bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Continuous attractors for dynamic memories

Davide Spalla[1,*], Isabel M. Cornacchia[1,2], Alessandro Treves[1]

1 SISSA – Cognitive Neuroscience, Via Bonomea 265, 34136 Trieste, Italy
2 University of Turin – Physics Department, Via Giuria 1, 10125 Torino, Italy

*** dspalla@sissa.it**

Abstract

Episodic memory has a dynamic nature: when we recall past episodes, we retrieve not only their content, but also their temporal structure. The phenomenon of replay, in the hippocampus of mammals, offers a remarkable example of this temporal dynamics. However, most quantitative models of memory treat memories as static configurations, neglecting the temporal unfolding of the retrieval process. Here we introduce a continuous attractor network model with a memory-dependent asymmetric component in the synaptic connectivity, that spontaneously breaks the equilibrium of the memory configurations and produces dynamic retrieval. The detailed analysis of the model with analytical calculations and numerical simulations shows that it can robustly retrieve multiple dynamical memories, and that this feature is largely independent on the details of its implementation. By calculating the storage capacity we show that the dynamic component does not impair memory capacity, and can even enhance it in certain regimes.

Introduction

The temporal unfolding of an event is an essential component of episodic memory. When we recall past events, or we imagine future ones, we do not produce static images but temporally structured movies, a phenomenon that has been referred to as "mental time travel" [1], [2].

The study of the neural activity of the hippocampus, known to be heavily involved in episodic memory, has provided many insights on the neural basis of memory retrieval and its temporal dynamics. A particularly interesting example is the phenomenon of hippocampal replay, i.e. the reactivation, on a compressed time scale, of sequences of cells active in previous behavioural sessions. Replay takes place during sharp wave ripples, fast oscillations of the hippocampal local field potential that are particularly abundant during sleep and restful wakefulness [3], [4]. Indeed, replay has been observed during sleep [5], [6], inter-trial rest periods [7], [8], and during still periods in navigational tasks [9], [10]. Replay activity has been hypothesized to be crucial for memory consolidation [11] and retrieval [12], as well as for route planning [10], [13].

A temporally structured activation takes place also before the exposure to an environment [14], a phenomenon known as *preplay*, and a recent study showed that this dynamical feature emerges very early during development, preceding the appearance of theta rhythm [15]. The fact that sequential activity can be present before the exposure to the environment suggests that their dynamical nature is not specific to a

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

role in spatial cognition, but is inherent to hippocampal operation in general. Moreover, in a recent study Stella et al. [16] have shown that the retrieved positions during slow wave sleep are not always replaying experienced trajectories, but are compatible with a random walk on the low dimensional manifold that represents the previously explored environment. This suggests that what is essential are not the sequences themselves, but the tendency to produce them: neural activity tends to move, constrained to abstract low-dimensional manifolds, which can then be recycled to represent spatial environments, and possibly non-spatial ones as well.

However, the activity is not constrained to a single subspace: replay in sleep can reflect multiple environments [17], [18], the content of awake replay reflects both the current and previous environments [12], and during behaviour fast hippocampal sequences appear to switch between possible future trajectories [19]. Further evidence comes from a recent study with human participants learning novel word pair associations [20]. The study shows that the same, pair-dependent neural sequences are played during the encoding and the retrieval phase.

A similar phenomenology – a dynamic activity on low dimensional manifolds – is present in memory schemata, cognitive frameworks that constrain and organize our mental activity [21], and have been shown to have a representation in the medial temporal lobe [22]. When we retrieve a schema, our memory follows a spontaneous flow guided by the schema. Yet another examples of dynamical continuous memories is offered by motor programs, which have been described as low-dimensional, temporally structured neural trajectories [23], [24], [25]. Fig.1 schematically illustrates the concept of dynamical continuous attractors and their possible role in some cognitive processes.

Mechanistic models of memory usually neglect the dynamic component, treating memories as static objects, either discrete [26] or continuous [27], [28], [29], [30]. The production of sequences of discrete memories can be implemented with a heteroassociative component [31], usually dependent on the time integral of the instantaneous activity, that brings the network out of equilibrium and to the next step in the sequence. A similar effect can be obtained with an adaptation mechanism in a coarse grained model of cortical networks [32], with the difference that in this case the transitions are not imposed, but driven by the correlations between the memories in so-called latching dynamics [33], [34]. Moreover, adaptation-based mechanisms have been used to model the production of random sequences on continuous manifolds [35], and shown to be crucial in determining the balance between retrieval and prediction in a network describing CA3-CA1 interactions [36]. In the case of continuous attractor networks, movement can be induced also by mechanisms that integrate an external velocity input and make use of asymmetric synaptic strengths [37], [38], [39]. In the simplest instantiation, this system does not necessarily reflect long-term memory storage: the activity is constrained on a single attractive manifold, which could well be experience independent.

Here we propose a continuous attractor network model able to store and retrieve multiple independent manifolds in a dynamical way. The model relies on a map-dependent asymmetric component in the connectivity, that produces a robust shift of the activity on the retrieved manifold. This connectivity profile is thought to be the result of a learning phase in which the mechanism of spike timing dependent plasticity (STDP) [40] produces the asymmetry. Crucially, the asymmetry is not treated here as a "pathological" feature, assumed to level out in the limit of long learning, but as a defining trait of the stored memories. The balance between two components – one symmetric and trajectory-averaged, the other asymmetric and trajectory-dependent – is explicit in the formulation of the model, and allows to study their effects on memory storage.

In what follows we develop an analytical framework that allows to derive the dependence of important features of the dynamics, such as the shift speed and the asymmetry of the activity cluster, as a function of the relevant parameters of the model. We show

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

with numerical simulations that the behaviour of the model is robust with respect to the details of the model, depending weakly on the shape of the interactions. Finally, we estimate the storage capacity for dynamical memories and we find it to be of the same order of the capacity for static continuous attractors, and even higher in some regimes.

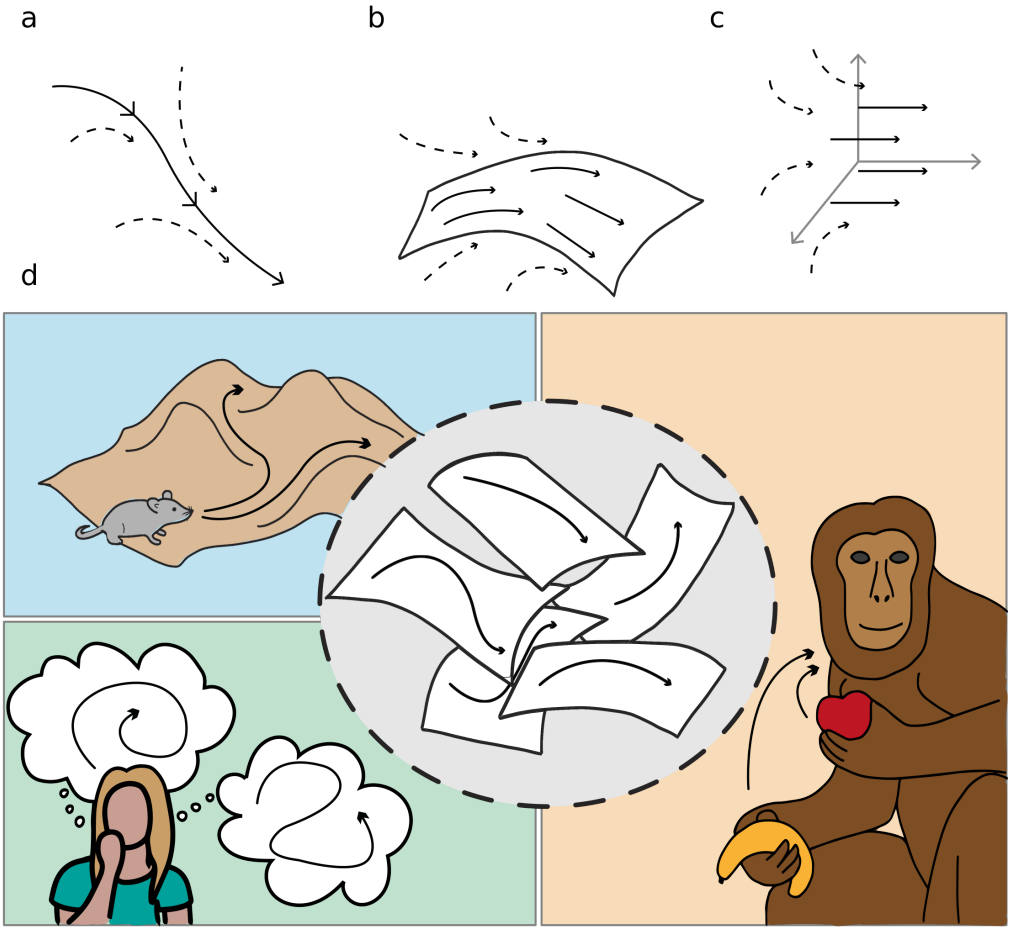


Figure 1: Schematic illustration of dynamic continuous attractors and their possible roles. Top row: a scheme of continuous attractive manifolds, with a dynamic component in 1D **(a)**, 2D **(b)** and 3D **(c)**. The neural activity quickly converges on the attractive manifold (dotted arrows), then slides along it (full arrows), producing a dynamics that is temporally structured and constrained to a low dimensional subspace. Bottom row: multiple dynamic memories could be useful for route planning (top left), involved in mind wandering activity (bottom left) or represent multiple learned motor programs (right).

