	–	•			
(1,2,4)			•							•	–	–	•		
(1,2,5)			•								•	–	–	•	
(1,3,4)				•					•	•					
(1,3,5)									•		•	–	–	–	•
(1,4,5)					•					•	•				
(2,3,4)				•								•	•		
(2,3,5)												•		•	•
(2,4,5)					•								•	•	
(3,4,5)				•	•	–	–	–	–	–	–	–	–	–	•

```
Deviation from ultrametricity (Lerman's H(\omega)): 0
```

Contents

The Hodgkin-Huxley model

neuron	ion	inside concentration (mmole)	outside concentration (mmole)
squid axon	K^+	410	10
squid axon	N_a^+	49	460
squid axon	Cl^{--}	40	540
cat spinal neuron	K^+	150	5.5
cat spinal neuron	N_a^+	15	150
cat spinal neuron	Cl^{--}	9	125

$$V = \frac{RT}{F} \log \left[\frac{P_K[K]_o + P_{Na}[Na]_o + P_{Cl}[Cl]_i}{P_K[K]_i + P_{Na}[Na]_i + P_{Cl}[Cl]_o} \right] \tag{30}$$

$$C\frac{dV}{dt} = g_{Na}(V - V_a) + g_K(V - V_K) + g_{Cl}(V - V_{Cl}) \tag{31}$$

$$g_{Na} = \overline{g_{Na}} m^3 h \tag{32}$$

$$g_K = \overline{g_K} n^4 \tag{33}$$

The Hodgkin-Huxley model

$$\frac{dm}{dt} = \alpha_m(1-m) - \beta_m m$$
$$\frac{dh}{dt} = \alpha_h(1-h) - \beta_h h$$
$$\frac{dn}{dt} = \alpha_n(1-n) - \beta_n n$$

$$m(t) = m_\infty + (m_0 - m_\infty)e^{-t/\tau_m}$$
$$h(t) = h_\infty + (h_0 - h_\infty)e^{-t/\tau_h}$$
$$n(t) = n_\infty + (n_0 - n_\infty)e^{-t/\tau_n}$$

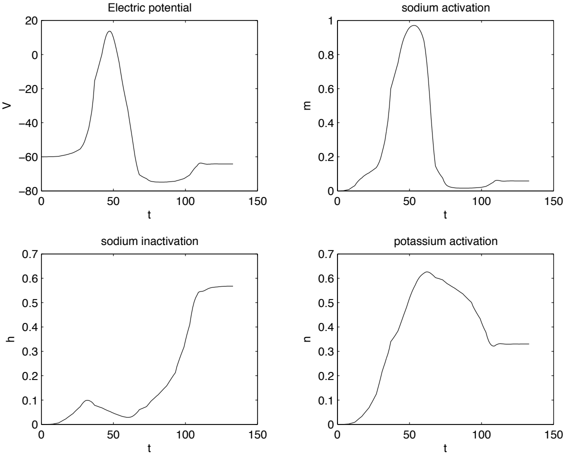
$$m_\infty = \alpha_m/(\alpha_m + \beta_m)$$
$$h_\infty = \alpha_h/(\alpha_h + \beta_h)$$
$$n_\infty = \alpha_n/(\alpha_n + \beta_n)$$
$$\tau_m = 1/(\alpha_m + \beta_m)$$
$$\tau_h = 1/(\alpha_h + \beta_h)$$
$$\tau_n = 1/(\alpha_n + \beta_n)$$

The Hodgkin-Huxley model

$$\alpha_m(V) = \frac{25 - V}{10[e^{(25-V)/10} - 1]}$$
$$\beta_m(V) = 4e^{-V/18}$$
$$\alpha_h(V) = \frac{7}{100}e^{-V/20}$$
$$\beta_h(V) = \frac{1}{e^{(30-V)/10} - 1}$$
$$\alpha_n(V) = \frac{10 - V}{100[e^{(10-V)/10} - 1]}$$
$$\beta_n(V) = \frac{1}{8}e^{-V/80}$$

