Electronic Journal of Qualitative Theory of Differential Equations
2021, No. **92**, 1–20; https://doi.org/10.14232/ejqtde.2021.1.92 www.math.u-szeged.hu/ejqtde/

Minimizing of the quadratic functional on Hopfield networks

Oleksandr Boichuk[✉], **Oleksandr Pokutnyi**, **Viktor Feruk** and **Dmytro Bihun**

Institute of Mathematics of the National Academy of Sciences,
3 Tereschenkivska Street, Kyiv, 01024, Ukraine

Received 12 May 2021, appeared 29 December 2021
Communicated by Michal Fečkan

Abstract. In this paper, we consider the continuous Hopfield model with a weak interaction of network neurons. This model is described by a system of differential equations with linear boundary conditions. Also, we consider the questions of finding necessary and sufficient conditions of solvability and constructive construction of solutions of the given problem, which turn into solutions of the linear generating problem, as the parameter ε tends to zero. An iterative algorithm for finding solutions has been constructed. The problem of finding the extremum of the target functions on the given problem solution is considered. To minimize a functional, an accelerated method of conjugate gradients is used. Results are illustrated with examples for the case of three neurons.

Keywords: boundary-value problem, Moore–Penrose pseudo-inverse matrix, differential equations, Hopfield networks, quadratic functional.

2020 Mathematics Subject Classification: 34B05, 49K15.

1 Introduction

The study of various natural and social phenomena is carried out today by building and investigating their mathematical models. Practical applications contributed to the birth and development of many mathematical disciplines. Among them, there is a theory of dynamic neural networks, which are used to solve various optimization problems, control theory and mathematical modelling. The variety of tasks to be solved led to the existence of several models of such networks. An important place among them takes Hopfield model (see [22,25,43]), a single layer neural network with general non-linear and additional internal linear connections among neurons. Hopfield nets have a large number of publications. Both models with discrete and continuous time are considered. In particular, such questions as stability (see [47]), absolute stability of neural nets (see [15]), modelling of closed control systems, asymptotics and

[✉]Corresponding author. Email: boichuk.aa@gmail.com

stability of relaxation self-oscillations in Hopfield nets with delay (see [19,20]) are considered. The vital phenomenon is flow invariance for such systems (see [34]). Ill-posed problems with fractional derivative (see [46]), optimization problems (see [17,26,45]), deep neural networks (see [35]), relativistic Hopfield model (see [1]), quantum generalization of Hopfield model (see [40]), as well as its discrete analogue (see [2]) are studied. Chaos is explored in the corresponding models (see [14]). The Hopfield model is considered as a model of memory (see [23]). The impulsive Hopfield model with boundary conditions is studied in [38]. In this work, the case of weak interaction of network neurons is considered, to the study of which for other models, for example, the papers are devoted [27,29]. Using the theory of pseudo-inverse matrices (see [3–8,41]), an approach allows establishing necessary and sufficient conditions for the solvability of boundary-value problem for a system of differential equations that describe Hopfield network for n neurons with weak interaction. We use hyperbolic tangent as the increasing activation function and symmetric matrix of weights as in [22, p. 690]). The application of the accelerated method of conjugate gradients (see [31,32]) for solving the problem of finding the extremum (minimum) of the loss function is explored on the solutions of the given problem in the form of a quadratic functional of synaptic communication scales.

2 Formulation of the problem

We consider a continuous Hopfield model with a weak interaction of the network neurons, the evolution in time of which is described by a system of n non-linear differential equations (see [22, p. 690], [43, p. 140])

$$x_j'(t) = -\frac{x_j(t)}{R_j} + \varepsilon\left(\hat{I}_j(t) + \sum_{i=1}^{n} w_{ij}\tanh\left(\frac{a_i x_i(t)}{2}\right)\right) + I_j(t), \quad j = \overline{1,n}, \tag{2.1}$$

where $x_j(t) \in \mathbf{C}^1[0,T]$ is the potential of the jth neuron; the real parameters a_j are gain coefficients of the jth neuron, and w_{ij} are the elements of a symmetric matrix W:

$$W = \begin{pmatrix} 0 & w_{12} & \dots & w_{1n} \\ w_{12} & 0 & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1n} & w_{2n} & \dots & 0 \end{pmatrix},$$

which consists of synaptic weights of the connection of the ith neuron with the jth neuron, R_j are the leakage resistances, $\hat{I}_j(t) \in \mathbf{C}^1[0,T]$, $I_j(t) \in \mathbf{C}^1[0,T]$ are external signals, $\varepsilon \ll 1$ is a small parameter characterizing the strength of the interaction of network neurons.