A mechanistic model for dynamic retrieval

The model we consider is a continuous attractor neural network, with an additional anti-symmetric component in the connectivity strength. We consider a population of N neurons, with recurrent connectivity described by an interaction matrix J_{ij}, whose entries represent the strength of the interaction between neuron i and j. The activation

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

function of the neurons is threshold-linear, i.e. the output V_i of neuron i, given the input h_i, is

$$V_i = [h_i]^+ = g(h_i - h_0)\theta(h_i - h_0) \tag{1}$$

where θ is the Heaviside step function and the gain g and threshold h_0 are global parameters of the network. The variables V_i are positive and continuous, and thought to represent the firing rates of the units.

The dynamic evolution of the network is regulated by the equations:

$$\tau \frac{\partial V_i}{\partial t} + V_i = \left[\left(\sum_{j \neq i} J_{ij} V_j + b \left(\frac{1}{N} \sum_i V_i \right) - h_0 \right) \right]^+ \tag{2}$$

The first term on the right hand side represent the excitatory inputs provided to neuron i from the rest of the network through recurrent connections. $b(.)$ is a global inhibition term that, together with h_0 and g, regulates the average activity and the sparsity of the activity pattern [41]. For the purpose of this work, the global inhibition term b can be reabsorbed in h_0, and will no longer be explicitly written.

A recurrent network of this kind can encode continuous maps in its connectivity matrix. In a basic model expressing static continuous attractors, each neuron is assigned a preferential firing location x_i in the stimulus space, and the strength of the interaction between pairs of neuron is a decreasing, symmetric function of the distance between their preferred firing locations

$$J_{ij} \sim K_S(x_i - x_j). \tag{3}$$

This formula is assumed to come from a time-averaged Hebbian plasticity rule: neurons with nearby firing fields will fire concurrently and strengthen their connections, while firing fields far apart will produce weak interactions. The symmetry of the function K, usually called **interaction kernel**, ensures that the network reaches a static equilibrium, where the activity of the neurons represents a certain position in the map and, if not pushed, remains still.

The shift mechanism

The assumption of symmetric interactions neglects any temporal structure in the learning phase. In case of learning a spatial map, for example, the order in which recruited neurons fire along a trajectory may produce an asymmetry in the interactions as a consequence of the phenomenon called Spike Timing Dependent Plasticity [40], that requires the postsynaptic neuron to fire *after* the presynaptic one in order to strengthen the synapse. This phenomenon can be accounted for in the definition of the interaction kernel. Any asymmetric kernel can be decomposed in two contributions:

$$K(x_i - x_j) = K_S(x_i - x_j) + \gamma K_A(x_i - x_j) \tag{4}$$

where K_S is the usual symmetric component and K_A is an anti-symmetric function ($K_A(x_i - x_j) = -K_A(x_j - x_i)$). The parameter γ regulates the relative strength of the two components. K_A generates a flow of activity along the direction of asymmetry: neuron i activates neuron j that, instead of reciprocating, will activate neurons downstream in the asymmetric direction. Mechanisms of this kind have been shown to produce a shift of the activity bump, without its disruption [37], [39], [38]. This effect is illustrated in Fig.2, and its quantitative properties are analyzed in detail in the next section.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Storing multiple dynamic memories

A network with the connectivity structure described in Eq. 4 has a single dynamical attractor. In order to model the autoassociative memory properties of the hippocampus, we want the system to be able to store and retrieve multiple manifolds, each with its own temporal structure. We construct an interaction matrix J_{ij} that is the sum of the contributions from p different, independent memories:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} K(x_i^\mu - x_j^\mu) \tag{5}$$

Here each x_i^μ represent the preferred firing location of neuron i in the manifold representing memory μ, and K, given by eq. 4, contains a symmetric and anti-symmetric component for each memory. The dynamic of this network, for low memory loads $\alpha = p/N$, evolves in two phases: a fast convergence to the retrieved manifold, and then a rigid movement along it, that replicates its temporal structure. The same activity, if projected on the other, unretrieved manifolds, appears as random noise. When the memory load α is increased above a certain value α_c, a phase transition occurs, and the network is not able to retrieve any memory, falling instead into a disordered state.

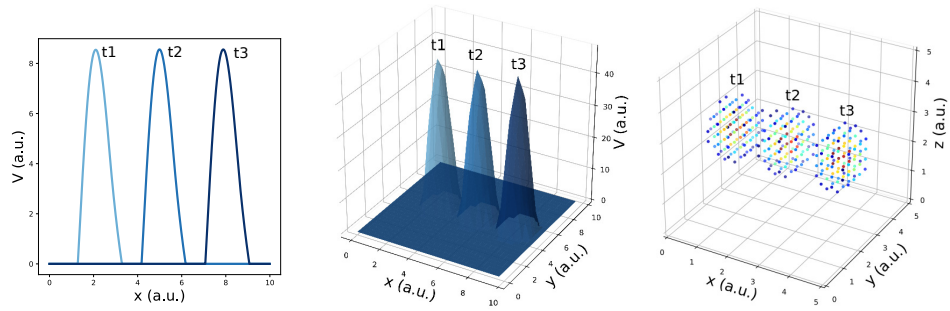
The value of α_c, the storage capacity of the network, is estimated in section and shown to be large enough to allow a network with biologically plausible connectivity density to store hundreds of different dynamical memories. Its dependence on the asymmetry parameter γ shows a non-trivial behaviour that depends crucially on the density of the connectivity of the network.

Before considering the multiple maps case, in the next section we present a quantitative study of the dynamics of the network in the case of a single map.

Dynamic retrieval

The presence of an asymmetry in the connection strengths prevents the system from reaching a stationary equilibrium. Instead, it generates a steady flow of activity in the direction of the asymmetry. This flow is illustrated in Fig. 2, obtained with numerical simulation of a network encoding a one-, two- or three-dimensional map, respectively. Note that the bump of activity translates without disruption, producing a steady flow in the asymmetric direction.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



(a) (b) (c)

Figure 2: Dynamic retrieval in different dimensions. Three snapshots of the network activity at three different times (t1, t2 and t3) are shown for a system encoding a one dimensional **(a)**, two dimensional **(b)** and three dimensional manifold **(c)**. In (c), activity is color-coded (blue represents low activity, red high activity, silent neurons are not plotted for better readability). In all cases the anti-symmetric component is oriented along the x axis.

The dynamic behaviour of the system and its features can be described analytically with a generalization of the framework developed by Battaglia & Treves [27]. For this purpose, it is easier to formulate the problem in the continuous limit, and describe the population activity by its profile $V(x)$ on the attractive manifold parametrized by the coordinate x, and the dynamical evolution as a discrete step map, equivalent to Eq. 2.

$$V(x, t+1) = g[h(x,t)]^+ \tag{6}$$

$$h(x,t) = \int_{-\infty}^{\infty} dx' K(x-x') V(x',t) - h_0 \tag{7}$$

The requirement of a rigid shift of population activity is then imposed by setting the activity at time $t+1$ to be equal at the activity at time t, but translated by an amount Δx, proportional to the speed of the shift. In this way we find the equation:

$$V(x+\Delta x) = g\left[\int_{-\infty}^{\infty} dx' K(x-x') V(x') - h_0\right]^+ \tag{8}$$

that we can rewrite as:

$$V(x+\Delta x) = \begin{cases} g\int_{\Omega} dx' K(x-x') V(x') - h_0, & \text{if } x \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where Ω is a compact domain for which there exist a solution of Eq. 8 that is zero at the boundary. This allows to exploit the fact that our threshold-linear system is, indeed, linear in the region in which $V(x) > 0$.