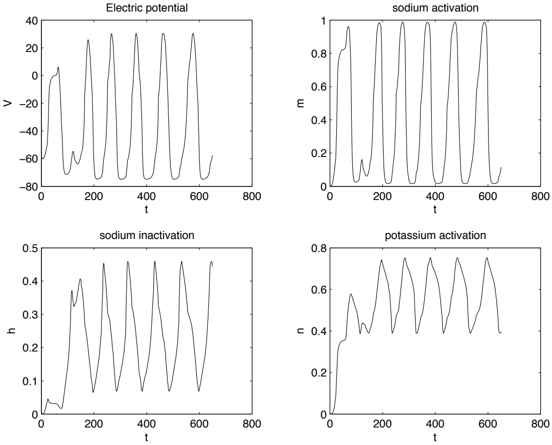
$$C\frac{dV}{dt} = \overline{g_{Na}}m^3h(V - V_a) + \overline{g_K}n^4(V - V_K) + g_{Cl}(V - V_{Cl}) + I$$

Potential and activation parameters for $I = 1$



No spiking

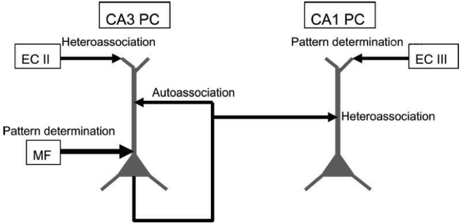
Potential and activation parameters for $I = 10$



Spiking

Contents

Associative memory in the hippocampus



Mossy fibre (MF) inputs from the dentate gyrus create pyramidal cell (PC) activity in CA3 that is stored autoassociatively by Hebbian modification of recurrent collateral synapses between coactive PCs. Patterns of activity in layer II of entorhinal cortex (EC II) may be heteroassociated with these CA3 patterns. At the same time, CA1 PCs receiving input both from layer III of entorhinal cortex and from CA3 PCs form a heterassociation with the active CA3 PCs through Hebbian modification of the Schaffer collateral synapses.

Hebbian pattern storage and recall

The elements of this matrix select the conductance strengths of AMPA synapses between CA1 and CA3 of the neurons which are active in the patterns.

Hebbian pattern storage and recall

A

$$\begin{bmatrix}1\\0\\0\\0\\1\\0\\0\\1\\0\\0\end{bmatrix} * \begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\end{bmatrix} = \begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix}$$

$$\begin{bmatrix}0\\0\\1\\1\\1\\0\\0\\0\\0\\0\end{bmatrix} * \begin{bmatrix}0\,0\,1\,1\,1\,0\,0\,0\,0\,0\end{bmatrix} = \begin{bmatrix}0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix}$$

$$\begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix} + \begin{bmatrix}0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix} = \begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\1\,0\,1\,1\,2\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix}$$

If $x(i,j) > 0$ then let $x(i,j) = 1$ $\longrightarrow$

$$\begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\1\,0\,1\,1\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix}$$

B

Pattern 1

$$\begin{bmatrix}1\\0\\0\\0\\1\\0\\0\\1\\0\\0\end{bmatrix} * \begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\0\,0\,1\,1\,1\,0\,0\,0\,0\,0\\1\,0\,1\,1\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\1\,0\,0\,0\,1\,0\,0\,1\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\\0\,0\,0\,0\,0\,0\,0\,0\,0\,0\end{bmatrix}$$

$$\begin{bmatrix}3\,0\,1\,1\,3\,0\,0\,3\,0\,0\end{bmatrix}$$

Treshold = 3

$$\begin{bmatrix}1\,0\,0\,0\,1\,0\,0\,1\,0\,0\end{bmatrix}$$

Hebbian pattern storage and recall

The elements of this matrix select the conductance strengths of AMPA synapses between CA1 and CA3 of the neurons which are active in the patterns.

A The individual weight matrices from the individual storage of the patterns are simply obtained as the outer product of the pattern with itself. The combined weight matrix is obtained by summing the individual matrices and then clipping entries to be 0 or 1.
B The cue pattern is multiplied with the weight matrix to give a vector of weighted input sums. This vector is thresholded to give the recalled binary vector. With the noiseless cue illustrated here, a suitable threshold is simply the number of active units in the cue pattern.

