ACADEMIC PRESS

J. Math. Anal. Appl. 273 (2002) 529–547

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.academicpress.com

A learning rule with generalized Hebbian synapses

Fernanda Botelho and James E. Jamison

Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, USA

Received 4 February 2002

Submitted by C.T. Leondes

Abstract

We study the convergence behavior of a learning model with generalized Hebbian synapses.

1. Introduction

Inspired by the anatomy and physiology of our brain, artificial neural networks are mathematical models that attempt to mimic specific brain functions, such as learning. A neural network architecture consists of several additive processing units (neurons) interconnected by channels (synapses). Quantifiable information flowing through the synapses changes via multiplicative factors, designated connecting weights. These weights reflect the relationship between pre- and post-synaptic neural activities, also referred as incoming and outgoing signals.

Learning is a brain process by which the connecting weights undergo a sequence of changes due to outside stimulation. This network's internal response enables an environmental adaptation and facilitates the development of problem-solving skills. In colloquial terms, we may say that, after being exposed to a list of problems and their respective solutions, we expect to be equipped to solve

E-mail addresses: mbotelho@memphis.edu (F. Botelho), jjamison@memphis.edu (J.E. Jamison).

0022-247X/02/$ – see front matter
PII: S0022-247X(02)00263-9

not only a familiar problem (one from the list) but also a spectrum of new ones. Algorithms, implementable as neural networks, that attempt to perform this function have been proposed by several researchers, among them we list Hebb, Kohonen, Oja, and Adams (cf. [1,6,12,19]).

According to the Hebbian postulate of learning, the synaptic adjustments are given by scaled products of the incoming and outgoing signals. Such an algorithm leads to an exponential growth and therefore synaptic saturation. In 1988, Kohonen introduced a forgetting term that limits the synaptic weight growth and consequently ameliorates this saturation problem (cf. [12]). The use of a nonlinear correcting term incorporated into the synaptic changes results in a negatively accelerated synaptic modification curve. Kohonen's learning rule aims to reduce the dimension of incoming signal patterns. Along this line of thought Oja, in [19], introduced a network that behaves as a selective learning process by filtering information and therefore adapting from an internally selected subset of inputs. This approach was based on the principal component analyzer statistical method (cf. [5]). Recently, Kingsley and Adams proposed a generalization of Oja's rule (cf. [7]) by incorporating a probabilistic component at the synaptic level (cf. [11]). Biological observations suggest that synaptic changes occur not only between co-active pairs of neurons but between a neuron from the pair and its immediate neighbors. This fact is referred as "volume learning" (cf. [16]).

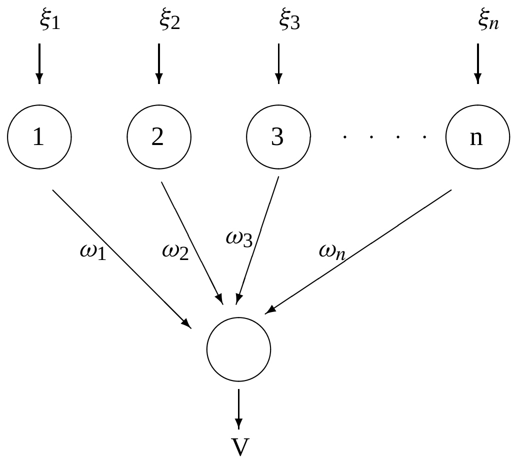
In this paper, we study Adams learning model following a similar approach to the one used by Oja (cf. [17,18]). This learning rule incorporates a Hebbian type of synapses where the updating relies upon a synaptic replication. We answer a question posed by P. Adams: "Will a stable weight vector emerge and how will it be related to the network's parameters?"

Convergence of this model is reduced to the stability behavior of the equilibrium points of a nonlinear system of differential equations. This study allows us to establish conditions for the algorithm to converge.

In Section 2, we define the learning rule, introduce notation to be used throughout the paper, and reduce the problem to solving a matrix equation. In Section 3, we solve a particular case where the weight matrix is a correlation matrix of independent random variables. The general case is studied in Section 4. The stability behavior is also analyzed in Section 4 (4.3), and sufficient conditions for convergence are derived. In Section 5, we summarize and interpret the main results of the paper.

2. Background and notation

Adams learning rule is implemented in a feedforward neural network with n input neurons and one output neuron whose architecture is shown in the picture below.



The input vector, denoted by ξ, represents an n-dimensional random vector following a joint probability distribution. The output value, denoted by V, is the outcome of the network's action on ξ. As the information travels through the synapses, it changes linearly via a multiplication by ω_j, the connecting weight for the synapse attached to the input neuron j. The connecting weight vector is denoted by ω. The summation of all these altered input components, $\sum_{j=1}^{n} \omega_j \xi_j$, is the value of V. The main goal is to define a converging algorithm that performs the network's adaptation without any outside interference. Such network is often called unsupervised learning rule. If convergence occurs, the network is fully characterized and ready to perform as an "educated" device. In [11], Kingsley and Adams consider that, in a learning process, new synapses may be created under a constant error rate, denoted by E. This learning rule relies upon a generalized type of Hebbian synapses that incorporates a synaptic replication and uses a correcting nonlinear term, cf. [19]. The updating for a Hebbian synapse is proportional to the product of the pre- and post-synaptic activities. More precisely, the change for the connecting weight ω_i is $\lambda \xi_i V$. Here, we allow synaptic changes to follow a probabilistic correlation rule between pre-synaptic and post-synaptic activities. In fact, as learning progresses, a synaptic strength may capture nearby activity. This is done by the creation of temporary synapses from the closest neurons to the output one. The synaptic corrections are therefore given by

$$\Delta\omega_i = V\left((1-E)\xi_i + \frac{1}{2}E\xi_{i+1} + \frac{1}{2}E\xi_{i-1} - V\omega_i\right),$$
$$\text{for } i = 2, \ldots, n-1,$$

or

$$\Delta\omega_i = V\left((1-E)\xi_i + \frac{1}{2}E\xi_{i\pm 1} - V\omega_i\right), \quad \text{for } i = 1 \text{ or } n,$$

respectively. As mentioned before, the network's outcoming signal is given by $V = \sum_{j=1}^{n} \omega_j \xi_j$. Substituting the value of V in the expression of $\Delta\omega_i$, we obtain:

$$\Delta\omega_i = \begin{cases} (1-E)\sum_{j=1}^{n} \omega_j \xi_j \xi_i + \frac{1}{2}E\sum_{j=1}^{n} \omega_j \xi_j \xi_{i+1} \\ \quad + \frac{1}{2}E\sum_{j=1}^{n} \omega_j \xi_j \xi_{i-1} - \sum_{j,k} \omega_j \omega_k \xi_j \xi_k \omega_i, \quad i \neq 1 \text{ and } n, \\ (1-E)\sum_{j=1}^{n} \omega_j \xi_j \xi_i + \frac{1}{2}E\sum_{j=1}^{n} \omega_j \xi_j \xi_{i\pm 1} \\ \quad - \sum_{j,k} \omega_j \omega_k \xi_j \xi_k \omega_i, \quad i = 1 \text{ or } n. \end{cases}$$