Equation 9 is valid in general, but we will focus here, to derive an analytical solution, on the one dimensional case and on an exponential kernel in the form

$$K(x-x') = e^{-|x-x'|} + \gamma \text{sign}(x-x') e^{-|x-x'|} \tag{10}$$

Differentiating twice Eq. 9, we obtain the differential equation

$$V''(x+\Delta x) = V(x+\Delta x) + 2gV(x) + 2g\gamma V'(x) + g\theta \tag{11}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

This is a second order linear ODE, with constant coefficients. The presence of the shit term Δx inside the unknown function makes the equation non-trivial to solve. To solve it, we proceed in the following way: first, we look for a particular solution, that is easily found in the constant function

$$V_c = \frac{g\theta}{1-2g} \tag{12}$$

Then, we consider the associated homogeneous equation, and look for a solution in the form $V(x) = e^{kx}$, where k is a solution of the characteristic equation $C(k) = 0$, with

$$C(k) = k^2 \mathrm{e}^{k\Delta x} + 2g\gamma k + 2g - \mathrm{e}^{k\Delta x}. \tag{13}$$

This trascendental equation has to be solved graphically in the complex domain, as shown in Fig. 3.

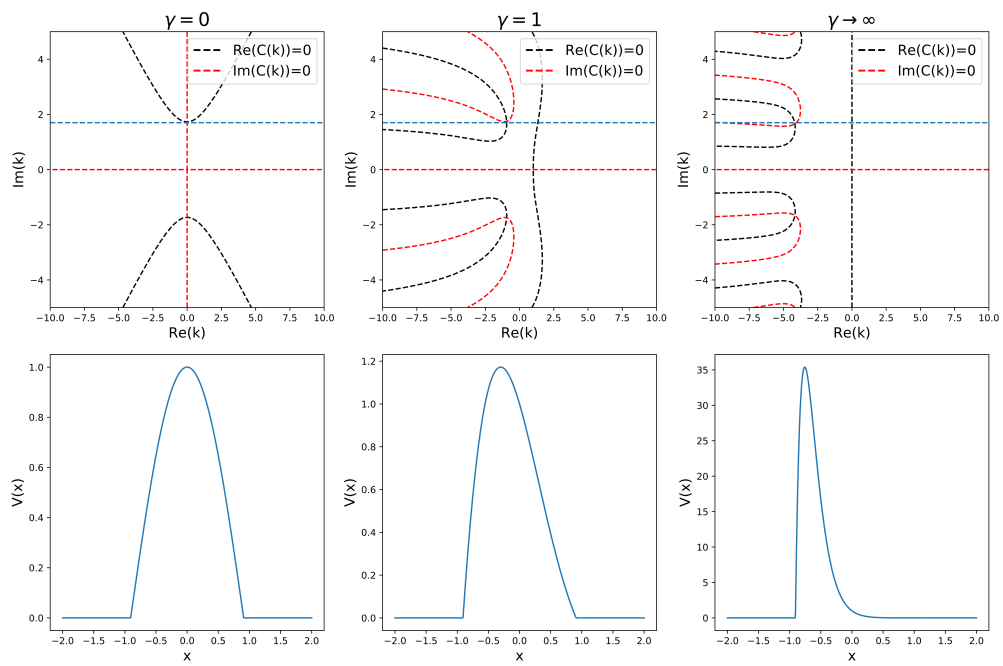


Figure 3: Analytic solution of equation (11).The top row shows the graphical procedure to find the complex zeros of the characteristic $C(k)$ given in (13), for three different values of γ. Black and red lines show the zeros of the real and imaginary part of $C(k)$, respectively. Their intersections are the complex solutions to $C(k) = 0$. The blue line represent the sparsity constraint $k_i = k_i^S$. The bottom row shows the corresponding solution shapes.

For each value of γ and Δx, the equation shows a pair of complex conjugate solutions

$$k^*_{1,2}(\gamma, \Delta x) = k_r(\gamma, \Delta x) \pm i k_i(\gamma, \Delta x) \tag{14}$$

The general solution of the equation will therefore have the form

$$V(x) = \begin{cases} Ce^{k_r x}\cos(k_i x) + \frac{g\theta}{1-2g} & \text{if } -R \le x \le R \\ 0 & \text{if } -R > x \text{ or } x > R \end{cases} \tag{15}$$

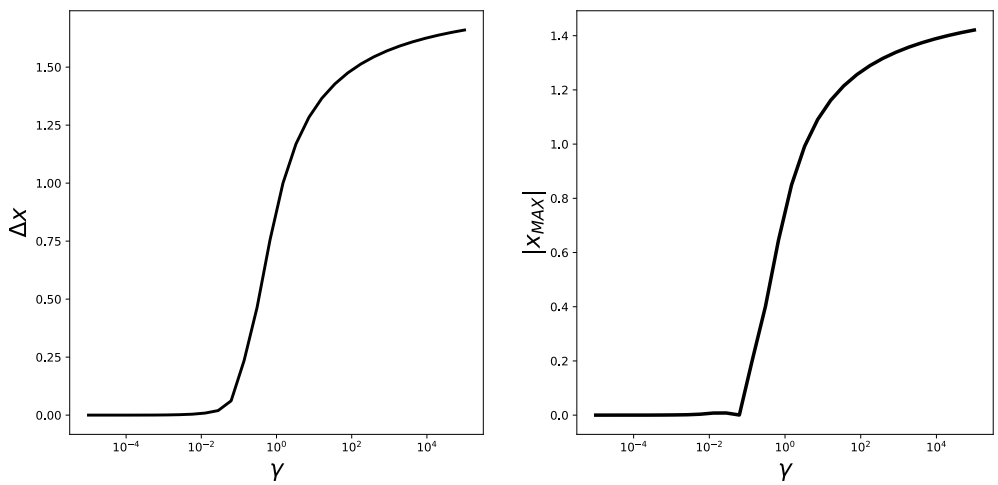
In the limit case $\gamma = 0$, $\Delta x = 0$ (Fig. 3, first column), the solutions are pure imaginary ($k_r = 0$), and we recover the solution of the symmetric case studied in [27]. From

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Eq.15 we can see that the absolute value of k_i is related to the width of the bump, and therefore to the sparsity of the solution, by the relation

$$R = \frac{\pi}{2k_i} \tag{16}$$

We focus here on the case in which R is kept constant when γ changes, i.e. the case in which the network is constrained to operate at a certain sparsity. This constraint is enforced by requiring that the zeros of 13 lie in the subspace $k_i = k_i^S$, i.e. we have the same sparsity of the solution k_i^S of the symmetric case. This imposes a relation between γ and both Δx (related to the speed of the shift) and k_r (related to the asymmetry of the shape of the solution). These relationships are shown in Fig. 4. The similarity between this two relationships can be understood intuitively by thinking that, for a fixed kernel shape, the larger the asymmetry of the solution, the more the bump with be translated by the evolution Eq. 8.

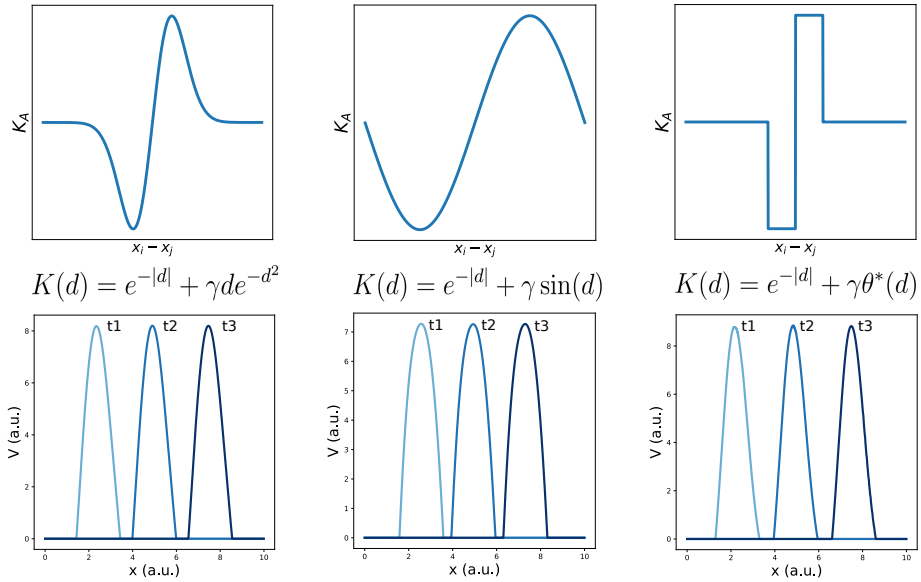


(a) (b)

Figure 4: Dependence of speed and shape on γ. **(a)** $\Delta x = -v\tau$ as a function of gamma. **(b)** Deviation of the position of the maximum of the bump from zero. This quantity is related to the value of the real part k_r of the solution of the characteristic equation by the relation $x_{max} = \arctan(k_r)$, and is one way to quantify the asymmetry in the bump shape produced by increasing γ. Note that the scale is logarithmic in γ.