Hebbian pattern storage and recall

It is easily seen that the highest sums (3) are all to the cells that belong to the stored cue pattern. Some other cells get a lower input of 1, since the two stored patterns overlap with each other. Recall proceeds by applying an activity threshold, and in this case a threshold of 3 is appropriate. The final output activity vector is determined by making active (vector entry 1) all those cells whose input sum is greater than or equal to the threshold i), else the vector entry is 0. The new output vector after the threshold setting is applied is [1 0 0 0 1 0 0 1 0 0]

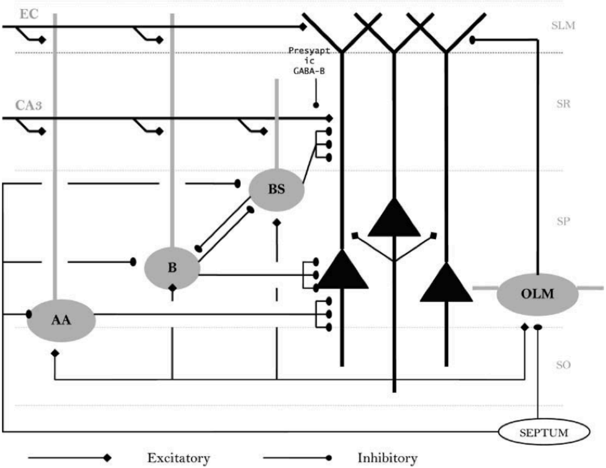
Heteroassociative memory in CA1

The principal excitatory cells of the CA1 region are pyramidal cells. These cells are driven by excitatory inputs from layer III of the entorhinal cortex and the CA3 Schaffer collaterals and an inhibitory input from the medial septum. Recurrent connectivity between pyramidal cells is negligible in CA1 (less than 1%).

Gamma cycles

Gamma frequency rhythms (30–100 Hz) are assumed to constitute a basic clock cycle such that patterns of activity for storage and recall correspond to PCs that are active in a particular gamma cycle.

Heteroassociative memory in CA1



Heteroassociative memory in CA1

Black filled triangles: pyramidal cells. *Grey filled circles*: CA1 inhibitory interneurons. EC: entorhinal cortex input; CA3: CA3 Schaffer collateral input; AA: axo-axonic cell; B: basket cell; BS: bistratified cell; OLM: oriens lacunosum-moleculare cell; SLM: stratum lacunosum-moleculare; SR: stratum radiatum; SP: stratum pyramidale; SO: stratum oriens. Open circles: Septal GABA inhibition. From Cutsuridis et al. (2010).

The CA1 network

The network contains 100 pyramidal cells (PC), 2 basket cells (BC), 1 bistratified cell (BSC), 1 axo-axonic cell (AAC) and 1 oriens lacunosum-moleculare (OLM) cell. All cell morphologies included a soma, apical and basal dendrites and a portion of axon (Poirazi et al., 2003a,b; Saraga et al., 2003; Santhakumar et al., 2005).

Synaptic properties

In the model, AMPA, NMDA, $GABA_A$ and $GABA_B$ synapses are considered. $GABA_A$ is present in all strata, whereas $GABA_B$ is present in medium and distal SR and SLM dendrites. AMPA synapses are present in strata LM (EC connections) and radiatum (CA3 connections), whereas NMDA are present only in stratum radiatum (CA3 connections).

The CA1 network

Model inputs

Inputs to the CA1 model come from the medial septum (MS), entorhinal cortex (EC) and CA3 Schaffer collaterals. The EC input is modelled as the firing of 20 entorhinal cortical cells at an average gamma frequency of 40 Hz (spike trains only modelled and not the explicit cells), and the CA3 input is modelled with the same gamma frequency spiking of 20 out of 100 CA3 pyramidal cells. PCs, BCs, AACs, BSCs received CA3 input in their medial SR dendrites, whereas PCs, BCs and AACs received also the EC layer III input in their apical LM dendrites. EC inputs preceded CA3 inputs by 9 ms on average (Soleng et al., 2003). MS input, which is modelled as the rhythmic firing of 10 septal cells, provides $GABA_A$ inhibition to all interneurons in the model (strongest to BC and AAC; Freund and Antal (1988)). MS input is phasic at theta rhythm and is on for 125 ms during the retrieval phase.

The CA1 network

Presynaptic $GABA_B$ inhibition

It has been shown that the strengths of the synaptic inputs from the EC perforant path and the CA3 Schaffer collaterals wax and wane according to the extracellular theta rhythm and 180° out of phase from each other (Wyble et al., 2000). These cyclical theta changes are likely due to the presynaptic $GABA_B$ inhibition to CA3 Schaffer collateral input to CA1 PCs' synapses, which is active during the storage cycle and inactive during recall (Molyneaux and Hasselmo, 2002). This is modelled simply as a reductive scaling during storage of the CA3-AMPA synaptic conductance, so that the effective conductance g' is

$$g' = g_s \cdot g \tag{34}$$

where g_s is the scaling factor (set to 0.4 in the presented simulations). During recall, g' is simply equal to g (the AMPA conductance determined by the connectivity weight matrix).

