

# Continuous Attractors with Morphed/Correlated Maps

Sandro Romani, Misha Tsodyks\*

Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel

## Abstract

Continuous attractor networks are used to model the storage and representation of analog quantities, such as position of a visual stimulus. The storage of multiple continuous attractors in the same network has previously been studied in the context of self-position coding. Several uncorrelated maps of environments are stored in the synaptic connections, and a position in a given environment is represented by a localized pattern of neural activity in the corresponding map, driven by a spatially tuned input. Here we analyze networks storing a pair of correlated maps, or a morph sequence between two uncorrelated maps. We find a novel state in which the network activity is simultaneously localized in both maps. In this state, a fixed cue presented to the network does not determine uniquely the location of the bump, i.e. the response is unreliable, with neurons not always responding when their preferred input is present. When the tuned input varies smoothly in time, the neuronal responses become reliable and selective for the environment: the subset of neurons responsive to a moving input in one map changes almost completely in the other map. This form of remapping is a non-trivial transformation between the tuned input to the network and the resulting tuning curves of the neurons. The new state of the network could be related to the formation of direction selectivity in one-dimensional environments and hippocampal remapping. The applicability of the model is not confined to self-position representations; we show an instance of the network solving a simple delayed discrimination task.

**Citation:** Romani S, Tsodyks M (2010) Continuous Attractors with Morphed/Correlated Maps. *PLoS Comput Biol* 6(8): e1000869. doi:10.1371/journal.pcbi.1000869

**Editor:** Karl J. Friston, University College London, United Kingdom

**Received:** April 29, 2010; **Accepted:** June 28, 2010; **Published:** August 5, 2010

**Copyright:** © 2010 Romani, Tsodyks. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded by the European Framework Programme 7, SPACEBRAIN. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: misha@weizmann.ac.il

## Introduction

The ability to keep an internal representation of a continuous variable in the absence of sensory stimuli, is a crucial requirement in order to succeed in what can be considered trivial day to day actions or experimenter designed tasks. For instance one may think about the eye position between successive saccades [1], the angle of stimulus presentation in an oculomotor delayed protocol [2], the spatial position or the head direction in a dark environment [3–5], or the phase of the recently discovered grid fields [6,7].

A widely used class of models for this kind of working memory is constituted by attractor neural networks. The temporary maintenance of an item in memory corresponds to a specific network pattern of activity which is stabilized via strengthened recurrent connections between the active neurons in the pattern [8–11]. These connections are usually imposed, or trained, as the outcome of some form of Hebbian learning. The attractor is called continuous when the stable states form a continuous manifold which can be parametrized by the state variables. This outcome is obtained under certain conditions on the synaptic connection, for example when the connections between neurons are lateral-inhibition like (e.g. Mexican hat) [12–14]. The underlying idea is that each neuron is assigned a location on an abstract *map*. The synaptic weights (*encoding*) depend on the location of the pre- and post-synaptic neurons. By means of Turing instability, the network dynamics creates a localized pattern of activity (or bump) on the map [15]. The external input links the position on the map to the state variable, forming a *representation*.

Continuous attractors have been used to explain the maintenance of various analog quantities close or far from the primary sensory and motor regions. For instance, the orientation tuning in the visual cortex [16,17], hippocampal place fields in one [18,19] and two dimensions [20,21], eye position [1,22], head direction tuning in the postsubiculum [23,24] and entorhinal grid fields [25,26].

The simple picture of a single continuous attractor can be naturally extended to the case of multiple attractors. The encoded maps can then be assumed to be either uncorrelated or correlated, and in particular to exhibit some structure (e.g. deriving from a morphing procedure). Assuming a complete lack of correlations between maps is not realistic, though useful for obtaining analytic results [27]. In this contribution, we analyze the network representations arising from the storage of two maps, with a varying degree of correlation between them, and from the storage of a morph sequence between two uncorrelated maps. We are interested in finding the conditions under which the network representation can provide some information about the state variables. Surprisingly, even when the correlation between two maps is very high, under conditions which will be clarified later it is possible for the network to maintain separate representations of the state variables.

## Multiple maps

Multiple state variables can be encoded in the same network. An example is offered by the place representations of several environments [20,21]. To each environment corresponds a neural map which is encoded in the synaptic efficacies. Sensory inputs

## Author Summary

How is your position in an environment represented in the brain, and how does the representation distinguish between multiple environments? One of the proposed answers relies on continuous attractor neural networks. Consider the web page of your campus map as a network of pixels. Every pixel is a neuron, and nearby pixels excite each other, while distant pairs are inhibited. As a result of their interactions, a bunch of close-by pixels will light up, indicating your current position as suggested by your web-cam (the sensory input). When you travel to another campus, the common assumption holds that pixels are completely scrambled and the excitatory/inhibitory pattern of connections is summed to the existing one. Now these connections and the sensory input will activate the pixels corresponding to your location in the new campus. The active pixels will look like noise in the old map. But what if the campuses are similar, i.e. the pixels are not completely scrambled? We show that the network has a novel way of distinguishing between the environments, by lighting up distinct subsets of pixels for each campus. This emergent selectivity for the environment could be a mechanism underlying hippocampal remapping and directional selectivity of place cells in 1D environments.

would then select the correct representation, i.e. both the environment and the position in the environment. The selected map wins the competition with the other maps stored in the network, and a localized pattern appears. In this case the network only maintains information about one of the several encoded state variables.

A more peculiar property of multiple continuous attractors, is their ability to represent simultaneously the values of several state variables. This property was explored in [28], where two partially overlapping neural populations (representing discrete features), are assigned two uncorrelated maps. Another example is provided in the study of [29], where a single network stores and represents simultaneously a continuous and discrete attractors.

In principle, given the existence of multiple representations in different brain regions (either one per region, or many in one region), a brain area downstream would necessarily encode several state variables. In light of a Hebbian interpretation on how this encoding takes place, it seems natural to distinguish between two cases. When multiple representations provide a simultaneous input to a region, the result is probably encoded multiplicatively [29], or, in general, non-linearly. For inputs happening non concurrently, as for instance when walking through several rooms sequentially, an additive encoding of each room is expected [21]. In the following we will analyze additive encoding.

## Correlations

The present contribution addresses the issue of encoding correlated maps. The motivations come from recent experimental results on place cells recording in morphed environments [30–32], where place fields remapping along a sequence of morphed arenas is experimentally tested, and from theoretical and experimental studies concerning the morphing of discrete attractors [33–35].

In general, we would consider the encoding of  $p$  manifolds  $X_\zeta$ , each of dimension  $d_\zeta$ , where  $\zeta = 1 \dots p$ . We will refer to a single manifold as a map, once a coordinate system  $x_\zeta$  is chosen. The use of uppercase (e.g.  $X$ ) or lowercase (e.g.  $x$ ) will distinguish between the whole map and a single point on it respectively. Given a pre-

synaptic neuron indexed by  $x_\zeta$ , and a post-synaptic  $x'_\zeta$ , the encoding of a single map is obtained using a synaptic matrix  $W_\zeta(x_\zeta, x'_\zeta)$ , and is such that a continuous attractor representation would arise if it were the only map. We assume, as mentioned above, that the complete encoding arises from a linear superposition of the  $p$  matrices,  $W = \frac{1}{p} \sum_\zeta W_\zeta$ . The statistical properties of the maps, and in particular the correlation between them, can be fully specified by providing the probability density  $v(\{x_\zeta\})$ .

The general problem is too difficult to be studied analytically. Some results can be obtained for the case of uncorrelated maps on the same manifold [27], though the system can be explored by simulating the full microscopic networks (see e.g. [21] for the uncorrelated case and [36] for simulation results of the correlated case).

In order to simplify the analysis, while retaining the basic structure of the problem, we focus on the case of  $p=2$  representations, on a 1-dimensional circular manifold (i.e. the ring model [16,37]). The correlation between the maps is constructed by limiting the distance between the single neuron locations on the two maps. We devise a simple method to generate a morph sequence between two uncorrelated maps, by linearly modifying the neurons locations between the original maps. This method also suggests a way to test the network response to the exposure of intermediate maps between the two stored correlated maps.

For concreteness, one could think about maps of two similar circular arenas, and reason in term of spatial coding. In this context, we are interested in clarifying how the information about the position in the current environment is represented by the network, when varying the constitutive parameters of the model; And how the representation changes when the network is exposed to environments along a morph sequence.

In the following we will describe with mean-field (MF) theory the attractor landscape of a network, i.e. the stable solutions in absence of any place specific input. We then consider the behavior of the solutions when a spatially tuned input is present. We will establish the approximate relationship between two strongly correlated maps and the encoding of a morph sequence between two reference rings, and study the behavior of the solutions in presence of a tuned input varying along the sequence. Finally we will verify the results with microscopic simulations of finite networks. The network properties can be tested experimentally to confirm (or falsify) the attractor hypothesis.

## Results

Let us consider two circular environments  $A$  and  $B$ , inducing two different maps  $\Theta_A$  and  $\Theta_B$  in the network. In the MF limit, we can imagine having a continuous manifold of neurons, where each neuron is identified by the pair of labels  $(\theta_A, \theta_B)$ , with  $\theta_\zeta \in [0, 2\pi)$ ,  $\zeta = (A, B)$ . In addition, a probability density  $v(\theta_A, \theta_B)$  tells us how likely is for a neuron to have the labels  $(\theta_A, \theta_B)$ . As mentioned in *Introduction*, we assume the resulting synaptic structure to be a linear superposition of ring models. Hence, the connection strength between two neurons  $(\theta_A, \theta_B)$  and  $(\theta'_A, \theta'_B)$  is

$$W(\theta_A - \theta'_A, \theta_B - \theta'_B) = \frac{J_1}{2} [\cos(\theta_A - \theta'_A) + \cos(\theta_B - \theta'_B)] + J_0.$$

The factor  $J_1$  is a measure of the amplitude of the map specific interaction, while  $J_0 (< 0)$  is a uniform inhibitory term. This form of connectivity can be thought as arising from the first two terms

of Fourier series of a more general coupling. The rate dynamics for the network activity  $\tilde{m}(\theta_A, \theta_B)$  is [38]:

$$\tau \dot{\tilde{m}}(\theta_A, \theta_B) = -\tilde{m}(\theta_A, \theta_B) + \left[ \int \int d\theta'_A d\theta'_B v(\theta'_A, \theta'_B) W(\theta_A - \theta'_A, \theta_B - \theta'_B) \tilde{m}(\theta'_A, \theta'_B) + I \right]_+$$

where we assumed a threshold-linear transfer function for the neurons,  $[x]_+ = x$  when  $x > 0$  and 0 otherwise. The external afferent current is denoted by  $I$ , and it is assumed uniform in the current Section.

We build the maps with a simple procedure which induces a correlation between them. First, we create a uniformly distributed map  $\Theta$  with coordinates  $\theta \in [0, 2\pi)$  and a uniformly distributed map  $R$  of distance values  $r \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ . Then we define the coordinates of the desired maps as

$$\theta_A = (\theta - \mu r) \text{ mod } 2\pi \tag{1}$$

$$\theta_B = (\theta + \mu r) \text{ mod } 2\pi.$$

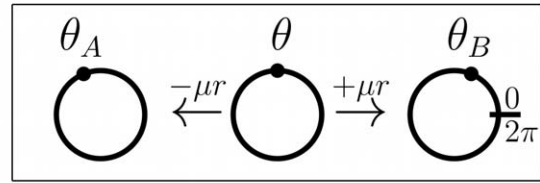
The parameter  $\mu \in [0, 1]$  is a measure of the distance between the two maps; the higher the distance between the maps, the lower the correlation between them. The coordinate  $\theta$  defines a “middle” map from which the coordinates of the environments  $A$  and  $B$  are constructed; each of them cannot be farther than  $\frac{\mu\pi}{2}$  from the middle map, hence they cannot differ more than  $\mu\pi$ . When  $\mu = 0$  the two maps are identical, and for  $\mu = 1$  the two maps are uncorrelated. As an example, let us fix the distance between the maps at  $\mu = \frac{1}{2}$  and consider the case of a neuron with  $\theta = \frac{\pi}{2}$ ; a choice of  $r$  for this neuron will yield the coordinates in the maps  $\Theta_A$  and  $\Theta_B$ . The range of possible values for  $r$  will generate  $\theta_A$  and  $\theta_B$  in the interval  $[\frac{\pi}{2} - \frac{\pi}{4}, \frac{\pi}{2} + \frac{\pi}{4}]$ , which shows how not all the possible pairs  $(\theta_A, \theta_B)$  are obtainable. An instance of this procedure is depicted in Fig. 1. A given angle in map  $\Theta_A$  or  $\Theta_B$  (corresponding to a given color in Fig. 1B) is represented by a straight line in the reference frame  $(\theta, r)$ . The effect of a decreasing  $\mu$  is to tilt this straight line toward the vertical direction (only identical angles in map  $\Theta_A$  and  $\Theta_B$  would be possible). Note that it is possible to define the inverse transformation  $(\theta_A, \theta_B) \rightarrow (\theta, r)$  (Eq. 21).