The analytical results are presented here for a specific choice of the kernel, but the qualitative behaviour of the model is extremely general. In fact, numerical simulations show that a shifting bump can be obtained with a wide variety of interaction kernels, without any relationship required, for example, between the symmetric and anti-symmetric components. Some examples are illustrated in Fig. 5.

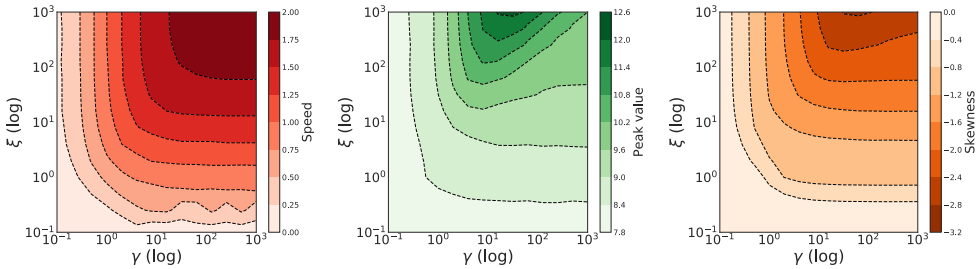
bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



(a) (b) (c)

Figure 5: Different interaction kernels produce similar behaviour. Three examples of dynamics with the same symmetric component and three different anti-symmetric components. Top row: shape of the anti-symmetric component K_A. Bottom row: three snapshots of the retrieval dynamics for the corresponding K_A. **(a)** Gaussian derivative; **(b)** Sinusoidal; **(c)** Anti-symmetric step function, $\theta^* = \theta(d)\theta(1-d) - \theta(-d)\theta(d-1)$.

Despite the robustness of the general features of the behaviour, the shape of the interaction kernel affects the details of the dynamics. Two parameters are particularly important: the relative strength γ between the symmetric and anti-symmetric components, and the characteristic length ξ of the anti-symmetric component. Their effect on the dynamics are shown in Fig. 6.



(a) (b) (c)

Figure 6: Effect of the kernel strength γ and its width ξ, in the case of the exponential kernel $K(d) = e^{-|d|} + \gamma sign(d) e^{-|d|/\xi}$.**(a)** Shift speed **(e)** Maximum value of the activity bump and **(e)** Skewness of the activity bump.

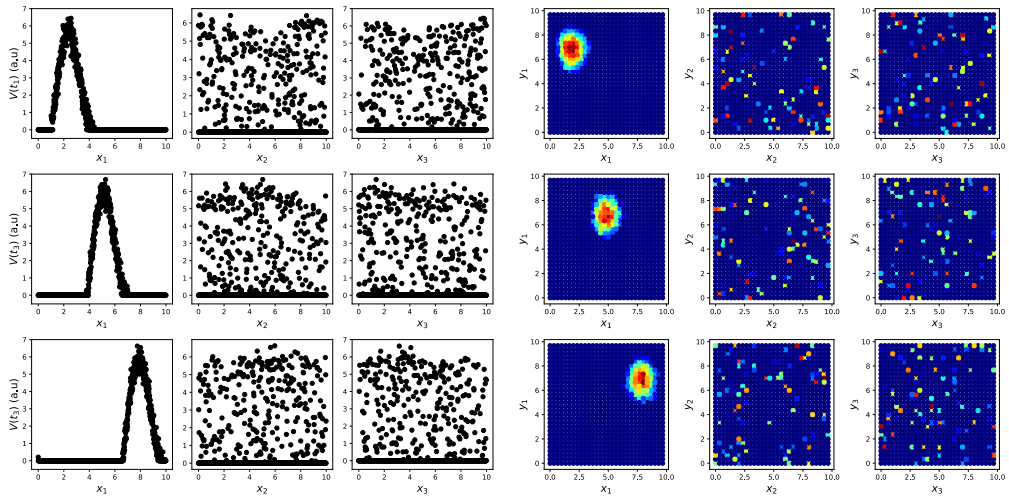
Taken together, these results show that the model can implement a dynamic memory system, whose dynamics is constrained to approach a memorized attractive manifold and to move along it at constant speed without disruption. This behaviour depends weakly on the details of the connectivity kernel and can be implemented with a rather general

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

type of asymmetric connections. However, in order to effectively work as a memory, the model has to be able to store and retrieve multiple manifolds. The estimation of the storage capacity is addressed in the next section.

Storage capacity for dynamic continuous attractors

A network with the connectivity profile described in section is able to store and retrieve multiple dynamic maps. The retrieval process, as in the single map case, unfolds in two phases: a fast transient in which the dynamics converges to one of the stored manifolds, and a subsequent stable shift. The second phase is illustrated in Fig. 7, obtained with numerical simulations.



(a) (b)

Figure 7: Dynamic retrieval in the presence of multiple memories in one dimension **(a)** and two dimensions **(b)**. Each row represents a snapshot of the dynamics at a point in time. The activity is projected on each of the three attractors stored in the network. In both cases, the first attractor is retrieved, and the activity organizes in a coherent bump that shifts in time. The same activity, projected onto the two non-retrieved maps looks like incoherent noise ((a) and (b), second and third columns).

The number of maps that can be stored and retrieved in this way is typically proportional to the number of inputs per neuron. Its magnitude, the storage capacity of the system, is crucial to determine if it can effectively operate as a memory.

To estimate the storage capacity for dynamic continuous attractors, we proceed along two complementary paths. For a fully connected network, where the analytical tools developed for equilibrium systems are not applicable, we take advantage of the fact that numerical simulations can be effective for the estimation of the capacity, since the number of connections per neuron C (the relevant parameters in the definition of the storage capacity $\alpha_c = p/C$) scales as the number of neurons. For a highly diluted system, on the other hand, the number of neurons is much larger than C, making the simulation of the system very difficult in practice. We then resort to an analytical formulation based on a signal to noise analysis [27], that exploits the vanishing correlation between inputs of different neurons in a highly diluted network, and does not require symmetry in the connectivity [42]. The quantification of the effect of loops in the dense connectivity

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

regime, developed in [43] and [44] for the case of static, discrete attractors, is beyond the scope of the present work and remains an interesting open direction.

In both the fully connected and the highly diluted case we study the dependence of the capacity on two important parameters: the map sparsity, i.e. the ratio between the width of the connectivity kernel (fixed to one without loss of generality) and the size L of the stored manifolds, and the asymmetry strength γ. We present the analytic solution for the diluted case in the next section, showing that a simple approximation yields a remarkably accurate estimation of the capacity and allows to decouple the effects of map sparsity and asymmetry. The capacity is found to be monotonically but gently decreasing with both sparsity and asymmetry.

In section we present the numerical results for the fully connected network. In this case, the presence of asymmetry can enhance the storage capacity, that is found to be maximal for finite values of γ and of map sparsity.

Analytical calculation of α_c in the highly diluted limit

We consider here the highly diluted limit, the case in which the number of connections per neuron C is much smaller than the total number of neurons N ($C/N \to 0$), and a number of maps $\alpha = p/C$ is stored. This scaling makes the system quite hard to simulate, but allows us to exploit its tree-like structure and the vanishing correlations between inputs to different neurons, and to study the network with a signal-to-noise approach that does not require its connectivity to be symmetric.

This approach, illustrated in details in [27], involves writing the local field h_i as the sum of two contributions: a signal term, due to the retrieved – "condensed" – map, and a noise term consisting of the sum of the contributions of the other, "uncondensed" maps. In the diluted regime these contributions are independent and can be summarized by a Gaussian term ρz, where z is a random variable with zero mean and unit variance. In the continuous limit, assuming that map $\mu = 1$ is retrieved we can write:

$$h(x^1) = g \int_L dx^{1\prime} K(x^1 - x^{1\prime}) V(x^{1\prime}) + \rho z \tag{17}$$

The noise will have variance:

$$\rho^2 = \alpha y L^2 \langle\langle K^2(x - x')\rangle\rangle \tag{18}$$

Where L is the size of the map, $\langle\langle K^2(x - x')\rangle\rangle$ is the spatial variance of the kernel and

$$y = \frac{1}{N} \sum_i V_i^2 \tag{19}$$

is the average square activity.