The expected synaptic updating follows the vector equation $\omega_{\text{new}} = \omega_{\text{old}} + \Delta\omega$. Therefore, it can be represented by the cubic vector-valued polynomial $TC\omega - (\omega \cdot C\omega)\omega$, where $C = [\xi_i \xi_j]_{ij}$, and T is the tridiagonal matrix

$$t_{ij} = \begin{cases} 1-E & \text{if } i = j, \\ E/2 & \text{if } |i-j| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

At the equilibrium points, where the algorithm converges, we expect the synaptic changes to average zero, $\langle \Delta\omega \rangle = 0$. Therefore, we reduce the problem to solving the equation

$$TC\omega - (\omega \cdot C\omega)\omega = 0, \tag{1}$$

with C now representing a correlation matrix of expected values, i.e. $C = [\langle \xi_i \xi_j \rangle]_{ij}$, and " $\cdot$ " the standard inner product in R^n.

The next two propositions state properties of the matrix T to be applied in subsequent results.

Proposition 2.1. (1) *The eigenvalues of matrix T are $\lambda_k = 1 - E + E\cos(k\pi/(n+1))$, with $k = 1, \ldots, n$.*

(2) *T is positive if and only if*

$$E \leqslant \frac{1}{2} \quad \text{or} \quad E > \frac{1}{2} \quad \text{and} \quad n \leqslant \frac{\cos^{-1}\left(1 - \frac{1}{E}\right)}{\pi - \cos^{-1}\left(1 - \frac{1}{E}\right)}.$$

Proof. (1) The matrix T is symmetric, therefore all its eigenvalues are real. We notice that $T = (1-E)I + \frac{1}{2}EA$, where I is the identity and A is the tridiagonal matrix with zeros along the diagonal, 1's right above and right below the main diagonal. It is shown in [3] that the eigenvalues of A are equal to $2\cos(\frac{k\pi}{n+1})$, for $k = 1, \ldots, n$. This implies that the eigenvalues of T are equal to $1 - E + E\cos(\frac{k\pi}{n+1})$, for $k = 1, \ldots, n$.

(2) T is positive if and only if all its eigenvalues are nonnegative. This is equivalent to $1 - E + E\cos(\frac{n\pi}{n+1}) \geqslant 0$ or $\cos(\frac{n\pi}{n+1}) \geqslant 1 - \frac{1}{E}$. This last inequality is clearly true for $E \leqslant \frac{1}{2}$. If $E > \frac{1}{2}$, then $\cos(\frac{n\pi}{n+1}) \geqslant 1 - \frac{1}{E}$ or equivalently $n \leqslant \cos^{-1}(1 - \frac{1}{E})/(\pi - \cos^{-1}(1 - \frac{1}{E}))$. $\square$

The next proposition collects some well-known results in linear algebra. For a proof we refer the reader to [4] or [9]. We remark that Q^{t} represents the transpose of Q.

Proposition 2.2. (1) *All eigenvalues of T are simple and if T is singular its kernel is one-dimensional.*

(2) *T is diagonalizable, i.e. there exists an orthonormal matrix of eigenvectors of T, Q, such that $Q^{\mathrm{t}}TQ = D$ with $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.*

(3) *If T is positive then its square root, $\sqrt{T}$, is a real matrix equal to*

$$Q \operatorname{diag}\bigl(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_k}\bigr) Q^{\mathrm{t}}.$$

(4) *If T is not positive then its square root is a complex-valued matrix equal to*

$$Q \operatorname{diag}\bigl(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_k}, i\sqrt{\lambda_{k+1}}, \dots, i\sqrt{\lambda_n}\bigr) Q^{\mathrm{t}},$$

where $\lambda_1, \dots, \lambda_k$ are the nonnegative eigenvalues of T.

3. A particular case: Independent inputs

In this section we solve the equation $TC\omega - (\omega \cdot C\omega)\omega = 0$ assuming that the input vector $\xi = [\xi_1, \dots \xi_n]$ is an n-dimensional random variable with independent components. This means that the joint distribution function for ξ factorizes into a product of n marginal density functions, $f(\xi_1, \dots, \xi_n) = \prod_{i=1}^{n} f_i(\xi_i)$. The correlation matrix C is now given by the product $\xi^{\mathrm{t}}\xi$. Given k vectors in R^n, $\{v_1, \dots, v_k\}$, we denote by $\overline{\{v_1, \dots, v_k\}}$ the vector space spanned by $\{v_1, \dots, v_k\}$, and by $\operatorname{Ker}(C)$ the kernel of C.

Proposition 3.1. *The matrix C defines an orthogonal invariant splitting of R^n, $R^n = \operatorname{Ker}(C) \oplus \overline{\{\xi\}}$.*

Proof. The matrix C is symmetric and has two eigenvalues $\lambda_0 = 0$ and $\lambda_1 = \xi_1^2 + \xi_2^2 + \dots + \xi_{n-1}^2 + \xi_n^2$. The eigenspace associated with λ_1 is spanned by $\{\xi\}$, and $\operatorname{Ker}(C)$ has dimension $n - 1$. $\square$

3.1. Solving $TC\omega - (\omega \cdot C\omega)\omega = 0$

Our goal is to determine the solution set, S, of the matrix equation above. We start by noticing that $\operatorname{Ker}(C) \subset S$. Vectors in S that are not in $\operatorname{Ker}(C)$ are of the form $\eta + t\xi$ with $\eta \in \operatorname{Ker}(C)$ and t a real number. Moreover, S is invariant under Z_2 actions, meaning that if $x \in S$ so is $-x$.

Let $\phi(\omega) = TC\omega - (\omega \cdot C\omega)\omega$, then

$$\phi(\eta + t\xi) = t\lambda_1 T(\xi) - t^3\lambda_1\|\xi\|^2\xi - t^2\lambda_1\|\xi\|^2\eta.$$

We conclude that $\phi(\eta + t\xi) = 0$ if and only if $t = 0$ or $T(\xi) = t\|\xi\|^2(\eta + t\xi)$. If $T(\xi) = t\|\xi\|^2(\eta + t\xi)$ then

$$\xi \cdot T(\xi) = \xi \cdot t\|\xi\|^2(\eta + t\xi) = t^2\|\xi\|^4 \quad \big(\text{where } \|\xi\|^2 = \xi \cdot \xi\big).$$

This implies that $t = \pm(\xi \cdot T(\xi))^{1/2}/\|\xi\|^2$ whenever $\xi \cdot T(\xi)$ is nonnegative which allows us to write the following set relations:

$$\mathrm{Ker}(C) \subseteq S \subseteq \mathrm{Ker}(C) \cup \left[\mathrm{Ker}(C) \pm \left(\frac{(\xi \cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right)\right].$$

Proposition 3.2. *If* $\xi \cdot T(\xi) > 0$ *then there exists a unique* $\eta \in \mathrm{Ker}(C)$ *such that*

$$S = \mathrm{Ker}(C) \cup \left\{\pm\left(\eta + \frac{(\xi \cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right)\right\}.$$

If $\xi \cdot T(\xi) \leqslant 0$ *then* $S = \mathrm{Ker}(C)$.