The new coordinates  $(\theta, r)$  are uniformly distributed by construction. We can then rewrite the dynamics of the network activity  $m(\theta, r) = \tilde{m}(\theta_A(\theta, r), \theta_B(\theta, r))$ , using Eq. 1, as

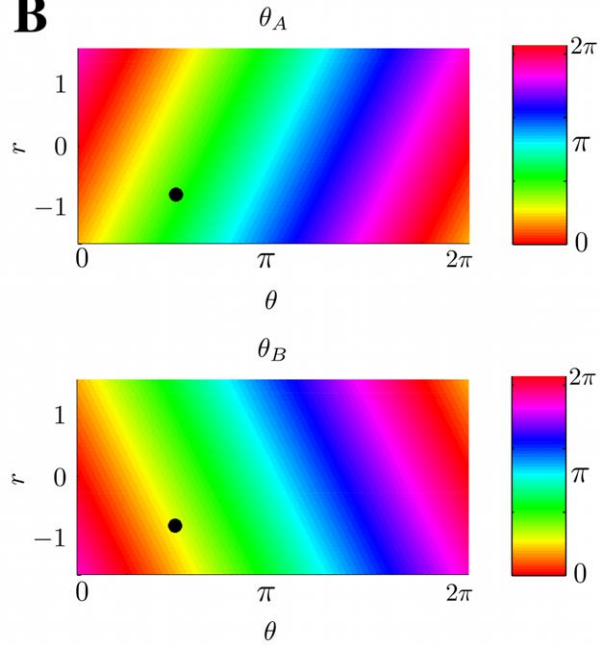
$$\begin{aligned} \tau \dot{m}(\theta, r) = & -m(\theta, r) + \\ & \left[ \frac{J_1}{2} \int \int \mathcal{D}\theta' \mathcal{D}r' [\cos(\theta - \theta' - \mu(r - r')) \right. \\ & + \cos(\theta - \theta' + \mu(r - r'))] m(\theta', r') + \\ & \left. J_0 \int \int \mathcal{D}\theta' \mathcal{D}r' m(\theta', r') + I \right]_+ \end{aligned} \tag{2}$$

The notations  $\int \mathcal{D}\theta$  and  $\int \mathcal{D}r$  are a shorthand for  $\frac{1}{2\pi} \int_0^{2\pi} d\theta$  and  $\frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} dr$  respectively. The use of the ring connectivity structure

**A**



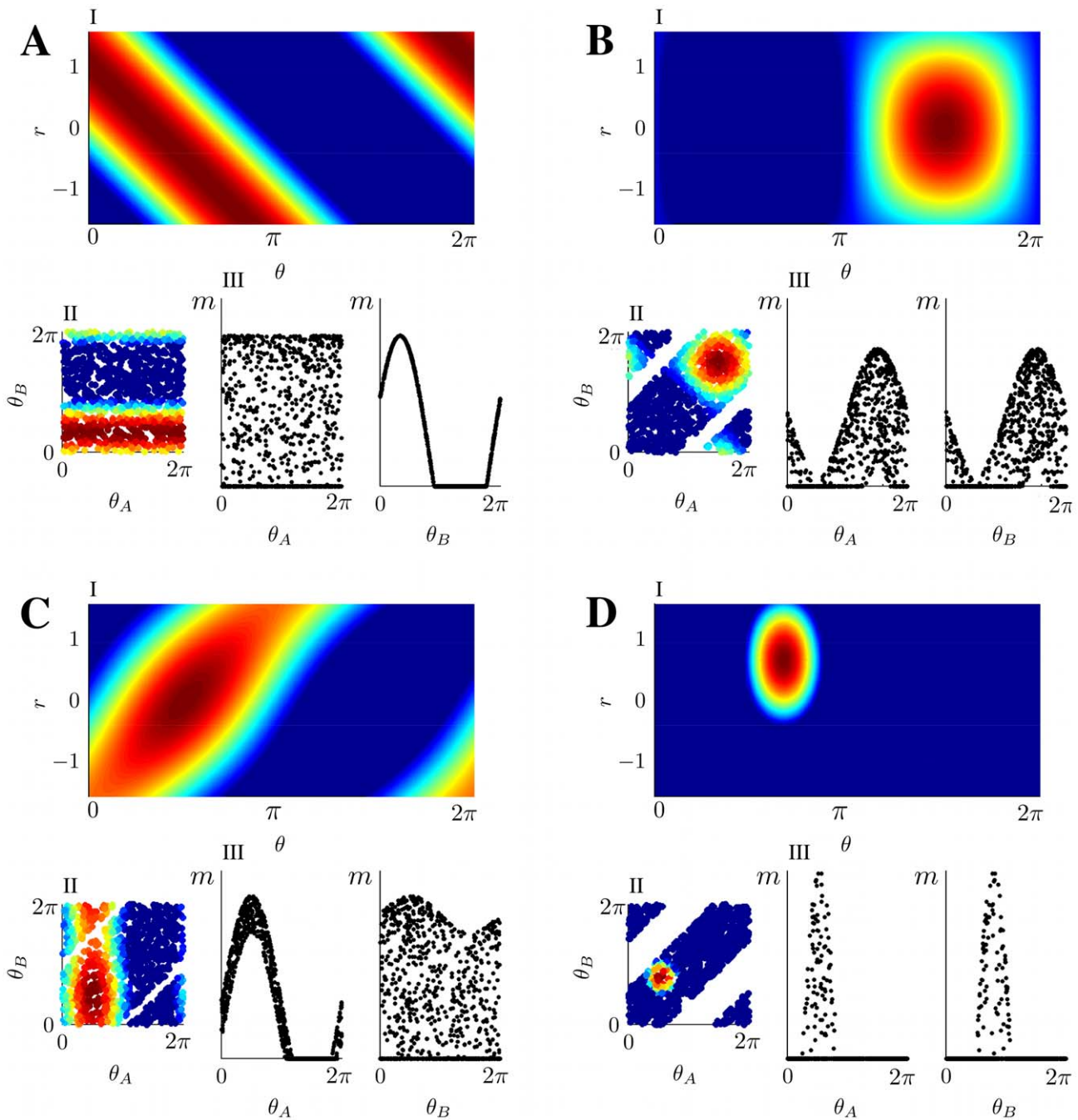
**B**



**Figure 1. Construction of correlated maps. A:** Cartoon showing how to generate two correlated maps  $A, B$  from neurons with randomly assigned index  $\theta$  on a reference ring, and distance value  $r \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ . Given the distance between maps  $\mu \in [0, 1]$  ( $0 \rightarrow$  identical,  $1 \rightarrow$  uncorrelated), the desired maps are created by adding and subtracting the value  $\mu r$  to the reference location  $\theta$ . The reference acts as an intermediate map, and the distance between maps limits the maximum distance between the location of the neurons on the maps  $A, B$ . **B:** Color codes corresponds to angles in map  $A$  (top panel) and  $B$  (bottom panel). The two plots show how neurons indexed by  $\theta_A, \theta_B$  are represented in  $\theta, r$  coordinates, for  $\mu = \frac{1}{2}$ . The particular neuron depicted in the cartoon is represented as a dot. doi:10.1371/journal.pcbi.1000869.g001

makes possible to reduce the dimensionality of the dynamical system to few *order parameters*. Five order parameters are necessary in order to describe the dynamics of the system:  $(\alpha, \gamma, \sigma, \psi_+, \psi_-)$  (see *Methods - Reduced Dynamics* for the details of the derivation, and the next Section for the equations describing their dynamics). Our choice of the order parameters exclude the analysis of the uniform solution of Eq. 2, i.e. a constant activity over the whole network. We will return to this solution in *Results - Phase diagram of the model*. After the reduction, the steady state activity profile in  $(\theta, r)$  space assumes the form:

$$m(\theta, r) = \alpha I [\cos(\theta - \psi_+) \cos(\mu r - \psi_-) - \gamma \sin(\theta - \psi_+) \sin(\mu r - \psi_-) + \sigma]_+ \tag{3}$$



**Figure 2. Network activity examples.** The top plots (numbered *I*) in each panel show the MF network activity (Eq. 3) in the 2D map  $(\theta, r)$ , corresponding to different choices of the order parameters  $(\gamma, \sigma, \psi_+, \psi_-)$  and the distance  $\mu$ . Each point in the plot corresponds to a neuron with labels  $(\theta, r)$ , the color is proportional to the activity level; warmer color represent higher activity, blue represents no activity. The activity scale is arbitrary, since it can be rescaled by a change in the external input  $I$ , without modifying the shape of the bump. Plots on the left of each panel (*II*) show the corresponding activity for a network of 1000 randomly chosen neurons in the 2D map  $(\theta_A, \theta_B)$ , see Eq. 1. The white stripes in the plots are due to the absence of neurons with labels in those regions, imposed by the correlation between the maps. The bottom-right pairs of plots (*III*) are projections of the activity on the individual maps. **A:**  $(\mu=1, \gamma=1, \sigma=0.2, \psi_+=\frac{4\pi}{3}, \psi_-=\pi)$  Activity localized in a single map, from a network storing uncorrelated maps. The locations of the peak activity in map *A* and *B* are  $\psi_{A,B}=\psi_+ \mp \psi_-$ . **B:**  $(\mu=0.8, \gamma=0, \sigma=.1, \psi_+=\frac{3\pi}{2}, \psi_-=0)$  Correlated maps, activity not favoring either *A* or *B*; since  $\psi_-=0$ , the bump is located at the same angle in both maps. **C:**  $(\mu=0.9, \gamma=-0.7, \sigma=0.3, \psi_+=\frac{\pi}{2}, \psi_-=0)$  Correlated maps, activity prefers map *A* ( $\gamma<0$ ). **D:**  $(\mu=.6, \gamma=0, \sigma=-0.8, \psi_+=\frac{2\pi}{3}, \psi_-=\frac{\pi}{8})$  Correlated maps; for this smaller sized bump, the location of the activity maximum is such that  $\psi_A \neq \psi_B$ .  
doi:10.1371/journal.pcbi.1000869.g002



Note that a change in the strength of the applied uniform input  $I$  produce no changes in the order parameters (see *Methods - Reduced dynamics* and Eqs. 5).

Several examples of network activity (Eq. 3), corresponding to different representative choices of the order parameters, are shown in Fig. 2. The various panels show the network activity in the two two-dimensional maps  $(\Theta_A, \Theta_B)$  and  $(\Theta, R)$ , and the one-dimensional projections of the activity to  $\Theta_A$  and  $\Theta_B$ . Note that not all the choices of order parameters corresponds to actual solutions of the dynamics (which are determined by the parameters  $(J_1, J_0, \mu)$  and the initial conditions), as will be shown later.

The meaning of the order parameters can be read out from Eq. 3. The variable  $\alpha$  represents a scaling factor for the amplitude of the network activity, which in turn is proportional to the uniform input  $I$ .

The variable  $\sigma$  is a measure of the spatial size of the activity profile, i.e. of the region in either  $(\Theta_A, \Theta_B)$  or  $(\Theta, R)$  in which the network activity  $m$  is strictly positive (Eq. 3). The activity profile is also referred to as a bump. For instance the case  $\sigma = -1$  would correspond to absence of activity (the current in the threshold-linear transfer function would always be negative), while  $\sigma = 1$  would make all the neurons in the network active.

The order parameter  $\gamma$  tells us how much the network representation “favors” one of the two maps. By its definition (*Methods - Reduced dynamics*, Eq. 24), the possible range for  $\gamma$  is  $[-1, 1]$ . The two extreme cases  $\gamma = \pm 1$  correspond to a network activity localized in either map  $\Theta_B$  or  $\Theta_A$ . For instance, the network activity Eq. 3 for  $\gamma = -1$  reads

$$\begin{aligned} m &= \alpha I [\cos(\theta - \psi_+) \cos(\mu r - \psi_-) + \sin(\theta - \psi_+) \sin(\mu r - \psi_-) + \sigma]_+ \\ &= \alpha I [\cos(\theta - \mu r - (\psi_+ - \psi_-)) + \sigma]_+ = \alpha I [\cos(\theta_A - \psi_A) + \sigma]_+, \end{aligned}$$

where in the last equality we used Eq. (1). From here we see that the position of the bump peak is located at  $\psi_A \equiv \psi_+ - \psi_-$ ; the same derivation, with  $\gamma = 1$ , would give us  $\psi_B \equiv \psi_+ + \psi_-$ . The network representation is in this case a bump of activity localized in one map, and does not have any spatial modulation in the other map, as exemplified in Fig. 2A,III. In the case  $\gamma = 0$ , from the explicit expression of the activity  $m$  we get

$$m = \alpha I [\cos(\theta - \psi_+) \cos(\mu r - \psi_-) + \sigma]_+. \quad (4)$$

This representation exhibits an equal amount of spatial modulation in both maps  $\Theta_A$  and  $\Theta_B$ , i.e. the solution represents equally the two stored maps, Fig. 2B,D. Depending on the value of  $\psi_-$  (see below), the location of the bumps in the maps  $\Theta_A$  and  $\Theta_B$  can be either the same ( $\psi_- = 0$ , Fig. 2B,III), or different ( $\psi_- \neq 0$ , Fig. 2D,III). Solutions with intermediate  $\gamma$  values ( $0 < |\gamma| < 1$ ) have a more localized projection in one of the two maps, depending on the sign of  $\gamma$  (see for instance Fig. 2C).

The quantities  $\psi_+$  and  $\frac{\psi_-}{\mu}$  identify respectively the location of the maximum of the network activity in the  $(\theta, r)$  coordinates, which is uniquely mapped to the maximum in  $(\theta_A, \theta_B)$  via Eqs. 1.

In the following, we will show that the network activity examples depicted in Fig. 2 are possible solutions of the dynamics described by Eq. 2. We refer to each of these classes as *double ring* (Fig. 2A,C), *single ring* (Fig. 2B) and *cylinder* (Fig. 2D), for reasons that will be clarified in the next Section. The cylinder class represents an interesting novel regime (simultaneous localized projections in both environments), and we will devote most of the paper to describe the properties of this particular solution.

## Phase diagram of the model

In this Section we analyze the fixed point solutions of the system, and heuristically describe the region of stability of these solutions. A more rigorous description of the stability can be found in *Methods - Stability*.