We can write the fixed point equation for the average activity profile $m^1(x)$, incorporating the dynamic shift with an argument similar to the one made for the single map case:

$$m^1(x + \Delta x) = g \int^+ Dz (h(x) - h_0) \tag{20}$$

Where $Dz = (e^{-z^2/2}/\sqrt{2\pi})dz$ and $\int^+ f(x)dx = \int f(x)\theta(x)dx$. The average square activity y, entering the noise term, reads

$$y = \frac{g^2}{L} \int dx \int^+ Dz (h(x) - h_0)^2 \tag{21}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introducing the rescaled variables

$$w = \frac{-h_0}{\rho} \tag{22}$$

$$v(x) = \frac{m^1(x)}{\rho} \tag{23}$$

And the functions

$$\mathcal{N}(x) = x\Phi(x) + \sigma(x) \tag{24}$$

$$\mathcal{M}(x) = (1+x^2)\Phi(x) + x\sigma(x) \tag{25}$$

where $\Phi(x)$ and $\sigma(x)$ are the Gaussian cumulative and the Gaussian probability mass function respectively, we can rewrite the fixed-point equation as

$$v(x+\Delta x) = g\mathcal{N}\left(\int dx' K(x-x')v(x') + w\right) \tag{26}$$

$$y = \rho^2 g^2 \int \frac{dx}{L}\mathcal{M}\left(\int dx' K(x-x')v(x') + w\right) \tag{27}$$

Substituting Eq.27 in the expression for the noise variance 18 we obtain

$$\frac{1}{\alpha} = g^2 L \left\langle\left\langle K^2\right\rangle\right\rangle \int dx \mathcal{M}\left(\int dx' K(x-x')v(x') + w\right) \tag{28}$$

If we are able to solve Eq. 26 for the rescaled activity profile $v(x)$, we can use Eq. 28 to calculate α. We can then maximize α with respect to g and w: this yields the maximal value α_c for which retrieval solutions can be found.

These equations are valid in general, but here we focus on the one dimensional case and the exponential kernel of Eq. 10. In this case we have

$$\langle\langle K^2(x-x')\rangle\rangle = (1+\gamma^2)\langle\langle K_S^2(x-x')\rangle\rangle. \tag{29}$$

where $K_S(x-x') = e^{-|x-x'|}$ is the symmetric component of the kernel. Eq. 26 can be transformed it into a non-linear, delayed differential equation, that we can solve numerically. This solution procedure is illustrated in appendix B. Plugging the obtained form of $v(x)$ into Eq. 28, we can calculate the capacity. The dependence of the capacity on γ is shown, for $L = 60$, in Fig. 8.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

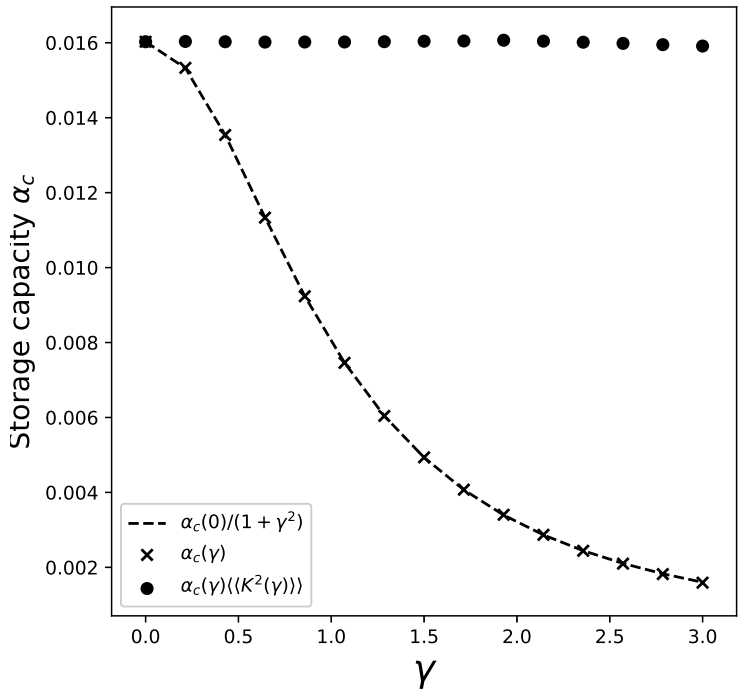


Figure 8: Dependence of the storage capacity on γ, for $L = 60$. The crosses show the full solution of Eq. 26 and 28. The dashed line is obtained by taking the value of the capacity $\alpha(0)$ obtained with full solution at $\gamma = 0$, and multiplying it by the scaling of the kernel variance $(1+\gamma^2)$. Full dots show the value of capacity obtained with the full solution and the contribution of the kernel variance factored out.

We can see from the full dots in the figure that the contribution of the integral in Eq.28 is remarkably constant in γ. This is due to the fact that the distortions of the bump shape induced by the presence of the asymmetry have a negligible effect on the average square activity y, whose value is dominated by the dependence on γ of the spatial variance of the kernel (Eq.18).

This allows us to approximate the value of the integral in Eq. 28 with its value in the $\gamma = 0$ case. We can then calculate the capacity as a function of γ and L by solving the symmetric case for different Ls, and then incorporating the dependence on γ given by the kernel variance:

$$\alpha_c(L,\gamma) \sim \alpha_c(L,0)/(1+\gamma^2) \tag{30}$$

The result is shown in Fig 9.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

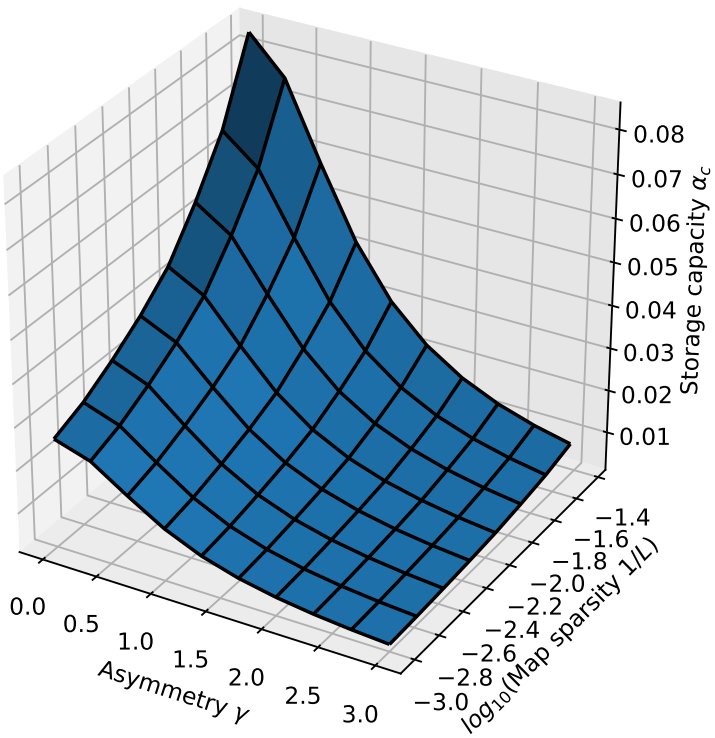


Figure 9: Dependence of the storage capacity on γ and $1/L$ (represented as $log_{10}(1/L)$).

With this approximate decoupling we see that, for sparse maps and small values of the asymmetry, the capacity scales as

$$\alpha_c \sim -\frac{1}{\ln(1/L)(1+\gamma^2)} \tag{31}$$

The scaling with $1/L$ is the same found by Battaglia & Treves [27] in the analysis of the symmetric case, as expected: for $\gamma = 0$ the two models are equivalent. The presence of asymmetry decreases the capacity, but does not have a catastrophic effect: the decrease is continuous and scales with a power of γ. There is therefore a large range of values of asymmetry and map sparsity in which a large number of dynamic maps can be stored and retrieved. We will see in the next section how this picture changes in a fully connected network, in which the asymmetry can enhance the capacity.

Numerical estimation of α_c for a fully connected network

To estimate the storage capacity for a fully connected network, we proceed with numerical simulations. For a network of fixed size N, and for given γ, L and number of maps p, we run a number of simulations D, letting the network evolve from a random initial

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

configuration. We consider a simulation to have performed a successful retrieval if the global overlap

$$m^\mu = \frac{1}{N^2} \sum_{i \neq j} V_i V_j K_S(x_i^\mu - x_j^\mu) \tag{32}$$

that quantifies the coherence of the activity with map μ, is large for one map μ^* (at least 95% of the overlap value obtained in the case of a single map) and low in all others maps $\mu \neq \mu*$. We then define the retrieval probability as $p_r = D_r/D$, where D_r is the number of observed retrievals.