Proof. The kernel of C is clearly contained in S so additional solutions are of the form $\eta + t\xi$ for some $\eta \in \mathrm{Ker}(C)$ and $t \neq 0$.

It was shown before that $t = \pm(T\xi \cdot \xi)^{1/2}/\|\xi\|^2$. We also notice that

$$\phi\left(\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\xi\right) \in \mathrm{Ker}(C)$$

since

$$\begin{aligned}\phi\left(\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\xi\right) \cdot \xi &= \left(\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\lambda_1 T(\xi) - \frac{(T\xi \cdot \xi)^{3/2}}{\|\xi\|^6}\lambda_1\|\xi\|^2\xi\right) \cdot \xi \\ &= 0.\end{aligned}$$

If

$$\phi\left(\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\xi\right) = 0$$

then

$$\phi\left(-\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\xi\right) = 0$$

and therefore

$$\begin{aligned}&\phi\left(\eta \pm \frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\xi\right) \\ &\quad = \pm\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\lambda_1 T(\xi) \mp \frac{(T\xi \cdot \xi)^{3/2}}{\|\xi\|^6}\lambda_1\|\xi\|^2\xi - \frac{(T\xi \cdot \xi)}{\|\xi\|^4}\lambda_1\|\xi\|^2\eta \\ &\quad = \phi\left(\pm\frac{(T\xi \cdot \xi)^{1/2}}{\|\xi\|^2}\xi\right) - (T\xi \cdot \xi)\eta = -(T\xi \cdot \xi)\eta \neq 0,\end{aligned}$$

for $\eta \neq O$. If

$$\phi\left(\frac{(T\xi\cdot\xi)^{1/2}}{\|\xi\|^2}\xi\right)\neq 0$$

then we determine $\eta_0 \in \operatorname{Ker}(C)$ such that

$$\phi\left(\eta_0 \pm \frac{(T\xi\cdot\xi)^{1/2}}{\|\xi\|^2}\xi\right)=0$$

or equivalently

$$\pm\phi\left(\frac{(T\xi\cdot\xi)^{1/2}}{\|\xi\|^2}\xi\right)-\bigl(T(\xi)\cdot\xi\bigr)\eta_0=0.$$

It follows that

$$\eta_0=\pm\frac{1}{(T\xi\cdot\xi)}\phi\left(\frac{(T\xi\cdot\xi)^{1/2}}{\|\xi\|^2}\xi\right)=\pm\frac{1}{(T\xi\cdot\xi)^{1/2}}\left(T\xi-\frac{(T\xi\cdot\xi)}{\|\xi\|^2}\xi\right),$$

and therefore

$$\eta=\frac{1}{(T\xi\cdot\xi)}\phi\left(\frac{(T\xi\cdot\xi)^{1/2}}{\|\xi\|^2}\xi\right). \qquad \square$$

3.2. Stability study

Following the ideas presented in [19], the stability of the discussed algorithm is studied by considering the differential equation

$$\frac{d\omega}{dt}=TC\omega-(\omega\cdot C\omega)\omega. \tag{2}$$

We recall that if $T\xi\cdot\xi>0$ then equilibrium set of the equation above is $\operatorname{Ker}(C)$ together with two additional isolated points, see Proposition 3.2.

First, we show that the hyperplane $\operatorname{Ker}(C)$ is repelling. In fact, the derivative of ϕ at a generic point $\eta\in\operatorname{Ker}(C)$ is a linear transformation in R^n given by $D\phi(\eta)(v)=TCv$, where v is an n-vector. The matrix TC has two eigenvalues. One eigenvalue is equal to 0 with an $(n-1)$-dimensional eigenspace, $\operatorname{Ker}(C)$. The additional eigenvalue, denoted by μ_1, has an associated eigenvector, denoted by $\eta_1+t_1\xi$, with $t_1\neq 0$. This is translated in the following equation

$$D\phi(\eta)(\eta_1+t_1\xi)=\mu_1(\eta_1+t_1\xi),$$

or equivalently

$$TC(\eta)(\eta_1+t_1\xi)=t_1\lambda_1T(\xi)=\mu_1(\eta_1+t_1\xi).$$

This implies that $\mu_1=T\xi\cdot\xi(>0)$.

We consider one of the isolated equilibrium points,

$$P=\eta+\frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi,$$

as described in the Proposition 3.2. The linearization of the equation in a neighborhood of this point is given by

$$\frac{d\omega}{dt} = D\phi(P)(\omega - P).$$

We now determine the linear transformation $D\phi(P)$. Let v be a generic n-vector, then

$$D\phi(P)(v) = \lim_{t\to 0} \frac{\phi\left(\eta + \frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi + tv\right) - \phi\left(\eta + \frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right)}{t}.$$

We apply the definition of ϕ, the properties of inner product, and the assumption that $\phi(P) = 0$ to conclude that

$$\begin{aligned} D\phi(P)(v) &= D\phi\left(\eta + \frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right)(v) \\ &= TCv - 2\frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\lambda_1(v\cdot\xi)\left(\eta + \frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right) \\ &\quad - \left(\eta + \frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right)\cdot C\left(\eta + \frac{(\xi\cdot T(\xi))^{1/2}}{\|\xi\|^2}\xi\right)v. \end{aligned}$$

We notice that if $v \in \mathrm{Ker}(C)$ then v is an eigenvector of $D\phi(P)$ with eigenvalue $-(T\xi\cdot\xi)$. Therefore, the eigenspace associated with $-(T\xi\cdot\xi)$ has dimension $n-1$. As before, we determine the missing real eigenvalue, denoted by μ_1. We represent by $\eta_1 + t_1\xi$ an associated eigenvector. We have the equation

$$D\phi(P)(\eta_1 + t_1\xi) = \mu_1(\eta_1 + t_1\xi),$$

straightforward computations imply that $\mu_1 = -2(T\xi\cdot\xi)$. These allow us to write the next proposition, but before we review the definition of attracting and repelling invariant space (cf. [2,8]).