As is known (see [22, p. 693], [43, p. 144]), in practice the property of a monotonic increasing of the activation function (in our case this is hyperbolic tangent) of the considered matrix W ($w_{ij} = w_{ji}$, $w_{ii} = 0$) and the asynchronous mode of network operation are often used. It provides the global asymptotic stability of the Hopfield network. These features persist in the case of a weak interaction of network neurons described by equation (2.1) and provide both practical and theoretical interest in Hopfield networks. For the convenience of further reasoning, we rewrite the equation (2.1) in the following form

$$x'(t,w,\varepsilon) = Ax(t,w,\varepsilon) + \varepsilon\left(\hat{I}(t) + WZ(x(t,w,\varepsilon))\right) + I(t), \tag{2.2}$$

where

$$x(t,w,\varepsilon) = \operatorname{col}\left(x_1(t,w,\varepsilon), \quad x_2(t,w,\varepsilon), \quad \dots, \quad x_n(t,w,\varepsilon)\ \right),$$

$$A = -\operatorname{diag}\left\{\tfrac{1}{R_1}, \quad \tfrac{1}{R_2}, \quad \ldots, \quad \tfrac{1}{R_n}\right\},$$

$$Z(x(t,w,\varepsilon)) = \operatorname{col}\left(\tanh\left(\tfrac{a_1 x_1(t,w,\varepsilon)}{2}\right), \quad \tanh\left(\tfrac{a_2 x_2(t,w,\varepsilon)}{2}\right), \quad \ldots, \quad \tanh\left(\tfrac{a_n x_n(t,w,\varepsilon)}{2}\right)\right),$$

$$\hat{I}(t) = \operatorname{col}\left(\hat{I}_1(t), \quad \hat{I}_2(t), \quad \ldots, \quad \hat{I}_n(t)\right), \quad I(t) = \operatorname{col}\left(I_1(t), \quad I_2(t), \quad \ldots, \quad I_n(t)\right),$$

w is a vector of dimension $M = n(n-1)/2$ formed from the elements of the matrix W in the following way:

$$w = \operatorname{col}\left(w_{12}, \quad w_{22}, \quad \ldots, \quad w_{1n}, \quad w_{23}, \quad w_{24}, \quad \ldots, \quad w_{2n}, \quad \ldots, \quad w_{(n-1)n}\right).$$

In some cases, the solutions of systems of equations describing the functioning of neural nets satisfy additional conditions due to particular properties of the modelled process. Various types of boundary-value problems are explored for such systems (see [12,38,44]). In our paper, we investigate the questions of finding conditions for the existence and effective construction of equation (2.2) solutions with m boundary conditions

$$lx(\cdot, w, \varepsilon) = \alpha, \tag{2.3}$$

$l = \operatorname{col}\left(l_1, \quad l_2, \quad \ldots, \quad l_m\right) : \mathbf{C}^1[0,T] \to \mathbb{R}^m$ is bounded linear vector functional, $l_\nu : \mathbf{C}^1[0,T] \to \mathbb{R}$, $\nu = \overline{1,m}$, $\alpha = \operatorname{col}\left(\alpha_1, \quad \alpha_2, \quad \ldots, \quad \alpha_m\right) \in \mathbb{R}^m$, which for $\varepsilon = 0$ turns into the solution of the generating problem

$$x'(t) = Ax(t) + I(t), \tag{2.4}$$

$$lx(\cdot) = \alpha. \tag{2.5}$$

These solutions will be called generating solutions of the boundary-value problem (2.2), (2.3). Note that the boundary-value problem (2.2), (2.3) includes both underdetermined ($m < n$) and overdetermined ($m > n$) boundary-value problems, the study of which for Hopfield models is not given enough attention, in our opinion.

