Analysis of free recall dynamics of an abstract working memory model[*]

Gianluca Villani, Matin Jafarian*, Anders Lansner, Karl Henrik Johansson

Abstract— This paper analyzes the *free recall* dynamics of a working memory model. Free recalling is the reactivation of a stored pattern in the memory in the absence of the pattern. Our free recall model is based on an abstract model of a modular neural network composed on N modules, *hypercolumns*, each of which is a bundle of *minicolumns*. This paper considers a network of N modules, each consisting of two minicolumns, over a complete graph topology. We analyze the free recall dynamics assuming a constant, and homogeneous coupling between the network modules. We obtain a sufficient condition for synchronization of network's minicolumns whose activities are positively correlated. Furthermore, for the synchronized network, the bifurcation analysis of one module is presented. This analysis gives a necessary condition for having a stable limit cycle as the attractor of each module. The latter implies recalling a stored pattern. Numerical results are provided to verify the theoretical analysis.

I. INTRODUCTION

Working Memory (WM) is a general-purpose cognitive system responsible for temporary holding information in service of higher-order cognition such as reasoning and decision making. The importance of understanding human memory functioning is evident from its central role in our cognitive functions [1] as well as its role as the main inspiration behind developments in artificial memory networks [5].

Among the most important features of working memory, is the spontaneous free recall process. The latter refers to reactivation of a stored pattern in the memory in the absence of the pattern. The precise mechanisms underlying free recall dynamics in the human brain is not yet fully understood [12]. Yet, several abstract neuro-computational [3] as well as more detailed spiking neural network models have been built based on different neurobiological hypotheses [11], at different levels of abstractions, to account for human experimental data on working memory.

Attractor neural networks, dynamical networks which evolves towards a stable pattern, have been employed for understanding the mechanisms of human memory, including working memory. Among the most studied models is the Hopfiled model [8] (and its several variations e.g. [6]) which represents a memory network with a fully connected graph, symmetric weights and capable of storing only binary values. The model has limitations on the capacity as well as recalling previously stored patterns [3]. The latter has motivated designs with more than two states including the modular recurrent networks of Potts type [10].

The model in this paper is originated from the biologically inspired modular model of Potts type in [12]. The network model in [12] is composed of N modules, hypercolumns, each of which consists of a bundle of elementary units, minicolumns, interacting via lateral inhibition, such that each hypercolumn module acts as a winner-take-all microcircuit. We refer the interested reader for a detailed biological rationale of this model to [12].

In this paper, we study the free recall dynamics of WM based on a simplification of the model in [12]. We consider a network of N hypercolumns, each consisting of two minicolumns. We assume that a few patterns have been encoded in this WM by means of fast Hebbian plasticity [4] using exogenous signals. In this paper, we study the post-training reactivation dynamics of such a multi-item WM assuming a constant and homogeneous coupling between the network modules. Using tools from stability theory (a Lyapunov-based argument, e.g. [9]), we obtain a sufficient condition for achieving synchronization of network units which are positively correlated assuming a complete graph topology. Furthermore, for the synchronized network, the bifurcation analysis of one module's dynamics is presented. This analysis gives a necessary condition for having a stable limit cycle as the attractor of each network module. The latter implies recalling a stored pattern.

To the best of our knowledge, the free recall dynamics of working memory has not been studied analytically before, in particular from a control theory perspective. As shown in this paper, such analysis is useful for a better understanding of the free recalling mechanism of working memory networks.

The paper is organized as follows. Section II, presents the model and problem formulation. Section III, presents a Lyapunov analysis for characterizing synchronization condition. This section also provides the bifurcation analysis of one module of the network assuming that the network is synchronized. In Section IV, simulation results are presented. Section V concludes the paper.

II. PROBLEM FORMULATION

In this section, we present our model and state our goal of analysis. We model the free-recall dynamics of WM based on a simplification of a (non-spiking) attractor neural network originated from a modular recurrent neural network model [12]. This network is composed of N modules, hypercolumns, each of which consists of m elementary units, m minicolumns. In this model, each memory input (or pattern) is encoded by a mechanism into N attributes (represented by hypercolumns) each of which is composed of m intervals (represented by minicolumns).
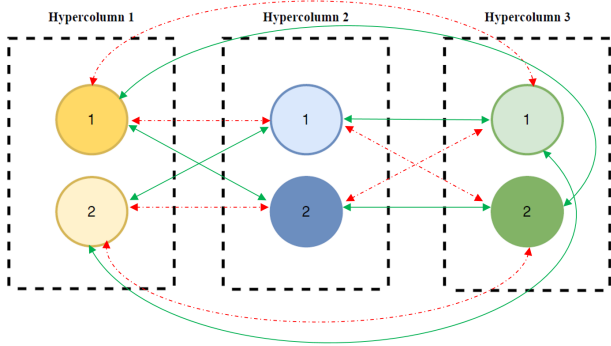
Fig. 1. A network composed of 3 hypercolumns. Each hypercolumn is composed of 2 minicolumns. Each minicolumn is connected via positive (excitatory, shown by solid lines) coupling to the minimcolumns in other hypercolumns if their activities are correlated. Otherwise the coupling is negative (inhibitory, shown by dashed lines).