We repeat the process varying the storage load, i.e. the number of stored manifolds p. As p is increased, the system reaches a transition point, at which the retrieval probability rapidly goes to zero. This transition is illustrated, for various values of γ, in Fig. 10.

The number of maps p_c at which the probability reaches zero defines as the storage capacity $\alpha_c(\gamma, L) = p_c(\gamma, L)/N$. Repeating this procedure for a range of values of γ and L, we obtain the plots shown in Fig. 11, for networks encoding one dimensional and two dimensional dynamical memories.

The first thing that can be noticed is that, also in the fully connected case, the network can store a large number of maps, for a wide range of γ and L. A network with size in the order of ten thousand neurons could store from tens up to hundreds of dynamical memories.

The capacity for one dimensional attractors is higher than the one for their two dimensional counterparts. This is in line with what was found for symmetric networks [27].

Finally, we see that the peak of the capacity is found not only for intermediate values of map sparsity – again in line with what is known from the symmetric case – but also for intermediate values of the coefficient γ. This shows that moderate values of asymmetry can be beneficial for the storage of multiple continuous attractors, a non-trivial phenomenon that may be crucial for the memory capacity of biological networks.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

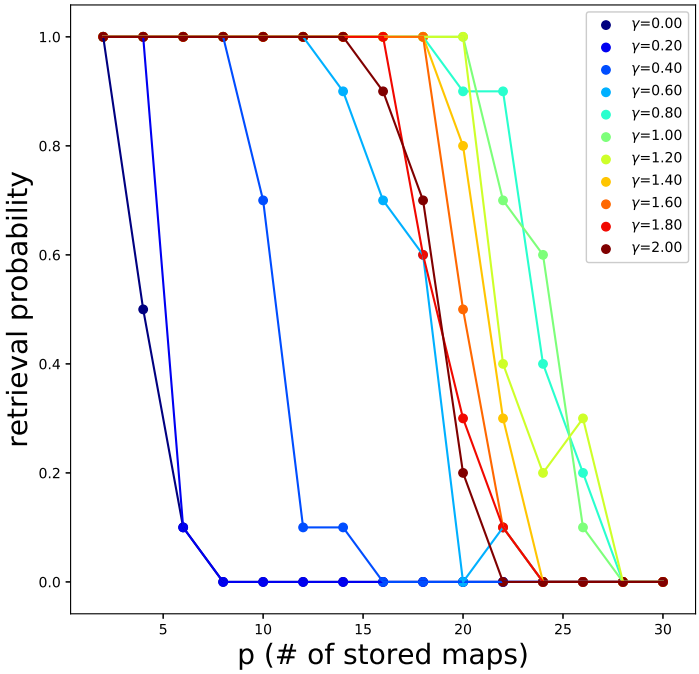
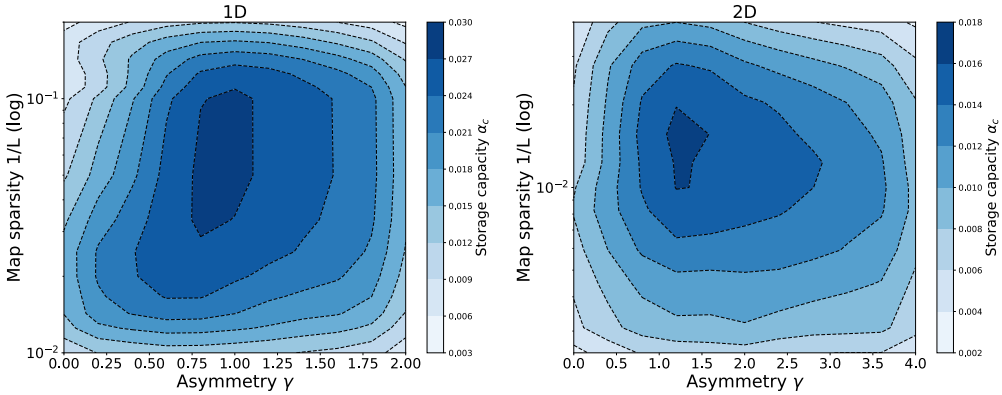


Figure 10: Retrieval / no retrieval phase transition for different values of γ, obtained from simulations with $N = 1000$, $D = 10$ and $L = 10$. The non-monotonic dependence of the capacity from γ can be appreciated here: the transition point moves towards the right with increasing γ up to $\gamma \sim 1$, then back to the left.



(a) (b)

Figure 11: Storage capacity as a function of map sparsity $1/L$ and asymmetry strength γ, for **(a)** one dimensional dynamic continuous attractors, **(b)** two dimensional dynamic continuous attractors.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Discussion

The results presented show how a continuous attractor neural network with memory-dependent asymmetric components in the connectivity can function as a dynamic memory. Our model is simple enough to be treated analytically, robustly produces dynamic retrieval for a large range of the relevant parameters and shows a storage capacity that is comparable to – and in some cases higher than – the capacity for static continuous attractors.

The analytical solution of the single attractor case shows that the interaction between the strength of the asymmetry and the velocity of the shift can be modulated by global features of the network activity such as its sparsity. This makes the network able to retrieve at different velocities in different regimes, without necessarily requiring short term synaptic modifications. The insensibility of the general features of the dynamics to the fine details of the shape of the interactions suggests that this mechanism could robustly emerge from learning or self organization processes in the presence of noise. The analysis of the storage capacity shows that the asymmetry does not heavily impair memory performance, and that, in densely connected networks, out of equilibrium effects can be beneficial for memory.

The storage capacity of out of equilibrium continuous attractors has been calculated, in a different scenario, by Zhong et al. [45]. The authors considered the case of an external signal driving the activity bump along the attractor, in a network of binary neurons, and proceeded to calculate the storage capacity with several assumptions that allowed to model the interference of multiple maps as thermal noise. Interestingly, their main result is broadly compatible with what we show here: in the highly diluted regime the velocity of the external signal has a mild – detrimental – effect on the capacity. This hints that out of equilibrium effects could show some form of universality across different network models and implementations of the shift mechanism.

The possibility of dynamic retrieval makes attractor models suitable for the description and the quantification of complex memory phenomena such as hippocampal replay. The model we propose suggests that tendency of the activity to move in the neural population is a natural feature of networks with asymmetric connectivity, when the asymmetry is organized along a direction in a low dimensional manifold, and that static memories could be the exception rather than the rule. Indeed, Mehta et al. [46] have shown that place fields can become more asymmetric in the course of spatial learning, demonstrating that the idea that symmetry emerges from an averaging of trajectory-dependent effects [47] does not always hold true. The model we presented can be useful for the quantification of the effects that symmetry and asymmetry in the interactions have on the acquisition, retention and retrieval of memory.

The dynamical retrieval of the model generalizes, in the framework of attractor networks, the idea of cognitive maps, incorporating a temporal organization in the low-dimensional manifold encoding the structure of the memory. This feature is reminiscent of the idea of memory schemata – constructs that can guide and constrain our mental activity when we reminisce about the past, imagine future or fictional scenarios or let our minds free to wander [48]. The use of the presented model to describe memory schemata will require further steps, such as an account of the interaction between hippocampus and neocortex, and a mechanism for the transition between different dynamical memories. Nevertheless, the idea of dynamic retrieval of a continuous manifold and the integration of the model presented here with effective models of cortical memory networks [49] open promising perspectives.

Our model can describe continuous attractors with more than one dimension; however, it is worth nothing that in the cases analysed here the asymmetry is constant along a single direction in each attractor. This can describe the situation in which the temporal evolution of the memory is structured along a certain dimension, and free to diffuse,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

without energy costs, in the remaining ones. The description of several one-dimensional trajectories, embedded in a two dimensional or three dimensional space and possibly intersecting, would instead require a position-dependent asymmetric component [50]: this is an interesting direction that will be pursued in future work.

Finally, the full analytical description of a densely connected, asymmetric attractor network is a challenge that remains open, and can yield valuable insights on the workings of the neural circuits underlying memory.

Acknowledgments

Work supported by the Human Frontier Science Program RGP0057/2016 collaboration. We are grateful for inspiring exchanges with Remi Monasson and others in the collaboration, and thank Silvia Girardi for her help with figure 1.