In *Methods - Reduced dynamics* we derive the dynamics of the order parameters from Eqs. 2. We report here the result

$$\begin{aligned} \tau \dot{\gamma} &= -J_1 \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) [\sin(\theta - \psi_+) \sin(\mu r - \psi_-) \\ &\quad + \gamma \cos(\theta - \psi_+) \cos(\mu r - \psi_-)] \end{aligned} \quad (5)$$

$$\tau \dot{\alpha} = \alpha \left( -1 + J_1 \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) \cos(\theta - \psi_+) \cos(\mu r - \psi_-) \right)$$

$$\begin{aligned} \tau \dot{\sigma} &= \frac{1}{\alpha} + J_0 \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) - \\ &\quad J_1 \sigma \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) \cos(\theta - \psi_+) \cos(\mu r - \psi_-) \end{aligned}$$

$$\begin{aligned} \tau \dot{\psi}_- &= \frac{J_1}{1 - \gamma^2} \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) [\cos(\theta - \psi_+) \sin(\mu r - \psi_-) \\ &\quad - \gamma \sin(\theta - \psi_+) \cos(\mu r - \psi_-)] \end{aligned}$$

$$\begin{aligned} \tau \dot{\psi}_+ &= \frac{J_1}{1 - \gamma^2} \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) [\sin(\theta - \psi_+) \cos(\mu r - \psi_-) \\ &\quad - \gamma \cos(\theta - \psi_+) \sin(\mu r - \psi_-)], \end{aligned}$$

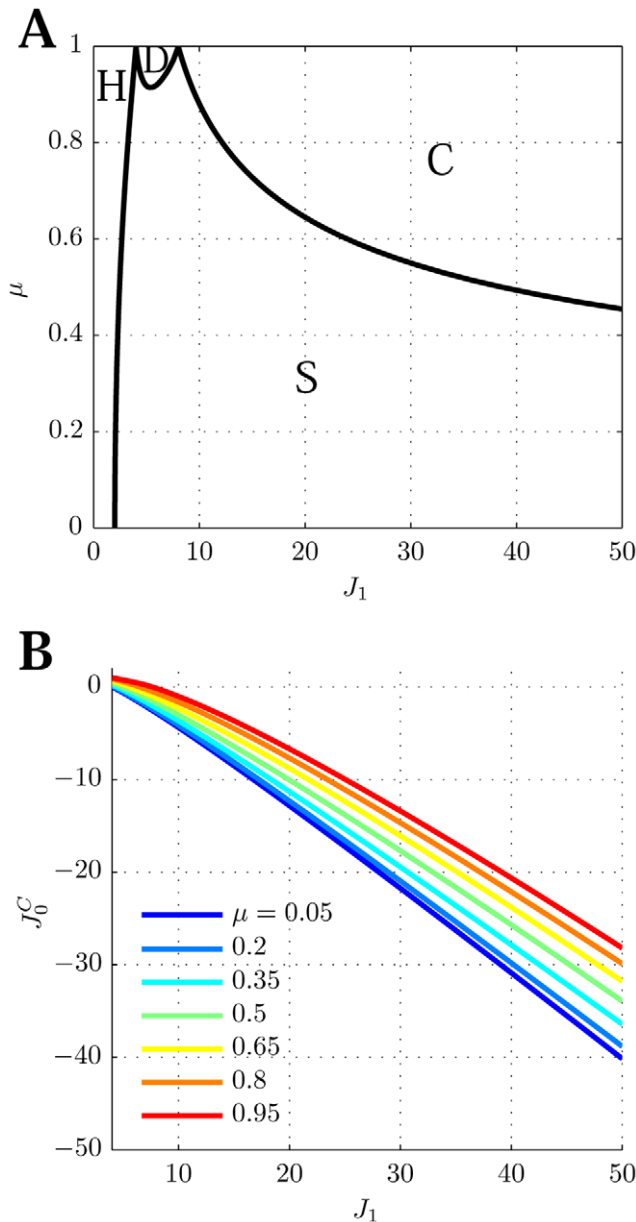
where the function  $g(\theta, r)$  is defined as

$$\begin{aligned} g(\theta, r) &\equiv [\cos(\theta - \psi_+) \cos(\mu r - \psi_-) \\ &\quad - \gamma \sin(\theta - \psi_+) \sin(\mu r - \psi_-) + \sigma]_+, \end{aligned} \quad (6)$$

i.e. the rescaled steady state activity profile Eq. 3. Note that  $\psi_+$  can be eliminated from the right hand sides of the Eqs. 5, rotating the integration variable  $\theta$ . This is possible because there is no spatial dependence in the external input to the network. The first four equations in Eqs. 5 can then be solved independently of the fifth one, since the right hand sides do not depend on  $\psi_+$ . We show in *Methods - Solutions properties* that, once we have the solution for the variables  $(\gamma, \alpha, \sigma, \psi_-)$ , the last equation reduces to  $\dot{\psi}_+ = 0$ . We can thus restrict the analysis to four out of five equations in Eqs. 5. The elimination of one angular degree of freedom is a consequence of the rotation invariant structure of the encoding, and is the hallmark of continuous attractors arising from spontaneous symmetry breaking. On the other hand, the integrals over  $r$  in Eqs. 5 are not over the whole circle and we cannot rotate  $\psi_-$  away.

## Homogeneous solution

Before analyzing the fixed point solutions of the system described by Eqs. 5, we briefly mention an uninteresting region in the parameters space which can be found also in the classical ring model. This region corresponds to the homogeneous solution, i.e. all the neurons in the network are active at a constant level,



**Figure 3. Phase diagram. A:** MF solutions in a network storing two correlated maps. The black curves in the  $J_1, \mu$  parameter space depict the separation between the qualitatively different solutions of the system. *Homogeneous* solution (denoted by H in the panel), all the neurons are active at a constant level. *Single Ring* (S) solution, localized activity in the middle map, the bump can be freely rotated in this map. *Double Ring* (D) solution, pair of solutions localized either in map *A* or *B*. *Cylinder* (C) solution, the network activity is localized in both maps; compared to the single and double ring solutions, there is an additional freedom in the choice of the location of the bump over  $r$  ( $\psi_-$ ). The term cylinder is used because the continuous attractor lives in the space defined by the angle  $\theta$  and a sub-segment of  $r$  (with the exception of  $\mu=1$ , see main text for details). **B:** Amplitude instability. Critical inhibition  $J_0^C$  corresponding to the onset of unstable solutions for varying  $J_1$ . Each curve corresponds to different distances between the maps  $\mu$ . doi:10.1371/journal.pcbi.1000869.g003

and can be obtained from Eq. 2. The expression corresponding to the line of separation in the plane  $(J_1, \mu)$  between the homogeneous solution and the spatially localized bump (see

Fig. 3A, curve surrounding the **H**omogeneous region), is

$$J_1 = \frac{4}{1 + \text{sinc}(\mu\pi)}, \quad (7)$$

where  $\text{sinc}(x) = \frac{\sin(x)}{x}$ . This result is obtained in *Methods - Stability*, see also below.

### Fixed point equations for the localized activity state

Let us start by imposing  $\psi_- = 0$ , a restriction that will be addressed later on. The first three equations at steady state from Eqs. 5 become then equations for the three order parameters  $(\alpha, \gamma, \sigma)$ :

$$\iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) \cos \theta \cos(\mu r) = \frac{1}{J_1} \quad (8)$$

$$\iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) [\sin \theta \sin(\mu r) + \gamma \cos \theta \cos(\mu r)] = 0$$

$$\frac{1}{\alpha} = \sigma - J_0 \iint \mathcal{D}\theta \mathcal{D}r g(\theta, r).$$

The first two equations determine the shape of the bump  $(\gamma, \sigma)$ . Given the map specific modulation in the coupling and the distance between the maps  $(J_1, \mu)$ , we can derive from the first two equations the size of the bump  $\sigma$  and the order parameter  $\gamma$ , representing how close the network representations are to the stored environments *A* and *B*. The last equation gives us the amplitude of the network activity  $\alpha$ , which also depends on the parameter  $J_0$ .

As mentioned in *Results - Phase diagram of the model*, the order parameter  $\psi_+$  can be chosen arbitrarily, due to the rotation invariance of the problem; for simplicity we choose  $\psi_+ = 0$ .

### Amplitude instability

We deal first with the equation concerning the amplitude of the solution. Given that the activity can be rescaled by changing the value of the applied external current  $I$ , we are not interested in actually solving the equation. The only requirement is that  $\alpha \geq 0$  in order for the solution to be meaningful, i.e. no negative amplitudes are allowed. This requirement translates to a constraint on the inhibition  $J_0$ :

$$J_0 \leq \frac{\sigma}{\iint \mathcal{D}\theta \mathcal{D}r g(\theta, r)}. \quad (9)$$

We show with stability analysis (*Methods - Stability*) that the critical value  $J_0^C$ , obtained by choosing the equality in the previous expression, corresponds to the onset of amplitude instability; given a choice for the parameters  $(J_1, \mu)$ , which specifies the bump shape  $(\gamma, \sigma)$ , for values of the inhibition weaker than  $J_0^C$  the solution grows to infinity. This qualitative behavior was present also in the classical ring model.

Fig. 3B shows the values of  $J_0^C$  as a function of  $J_1$  for various choices of  $\mu$ . In order to stabilize the solutions, the inhibition must grow with increasing  $J_1$  and decreasing  $\mu$ . Note that it is

reasonable to consider the previously mentioned homogeneous solution as a bump with maximal size ( $\sigma=1$ ). In this case the critical  $J_0$  can be explicitly computed, and turns out to be  $J_0^C=1$ .

### Single ring solution

Now we focus on the possible solution  $\gamma=0$ . It is easy to see that when  $\gamma=0$ , the second of Eqs. 8 is automatically satisfied due to the symmetry of the integrand in  $\theta$  (and  $r$ ); This means that the solution  $\gamma=0$  exists everywhere in the parameter space.

The steady state activity Eq. 3 with  $\gamma=0$  (and  $\psi_-=0$ , our initial assumption) reads

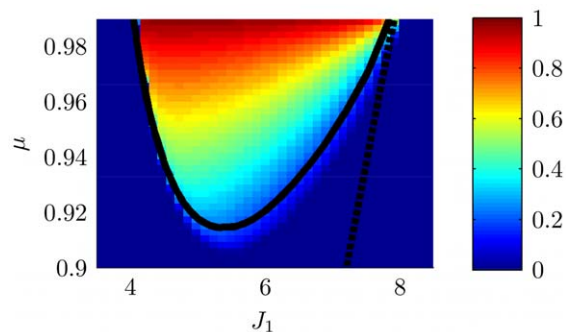
$$m = \alpha I [\cos(\theta - \psi_+) \cos(\mu r) + \sigma]_+, \quad (10)$$

which corresponds to a packet of activity localized in the  $\theta$  coordinate, and modulated in  $r$ , see Fig. 2B for a plot of the activity profile. The remaining fixed point equation can be used to obtain  $\sigma = \sigma(J_1, \mu)$ . We refer to the case ( $\gamma=0, \psi_-=0$ ) as a single ring solution; the ring is spanned by the freedom of choice in the angle  $\psi_+$ . In this regime of activity the network is not able to represent separately the environments  $A$  and  $B$ , but only the middle environment described by  $\theta$ . Even though the solution exists everywhere, it is destabilized in some regions of the parameter space, as shown in the phase diagram (Fig. 3A, Single ring region).

By looking at the maximal bump size  $\sigma=1$ , we can expect to reproduce the curve separating the homogeneous solution from the single ring. Inserting  $\sigma=1$  in the first of Eqs. 8, it is possible in this case to compute explicitly the integral, which in fact yields Eq. 7.

### Double ring solution

In order to find the region of existence of the solutions with  $\gamma \neq 0$ , we can solve numerically Eqs. 8 in the parameters plane ( $J_1, \mu$ ). The result is shown in Fig. 4, where the color code represents  $\gamma$  for a given choice of the parameters. It can be seen that there is only a narrow region of high  $\mu$  (low correlation) and low  $J_1$  where such a solution exists.



**Figure 4. Double ring solutions.** Region in the parameter space ( $J_1, \mu$ ), where the double ring solution  $\gamma \neq 0$  exists. The color correspond to the value of  $\gamma$  which solve Eqs. 8. Due to the symmetry of equations, both positive and negative  $\gamma$  values are allowed. Here only the positive solution is shown, corresponding to a localized solution in map  $B$ . **Black curve:** curve of separation between the null and positive  $\gamma$  solutions, obtained by finding the zeros of Eq. 12. Note that the method used to obtain the curve is more precise than the one used to estimate  $\gamma$ . **Dashed curve:** curve corresponding to a network activity size  $\sigma=0$ . Regions on the left of this curve have  $\sigma > 0$ , which is a limit size of the double ring solutions (reached at  $\mu=1$ ). doi:10.1371/journal.pcbi.1000869.g004

It is important to note that the equations used to find  $\gamma$  are invariant under the symmetry  $\gamma \rightarrow -\gamma$ . This means that both solutions ( $\pm \gamma^*$ ) representing map  $\Theta_A$  or  $\Theta_B$  are possible. The steady state activity profile in this case looks like:

$$m = \alpha I [\cos(\theta - \psi_+) \cos(\mu r) \pm \gamma^* \sin(\theta - \psi_+) \sin(\mu r) + \sigma]_+. \quad (11)$$

Given the freedom of choice for the phase  $\psi_+$ , each of this solutions lives on a ring; we call the solution ( $\gamma \neq 0, \psi_-=0$ ), double ring. An instance of the network activity in this regime is shown in Fig. 2C.

The curve separating representations preferring one of the two maps ( $\gamma \neq 0$ ), and  $\gamma=0$ , can be obtained by expanding the second of Eqs. 8 to first order in  $\gamma$ :

$$\gamma \iint \mathcal{D}\theta \mathcal{D}r H(\cos \theta \cos(\mu r) + \sigma) (\sin^2(\mu r) - \cos^2(\theta) - \sigma \cos(\theta) \cos(\mu r)) = 0, \quad (12)$$

where  $H(x)$  is the Heaviside step function,  $H(x > 0) = 1$  and  $H(x < 0) = 0$ . Dividing by  $\gamma$ , we get rid of the  $\gamma=0$  solution. By finding the zeros of the integral, we select the curve in the parameter space corresponding to the onset of existence of the double ring solution. This curve is shown in Fig. 4. We have found that the stability of the double ring solution coincides, empirically, with the region of existence of such solution (compare the phase diagram in Fig. 3A, Double ring region with Fig. 4).

### Cylinder solution

Finally, we examine the meaning of the equation for  $\psi_-$ , the order parameter linked to the location of the maximum of the bump in  $r$ . We have assumed  $\psi_- = 0$  for simplicity, given that a rotation in the integrands in Eqs. 5 is in general not viable due to the restricted range of integration in  $r$ . Note though, that when the size of the bump  $\sigma$  is small enough, it is possible to perform the rotation without affecting the value of the integrals; the only requirement is that the rotation keeps the bump from touching the boundaries  $r = \pm \frac{\pi}{2}$ .

In *Methods - Solutions properties* we verify that there are no solutions with both  $\gamma$  and  $\psi_-$  different from 0. We can therefore set  $\gamma=0$  in the steady state activity Eq. 3, and impose the activity itself to be zero on the boundary  $r = \pm \frac{\pi}{2}$  to find

$$\cos\left(\mu \frac{\pi}{2}\right) = -\sigma.$$

This equation corresponds to the curve of separation in the plane ( $\mu, J_1$ ) (using the relationship  $J_1 = J_1(\mu, \sigma)$ , Eq. 8) between the single ring solution and a *cylinder* solution (Fig. 3A, curve surrounding the **C** region). In this regime, in addition to the freedom of choice for the location of the bump in  $\Theta$ , the solution is also partially marginal in  $\psi_-$ . The bump can be freely moved on a segment and a circle, defining a cylinder; the activity profile in this case is described by Eq. 4, see an instance in Fig. 2D. This region extends in the high  $J_1$  limit and covers the whole range of correlations.

Despite the fact that each of the maps  $\Theta_A$  and  $\Theta_B$  defines a ring, it shouldn't come as a surprise that the topology of the attractor is a cylinder instead of a torus. The correlation between maps gives rise by definition to a cylinder structure, as can be seen for instance by inspecting Fig. 2B,II. It can be shown that when

$\mu=1$  the cylinder solution degenerates in a torus; the bump of activity can be in any location of the  $(\theta, r)$  coordinates (hence, also in  $(\theta_A, \theta_B)$ ). This regime is linked to the observation of an activity bump simultaneously localized in two environments in network simulations [39], and the study in [28].