Figure 1 shows a configuration of a network composed of 3 hypercolumns. We assume that the memory has been trained to learn some patterns using exogenous signals. In this paper, we study the post-training dynamics. Our goal is analyzing the network's dynamics in order to obtain conditions under which a stored pattern is recalled.

We represent the network by a connected, undirected complete graph composed of N nodes. Each node corresponds to one hypercolumn $i \in \{1, \ldots, N\}$ whose state is denoted by s_i. We study the case where $s_i \in \mathbb{R}^2$, i.e., $s_i = [s_{i1}, s_{i2}]$, where s_{ij} is the state of minicolumn j, $j \in \{1, 2\}$, of hypercolumn i.

We model the dynamics of each minicolumn s_{ij} based on a simplification of the model in [12], as follows

$$\dot{s}_{ij} = \sum_{\substack{k=1 \\ k \neq i}}^{N} \sum_{l=1}^{2} \omega_{kl,ij} o_{kl} - a_{ij} - s_{ij}, \quad (1)$$

$$\tau \dot{a}_{ij} = g_a o_{ij} - a_{ij}, \quad \tau > 1, \quad (2)$$

$$o_{ij} = \sigma(s_{ij}) = \frac{e^{s_{ij}}}{\sum_{k=1}^{m} e^{s_{ik}}}, \quad (3)$$

where $s_{ij} \in \mathbb{R}$, $a_{ij} \in \mathbb{R}$, and $o_{ij} \in (0, 1)$ represent the level of activation, the level of the adaptation, and the output of the minicolumn m_{ij} (minicolumn j in the hypercolumn i), receptively. The parameters $g_a > 0$ and $\tau > 1$ are constant. The coupling wight of the interconnection of the two minicolumns s_{kl} (in hypercolumn k) and s_{ij} (in hypercolumn i) is represented by $\omega_{kl,ij}$. We assume that the coupling weights are constant in the post-training dynamics. Here, we apply Hebb's rule in the way that minicolumns are interconnected. This implies that the mini-columns which were activated simultaneously in the training phase are connected with positive couplings in the post-taining dynamics, while the minicolumns with uncorrelated activities are connected with negative couplings (see Fig 1).

The role of variable a_{ij}, which models the biological mechanism of neural adaptation, is to deactivate its corresponding minicolumn m_{ij} in response to the changes in the activity of the other minicolumn in hypercolumn i. Owing to the definition of o_{ij} in (3), units within the same hypercolumn interact via *lateral inhibition*, such that each hypercolumn acts as a soft winner-take-all microcircuit. This means that what determines the output of each minicolumn is the relative level of its activation with respect to the activation of the other minicolumn in the same hypercolumn. This interaction is modeled by a *soft-max* function, σ as defined in Equation (3).

Assumption 1 All interconnection weights are equal such that $|\omega_{kl,ij}| = \frac{\omega}{2}, \omega > 0$.

Without loss of generality, we assume that any two minicolumns $i1$ and $k1$ ($i2$ and $k2$) are connected by positive coupling, whereas the interconnection of $i1$ to $k2$ is negative, $\forall i, k \in \{1, \ldots, N\}$. Since each hypercolumn is composed of two minicolumns, we can write $o_{i1} = 1 - o_{i2}, \forall i$ based on the definition of the σ function in (3).

Now, with Assumption 1, the dynamics of each minicolumn in (1) can be written as follows

$$\dot{s}_{ij} = \omega \sum_{\substack{k=1 \\ k \neq i}}^{N} \sigma(s_{kj}) - a_{ij} - s_{ij} - (N-1)\frac{\omega}{2}, \quad (4)$$

where $i \in \{1, \ldots, N\}, j \in \{1, 2\}$. The hypercolumn dynamics (the node dynamics) obeys

$$\dot{s}_i = -s_i - a_i - \frac{\kappa}{2} + \omega \sum_{\substack{k=1 \\ k \neq i}}^{N} \sigma(s_k), \quad \omega > 0, \quad (5)$$

$$\tau \dot{a}_i = g_a \sigma(s_i) - a_i, \quad \tau > 1, \quad (6)$$

where $s_i = [s_{i1}, s_{i2}]^T$, $a_i = [a_{i1}, a_{i2}]^T$, $\sigma(s_i) = [\sigma(s_{i1}), \sigma(s_{i2})]^T$, and $\kappa = (N-1)\omega$.