Definition 3.1. Suppose A is an invariant space of the differential equation $x' = f(x)$, where f is a differentiable map defined in R^n. The space A is said to be *attracting* if there exists an open neighborhood of A, U, such that every solution $x(t)$, with $x(0) \in U$, approaches A as t increases. The space A is said to be *repelling* if for every open neighborhood U of A there exist a neighborhood U_0 contained in U and $t_0 \in R_+$ such that every solution $x(t)$, with $x(0) \in U_0$, $x(t) \notin U$ for $t \geqslant t_0$. An invariant space that is neither attracting nor repelling is said to be of *saddle-type*. The basin of attraction of an attracting invariant space A is the set of all points that evolve toward A as time increases.

If A reduces to a single point then we refer to A as being an attracting, repelling, or saddle equilibrium point.

Proposition 3.3. *If* $\xi \cdot T(\xi) > 0$ *then system* (2) *has two attracting equilibrium points. The basin of attraction of each equilibrium point is the half space containing the point and bounded by* $\operatorname{Ker}(C)$.

Proof. The matrix C is symmetric so it is diagonalizable via an orthonormal matrix M (Antonne's theorem, see [9]). Let D be the diagonal matrix with zeros along the diagonal except for the last entry which is equal to $\lambda_1 = \|\xi\|^2$. The change of variables $\omega = My$ allows us to rewrite system (2) as follows:

$$\frac{dy}{dt} = BDy - (y \cdot Dy)y, \tag{3}$$

where

$$B = M^{\mathrm{t}}TM = [b_{ij}]_{i,j=1,\ldots,n}.$$

Componentwise, after a convenient time rescaling, this system becomes

$$\frac{dy_i}{dt} = b_{in}y_n - y_n^2 y_i \quad \text{if } i \neq n, \qquad \frac{dy_n}{dt} = b_{nn}y_n - y_n^3. \tag{4}$$

This system has the following qualitative dynamical behavior:

(1) y_n-direction is invariant.
(2) There are at most 3 equilibrium points and exactly three iff $b_{nn} > 0$.
(3) The hyperplane $y_n = \sqrt{b_{nn}}$ (and $y_n = -\sqrt{b_{nn}}$) is invariant and stable.
(4) Initial conditions with nonzero nth-component evolve toward one of the hyperplanes $y_n = \sqrt{b_{nn}}$ or $y_n = -\sqrt{b_{nn}}$ depending on the sign of the initial condition's nth-component.

Systems (2) and (3) are topologically equivalent or qualitatively similar (cf. [15]). This completes the proof of the proposition. □

Remark 3.1. It also follows from previous arguments that if $T\xi \cdot \xi < 0$ then $\operatorname{Ker}(C)$, the solution set of Eq. (1), is attracting. We notice that $T\xi \cdot \xi = b_{nn}$.

4. General case

This section is divided into three parts. First, we solve Eq. (1) with C symmetric and T positive. Second, we solve the given equation for a matrix T with negative eigenvalues together with additional symmetry assumptions. Finally, we study the stability behavior of solutions and establish conditions that assure convergence of the algorithm.

4.1. T is positive

We recall that T is said to be positive if and only if all its eigenvalues are nonnegative (cf. [20]). As stated in Proposition 2.2(3), T has real square root $\sqrt{T}$ given by

$$Q\,\mathrm{diag}\big(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}\big)\,Q^{\mathrm{t}}.$$

We notice that $\sqrt{T}$ is symmetric therefore the change of variables $\omega = \sqrt{T}y$ transforms Eq. (1) into

$$\sqrt{T}\big(\sqrt{T}C\sqrt{T}y - \big(y \cdot \sqrt{T}C\sqrt{T}y\big)y\big) = 0. \tag{5}$$

The solution set S of Eq. (5) consists of the set of all vectors y such that $\sqrt{T}C\sqrt{T}y - (y \cdot \sqrt{T}C\sqrt{T}y)y$ is in the $\mathrm{Ker}(\sqrt{T}\,)$.

The following lemma states a splitting of R^n naturally induced by a general symmetric matrix. For a proof we refer the reader to [4].

Lemma 4.1. *Let A be a symmetric n-dimensional matrix and s the number of its distinct eigenvalues. The matrix A defines an orthogonal splitting of R^n into a direct sum of invariant eigenspaces, $R^n = \bigoplus_{i=1}^s E_i$.*

The next theorem determines the solution set of Eq. (5) assuming T nonsingular.

Lemma 4.1 implies the existence of an invariant splitting of R^n, determined by $\sqrt{T}C\sqrt{T}$, into orthogonal eigenspaces E_i, $i = 1, \ldots, s$, where s is the number of distinct eigenvalues of $\sqrt{T}C\sqrt{T}$. Let α_i be the eigenvalue associated with E_i. We use the standard notation for the unit sphere in R^n, S^{n-1}.

Theorem 4.1. *If T is nonsingular and positive and C is symmetric then y is a solution of*

$$\sqrt{T}\big(\sqrt{T}C\sqrt{T}y - \big(y \cdot \sqrt{T}C\sqrt{T}y\big)y\big) = 0$$

if and only if

$$y \in \mathrm{Ker}\big(\sqrt{T}C\sqrt{T}\,\big) \cup \left(\bigcup_{i=1}^{s} E_i \cap S^{n-1}\right).$$

Proof. A vector y can be uniquely decomposed into the sum $\sum_{i=1}^s y_i$, with $y_i \in E_i$ for each i. Then, we have

$$\begin{aligned}
&\sqrt{T}\big(\sqrt{T}C\sqrt{T}y - \big(y \cdot \sqrt{T}C\sqrt{T}y\big)y\big)\\
&\quad = \sqrt{T}\left(\sqrt{T}C\sqrt{T}\left(\sum_{i=1}^{s} y_i\right) - \left(\left(\sum_{i=1}^{s} y_i\right) \cdot \sqrt{T}C\sqrt{T}\left(\sum_{i=1}^{s} y_i\right)\right)\sum_{i=1}^{s} y_i\right)\\
&\quad = 0
\end{aligned}$$

if and only if

$$\sum_{i=1}^{s} \alpha_i y_i = \left(\sum_{i=1}^{s} \alpha_i y_i \cdot y_i \right) \sum_{i=1}^{s} y_i .$$

The orthogonality of the eigenspaces $\{E_i\}_{i=1,\dots,s}$ implies that every i for which $\alpha_i \neq 0$ and $y_i \neq O$ we have

$$\alpha_i = \sum_{i=1}^{s} \alpha_i y_i \cdot y_i .$$

Therefore $y \in \operatorname{Ker}(\sqrt{T} C \sqrt{T})$ or there exists i, with $\alpha_i \neq 0$, such that $y \in E_i$. In this latter case, we have that $(\|y\|^2 =)\ y \cdot y = 1$. □