Pattern recall in CA1

Hasselmo et al. (2002a,b) have hypothesized that the hippocampal theta rhythm (4–7 Hz) contributes to memory formation by separating storage and recall into different functional subcycles. Recent experimental evidence has shown that different types of inhibitory interneurons fire at different phases of the theta rhythm (Klausberger et al., 2004; Somogyi and Klausberger, 2005; Klausberger and Somogyi, 2008). Here, we demonstrate how the recall performance of previously stored patterns is affected by the presence/absence of various types of inhibitory interneurons, which fire at different phases of the simulated theta rhythm (Paulsen and Moser, 1998).

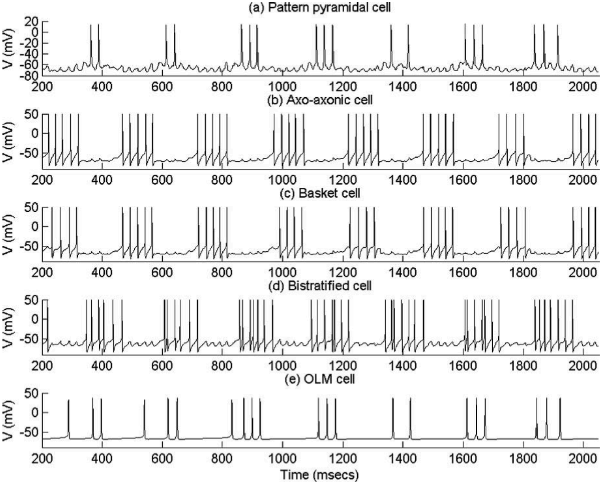
Pattern recall in CA1

As detailed previously, a set of patterns are stored by generating a weight matrix based on a clipped Hebbian learning rule, and using the weight matrix to prespecify the CA3 to CA1 PC connection weights. To test recall of a previously stored pattern, the associated input pattern is applied as a cue in the form of spiking of active CA3 inputs (those belonging to the pattern) distributed within a gamma frequency time window. The entire cue pattern is repeated at gamma frequency (40 Hz). At the same time, 20 EC inputs also fire randomly distributed within a 25 ms gamma window, but with mean activity preceding the CA3 activity by 9 ms. The CA3 spiking drives the CA1 PCs plus the B, AA and BS interneurons. The EC input also drives the B and AA interneurons.

Pattern recall in CA1

To test pure recall by the CA3 input cue, the EC input is disconnected from the CA1 PCs and no learning takes place at CA3 synapses on CA1 PCs. The CA3 synapses are suppressed during the “storage” phase of theta. Pattern recall only occurs during the “recall” half-cycle. Typical firing patterns of the different cell types across theta cycles are illustrated in figure.

Pattern recall in CA1

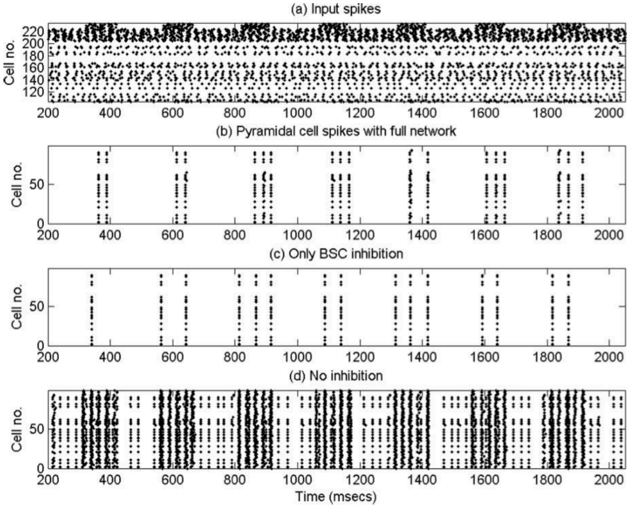


Pattern recall in CA1

The recall of the first pattern in a set of five is shown in the following figure.

The top subfigure shows a raster plot of the spiking of the septal (top 10 rows), EC (next 20 rows) and CA3 (bottom 100 rows) inputs. The remaining subplots show raster plots of CA1 PC activity for different configurations of network inhibition.