In the next section, we analyze the network dynamics (the free-recall dynamic) with the node dynamics as in (5), (6) and to answer the question that under which conditions a stored pattern is recalled. From a control theory perspective, this question is translated to characterizing coupling conditions under which the network modules synchronize [13] and oscillate according to the desired pattern.

III. ANALYSIS

In this section, we first derive the coupling condition under which synchronization, *i.e.* $s_{i,1} = s_{k,1}; s_{i,2} = s_{k,2}; a_{i,1} = a_{k,1}; a_{i,2} = a_{k,2}$, can be achieved for the network with the dynamics in (5), (6). We then analyze the behavior of one module (one hypercolumn) in the synchronized network using tools from the bifurcation theory. The analysis gives a coupling condition for having a stable limit cycle behaviour for each hypercolumn in the synchronized network.

Proposition 1 Consider the network dynamical system in (5)- (6). If $\frac{g_a}{\tau} < \omega < \frac{g_a}{\tau-1}$ holds, the solution to the system in (5)- (6) converges to the set $\mathcal{S} = \{s_i \in \mathbb{R}, a_i \in \mathbb{R} : s_i = s_k; a_i = a_k, \forall i, k \in \{1, \ldots, N\}\}$.

Proof: Define $\alpha = \frac{1}{\tau}$, $\alpha < 1$. Define $D_k = s_1 - s_k$ and $E_k = a_1 - a_k$. By multiplying the equation in (6) with α, the error dynamics for any two hypercolumns can be rewritten as:

$$\dot{D}_k = -D_k - E_k - \omega(\sigma(s_1) - \sigma(s_k))$$
$$\dot{E}_k = \bar{g}(\sigma(s_1) - \sigma(s_k)) - \alpha E_k, \quad \bar{g} = \alpha g_a.$$

Notice that the softmax function $\sigma(s_i)$ of the vector s_i as defined in (3) is a monotone function, *i.e.*

$$(a - b)^\top (\sigma(a) - \sigma(b)) \geq 0, \forall a, b \in \mathbb{R}^n.$$

Define the following Lyapunov function and denote by $\|\cdot\|_2$ the Euclidean norm:

$$V = \sum_{k=2}^{N} V_k, \quad V_k = \frac{1}{2} \|\bar{g} D_k + \omega E_k\|_2^2 + \frac{1}{2} \beta \omega \|D_k\|_2^2,$$

where $\beta = \bar{g} + (\alpha - 1)\omega$. By assuming $\omega < \frac{\alpha}{1-\alpha} g_a$, we guarantee that $\beta > 0$, thus the function in (7) is positive definite. We are interested in the conditions on ω such that the error dynamics is asymptotically stable, i.e. the conditions on ω such that $\dot{V} < 0$. We have,

$$\dot{V} = -\sum_{k=2}^{N} (\|\bar{g} D_k + \omega E_k\|_2^2 +$$
$$+ \beta \bar{g} D_k^\top E_k + \beta \omega \|E_k\|_2^2 + \omega \beta \|D_k\|_2^2 + \omega D_k^\top E_k +$$
$$+ \omega^2 \beta D_k^\top (\sigma(s_1) - \sigma(s_k))). \quad (7)$$

Rearranging the terms, we obtain

$$\dot{V} = -\sum_{k=2}^{N} (\|\bar{g} D_k + \omega E_k\|_2^2 +$$
$$+ \omega \beta \left(\|E_k\|_2^2 + \|D_k\|_2^2 + \left(1 + \frac{\bar{g}}{\omega}\right) D_k^\top E_k \right) +$$
$$+ \omega^2 \beta D_k^\top (\sigma(s_1) - \sigma(s_k))). \quad (8)$$

The quantity $\omega \beta \left(\|E_k\|_2^2 + \|D_k\|_2^2 + \left(1 + \frac{\bar{g}}{\omega}\right) D_k^\top E_k \right)$ is positive definite if $1 + \frac{\bar{g}}{\omega} < 2$, i.e. $\omega > \alpha g$. Therefore, a sufficient condition for synchronization is

$$\omega_{min} = g_a \alpha < \omega < \frac{\alpha}{1-\alpha} g_a = \omega_{max}. \quad (9)$$