Theorem 4.2. *If T is positive and C symmetric then y is a solution of*

$$\sqrt{T}\big(\sqrt{T} C \sqrt{T} y - \big(y \cdot \sqrt{T} C \sqrt{T} y\big) y\big) = 0$$

if and only if

$$y \in \operatorname{Ker}\big(\sqrt{T} C \sqrt{T}\big) \cup \left(\operatorname{Ker}(T) + \left(\bigcup_{i=1}^{s} E_i \cap S^{n-1} \right) \right).$$

Proof. Propositions 2.1 and 2.2 imply that $\operatorname{Ker}(T) = \operatorname{Ker}(\sqrt{T})$ is trivial or has dimension one. We consider the splitting of R^n described in the Lemma 4.1 and we follow the notation used in the previous theorem. If $\operatorname{Ker}(T)$ is nontrivial then we may assume, without loss of generality, that α_1 is equal to zero and u_1 is a unit vector spanning $\operatorname{Ker}(T)$, i.e. $\operatorname{Ker}(T) = \overline{\{u_1\}}$. The $\operatorname{Ker}(T)$ is therefore a subspace of E_1. A vector y has a unique decomposition $\sum_{i=1}^{s} y_i$, with $y_i \in E_i$. Consequently, a vector y, not in the $\operatorname{Ker}(\sqrt{T} C \sqrt{T})$, is a solution to the given equation if and only if

$$\sum_{i=1}^{s} \alpha_i y_i - \left(\sum_{i=1}^{s} \alpha_i y_i \cdot y_i \right) \sum_{i=1}^{s} y_i \in \operatorname{Ker}(T).$$

As in the proof of the Theorem 4.1, there exists a unique i_0 such that $y = y_1 + y_{i_0}$, where $y_1 \in \operatorname{Ker}(T)$, $y_{i_0} \in E_{i_0}$, and $y_{i_0} \cdot y_{i_0} = 1$. □

The next corollary states the relationship between the solution set of Eq. (1), denoted by S_ω, and the solution set of Eq. (5), denoted by S_y. The image of S_y under $\sqrt{T}$ is denoted by $\sqrt{T} S_y$.

Proposition 4.1. *If T is singular then $S_\omega = \operatorname{Ker} C \cup \sqrt{T} S_y$, otherwise $S_\omega = \sqrt{T} S_y$.*

Proof. We notice that if $y_0 \in S_y$ then $\sqrt{T}y_0 \in S_\omega$, and whenever $y_0 \in \operatorname{Ker} C$ we also have that $y_0 \in S_\omega$. If T is nonsingular then

$$S_y = \bigl\{y\colon \sqrt{T}C\sqrt{T}y - \bigl(y \cdot \sqrt{T}C\sqrt{T}y\bigr)y = 0\bigr\}.$$

Equation (1) is now equivalent to Eq. (5), since the change of variables $\omega = \sqrt{T}y$ is an isomorphism. Therefore, we have that $S_\omega = \sqrt{T}S_y$. If T is singular, $\operatorname{Ker} T$ has dimension 1 (cf. Proposition 2.2), $\operatorname{Ker} T = \overline{\{u\}}$ where $\|u\| = 1$. We recall that $\operatorname{Ker} T = \operatorname{Ker}\sqrt{T}$ and $R^n = \operatorname{Im}\sqrt{T} \oplus \operatorname{Ker}\sqrt{T}$. Let $\omega_0 \in S_\omega$, $\omega_0 = \omega_I + \omega_K$ where $\omega_I \in \operatorname{Im}\sqrt{T}$ and $\omega_K \in \operatorname{Ker}\sqrt{T}$. If $\omega_K = 0$ then $\omega_0 = \omega_I = \sqrt{T}y_0$, which implies that

$$\sqrt{T}\bigl(\sqrt{T}C\sqrt{T}y_0 - \bigl(y_0 \cdot \sqrt{T}C\sqrt{T}y_0\bigr)y_0\bigr) = 0, \quad y_0 \in S_y,$$

and $\omega_0 \in \sqrt{t}S_y$. If $\omega_K \neq 0$ then $TC\omega_0 - (\omega_0 \cdot C\omega_0)\omega_0 = 0$ implies that $TC\omega_0 \cdot \omega_K - (\omega_0 \cdot C\omega_0)(\omega_0 \cdot \omega_K) = 0$ and $\omega_0 \cdot C\omega_0 = 0$. Since $\omega_0 \in S_\omega$ and $\omega_0 \cdot C\omega_0 = 0$ we have that $C\omega_0 \in \operatorname{Ker} T$. Therefore $C\omega_0 = \lambda u$, for some scalar λ. We conclude that $\lambda = 0$ since the inner product $\omega_0 \cdot C\omega_0 = \omega_0 \cdot \lambda u = (\omega_K + \omega_I) \cdot (\lambda u) = \lambda(\omega_K \cdot u) = 0$. This proves that if $\omega_0 \in S_\omega$ and $\omega_0 \notin \operatorname{Im}\sqrt{T}$ then $\omega_0 \in \operatorname{Ker} C$. □

4.2. T is not positive

Throughout this section we assume that the matrix T has some negative eigenvalue. Proposition 2.2(4) states that there exits an orthonormal matrix of eigenvalues Q such that

$$Q^{\mathrm{t}}TQ = \operatorname{diag}\bigl(\lambda_1^2, \ldots, \lambda_k^2, -\lambda_{k+1}^2, \ldots, -\lambda_n^2\bigr).$$

We consider two n-dimensional diagonal matrices with real entries,

$$D_R = \operatorname{diag}(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0) \quad \text{and}$$
$$D_I = \operatorname{diag}(0, \ldots, 0, \lambda_{k+1}, \ldots, \lambda_n).$$

The square root of T is a complex matrix given by

$$\sqrt{T} = QD_RQ^{\mathrm{t}} + iQD_IQ^{\mathrm{t}}.$$

We set notation as follows: $R = QD_RQ^{\mathrm{t}}$, $I = QD_IQ^{\mathrm{t}}$, and

$$\mathcal{B} = \bigl\{Qe_1Q^{\mathrm{t}}, \ldots, Qe_kQ^{\mathrm{t}}, Qe_{k+1}Q^{\mathrm{t}}, \ldots, Qe_nQ^{\mathrm{t}}\bigr\},$$

an orthonormal basis for R^n, with e_i a unit n-vector with all its components equal to zero, except for the ith one. For simplicity of notation, we set $v_i = Qe_iQ^{\mathrm{t}}$, for each i. The next lemma collects properties of R and I to be used in forthcoming analysis.

Lemma 4.2. (1) *The range of R is spanned by $\{v_1, \ldots, v_k\}$ and its kernel by $\{v_{k+1}, \ldots, v_n\}$.*

(2) *The range of* I *is spanned by* $\{v_{k+1}, \dots, v_n\}$ *and its kernel by* $\{v_1, \dots, v_k\}$.
(3) $RI = IR = 0$.