### Phase diagram

Fig. 3 summarizes the results obtained so far. When  $J_1$  is low, the only solutions is a constant level of activity which spreads over the whole network (**H**omogeneous region). As  $J_1$  is increased, the interplay between the short range excitation and long range inhibition creates a pattern of localized activity in the middle map  $\Theta$  (**S**ingle ring, see also Fig. 2B) or, if the correlation between maps is small enough, a localized pattern in either  $\Theta_A$  or  $\Theta_B$  (**D**ouble ring, Fig. 2C). Intuitively, the network “remembers” the two maps separately ( $\psi_- = 0, \gamma \neq 0$ , two solutions) if they are weakly correlated ( $\mu > \sim 0.9$ ). When the maps are more similar, the network represents just an average between them ( $\psi_- = 0, \gamma = 0$ ).

The bump size decreases with increasing  $J_1$ . When  $J_1$  is further increased, instead of having a reduced size of the localized activity in just one of the maps, the presence of two stored maps in the synaptic structure and the inhibition  $J_0$  produce a packet of activity which looks localized in both maps (**C**ylinder solution, Fig. 2D).

Three particular values of the distance  $\mu$  deserve a special mention. The case  $\mu=0$ , corresponding to the encoding of two identical maps, can be shown to be identical to the ring model [37], as expected. In particular, besides the homogeneous solution and the amplitude instability region, the system can only exhibit the single ring solution.

The case  $\mu=1$ , corresponding to the encoding of two uncorrelated maps, does not have the single ring regime as a possible solution. The double ring solution in this case is depicted in Fig. 2A, where it can be seen that the bump is perfectly localized in either maps  $\Theta_A$  or  $\Theta_B$ , lacking any spatial tuning in the other map. This is the desired outcome in the “multi-chart” approach of [21].

The third case is  $\mu = \frac{1}{2}$ . We will see in *Results - Morphing maps* that this case is closely related to the behavior of a network storing a morph sequence between two uncorrelated maps. As can be seen in the phase diagram, the double ring solution is not possible in this regime.

How the environment, and the position in the environment, are represented by the network activity? For the single ring (Eq. 10) and the double ring (Eq. 11) solutions, both characterized by  $\psi_- = 0$ , it is evident that the position is coded by the order parameter  $\psi_+$ . The identity of the environment can only be represented with the ambiguity in the choice of the sign of  $\gamma$  when the network operates in the double ring regime.

In the cylinder regime, it is not clear how the information about the environment is represented in the network, since now the solution is described by  $\psi_+$  and  $\psi_-$ . The following Section is mainly devoted to explore the link between the state variable (eventually time-dependent)  $\Psi(t)$  in the active environment, and the behavior of the solution in this novel regime, by introducing a spatially tuned external input.

### Tuned external input

Until now we considered the condition in which the only external input to the network,  $I$ , was steady and uniform. Let us introduce a tuned input, for instance in map  $\Theta_A$  at position  $\Psi(t)$ :

$$I \rightarrow I \cdot (1 + \epsilon I_E(\theta_A - \Psi(t))) = I \cdot (1 + \epsilon I_E(\theta - \mu r - \Psi(t))).$$

For simplicity we assume the shape of the external input to be  $I_E(x) = \cos(x)$ . The parameter  $\epsilon$  measures the strength of the tuned component of the external input as a fraction of the constant baseline  $I$  we adopted so far. In general what we are interested in, and what is experimentally observable, are the tuning curves of the neurons i.e. their profile of activity as a function of the input angle in the active environment. It is easy to see the effect on the dynamics of the order parameters (Eqs. 5) when the location specific external current is inserted in the original dynamics for the network activity, Eq. 2. The dynamics keeps the same form as in Eq. 5, with the exception of the threshold-linear term in  $g(\theta, r)$ , which now reads

$$g(\theta, r) \rightarrow [\cos(\theta - \psi_+) \cos(\mu r - \psi_-) - \gamma \sin(\theta - \psi_+) \sin(\mu r - \psi_-) + \sigma + \frac{\epsilon}{\alpha} \cos(\theta - \mu * r - \Psi(t))]_+ \quad (13)$$

where  $\mu^* = +\mu$  correspond to the choice of map  $\Theta_A$  in the input, and  $\mu^* = -\mu$  for map  $\Theta_B$ .

### Tuning curves

With the input at a constant location  $\Psi(t) = \Psi$ , one can see that a solution of Eqs. 5 for the single and double ring regime ( $\psi_- = 0$ ), is  $\psi_+ = \Psi$ , i.e. the input pinpoints the location of the bump. This implies that, assuming a weak tuned input  $\frac{\epsilon}{\alpha} \ll 1$ , the tuning curve of a neuron  $(\theta, r)$  can be written in the single and double ring regime (from Eqs. 10,11) as

$$m(\theta, r | \Psi) \approx \alpha I [\cos(\theta - \Psi) \cos(\mu r) + \sigma]_+ \quad (14)$$

and

$$m(\theta, r | \Psi) \approx \alpha I [\cos(\theta - \Psi) \cos(\mu r) \pm \gamma \sin(\theta - \Psi) \sin(\mu r) + \sigma]_+ \quad (15)$$

respectively. The tuning curve in the single ring regime has a maximum for  $\Psi = \theta$  (hence  $\theta$  is the preferred angle for a neuron  $(\theta, r)$ ), independently of which map  $\mu^* = \pm \mu$  is being used in the external input, as can be seen from Eq. 14. This implies that each neuron has identical tuning curves in both environments, and that the preferred angle of a neuron does not coincide with either the assigned  $\theta_A$  or  $\theta_B$  but with their average.

For the double ring regime, the preferred angle assumes the form (maximizing Eq. 15 in  $\Psi$ )

$$\theta - \arctan(\pm \gamma \tan(\mu r)).$$

In this case each neuron has two different tuning curves according to the map used in the external input. The preferred angles coincide with the assigned ones  $(\theta_A, \theta_B)$  only when the stored maps are uncorrelated ( $\mu=1$ , hence  $\gamma = \pm 1$ ).

In the cylinder regime ( $\gamma=0$ ,  $\psi_-$  not necessarily 0), a solution for Eqs. 5 in presence of a tuned input is  $\psi_+ + \psi_- \frac{\mu^*}{\mu} = \Psi$ . For an input in map  $\Theta_A$ ,  $\mu^* = \mu$ , the tuning curve would then be proportional to (from Eq. 4)

$$[\cos(\theta - \Psi - \psi_-) \cos(\mu r - \psi_-) + \sigma]_+.$$

Note that the dependence on  $\psi_-$  means that the external stimulus does not determine completely the network activity, in contrast to what happens in the previously examined regimes. Neurons that



respond maximally to the tuned input are then  $\theta = \Psi + \psi_-$ , and  $\mu r = \psi_-$ , hence  $\theta - \mu r = \theta_A = \Psi$ . This means that the tuned external input pinpoints the location of the bump maximum in map  $\Theta_A$  but the bump is free to stabilize anywhere along the other map given the freedom of choice in  $\psi_-$  (see activity example in Fig. 2D). If several randomly selected external locations  $\Psi$  in one of the maps are presented to the network, once at time and starting from random initial conditions, the tuning curves would be an average over  $\psi_-$ :

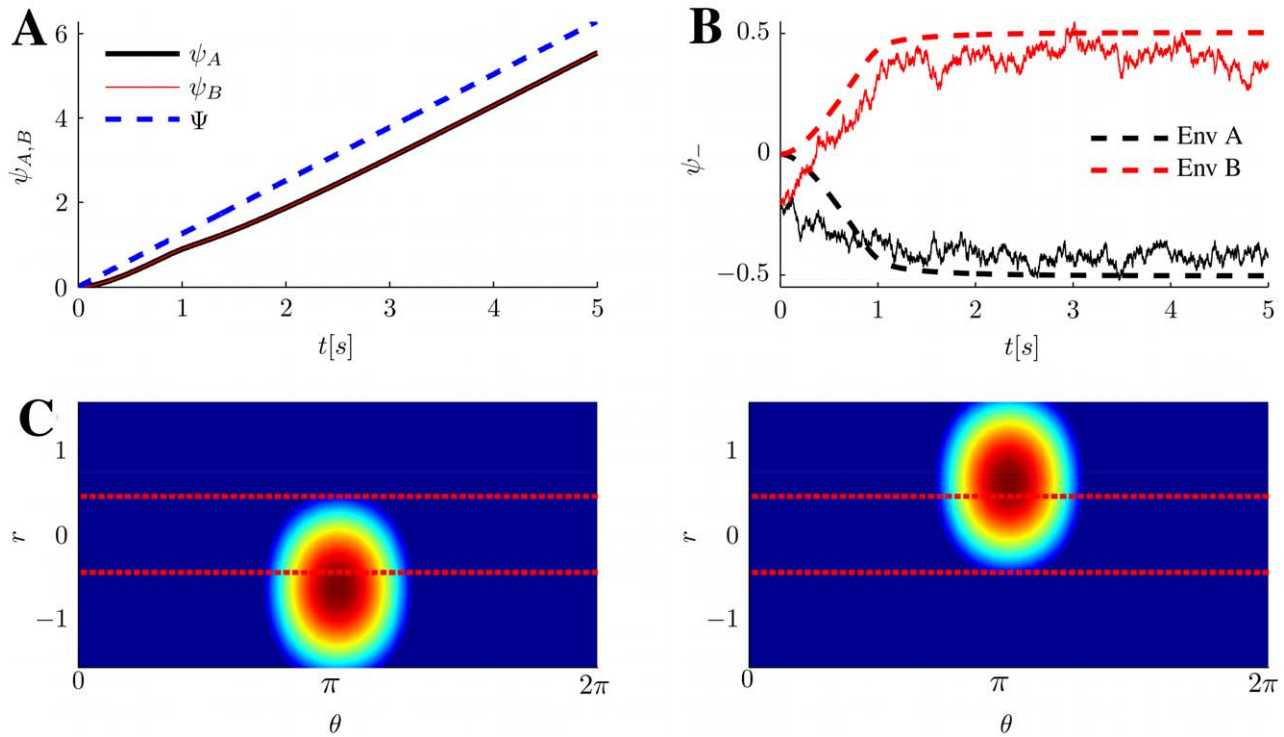
$$m(\theta, r | \Psi) \approx \alpha I \int d\psi_- [\cos(\theta - \Psi \mp \psi_-) \cos(\mu r - \psi_-) + \sigma]_+,$$

where the allowed range for  $\psi_-$  is  $\left[-\left(\frac{\pi}{2} - \arccos(-\sigma)\right), \left(\frac{\pi}{2} - \arccos(-\sigma)\right)\right]$ , see *Methods - Solutions properties*. The cylinder regime extends the region of existence of two tuning curves per neurons to an higher correlation between the stored maps; the difference is that the coding becomes *unreliable*: during a single exposure to a given value of the input angle  $\Psi$ , a neuron could remain silent even if its average tuning curve would predict a response.

When the representation refers to the location in an environment, it is natural to think about a smoothly varying location  $\Psi$ . With a moving input like  $\Psi(t) = v \cdot t, v > 0$ , the tuning curve depends as before on which map is stimulated, but in a novel

way. Assume for simplicity to start from a  $\psi_- = 0$  initial condition, corresponding to  $(\psi_A = \psi_B)$ . A moving input in the map  $\Theta_A$  would tend to move the bump along that map (i.e. increase the  $\psi_A$  of the solution), while keeping  $\psi_B$  constant (hence the bump will move to  $r < 0$ ). This movement is possible only until the bump reaches the part of configuration space not occupied by neurons due to the distance between maps  $\mu$ , see Fig. 2D. At that point, the bump will start to move equally along  $\theta_A$  and  $\theta_B$ , maintaining  $\psi_- < 0$ , which is proportional to  $\psi_B - \psi_A$ , and increasing  $\psi_+$  (proportional to  $\psi_B + \psi_A$ ). A similar scenario, but with  $\psi_- > 0$ , is obtained when stimulating the map  $\Theta_B$ .

If the size of the bump is sufficiently small, this effect has dramatic consequences. The small bump will move along neurons with  $r < 0$  when a moving stimulus is presented in environment  $A$ , and viceversa neurons with  $r > 0$  will be active only when the moving stimulus is presented in environment  $B$ . As a consequence, neurons will essentially just have a tuning curve (or field), only in one map, and will be silent in the other one. We refer to this phenomenon as dynamical pattern separation (see Fig. 5 for an example). The separation of the activity patterns is essentially a dynamical phenomenon, dependent on the history of the inputs. The figure shows also the robustness of the dynamical pattern separation behavior to the addition of Gaussian  $\delta$ -correlated noise in the external current (see *Methods - Numerical Methods*). Note that neurons characterized by  $r \sim 0$  (i.e.  $\theta_A = \theta_B$ ), will have tuning curves in the same location. The number of neurons with tuning curves in both environments grows with the size of the bump.



**Figure 5. Network representations in the cylinder regime.** **A:** MF time-course of  $\psi_{A,B}$ , when a moving input localized around  $\Psi(t) = \frac{2\pi t}{5}$  in either map  $A$  or  $B$ , is fed to the network. The location of the maximum of the external input in the stimulated map is shown with a dashed blue line. The order parameter  $\psi_{A,B}$  tracks the external input with a delay. **B:** The order parameter  $\psi_-$ , giving the position of the bump in  $r$ , with the same moving inputs. Depending on the stimulated map, the bump ends up in different positions, orthogonalizing the representations. Noisy curves from a simulation including a zero average white noise term in the external input, with  $s.d. = \epsilon_N I$ . **C:** A snapshot of the network activity at a given time  $t$  and corresponding angle  $\Psi(t)$ , when the stimulated map is  $A$ , left or  $B$ , right. Neurons with  $r$  values differing enough from 0 (depending on the size of the bump) will exhibit tuning curves only when stimulated in one of the two environments (above/below red dashed lines). The parameters for all the panels are ( $J_1 = 35, J_0 = -30, \mu = 0.8, \epsilon = 0.05$ ). doi:10.1371/journal.pcbi.1000869.g005

Note though that by changing the sign of the velocity in the moving input, the behavior would reverse; neurons with positive (negative)  $r$  would be active during a stimulation in map  $\Theta_A$  ( $\Theta_B$ ). In order to maintain the dynamical pattern separation and the analogy with place coding, one could think about two circular environments, as we did so far, with the additional constraint that the environments can only be traveled, for instance, in the counter-clockwise direction (CCW). As an alternative, the two environments may be thought as the same circular arena, but traveled clockwise (CW, environment  $A$ ) and CCW ( $B$ ); this interpretation would give rise to place fields with directional selectivity (see Discussion).