$\square$

A. Analysis of the dynamics of one hypercolumn in the synchronized network

Assuming that the network is synchronized, we can write the dynamics of the single module (hypercolumn) as

$$\dot{s}_{i1} = -s_{i1} - a_{i1} + \kappa o_{i1} - \frac{\kappa}{2},$$
$$\tau \dot{a}_{i1} = g_a o_{i1} - a_{i1},$$
$$\dot{s}_{i2} = -s_{i2} - a_{i2} + \kappa o_{i2} - \frac{\kappa}{2}, \quad (10)$$
$$\tau \dot{a}_{i2} = g_a o_{i2} - a_{i2},$$

where $\kappa = (N - 1)\omega$. Recall that $o_{i1} + o_{i2} = 1$. Define, $d_i = s_{i1} - s_{i2}$, and

$$f(d_i) = o_{i1} - o_{i2} = \frac{e^{d_i} - 1}{e^{d_i} + 1} = \tanh(\frac{d_i}{2}). \quad (11)$$

We now introduce the following change of variables that will help us to analyze the dynamics of each hypercolumn and earn a deeper insight into the behaviour of our system:

$$d_i = s_{i1} - s_{i2},$$
$$e_i = a_{i1} - a_{i2}. \quad (12)$$

Bifurcation Analysis

From (10), the dynamics of d_i, e_i obeys

$$\dot{d}_i = -d_i - e_i + \kappa f(d_i),$$
$$\dot{e}_i = \frac{g_a f(d_i) - e_i}{\tau}. \quad (13)$$

We start the analysis of the two dimensional system in (13) by computing the equilibria (d_i^*, e_i^*) and study their stability properties with the variations of the coupling parameter κ. The equilibria of the dynamics in (13) satisfy the following equations

$$(\kappa - g_a)f(d_i^*) - d_i^* = 0, \quad (14)$$
$$e_i^* = g_a f(d_i^*). \quad (15)$$

Recall from (11) that $f(d_i) = \tanh(\frac{d_i}{2})$. We can show that the system has a unique equilibrium if $\kappa < g_a + 2$, while for $\kappa > g_a + 2$ we have multiple equilibria.

Lemma 1 The origin is the unique equilibrium for the system in (13) if $\kappa < g_a + 2$.

Proof: For $f(d_i) = \tanh(\frac{d_i}{2})$, we have $|f(d_i)| \leq |\frac{d_i}{2}|$. Hence if $\kappa - g_a < 2$, the equality in (14) is satisfied only for $d_i = 0$. $\square$

Now, assume that $\kappa < g_a + 2$. The unique equilibrium is $(d_i^*, e_i^*) = (0, 0)$. The computed Jacobian matrix $J_{(0,0)}$ and the characteristic polynomial are the following:

$$J_{(0,0)} = \begin{bmatrix} -1 + \frac{\kappa}{2} & -1 \\ \frac{g_a}{2\tau} & -\frac{1}{\tau} \end{bmatrix}, \quad (16)$$
$$\rho(\lambda) = \lambda^2 - \sigma(\kappa)\lambda + \delta(\kappa), \quad (17)$$

where

$$\sigma(\kappa) = \frac{-2\tau - 2 + \tau\kappa}{2\tau},$$
$$\delta(\kappa) = \frac{-2\kappa + 2g_a + 4}{4\tau}.$$

Thus, the origin is asymptotically stable if the following two conditions hold

$$\sigma(\kappa) < 0 \quad \text{if} \quad \kappa < 2\left(1 + \frac{1}{\tau}\right), \quad (18)$$
$$\delta(\kappa) > 0 \quad \text{if} \quad \kappa < g_a + 2. \quad (19)$$

As a result, the origin is the asymptotically stable equilibrium point (attractor) for (13) if $\kappa < 2(1 + \frac{1}{\tau})$ holds. This condition

is not particularly interesting since it means that all the units in the network converge to the same value and their output is identical. In other words, the network is not able to recall any pattern (see Fig. 2-A). By increasing κ, the attractor of the dynamical system (13) changes. In fact, at the critical value $\kappa^* = 2\left(1 + \frac{1}{\tau}\right)$ a *supercritical Hopf* bifurcation [7] occurs. That is, a unique stable limit cycle bifurcates from the origin. This result is now formally presented below.