Appendix

A Numerical simulations

Numerical simulations were performed with python code, available at:

https://github.com/davidespalla/Dynamic-Continuous-Attractors

In the single map case, each of the N units ($N = 1000$ in 1D, $N = 1600$ in 2D) was assigned a preferential firing location x_i on a regular grid spanning the environment with linear dimension L. From this preferred firing locations the interaction matrix J_{ij} was constructed, with the formula:

$$J_ij = K_S(|x_i - x_j|) + \gamma K_A(|x_i - x_j|) \tag{33}$$

The precise shape of the symmetric and anti-symmetric parts of the kernel where chosen differently in different simulations, according to the feature the analysis focused on, as specified in the main text. Once the network was assembled, the dynamics was initialized either with a random assignment of activity values to each unit in the range $[0, 1]$, or with a gaussian bump centered in the middle of the environment (note that, due to the periodic boundary conditions and the translational invariance of the connectivity, the choice of the starting point does not influence the outcome). The dynamics was then evolved in discrete time steps, with the iteration of the following operations:

- Calculation of the local fields $h_i(t) = \sum_j J_{ij} V_j(t-1)$
- Calculation of the activity values $V_i(t) = g(h_i(t) - h_0)\theta(h_i(t) - h_0)$
- Dynamic adjustment of the threshold: $h_0 = f$ percentile of $\{V_i(t)\}$
- Dynamic adjustment of the gain: $g = g/\langle\{V_i(t)\}\rangle$
- Recalculation of the activity $V_i(t)$ with the adjusted gain and threshold

The adjustment of the parameters of the transfer function was enforced to constrain the network to operate at a fixed sparsity f and a fixed mean, set to one without loss of generality. The dynamics was iterated for a given number of steps (usually 200), large enough to assure the convergence to the attractive manifold (reached usually in < 5 steps).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In the case of multiple maps, the implemented dynamical evolution was the same, but the interaction matrix was constructed with multiple assignments of the preferred firing locations x_i, one for each of the p stored maps:

$$J_i j = \sum_{\mu=1}^{p} \left(K_S(|x_i^\mu - x_j^\mu|) + \gamma K_A(|x_i^\mu - x_j^\mu|) \right) \tag{34}$$

The multiple assignment of the preferred firing locations was performed by a random shuffling of the labels of the units before the assignment to the position on the regular grid spanning each map.

B Solution of the equation for the activity profile in the case of many maps

We illustrate here the procedure for the numerical solution of the equation 26:

$$v(x + \Delta x) = g\mathcal{N}\left(\int dx' K(x - x') v(x') + w \right) \tag{35}$$

We consider the one dimensional case and the exponential kernel

$$K(x - x') = e^{-|x-x'|} - \gamma \text{sign}(x - x') e^{-|x-x'|} \tag{36}$$

First, following [27] we rewrite it with the transformation

$$u(x) = \mathcal{N}^{-1}\left(\frac{v(x)}{g} \right) \tag{37}$$

obtaining

$$u(x + \Delta x) = g \int dx' K(x - x') \mathcal{N}(u(x')) + w \tag{38}$$

We then transform this integral equation in a differential one, by differentiating twice. We obtain

$$u''(x + \Delta x) + 2g\gamma\Phi(u(x))u'(x) + 2g\mathcal{N}(u(x)) - u(x + \Delta x) + w = 0 \tag{39}$$

where we used the fact that $\mathcal{N}'(x) = \Phi(x)$. Eq.39 is a second order, nonlinear delayed differential equation. To solve it, it is not sufficient to impose an initial condition on a single point for the solution and the first derivative (i.e. something like $u(x_0) = u_0, u'(x_0) = u'_0$): we have to specify the value of the function and its derivative in an interval $[x_0, x_0 + \Delta x]$.

To do so, we reason that, if we want a bump solution, $u(x)$ has to be finite in for $x \to \pm\infty$ and cannot diverge. We then require the function to be constant ($u(x) = u_0$, $u'(x) = 0$) before a certain value x_0, whose value can be set arbitrarily without loss of generality.

The value u_0, at $\gamma = 0$ and $\Delta x = 0$ determines the shape of $u(x)$, as shown by the numerical solution presented in Fig. 12 for $u_0 < u^*$ the solution will diverge for $x \to \infty$, for $u_0 > u^*$ it will oscillate. We are then left with a single value $u_0(g, w) = u^*(g, w)$ for which the solution has the required form.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

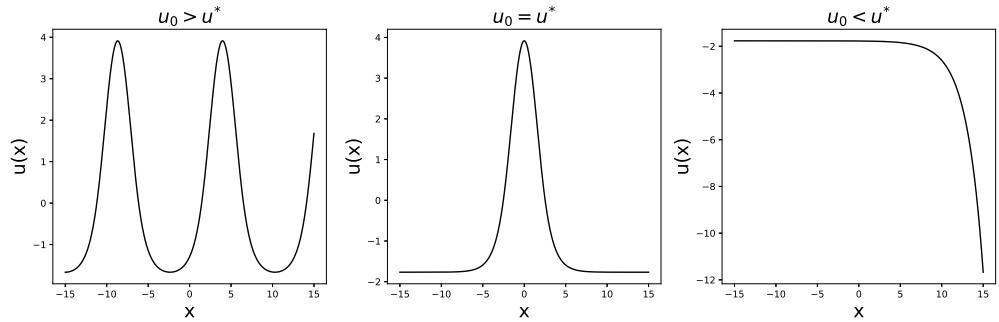


Figure 12: Solutions to Eq.39 for $g = 1$, $w = -1.8$, $\gamma = 0$, $\Delta x = 0$.

Then, keeping u_0 fixed, we can repeat a similar procedure to find Δx for different values of γ. Also in this case, the solution either diverges or oscillates, apart from a single value Δx^*, for which the solution has the desired shape (see Fig.13). This eliminates the arbitrariness in the choice of Δx since it imposes, for given g and w, a relation $\Delta x = \Delta x^*(\gamma)$.

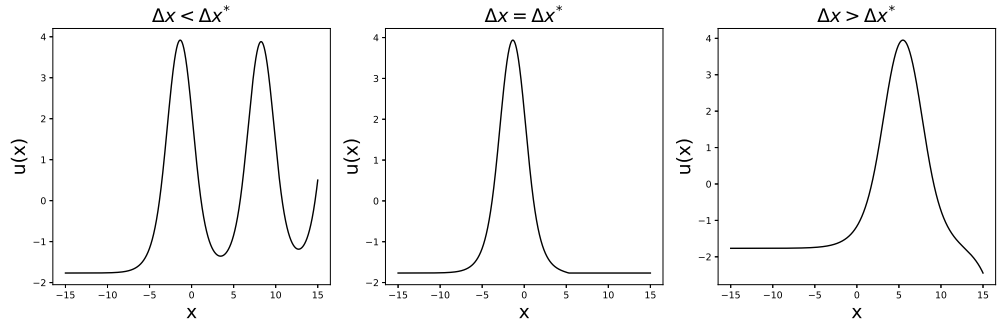


Figure 13: Solutions to Eq.39 for $g = 1$, $w = -1.8$, $\gamma = 0.2$.

We can then find the shape of the bump $u(x)$ for given values of g,w and γ, from which we can obtain the profile $v(x) = g\mathcal{N}(u(x))$ that we need for the calculation of the storage capacity. Some examples of the obtained profiles, for different values of γ, are shown in Fig. 14.

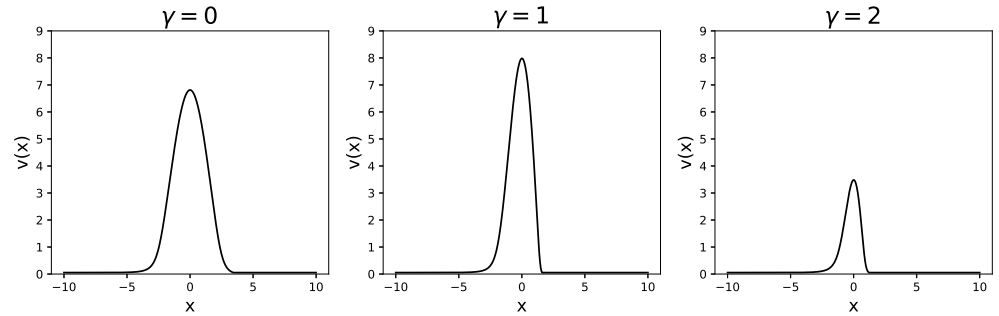


Figure 14: Activity profile $v(x)$, obtained for the same $g = 0.7$ and $w = -1.3$, at different values of γ.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

[1] Howard Eichenbaum and Neal J Cohen. *From conditioning to conscious recollection: Memory systems of the brain.* Number 35. Oxford University Press on Demand, 2004.