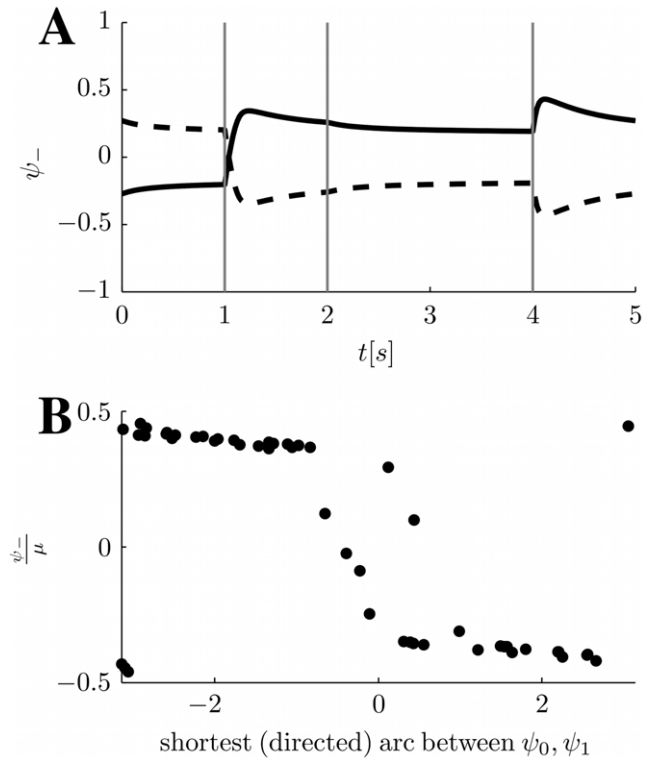
The dynamical pattern separation is basically dependent on the history of the input (positive or negative velocity), in addition to the identity of the map used in the stimulation. This history dependence is present also for non smooth time-dependent stimuli, as for instance the sequential presentation of stimuli with an intervening delay period. In this case the history dependence gives rise to a memory effect: the current location of the bump following a stimulation depends on the location attained after the previous stimulus presentation. Let us consider a basic example of this phenomenon, where the tuned external input is always presented in map  $\Theta_A$ . Consider for simplicity the state of the network being characterized by  $\psi_A = \psi_B = \psi_0$ , as a result of the presentation of stimulus  $\psi_0$  sometime in the past. If we now present a stimulus  $\psi_1$ , the bump will move, through the shortest arc on the map, to the new location  $\psi_A = \psi_1$ . Depending on the stimuli, this movement can happen in two ways. If the shortest arc from  $\psi_0$  to  $\psi_1$  is directed CCW, the bump will move with a positive velocity  $v > 0$  and will end up being located in the  $r < 0$  region (as we previously saw in the case of moving tuned input). If the shortest arc is directed CW, then the movement will happen with a negative velocity, and the final location of the bump will be in the  $r > 0$  region. Hence, by looking at the activity resulting from the presentation of  $\psi_1$ , we know whether the shortest way on the ring to it from  $\psi_0$  is CW or CCW. A similar result can be obtained if the stimulus presentation alternates between map  $\Theta_A$  and  $\Theta_B$ .

If we vary the manifolds on which the maps live, for example to segments instead of circles, the history dependence changes accordingly. For instance, on segments the activity would give us information about the second stimulus being greater/smaller than the first one (see Discussion). In the next section we present a simple (albeit artificial) delayed discrimination task which the network can perform by exploiting the memory effect.

### An application of the memory effect

Let us suppose to have a screen with a circle on it. A first stimulus (a dot) appears on the circle at some random location (described by an angle,  $\psi_0$ ), for the duration of 1s. This first stimulus is then removed for a delay period of 2s. Then a second stimulus appears at another random angle  $\psi_1$ ; the subject's task is to determine whether the shortest path on the circle from angle  $\psi_0$  to  $\psi_1$  is CW or CCW. The basic idea is that it is enough to look at the network activity (location of the bump in the  $r$  axis), to determine the relationship between the first and the second stimulus (see *Results - Tuned external input* for a description of the idea).

To test the ability of the network to solve this task, we numerically solve the dynamics for the order parameter (*Results - Phase diagram of the model*) with an external input (*Results - Tuned external input*) mimicking the presentation of the stimuli, for a sequence of 50 trials. We used no inter-trial interval, i.e. the presentation of the second stimulus in the  $k$ -th trial is immediately followed by the presentation of the first stimulus in trial  $k + 1$ . The



**Figure 6. Solving the shortest path task. A:** Two instances of MF dynamics from a network storing two correlated maps, during the presentation of two angles ( $\psi_0, \psi_1$ ) with an intervening delay. The network operates in the cylinder regime. The plot shows the time course of the order parameter  $\psi_-$ ; 0–1s: pre-stimulus period, random starting  $\psi_+, \psi_-$ , no tuned input. 1–2s: first stimulus (presented in map  $\Theta_A$ , at  $\psi_0 = 0$  (solid and dashed). 2–4s: delay period, no tuned input. 4–5s: second stimulus  $\psi_1 = \frac{3\pi}{2}$  (solid) and  $\psi_1 = \frac{\pi}{2}$  (dashed). By looking at the position of the bump in  $r$  ( $\psi_-$ ) during the second stimulus presentation, it is possible to decide whether the shortest path between the first and second stimulus is CW ( $\psi_- > 0$ ) or CCW ( $\psi_- < 0$ ). **B:** Sequential repetition of the task, 50 trials. The location of the bump maximum along  $r$  is plotted vs the oriented distance on the circle between the second and the first stimulus. Parameters ( $\mu = 0.7, J_1 = 30, J_0 = 30, \epsilon = 0.1$ ). doi:10.1371/journal.pcbi.1000869.g006

time courses of the bump location on the  $r$  axis ( $\psi_-$ ) in two example trials for which  $\psi_0 = 0$ , are shown in Fig. 6A. When looking at the location of the bump in the  $r$  axis at the end of a trial, there is a clear difference between the two cases of shortest CW, corresponding to positive  $\psi_-$  (in the specific example  $\psi_1 = \frac{3\pi}{2}$ ), or CCW arcs ( $\psi_- < 0$ , where  $\psi_1 = \frac{\pi}{2}$ ). Fig. 6B shows that the bump location at the end of trial, can be used to easily discriminate between the two possible answers (except for the cases in which the first and second stimuli are relatively close to each other). Note that this result has been obtained without any activity reset to new initial conditions during the inter-trial intervals.

### Morphing maps

How do the results described so far change when, instead of storing just two correlated maps, the network encodes a sequence of maps gradually morphed between two uncorrelated ones? Let us start by constructing two random uncorrelated maps,  $\Theta_{-1}$  and  $\Theta_1$ . We would like to define the intermediate maps as gradual rotations between the two extreme ones; since we are dealing with circles, the rotation should be performed along the shortest arc

between  $\theta_{-1}$  and  $\theta_1$  (see Eq. 21, *Methods - Inverse transformation*). We assume here to have already transformed the variables in such a way that we can write directly

$$\theta_\zeta = \frac{1-\zeta}{2}\theta_{-1} + \frac{1+\zeta}{2}\theta_1 \quad (16)$$

where  $\zeta \in [-1,1]$  indexes the maps along the morph sequence. Hence a neuron with label  $\theta_{-1}$  in the first map, will rotate along the sequence to its location  $\theta_1$  on the last map, following the shortest path on the circle. With this choice of the morphing procedure, each neuron is still characterized by just two quantities, its labels in the extreme maps.

We store the whole morph sequence by a superposition of the synaptic structures generated in each map separately, as for the case of two correlated maps previously described. For the sake of analytical tractability, we study the resulting coupling in the limit  $p \rightarrow \infty$

$$\frac{1}{p} \sum_{\zeta} W(\theta_\zeta - \theta'_{\zeta}) \rightarrow J_1 \int_{-1}^1 d\zeta \cos(\theta_\zeta - \theta'_{\zeta}) + J_0. \quad (17)$$

Introducing the definition of two uncorrelated maps (Eq. (1) with  $\mu=1$ ) into Eq. (16), we can rewrite the angles in the intermediate maps as  $\theta_\zeta = \theta - \zeta r$ . We can now integrate Eq. (17)

$$W(\theta - \theta', r - r') = J_1 \cos(\theta - \theta') \text{sinc}(r - r') + J_0. \quad (18)$$

Making use of the Euler formula for the *sinc*() function

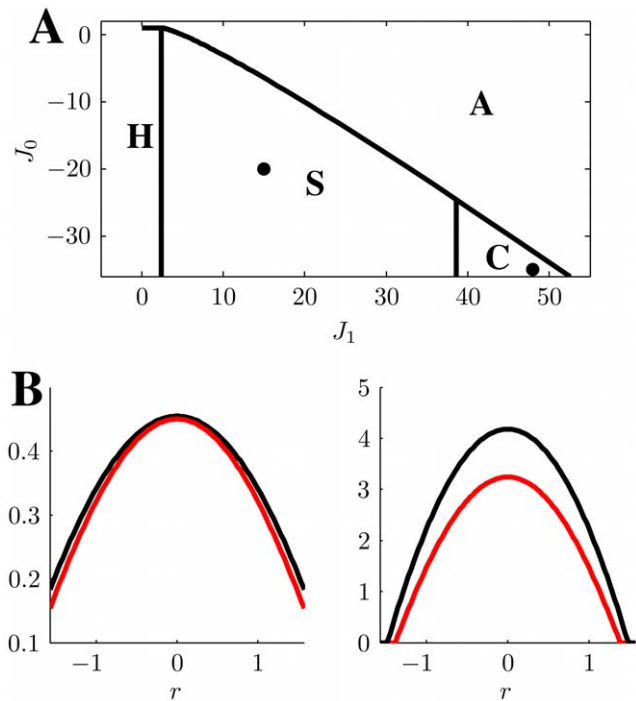
$$\text{sinc}(x) = \prod_{k=1}^{\infty} \cos\left(\frac{x}{2^k}\right)$$

it is possible to derive

$$\text{sinc}(x) = \lim_{n \rightarrow \infty} \frac{1}{2^{n-1}} \sum_{k=1}^{2^{n-1}} \cos\left(\frac{x}{2^{n-k+1}}\right).$$

The first term of the infinite product in the Euler formula, or the first term in the limit sum, gives us  $\cos\left(\frac{r-r'}{2}\right)$ . Comparing the coupling in Eq. (18), and the one derived for two maps, Eq. (2), we see that to first order, the synaptic coupling induced by the storage of the whole morph sequence, is equivalent to the storage of two correlated maps with  $\mu = \frac{1}{2}$ .

In Fig. 7, we compare the network activity generated by the approximated coupling and the full result of Eq. 18, when the external input is constant. The results are qualitatively similar but the full morph case reaches the cylinder regime for lower  $J_1$  compared to the  $\mu = \frac{1}{2}$  case. Note that the network storing the morph sequence shows the same dynamical pattern separation observed in the two maps case (Fig. 8), see next Section for a simulation example in a finite network with a finite number of encoded maps. The important difference, is that while the very correlations between maps forced the absence of neurons with certain labels, hence constraining the permissible region for a marginal solution in  $\psi_-$ , here the neurons cover the entire  $(\theta_A, \theta_B)$  space. The result is purely due to the process of storing multiple maps along the morph sequence.



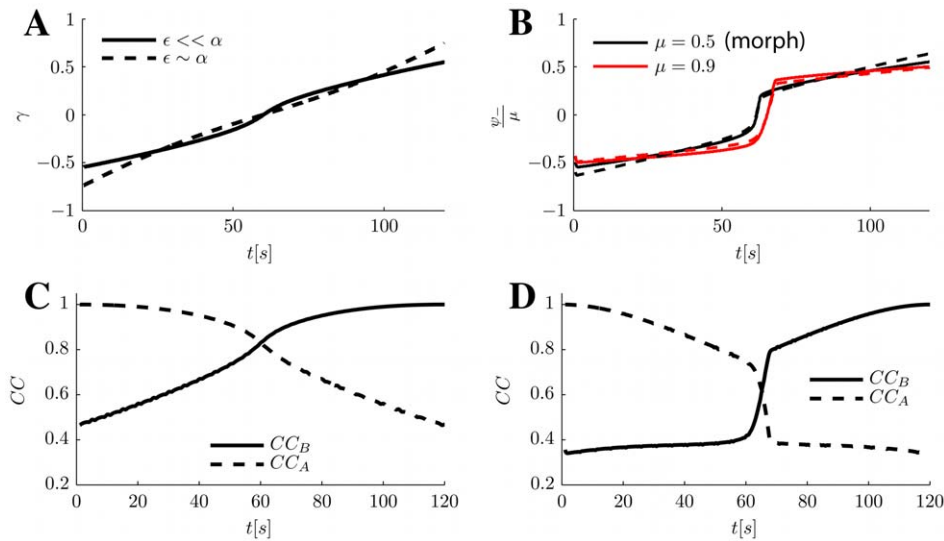
**Figure 7. Comparison between whole morph sequence storage and two correlated maps with  $\mu = \frac{1}{2}$ .** **A:** Phase diagram of the model storing two correlated maps with  $\mu = \frac{1}{2}$ , in the plane  $(J_1, J_0)$ . Note that with this value of the distance, the double ring regime is not achievable. (A) region correspond to amplitude instability. **B:** Slice of the steady activity along  $r$ , in correspondence of the maximum in  $\theta$ . Left: parameters  $(J_1=15, J_0=-20)$  (dot in the S region). Right:  $(J_1=48, J_0=-35)$  (C region); black curve, two stored maps; red curve, whole morph sequence stored. The bumps do not cover the whole range in both cases. External current  $I=1$ . doi:10.1371/journal.pcbi.1000869.g007

This morphing algorithm also yields a way of stimulating the network with positions in environments intermediate between *A* and *B* (with or without the intermediate maps encoded in the network). It is sufficient to use as a place specific input what we had in Eq. 13

$$\frac{\epsilon}{\alpha} \cos(\theta + \mu^* r - \Psi(t))$$

This time, the suitable range for the variable indexing the morph sequence ( $\mu^*$ ) is the whole range  $[-1,1]$ , if using  $\mu = \frac{1}{2}$  as an approximation for the morphed case, or the restricted  $\mu^* \in [-\mu, \mu]$  if the network is storing just two correlated maps. In the reference frame defined by the original coordinates  $(\theta_A, \theta_B)$ , a change in the stimulated environment  $\mu^*$  corresponds to a rotation of the axis representing the maximal external input; between a vertical axis (stimulus localized in environment *A*, to an horizontal axis, stimulus localized in environment *B*).