Proposition 2 For the system in (13), a unique stable limit cycle bifurcates from the fixed point $(d_i, e_i) = (0,0)$ into the region $\kappa > 2(1 + \frac{1}{\tau})$ if $g_a > \frac{2}{\tau}$ holds.

Proof: To prove the above statement, we show that all of the conditions of [7, Theorem 3.4.2] are satisfied. Denote the eigenvalues of $J_{(0,0)}$ in (16) as a function of the parameter κ by $\lambda_{1,2}(\kappa)$. From (17), We have

$$\lambda_{1,2}(\kappa) = \frac{1}{2}\left(\sigma(\kappa) \pm \sqrt{\sigma(\kappa)^2 - 4\delta(\kappa)}\right).$$

At the critical value $\kappa^* = 2\left(1 + \frac{1}{\tau}\right)$, the following conditions should be satisfied:

1) The Jacobian matrix, $J_{(0,0)}$ in (16), has a conjugate pair of imaginary eigenvalues. That is,

$$\sigma(\kappa^*) = 0, \quad \delta(\kappa^*) > 0. \tag{20}$$

   The above condition is satisfied if $g_a > \frac{2}{\tau}$ holds.
2) Eigenvalues of $J_{(0,0)}$ in (16) vary smoothly by κ, *i.e.*

$$\left.\frac{\partial \sigma(\kappa)}{\partial \kappa}\right|_{\kappa = \kappa^*} = \frac{1}{2} \neq 0. \tag{21}$$

3) The last condition is to prove bifurcation of a stable limit cycle. We first rewrite the system in (13) in the follwoing form

$$\begin{bmatrix} \dot{d}_i \\ \dot{e}_i \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{\tau} & -1 \\ \frac{g_a}{2\tau} & -\frac{1}{\tau} \end{bmatrix}}_{A} \begin{bmatrix} \dot{d}_i \\ \dot{e}_i \end{bmatrix} + \begin{bmatrix} F_1(d_i) \\ F_2(d_i) \end{bmatrix}, \tag{22}$$

   with $F_1(d_i) = (-1 - \frac{1}{\tau})d_i + \kappa \tanh(\frac{d_i}{2})$ and $F_2(d_i) = -\frac{g_a}{2\tau}d_i + \frac{g_a}{\tau}\tanh(\frac{d_i}{2})$. Denote the eigenvalues of the matrix A in (22) by $\pm\beta i$, $\beta > 0$. Notice that the latter are also the eigenvalues of $J_{(0,0)}$ at $\kappa = \kappa^*$. We now consider a coordinate transformation $\begin{bmatrix} d_i \\ e_i \end{bmatrix} = E \begin{bmatrix} u_i \\ v_i \end{bmatrix}$ such that

$$E = \begin{bmatrix} 0 & 1 \\ -\beta & \frac{1}{\tau} \end{bmatrix}, \begin{bmatrix} u_i \\ v_i \end{bmatrix} = E^{-1} \begin{bmatrix} d_i \\ e_i \end{bmatrix} = \begin{bmatrix} \frac{d_i}{\tau\beta} - \frac{e_i}{\beta} \\ d_i \end{bmatrix}.$$

   Notice that in the above $d_i = v_i$, hence $F_1(d_i) = F_1(v_i), F_2(d_i) = F_2(v_i)$. We write the dynamics of (13) in (u_i, v_i) coordinate which gives

$$\begin{bmatrix} \dot{u}_i \\ \dot{v}_i \end{bmatrix} = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \end{bmatrix} + E^{-1} \begin{bmatrix} F_1(v_i) \\ F_2(v_i) \end{bmatrix}. \tag{23}$$

   We now can calculate the cubic coefficient of the Taylor series of degree 3 for the above system based on

the formula (3.4.11) of [7]. The computation gives a negative value, $-\frac{\kappa^*}{64}$, which ends the proof.
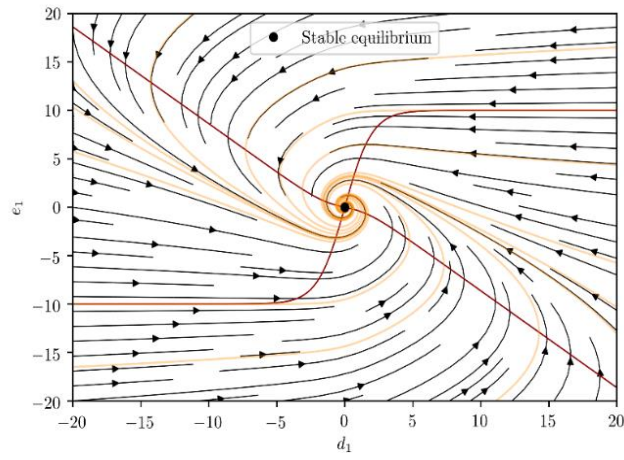
$\square$

Proposition 1 gives a sufficient condition for synchronization of the network in (5)-(6) and Proposition 2 presents a necessary condition such that the synchronous state is a stable limit cycle, *i.e.* a stored pattern is recalled. We now combine the two results below.