Proof. Statements (1)–(3) are straightforward consequences of the following:

$$R(v_i) = \begin{cases} \lambda_i v_i & \text{for } i = 1, \dots, k, \\ 0 & \text{for } i = k+1, \dots, n, \end{cases}$$

and

$$I(v_i) = \begin{cases} 0 & \text{for } i = 1, \dots, k, \\ \lambda_i v_i & \text{for } i = k+1, \dots, n. \end{cases} \qquad \square$$

Remark 4.1. The matrices R and I are respectively represented by the diagonal matrices D_R and D_I, relatively to the basis $\mathcal{B}$.

Now, we introduce the change of variables $\omega = (R + iI)y$, where y is a complex n-vector. Let $y = y_1 + iy_2$, with y_1 and y_2 in R^n. It is entrained in this change of variables that $\omega = Ry_1 - Iy_2$ and $Iy_1 + Ry_2 = 0$. The relation $Iy_1 + Ry_2 = 0$ implies that $Iy_1 = Ry_2 = 0$, since

$$Iy_1, Ry_2 \in \overline{\{v_1, v_2, \dots, v_k\}} \cap \overline{\{v_{k+1}, \dots, v_n\}} \quad \bigl(= \{O\}\bigr).$$

This shows that y has the following representation in $\mathcal{B}$:

$$y = \sum_{j=1}^{k} y_j' v_j + i \sum_{j=k+1}^{n} y_j'' v_j.$$

Let C_0 be the set of all complex vectors whose real part and imaginary part are in $\overline{\{v_1, v_2, \dots, v_k\}}$ and in $\overline{\{v_{k+1}, \dots, v_n\}}$, respectively. Similar reasoning also shows that $Ry_1 - Iy_2 \neq 0$ is equivalent to $\operatorname{Ker}(\sqrt{T}\,) \cap C_0 = \{O\}$.

The next theorem reduces Eq. (1) to a real system in y_1 and y_2.

Theorem 4.3. *If* C *and* ICR *are symmetric matrices and* T *is a nonsingular matrix then Eq.* (1) *is equivalent to*

$$\begin{cases} RCRy_1 - (y_1 \cdot RCRy_1 - y_2 \cdot ICIy_2)y_1 = 0, \\ -ICIy_2 - (y_1 \cdot RCRy_1 - y_2 \cdot ICIy_2)y_2 = 0, \end{cases} \tag{6}$$

where $y_1 \in \overline{\{v_1, v_2, \dots, v_k\}}$ *and* $y_2 \in \overline{\{v_{k+1}, \dots, v_n\}}$.

Proof. Equation (1) becomes

$$\sqrt{T}\bigl(\sqrt{T}C\sqrt{T}y - \bigl(y \cdot \sqrt{T}C\sqrt{T}y\bigr)y\bigr) = 0,$$

when $\omega = \sqrt{T}y$ and $y \in C_0$. We recall that $\operatorname{Ker}\sqrt{T} \cap C_0 = \{O\}$. Therefore, we are reduced to solve

$$\sqrt{T}C\sqrt{T}y - \bigl(y \cdot \sqrt{T}C\sqrt{T}y\bigr)y = 0, \quad \text{for } y \in C_0. \tag{7}$$

On the other hand, we have that

$$\sqrt{T}C\sqrt{T} = (RCR - ICI) + i(ICR + RCI)$$

and

$$\sqrt{T}C\sqrt{T}y = (RCRy_1 - RCIy_2) + i(ICRy_1 - ICIy_2).$$

Therefore

$$y \cdot \sqrt{T}C\sqrt{T}y = y_1 \cdot (RCRy_1 - ICIy_2) + y_2 \cdot (ICRy_1 - ICIy_2).$$

Equation (7) is equivalent to the system

$$\begin{cases} RCRy_1 - RCIy_2 \\ \quad - \bigl(y_1 \cdot (RCRy_1 - RCIy_2) + y_2 \cdot (ICRy_1 - ICIy_2)\bigr)y_1 = 0, \\ ICRy_1 - ICIy_2 \\ \quad - \bigl(y_1 \cdot (RCRy_1 - RCIy_2) + y_2 \cdot (ICRy_1 - ICIy_2)\bigr)y_2 = 0. \end{cases} \tag{8}$$

The matrix ICR is symmetric, then $RCIy_2 = ICRy_2 = 0$ and $ICRy_1 = RCIy_1 = 0$. Consequently, the system above simplifies to the one in the statement of the theorem. □

We are ready to solve system (6). We represent by p and q the number of distinct nonzero eigenvalues of RCR and ICI, respectively. Let E_i with $i = 1, \dots, p$ and F_i with $i = 1, \dots, q$ be the associated eigenspaces as described in Lemma 4.1. Let α_i and β_i be the eigenvalues associated to E_i and to F_i, respectively.

Theorem 4.4. *The pair* (y_1, y_2) *is a solution of system* (6) *if and only if one of the following statements holds*:

(1) $y_1 \in \mathrm{Ker}(RCR)$ *and* $y_2 \in \mathrm{Ker}(ICI)$,
(2) $y_1 \in E_i \cap S^{n-1}$, *for some* $i = 1, \dots, p$, *and* $y_2 = 0$,
(3) $y_1 = 0$ *and* $y_2 \in F_i \cap S^{n-1}$, *for some* $i = 1, \dots, q$,
(4) $(y_1, y_2) \in E_i \times F_j$, *for some* $i = 1, \dots, p$ *and* $j = 1, \dots, q$, *with* $\alpha_i + \beta_j = 0$, $\alpha_i\beta_j \neq 0$, *and* $y_1 \cdot y_1 + y_2 \cdot y_2 = 1$.

Proof. Let (y_1, y_2) be a solution of system (6). We apply Lemma 4.1 to define a splitting of R^n associated to RCR and ICI, more precisely $R^n = \bigoplus_{i=1}^{p} E_i$ and $R^n = \bigoplus_{i=1}^{q} F_i$. Let $\alpha_1, \dots, \alpha_p$ be the (pairwise distinct) eigenvalues of RCR and $\beta_1, \dots, \beta_q$ be the (pairwise distinct) eigenvalues of ICI. Both y_1 and y_2 have unique decompositions $y_1 = \sum_{i=1}^{p} y_i'$ and $y_2 = \sum_{i=1}^{q} y_i''$, respectively. Since

$$y_1 \cdot RCRy_1 - y_2 \cdot ICIy_2 = \sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i'',$$

we rewrite system (6) as follows:

$$\begin{cases} \sum_{i=1}^{p} \alpha_i y_i' - \left(\sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i''\right) \sum_{i=1}^{p} y_i' = 0, \\ -\sum_{i=1}^{q} \beta_i y_i'' - \left(\sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i''\right) \sum_{i=1}^{q} y_i'' = 0. \end{cases} \tag{9}$$