In the experiment of [32], the rat is trained until it develops two separate place coding for a single arena with different light configurations (representing two distinct environments). The advantage of this setup is that it allows, for instance, to slowly morph the light configuration between the two environments familiar to the rat. The experimental results shows a sharp transition around the middle of the light morphing (lasting



**Figure 8. Representation switch along the morph sequence.** Slow morph experiment between two reference environments, over  $T = 120s$  of simulation. The identity of the environment varies as  $\mu^*(t) = \mu_A^* + \frac{t}{T} \cdot (\mu_B^* - \mu_A^*)$  (linear change from environment  $A$  to  $B$ ), where  $\mu_{A,B}^*$  is either  $\mp 1$  for the morph sequence, or  $\mp \mu$  for two correlated environments. The input location is also time-varying,  $\Psi(t) = \frac{2\pi t}{5}$  (a full circle in  $5s$ ). **A:** Double ring solution. Order parameter  $\gamma$  with time, both for weak  $\epsilon \sim 0.1\alpha$  (solid curve) and strong external input  $\epsilon \sim \alpha$  (dashed curve). The result does not depend strongly on the amplitude of the external input. Parameters ( $\mu = 0.95, J_1 = 6, J_0 = -5, \epsilon = 0.05, 0.5$ ). **B:** Cylinder regime. Order parameter  $\frac{\psi_-}{\mu}$  (location of the bump maximum in  $r$ ) for the approximated storage of the morph sequence  $\mu = \frac{1}{2}$  and for two correlated maps with  $\mu = .9$ , both for weak and strong input. Note how the sharpest transition, is exhibited by the weakly correlated maps. For the cylinder regime, the result is only mildly dependent on the strength of the external modulation. Parameters ( $\mu = \frac{1}{2}, J_1 = 50, J_0 = -50, \epsilon = 0.3, 3$ ) and ( $\mu = 0.9, J_1 = 18, J_0 = -15, \epsilon = 0.15, 1$ ). **C:** Instantaneous correlation coefficients between the double ring regime network activity during the morphing procedure, and network activity in environment  $A$  (dashed curve) or  $B$  (black curve). Simulations with weak external input. **D:** As in C, for the cylinder solution. Note how the sharp transition is delayed compared to  $\frac{T}{2}$ .  
doi:10.1371/journal.pcbi.1000869.g008

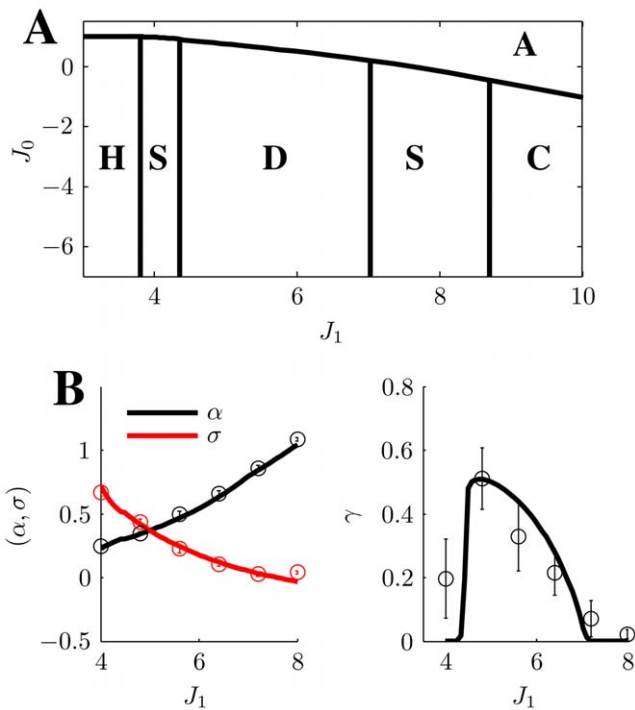
$T \sim 120s$ ) between the place representation in light configuration  $A$  and  $B$ . A link to these experimental results is provided by the use of time-varying external environment  $\mu^* = \mu^*(t) = \mu_A^* + \frac{t}{T} \cdot (\mu_B^* - \mu_A^*)$ , where  $T$  represents the duration of the morphing and  $\mu_{A,B}^*$  denote the upper and lower bounds of the range. An example usage of this protocol is shown in Fig. 8 for the approximated whole morph sequence storage, for two slightly correlated maps in the cylinder region of the parameter range and for the double ring regime. For each run we show the dynamics of the relevant order parameter for the regime under consideration,  $\gamma$  for the double ring case and  $\psi_-$  for the cylinder solution. In addition, we numerically solve the dynamics for a moving stimulus in either environment  $A$  or  $B$ . We use this as a reference for computing, at each time step, the correlation coefficient between the network activity during the morphing protocol and the activity in the fixed environment. The transition is sharpest for the storage of two slightly correlated maps. Note that similar results would be obtained by testing the network separately in each environment of the sequence (see e.g. [31]). The sharp transition is maintained when increasing the amplitude of the external tuned input, because a small tilt in the tuned input towards either map  $A$  or  $B$  is sufficient to generate the dynamical pattern separation described in the previous Section. The transition in the cylinder regime occurs few seconds later than the one occurring in the double ring regime, which in turn happens in the middle of the morphing ( $\frac{T}{2}$ ). This delay is due to the time required for the bump to move from

the region of  $r < 0$  to  $r > 0$ , or viceversa (see also Fig. 5B). This result could be compared with the experimental results of [32]. The delay does not occur when testing the network in separate environments along the morph sequence.

There are two additional observations to be made (data not shown). The first one is related to the sharpness of the transition in the double ring regime; by further reducing the amplitude of the external input, the mean-field dynamics can produce a sharp transition between the environments representations, which is also delayed compared to the middle of the morphing period. The delay gets longer as the external input gets weaker, in extreme cases it happens just before the end of the morphing procedure. This sharp and delayed transition is not observed in microscopic simulations with up to 2000 neurons, since the weak input is not able to overcome the local inhomogeneities in which the bump is trapped (see e.g. [18]). It is possible that in larger networks the transition can be observed. The fine-tuning of the external input strength required to have the transition around the middle of the sequence, makes the double ring regime a weaker candidate explanation for the experimental results of [32] compared to the cylinder regime.

The second observation concerns the dependency of the transition parameters on the velocity of the moving external input. We have noticed that the transition becomes smoother and closer to the middle as the velocity of the simulated animal is reduced. The details of the transition in a realistic setting would depend on the velocity history of the animal.





**Figure 9. Comparison between MF solutions and microscopic network dynamics.** **A:** Phase diagram from MF analysis for  $\mu=0.95$ . **B:** Order parameters from the numerical solution of MF equations (solid curves) with varying  $J_1$ ; estimated order parameters from simulation results; dots with error bars, average and s.d. over 6 simulations with different realizations of the neural maps.  $N=5000$  neurons.  $J_0 = -5$ . doi:10.1371/journal.pcbi.1000869.g009

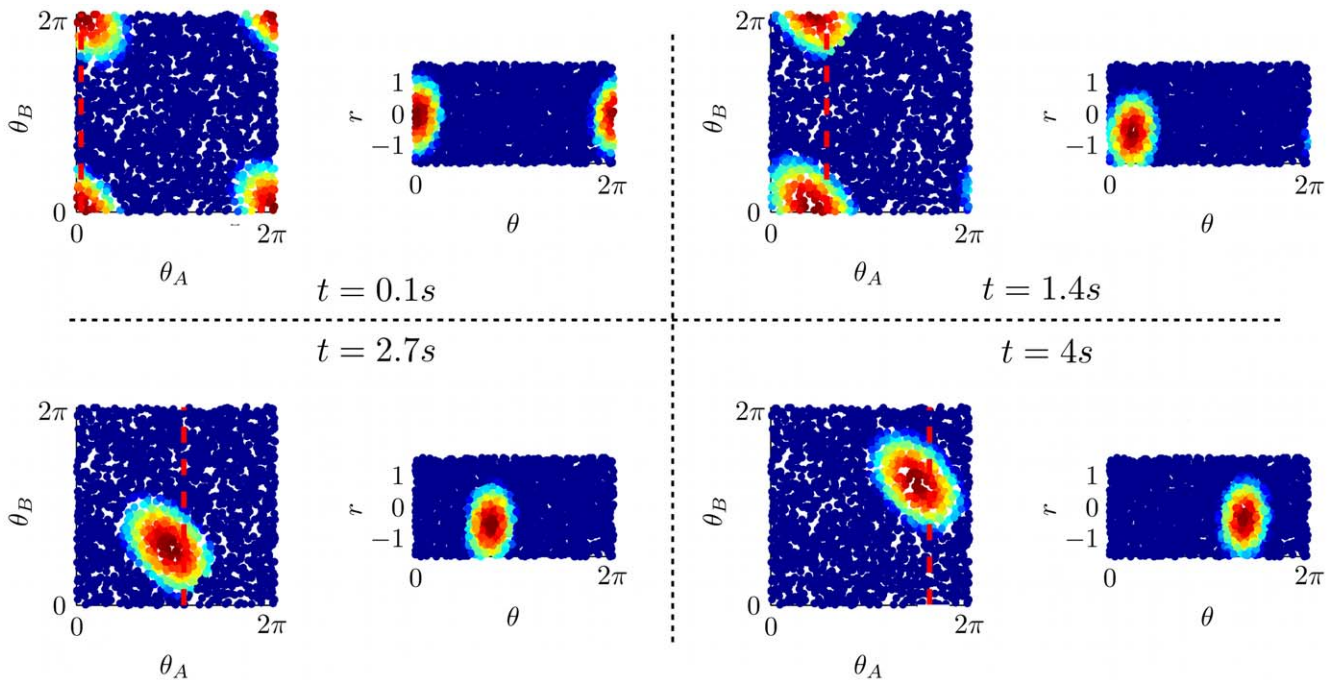
Comparison with simulations

In order to verify that the results obtained in the previous Sections are not artifacts coming from our assumptions of having an infinite number of neurons (and maps, referred to the morphing procedure) we compare some of the MF predictions to simulations of networks with a finite number  $N$  of neurons. Each neuron is assigned a random pair of labels  $(\theta^{(i)}, r^{(i)})$ , for the  $i$ -th neuron), from which we create either two maps with distance  $\mu$ , or a finite number  $p$  of maps  $(\theta_{c_k}^{(i)})$ , for the  $k$ -th map) along the morph sequence between two uncorrelated references (see *Methods - Numerical Methods*).

In Fig. 9 we compare the order parameters from MF and estimated from simulations, at a fixed value of the distance between the maps and inhibition. Varying  $J_1$ , the solution goes through the double ring, single ring and cylinder regime. The order parameter  $\gamma$  is particularly sensitive to the finite size of the network (and the randomized maps, see [18]).

Fig. 10 shows the time evolution of a network storing few maps from a morph sequence. This is the best example to show dynamical pattern separation at finite size, since it is less intuitive than the case of two correlated maps. From an arbitrary initial position, the bump of activity starts moving first towards negative  $r$  (increasing angles in map  $A$ ), then along increasing  $\theta$  without changing its location in  $r$ . Note that, despite the presence of neurons everywhere in the  $(\theta_A, \theta_B)$  plane, the bump moves along an invisible barrier resulting from the storage of the morph sequence.

We have also verified that all the qualitative behaviors, number and type of solutions, unreliable coding, dynamical pattern separation and memory effect, are maintained when moving from maps on rings, to segments (either two correlated maps or morphed), as studied e.g. in [18,37] for the single map (data not



**Figure 10. Dynamical pattern separation for a network storing a finite number of maps along the morph sequence.** Sequence of snapshots of the network activity, in  $(\theta_A, \theta_B)$  and  $(\theta, r)$  coordinates. The network with  $N=2000$  neurons is storing  $p=11$  maps, equally spaced along the morph sequence. The external input is localized in map  $A$  and evolves as  $\Psi(t) = \frac{2\pi t}{5}$ . Each dot represent a neuron, with color coded activity. The dashed red line represents the maximum of the external tuned input. As the activity evolves in time from a random configuration, the subset of active neurons moves towards the negative  $r$  region and stays there moving only along  $\theta$ . The parameters are  $(J_1 = 30, J_0 = -30, \epsilon = 0.1)$ . doi:10.1371/journal.pcbi.1000869.g010



shown). Instead of having neurons arranged on a cylinder in the  $(\theta_A, \theta_B)$  coordinates, as for the ring case (see e.g. Fig. 2B,II), the geometry resulting from two correlated linear maps would be an infinite strip. A strong enough map-specific interaction would produce a bump localized in both maps. An external moving input in one of the maps would move the bump on the strip up to the boundary, and then the bump would crawl along such boundary. Depending on the direction of the moving input or the identity of the stimulated environment, the bump can settle either in “upper” or “lower” part of the strip as in the cylinder regime.

## Discussion

We have studied a continuous attractor network model storing a pair of correlated maps. The storage of a morph sequence between two uncorrelated maps falls in this class of model, since it is approximately equivalent to the storage of two strongly correlated maps. The other relevant parameter for describing the possible network behaviors, beside the correlation between the maps, is the strength of map-specific interaction between neurons.

The analysis of the solutions of the system with a weak tuned external input, reveals several interesting behaviors. When the correlation between the maps is weak, neurons have two different tuning curves corresponding to the stimulus presentation in different maps. The representation is reliable, in that the single neuron response is consistent between presentations. This is the operating regime which is usually considered useful in place coding applications.

For higher correlations between the maps and weak map-specific interactions, each neuron possesses only one tuning curve, irrespectively of the stimulated map. In contrast to the previous regime, this one is rendered useless by the inability to represent fully the state of the external world, i.e. the identity of the environment in the context of place coding analogy.

We find another, novel regime for strong interactions and for *any* amount of correlation between maps. The surprising aspect of this regime is that the state of the world does not uniquely determine the state of the network; there is an additional degree of freedom in the network representation.

To a closer look, this additional freedom found in the novel regime is rich of consequences. When the external input location is randomly varied between presentations in one map, we can define the response of a neuron to a particular location as an average of the neuron activity over external input presentations in that location. In this context each neuron has different tuning curves relative to the different maps used in the stimulation, but the price to pay is unreliable coding; a neuron which should be active during a particular state of the world, could remain silent.

When the location of the external input changes smoothly in time on one map, some neurons develop a selectivity to the direction of change. When the increase happens on the other map, another subset of neurons fires. The overlap between the two subsets may be arbitrarily small, depending on the parameters choice. Neurons active in both maps would have tuning curves around similar values of the external input location. We refer to this phenomenon as dynamical pattern separation. There is an ambiguity in the network representation, due to the fact that the subset of neurons activating with the increase of the external location in map *A*, will also activate with a decrease of the location in map *B*. There are three possible experimental contexts in which this ambiguity does not arise.

A simple experimental context would arise if the input is tuned in only one of the two maps and the only parameter changing is the location of the external input. Given some state variable, like

size and orientation of objects, or frequency of sound waves for instance, our model would produce respectively tuning for expansion/contraction, CW/CCW rotation and upward/downward frequency sweeps (all experimentally observed, see e.g. [40,41]).