Corollary 1 For the network dynamical system in (5)- (6), $\frac{g_a}{\tau} < \omega < \frac{g_a}{\tau - 1}$ is a sufficient condition for the network to converge to the set $\mathcal{S} = \{s_i \in \mathbb{R}, a_i \in \mathbb{R} : s_i = s_k; a_i = a_k \ \forall i, k \in \{1, \ldots, N\}\}$. Moreover if $\omega < \frac{g_a + 2}{N - 1}$ holds, then the origin is the unique equilibrium for the relative state, (d_i, e_i) with $d_i = s_{i,1} - s_{i,2}, e_i = a_{i1} - a_{i2}, \forall i$. Furthermore, the two conditions $g_a > \frac{2}{\tau}$, and $\omega > \frac{2}{N - 1}(1 + \frac{1}{\tau})$ are necessary for achieving a stable limit cycle as the attractor of the relative state (d_i, e_i) of each s_i on the set $\mathcal{S}$.
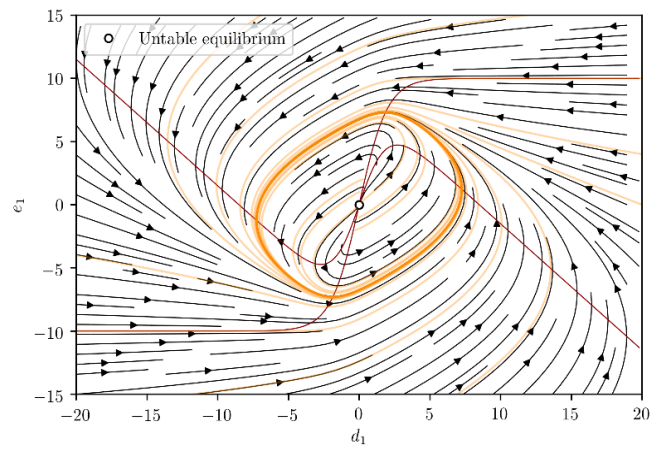
By increasing κ, the attractor of the dynamics in (13) changes. In what follows, we present a numerical example explaining these changes. The numerical example is also presented in Figure 2 (plotted using MATCONT [2] and Python simulations).