Pattern recall in CA1



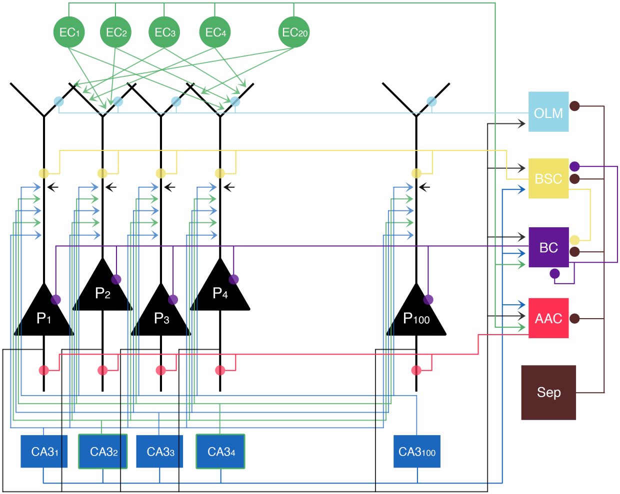
Contents

Network architecture

All simulations were performed using NEURON (Hines and Carnevale, 1997) running on a cluster of 8 nodes with MPI (Message Passing Interface).

The model consists of 100 pyramidal (P) cells, 2 basket (B) cells, 1 bistratified (BS) cell, 1 axo-axonic (AA) cell, and 1 oriens lacunosum moleculare (OLM) cell.

Network architecture

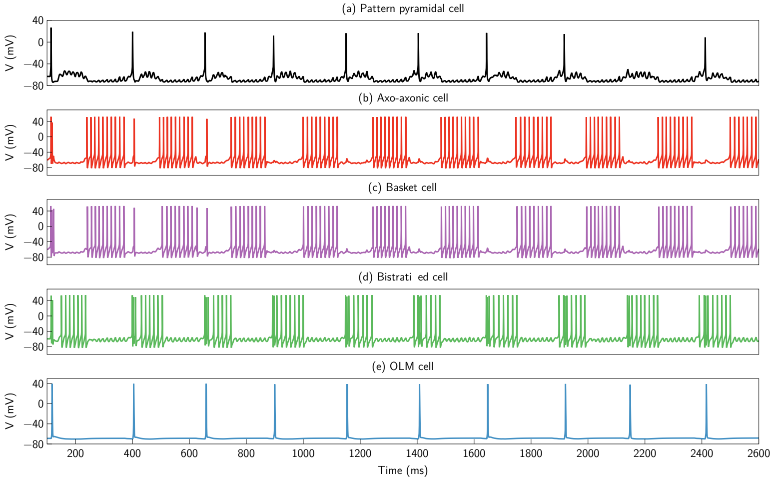


Network architecture

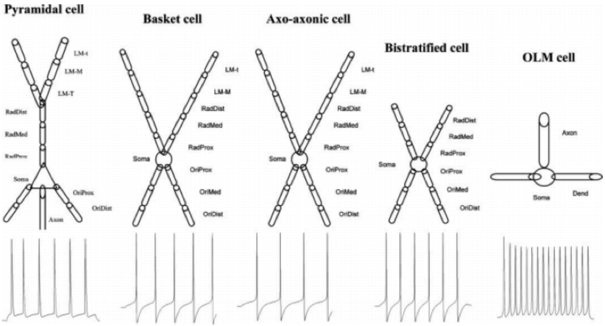
Arrows stands for excitatory synapses, *circles* for inhibition. $P_1 \ldots P_{100}$: pyramidal cells. $EC_1 \ldots EC_{20}$: entorhinal cortex input; $CA3_1 \ldots CA3_{100}$: CA3 Schaffer collateral input; OLM: oriens lacunosum-moleculare cell; BSC: bistratified cell; BC: basket cell; AAC: axo-axonic cell; Sep: Septal GABA inhibition.

Each kind of interneuron has a specific function in modulating not only the overall network functions, but also the I/O properties of the principal neurons (the CA1 pyramidal neurons) and, especially, the synaptic plasticity processes leading to memory storage. For the OLM, BC, BSC and AAC the models defined for the networks of Cutsuridis et al. (2010) are employed. The CA1 model employed, instead, has the same morphology of the CA1 template used by Cutsuridis et al. (2010) but different distributions of the ionic currents.