Our model provides a unique way for producing selectivity for the *direction of change* of a state variable, given a selectivity for the variable itself. Both kind of responses give rise to another interesting phenomenon: The current representation of the state of the world is influenced by the preceding one, even with an intervening delay. It is possible to read out from the network the direction of change of the state variable. This property may be exploited when solving delayed discrimination tasks (see [42] for data analysis and modeling in terms of remapping for a somatosensory discrimination task).

A second experimental context is related to place coding; the two environments should be considered as two distinct circular arenas which can be traveled only in one direction. Experimental observations show that when an animal is exposed to two environments, the majority of place cells have a place field in only one of the two environments (see e.g. [43,44]). A possible experiment to test the model would consist in training the animals in two well differentiated environments. After measuring the distance between preferred locations for neurons having tuning curves (place fields) in both environments, one could train the animals in intermediate environments, which would correspond to the storage of the morph sequence in the model. For the novel regime of the model, the disappearance of the place fields in one of the environment would be predicted for neurons with very different preferred locations, and the remaining fields will converge to a common representation. Alternatively the training could be performed by using the initial two environments, and then slowly changing them across several training days to increase their similarity. This would correspond to the storage of two correlated environments.

A third experimental context is related to direction selectivity in place cells. Animals trained to shuttle back and forth in a one-dimensional track (a segment or a circle), have place cells showing selectivity to the direction of motion. For instance a cell could be active in a certain region of the circular environment when the animal is moving clockwise, while being completely silent when the animal moves counterclockwise. The link with our model is provided by the simple observation that the same 1D track, but walked in opposite directions, correspond to two different environments. Dynamical pattern separation would produce directional selective neurons, while a neuron having place fields in both environments would have similar preferred locations. In [45], place cells recorded from rats trained in a circular environment indeed showed bi-directional place fields in similar locations. There was however a systematic bias in the difference between the preferred locations in the CW and CCW directions of the majority of the bi-directional cells: place fields were displaced backward with respect to the direction of motion of the animal. We believe that this result, termed by the authors “prospective misalignment”, could be obtained in the context of our model in more than one way. One possibility is the introduction of an asymmetry in the synaptic connections (following [46]), with the asymmetry determined by the emerging direction selectivity of the neurons. The spread of activity due to the asymmetry would activate neurons earlier compared to the symmetric case, reproducing the prospective misalignment. A similar result could be obtained with short-term synaptic plasticity, which is known to produce a moving bump of activity ([47]). A third option could be the introduction of a systematic shift between the maps, possibly

resulting from Hebbian learning of the configurations generated by the suggested asymmetry mechanisms.

In the experiments of [32], two environments correspond to two different light configurations in the same arena. A slow linear morph between light configurations results in a sharp transition from the population representation for one environment to the other. This is a promising experimental technique which is able to probe with unprecedented flexibility the dynamics of remapping between two environments or along a morph sequence [32], and could serve as a fertile ground for our model’s predictions, hence for testing the attractor hypothesis. We show that, in agreement with the experiment, the slow morph protocol produces sharp transitions due to dynamical pattern separation. This result is even more significant considering the acknowledged difficulties in reproducing sharp transitions between correlated maps in a “traditional” setting [36]. The model predicts a transition between representations slightly delayed compared to half of the morphing period; it remains to be seen whether this occurs also in the experiment.

Our results can be related to experimental observations about changes in place representation between distinct environments. Two major classes of remapping have been observed when an animal is tested in two distinct environments: rate remapping, in which cells maintain the positions of their firing fields while differentially changing their amplitudes, and global remapping, where changes in firing location are observed in addition to firing rate modifications (see e.g. [43]). Based on these properties, we could associate the double ring regime to the global remapping and the cylinder regime to the rate remapping.

The model results can also be compared to experiments with sequences of continuously morphed environments. When animals explored intermediate environments, both sharp and smooth transitions in representations were observed in different experiments (see [31] and [30] correspondingly). Our model exhibits both sharp transitions between the place representations corresponding to intermediate environments (cylinder regime) and smooth transitions (double ring regime).

The linkage of cylinder and double ring regimes to sharp and smooth transitions respectively, taken together with the above mentioned association between these two model regimes with global and rate remapping, would be against the hypothesis made in [30] that related global remapping and sharp transitions on one hand, and rate remapping with smooth transitions on the other. In the present form, our model cannot be made compatible with this hypothesis. Since both the recordings of [31] and [30] contained populations of neurons exhibiting different transition behaviors, we speculate that the introduction of an additional selectivity for the environments (see below) could help in resolving the contradiction. Rate remapping would then correspond to a mixed single ring-cylinder regime (different subsets of the network would exhibit the different regimes), while global remapping would resemble a mix of the double ring and cylinder regimes.

A future extension of the model would include neurons with some form of selectivity for the context; each neuron would then be characterized not only by its location on the two maps, but also by selectivity indexes measuring its “preference” for the maps (e.g. [17]). This more realistic setting including selectivity would produce silent neurons and place fields with variable peak rates/widths even when storing a single map.

A second issue to be addressed is how the network can learn the synaptic structure from its inputs. The long-term plasticity (e.g. [33,34]), could bring the network through various operating regimes depending on the training protocol. This could impose

additional constraints on the model and provide additional predictions.

Finally, with the introduction of short-term plasticity [48–53], the network could exhibit an even richer repertoire of dynamics. This extension of the model would be an important step towards the experimental results of [32]. In this study, it was observed that when there is a fast switch between the two light configurations, the population vector sometime oscillates between the place representation of the environments, before settling on the current one. Preliminary results coming from the introduction of short term facilitation and depression in a network exhibiting a double ring solution, show that is indeed possible to observe oscillations between place representations. A detailed analysis of this behavior will be matter for a future report.

## Methods

### Numerical methods

To solve numerically the MF dynamics described by Eq. 5, we discretized  $(\theta, r)$  on  $100 \times 100$  regular grid in  $[0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ . The integrals in the rhs of the equations were estimated using a trapezoidal method. The system of ODEs were integrated with an adaptive 4-th order Runge Kutta scheme.

The simulation of the microscopic networks, whose results are reported in Figs. (9,10), were performed by solving numerically the system of ODEs

$$\tau \dot{m}_j = -m_j + \left[ \frac{1}{N} \sum_j W_{ij} m_j + I \right]_+, \quad (19)$$

where  $j = 1 \dots N$  indexes the neurons. The matrix  $W_{ij}$  is built by summing the single map encoding  $W_{ij} = \frac{1}{p} \sum_{\zeta} W_{ij}^{\zeta}$ , where

$$W_{ij}^{\zeta} = J_1 \cos(\theta_i^{\zeta} - \theta_j^{\zeta}) + J_0.$$

To obtain the  $p$  labels  $\theta_i^{\zeta}$  characterizing each neuron, we first randomly generated a  $\theta_i, r_i$  and used Eq. 1 for  $p = 2$  or Eq. 16 for  $p > 2$ . For the comparison of the simulation with the MF results in Fig. 9, we estimated from the steady state activity (compare with Eq. 22)

$$Z_A^E = \frac{1}{N} \sum_j e^{i\theta_j^A} m_j \equiv \rho_A^E e^{i\psi_A^E} \quad (20)$$

$$Z_B^E = \frac{1}{N} \sum_j e^{i\theta_j^B} m_j \equiv \rho_B^E e^{i\psi_B^E}$$

$$\eta^E = \frac{1}{N} \sum_j m_j \equiv \rho_A^E e^{i\psi_A^E},$$

from which we constructed the estimates for the order parameters, using Eq. 24.

For the noisy simulations shown in Fig. 5B, we used a current-based version of the dynamics described by Eqs. 19:

$$\tau \dot{h}_j = -h_j + \frac{1}{N} \sum_j W_{ij} [h_j]_+ + I.$$

We then estimated the order parameters via Eqs. 20, using the firing rates  $m_j = [h_j]_+$ . The noise was introduced as an additional term in the current

$$I \rightarrow I \cdot (1 + \epsilon_N B(t))$$

where  $B(t)$  is a zero average, unit variance Gaussian  $\delta$ -correlated noise. We used  $\epsilon_N = 0.05$  for the results in Fig. 5. The numerical solution was obtained using the Euler-Maruyama integration scheme.

The simulations performed in Fig. 7, for a network storing the whole morph sequence, were carried out as follows. Substituting the synaptic coupling obtained in Eq. 18 with the one in Eq. 2, it is possible to derive a dynamics for the ‘‘order function’’

$$Z(r') = \iint \mathcal{D}\theta \mathcal{D}r e^{i\theta} \text{sinc}(r' - r) m(\theta, r) \equiv \rho(r') e^{i\psi(r')},$$

following the same procedure of *Methods - Reduced dynamics*. An order parameter  $\eta$  is defined exactly as in Eq. 22. The steady state activity of such dynamics

$$m = [J_1 \rho(r) \cos(\theta - \psi(r)) + J_0 \eta + I]_+,$$

was compared with Eq. 4, for  $\theta = \psi_-$  and  $\theta = \psi(r)$  (in absence of a spatially tuned input  $\psi(r)$  is constant).

The time constant  $\tau$  was set to 10ms everywhere.

### Inverse transformation

The inverse transformation  $(\theta_A, \theta_B) \rightarrow (\theta, r)$  can be obtained from Eq. 1, defining

$$\theta'_B = \theta_B - 2\pi \text{sign}(\theta_B - \theta_A) H(|\theta_B - \theta_A| - \pi) \quad (21)$$

$$\theta = \left( \frac{\theta'_B + \theta_A}{2} \mod 2\pi \right)$$

$$r = \frac{\theta'_B - \theta_A}{2\mu}.$$

The  $2\pi$  rotation in the first equation is just needed to select the shortest distance between two maps on a ring, and it is transparent for the connectivity given its periodicity. This rotation was implicitly assumed when defining the neurons locations along the morph sequence, Eq. 16.

### Reduced dynamics

A first reduction of the dynamics described by Eq. 2 is done using the first two Fourier components of the activity  $m$  with respect to the two correlated maps  $\theta_A$  and  $\theta_B$ , rewritten in terms of center map  $\theta$  and the distance  $r$  using Eq. 1. In line with [37] we define the following variables

$$Z_A \equiv \iint \mathcal{D}\theta \mathcal{D}r e^{i(\theta - \mu r)} m(\theta, r) \equiv \rho_A e^{i\psi_A} \quad (22)$$

$$Z_B \equiv \iint \mathcal{D}\theta \mathcal{D}r e^{i(\theta + \mu r)} m(\theta, r) \equiv \rho_B e^{i\psi_B}$$

$$\eta \equiv \iint \mathcal{D}\theta \mathcal{D}r m(\theta, r).$$

The variable  $\eta$  is just the average activity, while  $Z_A$  and  $Z_B$  measure the spatial modulation of the network activity, in the map  $A$  and  $B$  respectively. Intuitively their values tell us which angle of which map is instantaneously represented by the network.

The dynamics of the network activity  $m$ , and of the order parameters  $\eta$ ,  $Z_A$ ,  $Z_B$ , becomes

$$\tau \dot{m} = -m + \tilde{g}(\theta, r) \quad (23)$$

$$\tau \dot{Z}_{B,A} = -Z_{B,A} + \iint \mathcal{D}\theta \mathcal{D}r e^{i(\theta \pm \mu r)} \tilde{g}(\theta, r)$$

$$\tau \dot{\eta} = -\eta + \iint \mathcal{D}\theta \mathcal{D}r \tilde{g}(\theta, r),$$

with

$$\tilde{g}(\theta, r) = \left[ \frac{J_1}{2} (\rho_A \cos(\theta - \mu r - \psi_A) + \rho_B \cos(\theta + \mu r - \psi_B)) + J_0 \eta + I \right]_+$$

It is convenient to introduce dimensionless combinations of the order parameters to better expose the structure of the solutions, and then derive the dynamics of these new order parameters. From the two complex variables and the real one, we construct five new variables

$$\gamma = \frac{\rho_B - \rho_A}{\rho_B + \rho_A} \quad (24)$$

$$\alpha = \frac{J_1}{I} \frac{\rho_B + \rho_A}{2}$$

$$\sigma = \frac{I + J_0 \eta}{\alpha I}$$

$$\psi_+ = \frac{\psi_B + \psi_A}{2}$$

$$\psi_- = \frac{\psi_B - \psi_A}{2}.$$

From Eqs. 23, after some algebra, it is finally possible to obtain the dynamics of the new order parameters, Eq. 5.

### Solutions properties

In *Methods - Phase diagram of the model*, we mentioned that the equation  $\psi_+ = 0$  from Eqs. 5, i.e.

$$\iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) [\sin(\theta - \psi_+) \cos(\mu r - \psi_-) - \gamma \cos(\theta - \psi_+) \sin(\mu r - \psi_-)] = 0$$

is automatically satisfied once the solution for the other four order parameters has been found. This can be seen using the fact that, by definition, the imaginary part of the real numbers  $\rho_A$  and  $\rho_B$  is 0. Since at steady state  $m(\theta, r) = g(\theta, r)$ , and  $\rho_{A,B} \equiv Z_{A,B} e^{-i\psi_{A,B}}$  (Eqs. 22), by computing  $Im(\rho_A) \pm Im(\rho_B) = 0$  we can prove the property.

Another statement mentioned in *Methods - Phase diagram of the model*, is that fixed points solutions of Eqs. 5 with  $(\gamma \neq 0, \psi_- \neq 0)$  do not exist. Observing the shape of the network activity at steady state (Eq. 3) (setting the phase  $\psi_+ = 0$  for convenience) which we rewrite here

$$m = \alpha I [\cos \theta \cos(\mu r - \psi_-) - \gamma \sin \theta \sin(\mu r - \psi_-) + \sigma]_+.$$

we would like to know, given the correlation between the stored maps  $\mu$  and the bump size  $\sigma$ , how much we can move the bump along  $r$  by increasing  $|\psi_-|$  without having active neurons at  $r = \pm \frac{\pi}{2}$ . We first analyze the onset of the freedom of choice of  $\psi_-$ , by requiring the bump to “fit” exactly the  $r$  range; with a bigger bump, the only possible choice for  $\psi_-$  would be  $\psi_- = 0$ , with a smaller bump it would be possible to move it along  $r$ . Hence, posing  $\psi_- = 0$ , the activity at  $r = \frac{\pi}{2}$  would be

$$m = \alpha I \left[ \cos \theta \cos\left(\mu \frac{\pi}{2}\right) - \gamma \sin \theta \sin\left(\mu \frac{\pi}{2}\right) + \sigma \right]_+.$$

The angle  $\theta^*$  at which this activity is maximal is

$$\theta^* = \arctan\left(-\gamma \tan\left(\mu \frac{\pi}{2}\right)\right),$$

so the maximal activity at the  $r$  boundaries is

$$m = \alpha I \left[ \frac{\cos\left(\mu \frac{\pi}{2}\right) + \gamma^2 \sin\left(\mu \frac{\pi}{2}\right) \tan\left(\mu \frac{\pi}{2}\right)}{\sqrt{1 + \gamma^2 \tan^2\left(\mu \frac{\pi}{2}\right)}} + \sigma \right]_+.$$

We recognize the first term inside the transfer function to be positive, so the only way to obtain a vanishing activity is to have  $\sigma < 0$ . From Fig. 4 it is possible to see that the double ring solutions  $\gamma \neq 0$  have always size  $\sigma > 0$ .