This implies that if $\alpha_i \neq 0$ and $y_i' \neq O$ ($\beta_i \neq 0$ and $y_i'' \neq O$) then

$$\alpha_i = \sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i''$$

($-\beta_i = \sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i''$, respectively). Therefore, there exist at most two indices $i_0 \in \{1, \dots, p\}$ and $j_0 \in \{1, \dots, q\}$ such that $\alpha_{i_0} \beta_{j_0} \neq 0$ and $\|y_{i_0}'\| \|y_{j_0}''\| \neq 0$. This implies that $y_1 = y_1' + y_{i_0}'$ and $y_2 = y_1'' + y_{j_0}''$ where $y_1' \in \operatorname{Ker}(RCR)$ and $y_1'' \in \operatorname{Ker}(ICI)$. Therefore system (9) becomes

$$\begin{cases} \alpha_{i_0} y_{i_0}' - \left(\sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i''\right)\left(y_1' + y_{i_0}'\right) = 0, \\ -\beta_{j_0} y_{j_0}'' - \left(\sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i''\right)\left(y_1'' + y_{j_0}''\right) = 0. \end{cases} \tag{10}$$

Then, clearly $\|y_1'\| \|y_{i_0}'\| = 0$ since otherwise $\sum_{i=1}^{p} \alpha_i y_i' \cdot y_i' - \sum_{i=1}^{q} \beta_i y_i'' \cdot y_i'' = 0$ and $\alpha_{i_0} = 0$ which contradicts $\alpha_{i_0} \neq 0$. Similar reasoning shows that we also have $\|y_1''\| \|y_{j_0}''\| = 0$. Therefore, if $\|y_1'\| = 0$ and $\|y_1''\| = 0$ then $y_1 = y_{i_0}'$ and $y_2 = y_{j_0}''$, where $\alpha_{i_0} + \beta_{j_0} = 0$ and $y_1 \cdot y_1 + y_2 \cdot y_2 = 1$, as stated in case (4). If $\|y_1'\| \neq 0$ then $y_1 = y_1'$ and either $y_2 = y_1''$ or $y_2 = y_{j_0}''$. These are listed in case (1) or in case (3), respectively. In fact, if $y_1 = y_1'$ and $y_2 = y_{j_0}''$ then system (10) reduces to

$$\begin{cases} \beta_{j_0} \|y_{j_0}''\|^2 y_1' = 0, \\ -\beta_{j_0} y_{j_0}'' + \beta_{j_0} \|y_{j_0}''\|^2 y_{j_0}'' = 0. \end{cases}$$

Consequently, we have $y_1' = 0$ and $\|y_{j_0}''\| = 1$.

The remaining possibility, i.e. $\|y_1''\| \neq 0$ and $y_1 = y_{i_0}'$, is equivalent to case (2). These considerations prove that if (y_1, y_2) is a solution of system (6) then it satisfies one of the 4 listed cases. The other implication follows from straightforward calculations. $\square$

Remark 4.2. (1) We notice that if either R or I commutes with C then the condition of Theorem 4.3 is satisfied. This is so because $RI = IR$. Moreover, we can say that if T commutes with C, C can be written as a polynomial in T, therefore in $\sqrt{T}$. This implies that C commutes with both R and I.

(2) We point out that the general system (8) remains to be solved however the same previous techniques may still be applied under some restrictions. For example, if we assume $y_1 = 0$, system (8) reduces to

$$\begin{cases} RCIy_2 = 0, \\ ICIy_2 - (y_2 \cdot ICIy_2)y_2 = 0. \end{cases}$$

We apply Theorem 4.1 to conclude that the solution set of the system is given by the intersection

$$\mathrm{Ker}(RCI) \cap \bigcup_{i=1}^{s} \big(E_i \cap S^{n-1}\big),$$

where s is the number of distinct eigenvalues of ICI. We encounter a similar situation if we assume $y_2 = O$.

4.3. *Stability study*

The zeros of Eq. (1) are those weight values expected to remain unchanged under the algorithm's action. This is due to the fact that $TC\omega - (\omega \cdot C\omega)\omega$ represents the expected weight change. Given an initial weight ω_0, the new value is determine by the formula

$$\omega_1 = \omega_0 + \Delta\omega_0,$$

where $\Delta\omega_0 = TC\omega_0 - (\omega_0 \cdot C\omega_0)\omega_0$. To identify those initial weights for which the iterative process stabilizes is desirable since this amounts to saying that the network has learned and may perform as an educated device. It is a standard procedure to consider the differential equation

$$\frac{d\omega}{dt} = TC\omega - (\omega \cdot C\omega)\omega,$$

and study the stability behavior of each equilibrium point. The equilibria of this equation, under some symmetry assumptions, were determined in Theorems 4.1, 4.2, and 4.4. The strategy is to linearize the system around each equilibrium point and apply classical results in dynamical systems to conclude its stability behavior. As followed before, we first consider T positive and nonsingular. The differential equation to be studied is

$$\frac{dy}{dt} = \big(\sqrt{T}C\sqrt{T}y - \big(y \cdot \sqrt{T}C\sqrt{T}y\big)y\big), \quad \text{where } \omega = \sqrt{T}y. \tag{11}$$

The equilibria of Eq. (11) is given in Theorem 4.1. As in Section 3.1, let

$$\phi(y) = \sqrt{T}C\sqrt{T}y - \big(y \cdot \sqrt{T}C\sqrt{T}y\big)y$$

and y_0 some vector in R^n, the derivative of ψ at y_0 in the direction of a vector $v \in R^n$ is given by

$$D\phi(y_0)(v) = \sqrt{T}C\sqrt{T}v - 2\big(v \cdot \sqrt{T}C\sqrt{T}y_0\big)y_0 - \big(y_0 \cdot \sqrt{T}C\sqrt{T}y_0\big)v.$$

Remark 4.3. If $y_0 \in \mathrm{Ker}(\sqrt{T}C\sqrt{T})$ then $D\phi(y_0) = \sqrt{T}C\sqrt{T}$. Therefore, $\mathrm{Ker}(\sqrt{T}C\sqrt{T})$ is attracting, repelling, or saddle-type if and only if the eigenvalues of $\sqrt{T}C\sqrt{T}$ are all negative, all positive, or there are two eigenvalues whose product is negative, respectively.