Voltage response of the different neurons of the network



Compartmental structure models for the different cell types



Synaptic plasticity

Recent results have elaborated on the timing dependence of LTP by showing that long-term plasticity depends critically on the millisecond timing of pre- and postsynaptic spikes. Typically, if the presynaptic cell fires an AP a few milliseconds before the postsynaptic cell, LTP is produced, whereas the opposite temporal order results in LTD, a notion called spike timing-dependent plasticity (STDP). Interestingly, the rules of STDP vary widely within brain region, cell, and synapse type.

During storage an STDP learning rule (based on the experimental findings by Nishiyama et al. (2000)) was applied at CA3-AMPA synapses on P cells medial SR dendrites, where presynaptic CA3 input spike times were compared with the postsynaptic voltage response to determine an instantaneous change in the peak synaptic conductance.

Synaptic plasticity

$$g_{\text{peak}}(t) = g^0_{\text{peak}} + A(t) \tag{35}$$

with

$$A(t) = \begin{cases} A(t-1)\left(1 - d\,\frac{e^{(\Delta t - M_d)^2/2V_d^2}}{V_d\sqrt{2\pi}}\right) & \text{if } \Delta t < 0 \\ A(t-1) + \left(g^{\max}_{\text{peak}} - g^0_{\text{peak}} - A(t-1)\right)\, p\, e^{-\Delta t/\tau_p} & \text{if } \Delta t > 0 \end{cases} \tag{36}$$

where $\Delta t = t_{\text{post}} - t_{\text{pre}}$. $M_d = -22\,\text{ms}$, $V_d = 5\,\text{ms}$, $\tau_p = 10\,\text{ms}$ are set in order to reproduce the critical time window found by Nishiyama et al. (2000), g^0_{peak} is the initial peak conductance, $g^{\max}_{\text{peak}}$ is the maximum value which g can reach. The parameters p and d are chosen in such way that, in the same time of the protocol of Nishiyama et al., i.e. 16 ms, the conductance peak of synapses under plasticity lies in a range near to maximum value.

Synaptic plasticity

In addition, during storage the CA3-AMPA synaptic conductance suppression by the putative $GABA_B$ inhibition present during this phase was implemented simply by scaling so that effective conductance g' was:

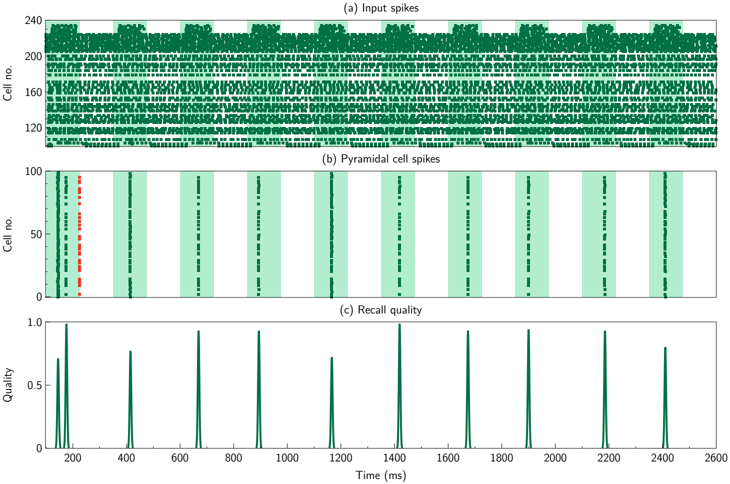
$$g' = g_s \times g \tag{37}$$

where g_s is the scaling factor (set to 0.4). During recall, g' was simply equal to g.

Preliminary results

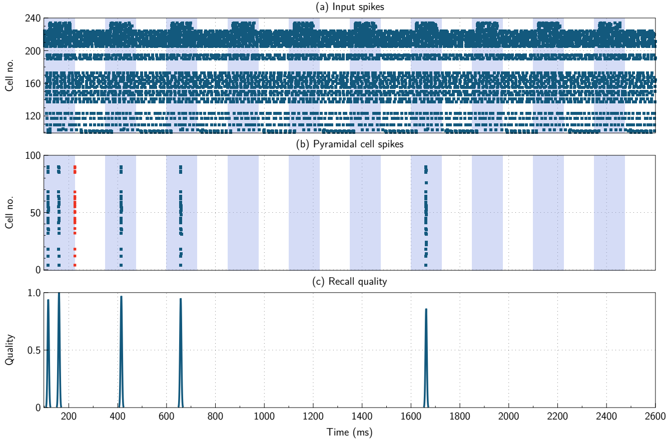
Here we show the recall of one of 6 **ultrametric** patterns (in the sense of the hierarchy introduced before) stored by the network. This is the case already described in the preorder example, thus the Rammal index of the deviation from ultrametricity is -0.0079. The green highlighted areas are the theta half-cycles in which recall occurs. The red pattern is the input pattern, shown for clarity.