In order to obtain the range of integration for  $\psi_-$  used to compute the average tuning curve in the cylinder regime (*Results - Tuned external input*), it is enough to consider the activity  $m$  at its maximum in  $\theta$

$$m \sim [\cos(\mu r - \psi_-) + \sigma]_+.$$

We want this bump in  $r$  to at most touch the endpoints  $r = \pm \frac{\pi}{2}$ . Given that the half-width of the bump is  $\arccos(-\sigma)$ , the allowed range for  $\psi_-$  is  $\left[-\left(\frac{\pi}{2} - \arccos(-\sigma)\right), \left(\frac{\pi}{2} - \arccos(-\sigma)\right)\right]$ .

### Stability

In order to study the stability of the homogeneous solution, corresponding to  $\rho_A = \rho_B = 0$  in Eq. 23 (i.e.  $\alpha = 0$  from 24), we can either linearize Eq. 23, or take a step back from the MF reduction which lead to Eq. 5, so to avoid division by  $\alpha$ . We take the second approach and redefine one of the order parameter,  $\beta = \alpha\sigma$ . To study the stability of the solution, it is sufficient to look at the dynamics of  $\alpha$  and  $\beta$ . Posing  $\gamma = \psi_+ = \psi_- = 0$ , it is easy to verify that

$$\tau \dot{\beta} = -\beta + 1 + J_0 \iint \mathcal{D}\theta \mathcal{D}r f(\theta, r)$$

$$\tau \dot{\alpha} = -\alpha + J_1 \iint \mathcal{D}\theta \mathcal{D}r f(\theta, r) \cos \theta \cos(\mu r),$$

where the function  $f$  is defined as

$$f(\theta, r) = [\alpha \cos \theta \cos(\mu r) + \beta]_+.$$

The matrix describing the linear dynamics for the vector of small perturbations  $(\delta\beta, \delta\alpha)$  around the solution  $\alpha = 0, \beta = \beta^*$  reads

$$\begin{pmatrix} J_1 C - 1 & 0 \\ 0 & J_0 - 1 \end{pmatrix},$$

where

$$C = \iint \mathcal{D}\theta \mathcal{D}r \cos^2 \theta \cos^2(\mu r).$$

Therefore, two conditions must be satisfied for the solution to be stable:  $J_0 < 1$  (amplitude instability) and  $J_1 < \frac{1}{C}$ . Evaluating the integral in  $C$  explicitly, we get the line of separation between the homogeneous solution and the localized bump (Turing instability), expressed in Eq. 7.

For the single, double ring and cylinder solution we can linearize directly Eq. 5, posing  $\psi_+ = \psi_- = 0$ . The matrix associated with the dynamics of the vector  $(\delta\gamma, \delta\alpha, \delta\sigma)$ , after using the fixed points equations (Eq. 8), is

$$\begin{pmatrix} [-1 + J_1(A_{S2} + \gamma^* A_{SC})] & 0 & -J_1(A_S + \gamma^* A_C) \\ -\alpha^* J_1 A_{SC} & 0 & \alpha^* J_1 A_C \\ [-J_0 A_S + J_1 \sigma^* A_{SC}] & -\frac{1}{(\alpha^*)^2} & [-1 + J_0 A - J_1 \sigma^* A_C] \end{pmatrix}. \quad (25)$$

We define

$$A = \iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) \quad (26)$$

$$A_S = \iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) \sin \theta \sin(\mu r)$$

$$A_C = \iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) \cos \theta \cos(\mu r)$$

$$A_{S^2} = \iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) \sin^2 \theta \sin^2(\mu r)$$

$$A_{C^2} = \iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) \cos^2 \theta \cos^2(\mu r)$$

$$A_{SC} = \iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) \sin \theta \sin(\mu r) \cos \theta \cos(\mu r)$$

where

$$h(\theta, r) = H[\cos \theta \cos(\mu r) - \gamma^* \sin \theta \sin(\mu r) + \sigma^*].$$

Using the identity  $[x]_+ = xH(x)$ , to write the function  $g$  (Eq. 6), we see that

$$\iint \mathcal{D}\theta \mathcal{D}r g(\theta, r) = A_C + \sigma^* A.$$

The fixed points equations can thus be rewritten in term of the quantities in Eq. 26

$$\frac{1}{J_1} = A_{C^2} - \gamma^* A_{SC} + \sigma^* A_C \tag{27}$$

$$\gamma^* = \frac{A_{SC} + \sigma^* A_C}{A_{S^2} - \frac{1}{J_1}}$$

$$\frac{1}{\alpha^*} = \sigma^* - J_0(A_C + \sigma^* A).$$

Let us examine the single ring and cylinder solution,  $\gamma^* = 0$ . Given the symmetry in the integrand,  $A_C = A_{SC} = 0$  in this case. The stability matrix from Eq. 25 becomes then

$$\begin{pmatrix} [-1 + J_1 A_{S^2}] & 0 & 0 \\ 0 & 0 & \alpha^* J_1 A_C \\ 0 & -\frac{1}{(\alpha^*)^2} & [-1 + J_0 A - J_1 \sigma^* A_C] \end{pmatrix}. \tag{28}$$

## References

- Seung HS (1996) How the brain keeps the eyes still. *Proc Natl Acad Sci U S A* 93: 13339–13344.
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61: 331–349.
- O'Keefe J (1976) Place units in the hippocampus of the freely moving rat. *Exp Neurol* 51: 78–109.
- Taube J, Muller R, Ranck JJB (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *J Neurosci* 10: 420–435.
- Taube J, Muller R, Ranck JJB (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. ii. effects of environmental manipulations. *J Neurosci* 10: 436–447.
- Fyhn M, Molden S, Witter MP, Moser EI, Moser MB (2004) Spatial representation in the entorhinal cortex. *Science* 305: 1258–1264.
- Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436: 801–806.
- Willshaw D, Buneman OP, Longuet-Higgins H (1969) Non-holographic associative memory. *Nature* 222: 960–962.
- Amari SI (1972) Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Transactions on Computers* 21: 1197–1206.
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A* 79: 2554–2558.
- Amit DJ (1989) *Modeling brain function: the world of attractor neural networks*. Cambridge: Cambridge University Press.
- Amari SI (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol Cybern* 27: 77–87.
- Ermentrout GB, Cowan JD (1979) A mathematical theory of visual hallucination patterns. *Biol Cybern* 34: 137–150.

It is immediately seen that the eigenvalue corresponding to a destabilization of  $\gamma^* = 0$  changes sign when  $J_1 A_{S^2} = 1$ . Substituting for  $J_1$  the expression in Eq. 27, it is easy to verify that this reproduces the curve of separation between the single and double ring regime described by Eq. 12. We analyze the remaining two eigenvalues by looking at the trace  $T$  and the determinant  $D$  of the  $2 \times 2$  sub-matrix in the  $\alpha, \sigma$  subspace (from 28):

$$T = -1 + J_0 A - J_1 \sigma^* A_C$$

$$D = \frac{J_1 A_C}{\alpha^*}.$$

Given that  $A_C \geq 0$ , we see immediately that the eigenvalues have the same sign for  $\alpha^* > 0$ , and one of them changes sign when  $\alpha^* < 0$ . Recall that  $\alpha^* = 0$  is the onset of amplitude instability we introduced without proof in Eq. 9. If we find that when  $D > 0$  the trace is negative, then we know that  $\alpha^* = 0$  correspond to a destabilization of the solution.

Using Eq. 27, we see that imposing  $D > 0$  is equivalent to  $J_0 < \frac{\sigma^*}{A_C + \sigma^* A}$ , and that the trace  $T$  satisfies

$$T < -\frac{A_C}{A_C + \sigma^* A} \frac{A_{C^2} + 2\sigma^* A_C + (\sigma^*)^2 A}{A_C + \sigma^* A}.$$

The numerator in the first term is non-negative ( $A_C \geq 0$ ). The denominator is simply  $\iint \mathcal{D}\theta \mathcal{D}r g$ , non-negative by definition. The denominator in the second term is  $J_1^{-1} > 0$ , and we can write the numerator as

$$\iint \mathcal{D}\theta \mathcal{D}r h(\theta, r) (\cos \theta \cos(\mu r) + \sigma^*)^2 > 0.$$

Finally, we numerically verified that the region of existence of the double ring solution coincides with its stability region.

## Acknowledgments

We thank Barak Blumenfeld, Omri Barak and David Hansel for precious analytical insights on the problem, and we are grateful to Alon Rubin and Nachum Ulanovsky for useful discussions on hippocampal remapping.

## Author Contributions

Conceived and designed the experiments: SR MT. Performed the experiments: SR MT. Analyzed the data: SR MT. Contributed reagents/materials/analysis tools: SR MT. Wrote the paper: SR MT.



14. Ermentrout B (1998) Neural networks as spatio-temporal pattern-forming systems. *Rep Prog Phys* 61.
15. Murray JD (2003) *Mathematical biology*. New York: Springer, 3rd ed edition.
16. Ben-Yishai R, Bar-Or R, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A* 92: 3844–3848.
17. Blumenfeld B, Bibitchkov D, Tsodyks M (2006) Neural network model of the primary visual cortex: From functional architecture to lateral connectivity and back. *J Comput Neurosci* 20: 219–241.
18. Tsodyks M, Sejnowski TJ (1995) Associative memory and hippocampal place cells. *Int J Neural Syst* 6: 81–86.
19. Tsodyks M (1999) Attractor neural network models of spatial maps in hippocampus. *Hippocampus* 9: 481–489.
20. McNaughton B, Barnes C, Gerrard J, Gothard K, Jung M, et al. (1996) Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J Exp Biol* 199: 173–185.
21. Samsonovich A, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. *J Neurosci* 17: 5900–5920.
22. Seung HS, Lee DD, Reis BY, Tank DW (2000) Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26: 259–271.
23. Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci* 16: 2112–2126.
24. Redish A, Elga A, Touretzky DS (1996) A coupled attractor model of the rodent head direction system. *NETWORK* 7: 671–685.
25. Fuhs MC, Touretzky DS (2006) A spin glass model of path integration in rat medial entorhinal cortex. *J Neurosci* 26: 4266–4276.
26. Burak Y, Fiete IR (2009) Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput Biol* 5: e1000291.
27. Battaglia FP, Treves A (1998) Attractor neural networks storing multiple space representations: A model for hippocampal place fields. *Phys Rev E* 58: 7738–7753.
28. Stringer SM, Rolls ET, Trappenberg TP (2004) Self-organising continuous attractor networks with multiple activity packets, and the representation of space. *Neural Networks* 17: 5–27.
29. Roudi Y, Treves A (2008) Representing where along with what information in a model of a cortical patch. *PLoS Comput Biol* 4: e1000012.
30. Leutgeb JK, Leutgeb S, Treves A, Meyer R, Barnes CA, et al. (2005) Progressive transformation of hippocampal neuronal representations in “morphed” environments. *Neuron* 48: 345–358.
31. Wills TJ, Lever C, Cacucci F, Burgess N, O’Keefe J (2005) Attractor dynamics in the hippocampal representation of the local environment. *Science* 308: 873–876.
32. Jezek K, Treves A, Moser M, Moser E (2008) Abrupt and stuttering transitions between hippocampal representations. In: Program No. 94.11/TT52 2008 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, Online.
33. Blumenfeld B, Preminger S, Sagi D, Tsodyks M (2006) Dynamics of memory representations in networks with novelty-facilitated synaptic plasticity. *Neuron* 52: 383–394.
34. Bernacchia A, Amit DJ (2007) Impact of spatiotemporally correlated images on the structure of memory. *Proc Natl Acad Sci U S A* 104: 3544–3549.
35. Preminger S, Sagi D, Tsodyks M (2007) The effects of perceptual history on memory of visual objects. *Vision Res* 47: 965–973.
36. Papp G, Witter MP, Treves A (2007) The ca3 network as a memory store for spatial representations. *Learn Mem* 14: 732–744.
37. Hansel D, Sompolinsky H (1998) *Methods in Neuronal Modeling. From Synapse to Networks*. Koch C and Segev I, editors, MIT Press, Cambridge, MA. Chapter Modeling Feature Selectivity in Local Cortical Circuits. 13 p.
38. Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13: 55–80.
39. Samsonovich A (1998) Hierarchical multichart model of the hippocampal cognitive map. In: *Proceedings of the fifth joint symposium on neural computation*.
40. Tanaka K, Saito H (1989) Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62: 626–641.
41. Zhang LI, Tan AYY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature* 424: 201–205.
42. Barak O, Tsodyks M, Hernandez A, Romo R (2010) Neuronal population coding of parametric working memory. *J Neurosci*, in Press.
43. Fyhn M, Hafting T, Treves A, Moser MB, Moser EI (2007) Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446: 190–194.
44. Colgin LL, Moser EI, Moser MB (2008) Understanding memory through hippocampal remapping. *Trends Neurosci* 31: 469–477.
45. Battaglia FP, Sutherland GR, McNaughton BL (2004) Local sensory cues and place cell directionality: Additional evidence of prospective coding in the hippocampus. *J Neurosci* 24: 4541–4550.
46. Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL (1996) Population dynamics and theta rhythm phase precession of hippocampal place cell firing: a spiking neuron model. *Hippocampus* 6: 271–280.
47. York L, van Rossum M (2009) Recurrent networks with short term synaptic depression. *J Comput Neurosci* 27: 607–620.
48. Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci U S A* 94: 719–723.
49. Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. *Science* 275: 220–224.
50. Tsodyks M, Pawelzik K, Markram H (1998) Neural networks with dynamic synapses. *Neural Comput* 10: 821–835.
51. Romani S, Amit D, Mongillo G (2006) Mean-field analysis of selective persistent activity in presence of short-term synaptic depression. *J Comput Neurosci* 20: 201–217.
52. Barak O, Tsodyks M (2007) Persistent activity in neural networks with dynamic synapses. *PLoS Comput Biol* 3: e35.
53. Mongillo G, Barak O, Tsodyks M (2008) Synaptic theory of working memory. *Science* 319: 1543–1546.