[2] Endel Tulving. Episodic memory: From mind to brain. *Annual review of psychology*, 53(1):1–25, 2002.

[3] György Buzsáki, Cornelius H Vanderwolf, et al. Cellular bases of hippocampal EEG in the behaving rat. *Brain Research Reviews*, 6(2):139–171, 1983.

[4] Gyorgy Buzsaki, Zsolt Horvath, Ronald Urioste, Jamille Hetke, and Kensall Wise. High-frequency network oscillation in the hippocampus. *Science*, 256(5059):1025–1027, 1992.

[5] William E Skaggs and Bruce L McNaughton. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, 271(5257):1870–1873, 1996.

[6] Zoltán Nádasdy, Hajime Hirase, András Czurkó, Jozsef Csicsvari, and György Buzsáki. Replay and time compression of recurring spike sequences in the hippocampus. *Journal of Neuroscience*, 19(21):9497–9507, 1999.

[7] David J Foster and Matthew A Wilson. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–683, 2006.

[8] Jadin C Jackson, Adam Johnson, and A David Redish. Hippocampal sharp waves and reactivation during awake states depend on repeated sequential experience. *Journal of Neuroscience*, 26(48):12415–12426, 2006.

[9] David Dupret, Joseph O'Neill, Barty Pleydell-Bouverie, and Jozsef Csicsvari. The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature neuroscience*, 13(8):995, 2010.

[10] Brad E Pfeiffer and David J Foster. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79, 2013.

[11] Joseph O'Neill, Barty Pleydell-Bouverie, David Dupret, and Jozsef Csicsvari. Play it again: reactivation of waking experience and memory. *Trends in neurosciences*, 33(5):220–229, 2010.

[12] Mattias P Karlsson and Loren M Frank. Awake replay of remote experiences in the hippocampus. *Nature neuroscience*, 12(7):913–918, 2009.

[13] H Freyja Ólafsdóttir, Daniel Bush, and Caswell Barry. The role of hippocampal replay in memory and planning. *Current Biology*, 28(1):R37–R50, 2018.

[14] George Dragoi and Susumu Tonegawa. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*, 469(7330):397–401, 2011.

[15] Usman Farooq and George Dragoi. Emergence of preconfigured and plastic time-compressed sequences in early postnatal development. *Science*, 363(6423):168–173, 2019.

[16] Federico Stella, Peter Baracskay, Joseph O'Neill, and Jozsef Csicsvari. Hippocampal reactivation of random trajectories resembling Brownian diffusion. *Neuron*, 102(2):450–461, 2019.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[17] George Dragoi and Susumu Tonegawa. Distinct preplay of multiple novel spatial experiences in the rat. *Proceedings of the National Academy of Sciences*, 110(22):9100–9105, 2013.

[18] Igor Gridchyn, Philipp Schoenenberger, Joseph O'Neill, and Jozsef Csicsvari. Assembly-specific disruption of hippocampal replay leads to selective memory deficit. *Neuron*, 2020.

[19] Kenneth Kay, Jason E Chung, Marielena Sosa, Jonathan S Schor, Mattias P Karlsson, Margaret C Larkin, Daniel F Liu, and Loren M Frank. Constant sub-second cycling between representations of possible futures in the hippocampus. *Cell*, 180(3):552–567, 2020.

[20] Alex P Vaz, John H Wittig, Sara K Inati, and Kareem A Zaghloul. Replay of cortical spiking sequences during human memory retrieval. *Science*, 367(6482):1131–1134, 2020.

[21] Vanessa E Ghosh and Asaf Gilboa. What is a memory schema? a historical perspective on current neuroscience literature. *Neuropsychologia*, 53:104–114, 2014.

[22] Pierre Baraduc, Jean-René Duhamel, and Sylvia Wirth. Schema cells in the macaque hippocampus. *Science*, 363(6427):635–639, 2019.

[23] Krishna V Shenoy, Maneesh Sahani, and Mark M Churchland. Cortical control of arm movements: a dynamical systems perspective. *Annual review of neuroscience*, 36, 2013.

[24] Juan A Gallego, Matthew G Perich, Lee E Miller, and Sara A Solla. Neural manifolds for the control of movement. *Neuron*, 94(5):978–984, 2017.

[25] Erhan Oztop and Michael A Arbib. Schema design and implementation of the grasp-related mirror neuron system. *Biological cybernetics*, 87(2):116–140, 2002.

[26] Daniel J Amit, Hanoch Gutfreund, and Haim Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, 55(14):1530, 1985.

[27] Francesco P Battaglia and Alessandro Treves. Attractor neural networks storing multiple space representations: a model for hippocampal place fields. *Physical Review E*, 58(6):7738, 1998.

[28] Rémi Monasson and Sophie Rosay. Crosstalk and transitions between multiple spatial maps in an attractor neural network model of the hippocampus: Phase diagram. *Physical review E*, 87(6):062813, 2013.

[29] Alexei Samsonovich and Bruce L McNaughton. Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, 17(15):5900–5920, 1997.

[30] Davide Spalla, Alexis Dubreuil, Sophie Rosay, Remi Monasson, and Alessandro Treves. Can grid cell ensembles represent multiple spaces? *Neural Computation*, 31(12):2324–2347, 2019.

[31] Haim Sompolinsky and Ido Kanter. Temporal association in asymmetric neural networks. *Physical review letters*, 57(22):2861, 1986.

[32] Emilio Kropff and Alessandro Treves. The storage capacity of Potts models for semantic memory retrieval. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(08):P08010, 2005.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[33] Eleonora Russo, Vijay MK Namboodiri, Alessandro Treves, and Emilio Kropff. Free association transitions in models of cortical latching dynamics. *New Journal of Physics*, 10(1):015008, 2008.

[34] Chol Jun Kang, Michelangelo Naim, Vezha Boboeva, and Alessandro Treves. Life on the edge: latching dynamics in a Potts neural network. *Entropy*, 19(9):468, 2017.

[35] Amir Hossein Azizi, Laurenz Wiskott, and Sen Cheng. A computational model for preplay in the hippocampus. *Frontiers in computational neuroscience*, 7:161, 2013.

[36] Alessandro Treves. Neural phase transitions that made us mammals. In Péter Érdi, Anna Esposito, Maria Marinaro, and Silvia Scarpetta, editors, *Computational Neuroscience: Cortical Dynamics*, pages 55–70, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[37] Kechen Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *Journal of Neuroscience*, 16(6):2112–2126, 1996.

[38] Mark C Fuhs and David S Touretzky. A spin glass model of path integration in rat medial entorhinal cortex. *Journal of Neuroscience*, 26(16):4266–4276, 2006.

[39] Yoram Burak and Ila R Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput Biol*, 5(2):e1000291, 2009.

[40] Henry Markram, Joachim Lübke, Michael Frotscher, and Bert Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science*, 275(5297):213–215, 1997.

[41] Alessandro Treves. Threshold-linear formal neurons in auto-associative nets. *Journal of Physics A: Mathematical and General*, 23(12):2631, 1990.

[42] Bernard Derrida, Elizabeth Gardner, and Anne Zippelius. An exactly solvable asymmetric neural network model. *EPL (Europhysics Letters)*, 4(2):167, 1987.

[43] Masatoshi Shiino and Tomoki Fukai. Self-consistent signal-to-noise analysis and its application to analogue neural networks with asymmetric connections. *Journal of Physics A: Mathematical and General*, 25(7):L375, 1992.

[44] Yasser Roudi and Alessandro Treves. An associative network with spatially organized connectivity. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(07):P07010, 2004.

[45] Weishun Zhong, Zhiyue Lu, David J Schwab, and Arvind Murugan. Nonequilibrium statistical mechanics of continuous attractors. *Neural Computation*, 32(6):1033–1068, 2020.

[46] Mayank R Mehta, Carol A Barnes, and Bruce L McNaughton. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, 94(16):8918–8921, 1997.

[47] Patricia E Sharp. Computer simulation of hippocampal place cells. *Psychobiology*, 19(2):103–115, 1991.

[48] Elisa Ciaramelli and Alessandro Treves. A mind free to wander: neural and computational constraints on spontaneous thought. *Frontiers in psychology*, 10:39, 2019.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.08.373084; this version posted November 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[49] Vezha Boboeva, Romain Brasselet, and Alessandro Treves. The capacity for correlated semantic memories in the cortex. *Entropy*, 20(11):824, 2018.

[50] Kenneth I Blum and Larry F Abbott. A model of spatial map formation in the hippocampus of the rat. *Neural computation*, 8(1):85–93, 1996.