Preliminary results



Preliminary results

Here instead we show the recall of one of 6 **random** patterns stored by the network, with a Rammal index of deviation from ultrametricity of 0.02.



References I

Amit, D., Gutfreund, H., and Sompolinsky, H. (1985). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.*, 55:1530–1533.

Cutsuridis, V., Cobb, S., and Graham, B. P. (2010). Encoding and retrieval in a model of the hippocampal ca1 microcircuit. *Hippocampus*, 20(3):423–446.

Freund, T. and Antal, M. (1988). Gaba-containing neurons in the septum control inhibitory interneurons in the hippocampus. *Nature*, 336(6195):170–173.

Hasselmo, M., Bodelón, C., and Wyble, B. (2002a). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural computation*, 14(4):793–817.

Hasselmo, M., Hay, J., Ilyn, M., and Gorchetchnikov, A. (2002b). Neuromodulation, theta rhythm and rat spatial navigation. *Neural Networks*, 15(4):689–707.

Hines, M. and Carnevale, N. (1997). The neuron simulation environment. *Neural computation*, 9(6):1179–1209.

Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558.

Klausberger, T., Márton, L., Baude, A., Roberts, J., Magill, P., and Somogyi, P. (2004). Spike timing of dendrite-targeting bistratified cells during hippocampal network oscillations in vivo. *Nature neuroscience*, 7(1):41–47.

Klausberger, T. and Somogyi, P. (2008). Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science*, 321(5885):53–57.

Lerman, I. C. (1981). *Classification et analyse ordinale des données*. Dunod, Paris.

References II

McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biology*, 5(4):115–133.

Molyneaux, B. and Hasselmo, M. (2002). Gaba$_B$ presynaptic inhibition has an in vivo time constant sufficiently rapid to allow modulation at theta frequency. *Journal of neurophysiology*, 87(3):1196–1205.

Murtagh, F. (1983). A survey of recent advances in hierarchical clustering algorithms. *The Computer Journal*, 26(4):354–359.

Nishiyama, M., Hong, K., Mikoshiba, K., Poo, M., and Kato, K. (2000). Calcium stores regulate the polarity and input specificity of synaptic modification. *Nature*, 408(6812):584–588.

Paulsen, O. and Moser, E. (1998). A model of hippocampal memory encoding and retrieval: Gabaergic control of synaptic plasticity. *Trends in neurosciences*, 21(7):273–278.

Poirazi, P., Brannon, T., and Mel, B. (2003a). Arithmetic of subthreshold synaptic summation in a model ca1 pyramidal cell. *Neuron*, 37(6):977–987.

Poirazi, P., Brannon, T., and Mel, B. (2003b). Pyramidal neuron as two-layer neural network. *Neuron*, 37(6):989–999.

Prim, R. C. (1957). Shortest connection networks and some generalizations. *Bell system technical journal*, 36(6):1389–1401.

Santhakumar, V., Aradi, I., and Soltesz, I. (2005). Role of mossy fiber sprouting and mossy cell loss in hyperexcitability: a network model of the dentate gyrus incorporating cell types and axonal topography. *Journal of neurophysiology*, 93(1):437–453.

References III

Saraga, F., Wu, C., Zhang, L., and Skinner, F. (2003). Active dendrites and spike propagation in multi-compartment models of oriens-lacunosum/moleculare hippocampal interneurons. *The Journal of physiology*, 552(3):673–689.

Soleng, A., Raastad, M., and Andersen, P. (2003). Conduction latency along ca3 hippocampal axons from rat. *Hippocampus*, 13(8):953–961.

Somogyi, P. and Klausberger, T. (2005). Defined types of cortical interneurone structure space and spike timing in the hippocampus. *J Physiol*, 562:9–26.

Wyble, B., Linster, C., and Hasselmo, M. (2000). Size of ca1-evoked synaptic potentials is related to theta rhythm phase in rat hippocampus. *Journal of Neurophysiology*, 83(4):2138–2144.