Lemma 4.3. *If C is positive then $\sqrt{T}C\sqrt{T}$ is positive.*

Proof. For every vector y we have $y \cdot \sqrt{T}C\sqrt{T}y = \sqrt{T}y \cdot C\sqrt{T}y \geqslant 0$. $\square$

Proposition 4.2. *If T and C are positive, T is nonsingular, and C is symmetric then y_0 is an attracting equilibrium point of Eq.* (11) *if and only if it is a unit eigenvector associated to the largest simple eigenvalue of $\sqrt{T}C\sqrt{T}$.*

Proof. Let $\{\alpha_1, \ldots, \alpha_t\}$ be the simple eigenvalues of $\sqrt{T}C\sqrt{T}$ and $\{\beta_1, \ldots, \beta_k\}$ those eigenvalues with higher multiplicity, $\{n_1, \ldots, n_k\}$, respectively. Attached to the set of eigenvalues we define an orthonormal basis of eigenvectors denoted by

$$\mathcal{B} = \left\{u^{\alpha_1}, \ldots, u^{\alpha_t}, u_1^{\beta_1}, \ldots, u_{n_1}^{\beta_1}, \ldots, u_1^{\beta_k}, \ldots, u_{n_k}^{\beta_k}\right\}.$$

We assume that y_0 is one of the eigenvectors associated to a simple eigenvalue, say $y_0 = u^{\alpha_j}$. The matrix representing $D\phi(y_0)$, relatively to $\mathcal{B}$, is diagonal, with diagonal entries given by

$$d_{ii} = \begin{cases} \alpha_i - \alpha_j, & \text{if } i \neq j \text{ and } i = 1, \ldots, t, \\ -2\alpha_j, & \text{if } i = j, \\ \beta_i - \alpha_j, & \text{if } i = t+1, \ldots, n. \end{cases}$$

Therefore, y_0 is an attracting equilibrium point if and only if α_j is the largest eigenvalue. Let $y_0 = \sum_{i=1}^{n_1} a_i u_i^{\beta_1}$ where $\sum_{i=1}^{n_1} a_i^2 = 1$ (or $y_0 \cdot y_0 = 1$). The matrix representing $D\phi(y_0)$, relatively to $\mathcal{B}$, is a block matrix of the following form:

$$\begin{bmatrix} A & O & O \\ O & B & O \\ O & O & C \end{bmatrix}.$$

The t-matrix A is diagonal whose ith-diagonal entry is equal to $\alpha_i - \beta_1$, B is a n_1-matrix given by

$$-2\beta_1 \begin{bmatrix} a_1 \\ \vdots \\ a_{n_1} \end{bmatrix} [a_1, \ldots, a_{n_1}],$$

and C is an diagonal matrix, of dimension $n - n_1 - t$, composed by $k-1$ diagonal blocks C_i with $i = 2, \ldots, k$. Each C_i has all the diagonal entries equal to $\beta_i - \beta_1$, with $i = 2, \ldots, k$. The linear transformation $D\phi(y_0)$ has an eigenvalue equal to 1, namely $\sum_{i=1}^{n_1} a_i^2 = 1$. This implies that y_0 is not stable. This completes the proof of the theorem. $\square$

Remark 4.4. From arguments presented in the proof above, the invariant sphere defined by $\sum_{i=1}^{n_1} a_i u_i^{\beta_1}$, where $\sum_{i=1}^{n_1} a_i^2 = 1$, is stable if and only if β_1 is the largest eigenvalue of $\sqrt{T}C\sqrt{T}$.

5. Conclusions

In this paper we have studied a learning rule proposed by P. Adams. This rule involves synaptic adjustments that incorporate both, a probabilistic component representing synaptic replication and a nonlinear forgetting term.

A learning rule is interpreted as an algorithm that searches for a natural weight vector whose components are assigned as synaptic weights to the underlying network. This algorithm relies upon a selection based on an iterative action on a collection of inputs. A selection may be achieved when an iterative rule stabilizes. We deduce the rule in Section 2 and reduce the problem to solving the matrix equation (1) and studying the stability type of its solutions. The solution set of Eq. (1), generically, consists of an hyperplane (kernel of resulting correlation matrix), a set of isolated points (unit eigenvectors of the same resulting matrix), and a set of spheres (nonisolated unit eigenvectors of the same matrix). At each one of these vectors the algorithm halts. The identification of those stable ones is desirable since the probability of choosing one such vector is extremely small. The hope is to increase such probability allowing the algorithm to run for sometime and observe convergence. This study is done in Section 4.3 where we show that under some conditions we observe convergence to a principal component. In addition, this allows us to conclude that, Adams' model acts as an information filter as in Oja's proposed model. Furthermore, when spheres exist, considered as invariant spaces, the analysis of their stability follows a similar behavior as the one presented for isolated vectors. The biological interpretation of this seems unclear at this point, but this fact might bring an enriched filtering performance.

References

[1] P. Adams, Hebb and Darwin, J. Theoret. Biol. 195 (1998) 419–438.
[2] D. Arrowsmith, C. Place, An Introduction to Dynamical Systems, Cambridge University Press, Cambridge, 1990.
[3] U. Grenander, G. Szegö, Toeplitz Forms and Their Applications, Chelsea, New York, 1984.
[4] P. Halmos, Finite Dimensional Vector Spaces, in: Undergrad. Texts Math., Springer-Verlag, New York, 1970.
[5] S. Haykin, Neural Networks: A Comprehensive Foundation, Macmillan, New York, 1994.
[6] D. Hebb, The Organization of Behavior: A Neurophysiological Theory, Wiley, New York, 1949.
[7] J. Hertz, A. Krogh, R. Palmer, Introduction to the Theory of Neural Computation, Addison–Wesley, Reading, MA, 1991.
[8] M. Hirsch, S. Smale, Differential Equations, Dynamical Systems, and Linear Algebra, Academic Press, New York, 1974.
[9] R. Horn, C. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[10] R. Linsker, Perceptual neural organization: Some approaches based on network models and information theory, Ann. Rev. Neurosci. 13 (1990) 181–257.
[11] K. Kingsley, P. Adams, Formation of new connections by a generalisation of Hebbian learning, preprint (2001).
[12] T. Kohonen, An introduction to neural computing, Neural Networks 1 (1988) 3–16.

[13] T. Kohonen, Self-Organization and Associative Memory, Springer-Verlag, Berlin, 1988.

[14] T. Kohonen, Self-organized formation of topologically correct feature maps, Biol. Cybernet. 43 (1982) 59–69.

[15] Y. Kuznetsov, Elements of Applied Bifurcation Theory, in: Appl. Math. Sci., Vol. 112, Springer-Verlag, New York, 1998.

[16] P. Montague, J. Gally, G. Edelman, Spatial signalling in the development and function of neural connecting, Cereb. Cortex 1 (1991) 199–220.

[17] E. Oja, A simplified neuron model as a principal component analyzer, J. Math. Biol. 15 (1982) 267–273.

[18] E. Oja, J. Karhunen, A stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix, J. Math. Anal. Appl. 106 (1985) 69–84.

[19] E. Oja, Principal components, minor components, and linear neural networks, Neural Networks 5 (1992) 927–936.

[20] J. Scheick, Linear Algebra with Applications, in: Internat. Ser. Pure Appl. Math., McGraw–Hill, New York, 1997.