Example 1 (Numerical bifurcation analysis) For the system (representing one hypercolumn) in (10), set $\tau = 2$, $g_a = 10$. The relative states (d_i, e_i) with the dynamics in (13) converge to
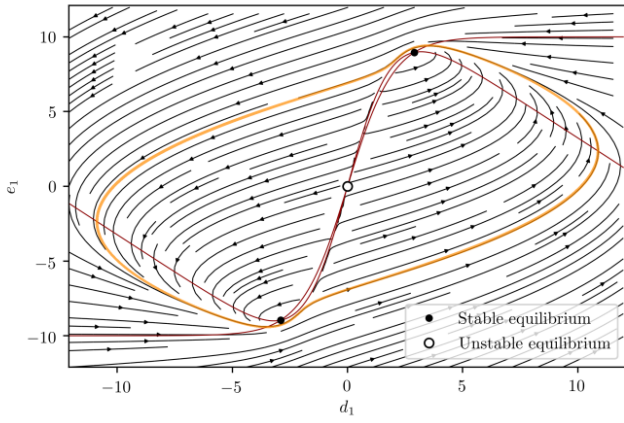
- a stable equilibrium point for $\kappa < 2(1 + \frac{1}{\tau}) = 3$. In this case, no pattern is recalled, *i.e.* both minicolumns reach the same level of activation, Fig. (2-A),
- a stable limit cycle with $2(1 + \frac{1}{\tau}) < \kappa < \kappa^{oo}$, $g_a + 2 = 12 < \kappa^{oo} \approx 13.11$. In this case the two minicolumns are activated in turn, that is the two stored patterns are reactivated in turn, Fig. (2-B),
- a limit cycle and two locally stable equilibria points with $13.11 \lessapprox \kappa \lessapprox 13.24605$. In this case, both oscillatory behaviour (cyclic activation) and fixed point behaviours are present. Depending on the initial conditions, the system converges to one of the three attractors, Fig. (2-C, 2-D),
- two locally stable equilibria with $\kappa \gtrapprox 13.24605$. In this case, the system does not show any oscillatory behaviour but it locally converges to one of the two stored patterns, *i.e.* one of the two minicolumns locally reaches and stays at a higher level of activation, Fig. (2-E).
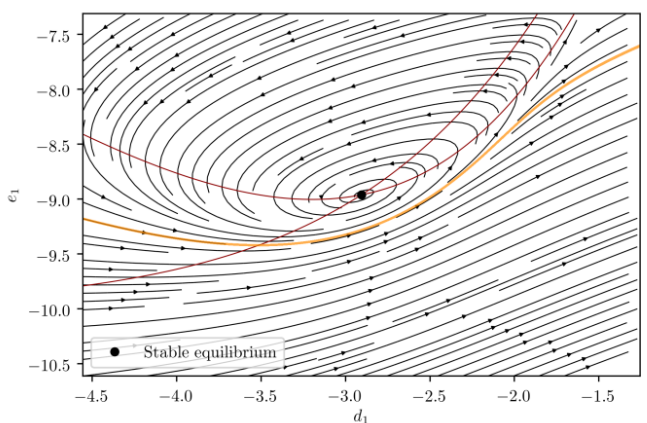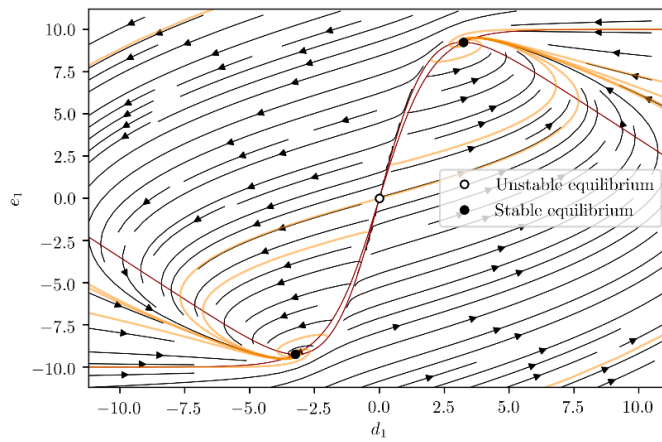
Fig. 2. Phase portrait of system in (13) for $g_a = 10$, $\tau = 2$. The nullclines are shown in red; the streamlines are shown in black. Several trajectories of the system with different initial conditions are shown in orange. The plots show changes of the attractor for the dynamics in (13) based on the increase of κ. In plots C, D where both fixed points and limit cycles (LC) exist, the LCs are plotted in Orange. Plot (D) shows a magnification of Plot (C).

IV. Simulation results

We consider a network of $N = 12$ hypercolumns, each composed of two minicolumns. We set, $g_a = 97$, $\tau = 54$ (implemented as $\frac{\tau_a}{\tau_m} = \frac{27}{0.05}$, see [12]). Based on Corollary 1, considering both the sufficient condition for achieving synchronization and the necessary condition such that each hypercolumn attractor is a limit cycle, we calculate $\omega = 1.8$, hence $\kappa = 19.8$. Figure 3 shows the relative state of positively correlated minicolumns. For clarity of presentation, the relative states with respect to hypercolumn 1, *i.e.* variables $D_{1,\ell} = s_{1,1} - s_{\ell,1}$ and $E_{1,\ell} = a_{1,1} - a_{\ell,1}$ with $\ell \in \{2,..,N\}$, are plotted. As shown the relative states of positively correlated minicolumns in all hypercolumns converge to zero. The latter implies synchronization. Figures 4 and 5 show the dynamics of one hypercolumn. As shown in Figure 4, the two minicolumns oscillate in turn. Figure 5 shows the output of the two minicolumns and the level of their corresponding adaptation variables. As shown, the role of adaptation variable is to inhibit the minicolumn which is most active and reactivate it in turn.
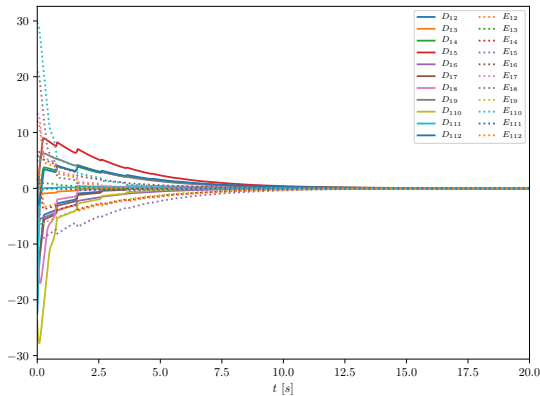

Fig. 3. The plot of the relative states of positively correlated minicolumns with respect to hypercolumn 1.
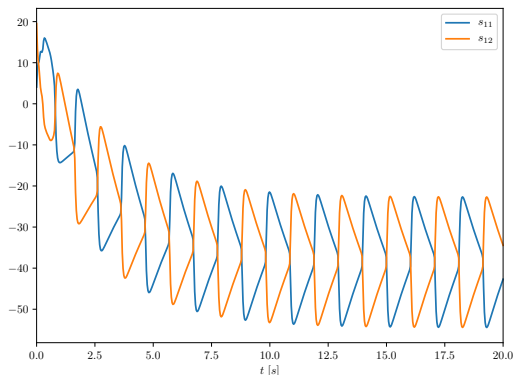

Fig. 4. The state of hypercolumn 1. Variables s_{11} and s_{12} achieves alternative activation and deactivation.

V. Conclusions

This paper has studied the free recall dynamics of a multi-item working memory network modeled as a simplified
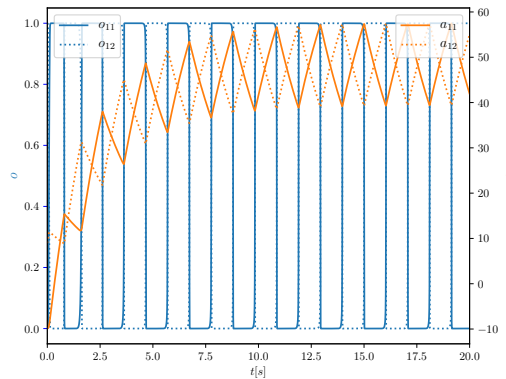

Fig. 5. Output and adaptation variables of a single hypercolumn. The adaptation variables are plotted in orange, the output of the units in blue. The inhibition of one unit allows the other unit to activate.

modular attractor neural network. The network consisted of N modules each one composed of two units. We analyzed the free recall dynamics assuming a constant, and homogeneous coupling between the network modules. A sufficient condition for synchronization of the network was obtained. Furthermore, assuming a synchronized network, the behavior of one module was analyzed using bifurcation analysis tools. Based on this analysis, a necessary coupling condition was given for achieving a stable limit cycle behavior for each module in the synchronized network. The latter implies the memory is recalling a stored pattern.

References

[1] M. D'Esposito and B.R. Postle. The cognitive neuroscience of working memory. *Annual review of psychology*, 66:115–142, 2015.

[2] A. Dhooge, W. Govaerts, and Y.A. Kuznetsov. Matcont: a matlab package for numerical bifurcation analysis of odes. *ACM Transactions on Mathematical Software (TOMS)*, 29(2):141–164, 2003.

[3] F. Fiebig. *Active Memory Processing: on Multiple Time-scalesin Simulated Cortical Networkswith Hebbian Plasticity*. PhD thesis, KTH Royal Institute of Technology, 2018.

[4] F. Fiebig and A. Lansner. A spiking working memory model based on hebbian short-term potentiation. *Journal of Neuroscience*, 37(1):83–96, 2017.

[5] A. Graves et al. Hybrid computing using a neural network with dynamic external memory. *Nature*, 538(7626):471–476, 2016.

[6] R. Gray, A. Franci, V. Srivastava, and N.E. Leonard. An agent-based framework for bio-inspired, value-sensitive decision-making. *IFAC-PapersOnLine*, 50(1):8238–8243, 2017.

[7] J. Guckenheimer and P. Holmes. *Local Bifurcations*, pages 117–165. Springer New York, New York, NY, 1983.

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

[9] M. Jafarian, X. Yi, M. Pirani, H. Sandberg, and K.H. Johansson. Synchronization of kuramoto oscillators in a bidirectional frequency-dependent tree network. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 4505–4510.

[10] I. Kanter. Potts-glass models of neural networks. *Physical Review A*, 37(7):2739, 1988.

[11] A. Lansner. Associative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. *Trends in neurosciences*, 32(3):178–186, 2009.

[12] A. Lansner, P. Marklund, S. Sikström, and L.G. Nilsson. Reactivation in working memory: an attractor network model of free recall. *PloS one*, 8(8):e73776, 2013.

[13] R. Sepulchre. Oscillators as systems and synchrony as a design principle. In *Current trends in nonlinear systems and control*, pages 123–141. Springer, 2006.