3 Necessary condition for the solvability of the problem (2.2), (2.3)

Let us first consider the question of solution existence to the problem (2.2), (2.3). For this purpose, we use the general scheme for the exploration of boundary-value problems studied in detail in [7], which allows finding effective coefficients which are necessary and sufficient for the solvability of problem (2.2), (2.3). In particular, for the generating problem (2.4), (2.5), the following criterion holds (see [7]).

Theorem 3.1. *The homogeneous problem* (2.4), (2.5) ($I(t) = 0, \alpha = 0$) *has an r-parametric ($r \leq n$) family of solutions $x(t, c_r) \in \mathbf{C}^1[0,T]$*

$$x(t, c_r) = U(t) P_{Q_r} c_r \quad \forall c_r \in \mathbb{R}^r.$$

The inhomogeneous problem (2.4), (2.5) *is solvable if and only if g satisfies d ($d \leq m$) linearly independent conditions:*

$$P_{Q_d^*} g = 0. \tag{3.1}$$

In this case, the inhomogeneous problem (2.4), (2.5) *has an r-parameter family of solutions $x(t, c_r) \in \mathbf{C}^1[0,T]$ of the following form:*

$$x(t, c_r) = U(t) P_{Q_r} c_r + (G[I, \alpha])(t) \quad \forall c_r \in \mathbb{R}^r, \tag{3.2}$$

where

$$(G[I,\alpha])\,(t) := U(t)\left(Q^{+}g + \int_0^t U^{-1}(\tau)I(\tau)d\tau\right)$$

is the generalized Green operator.

Here

$$U(t) = \operatorname{diag}\left\{e^{-\frac{t}{R_1}},\ \ e^{-\frac{t}{R_2}},\ \ \ldots,\ \ e^{-\frac{t}{R_n}}\right\}$$

is a fundamental decision matrix of the linear homogeneous system (2.4),

$$g = \alpha - l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)I(\tau)d\tau,$$

$Q = lU(\cdot)$ is a matrix of dimension $(m \times n)$, P_{Q_r} $\left(P_{Q_d^*}\right)$ is the matrix which consists of the complete system r (d) of linearly independent columns (rows) of the projector matrix P_Q (P_{Q^*}), where P_Q (P_{Q^*}) is projector onto the kernel (cokernel) of the matrix Q, Q^{+} is the Moore–Penrose pseudo-inverse (see [37]) to the Q matrix.

Let us find necessary conditions for the existence of a solution $x(t,w,\varepsilon)$ to the problem (2.2), (2.3), which for $\varepsilon = 0$ turns into one of the solutions $x(t,c_r)$ of the generating problem (2.4), (2.5). According to Theorem 3.1, boundary-value problem (2.2), (2.3) is solvable if and only if d linearly independent conditions are satisfied

$$P_{Q_d^*}\left(g - \varepsilon l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)\left(\hat{I}(\tau) + WZ(x(\tau,w,\varepsilon))\right)d\tau\right) = 0. \tag{3.3}$$

Taking into account (3.1), we obtain that condition (3.3) is equivalent to the following

$$P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)\left(\hat{I}(\tau) + WZ(x(\tau,w,\varepsilon))\right)d\tau = 0. \tag{3.4}$$

Considering the limit for (3.4) as $\varepsilon \to 0$ and also taking into account that $x(t,w,\varepsilon) \to x(t,c_r)$ in this case, we obtain the following solvability condition

$$F(c_r) := P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)\left(\hat{I}(\tau) + WZ(x(\tau,c_r))\right)d\tau = 0. \tag{3.5}$$

Note that in the case of the periodic boundary-value problem (2.2), (2.3) ($lx(\cdot,w,\varepsilon) = x(0,w,\varepsilon) - x(T,w,\varepsilon) = \alpha = 0$) equation (3.5) corresponds to that known in the theory of nonlinear oscillations of the equation for the generating amplitudes (see [21,33]). Therefore, we will call the equation (3.5) the equation for the generating vectors of boundary-value problem (2.2), (2.3). If equation (3.5) has a solution $c_r = c_r^0 \in \mathbb{R}^r$, then c_r^0 defines the solution

$$x\left(t,c_r^0\right) = \operatorname{col}\left(x_1\left(t,c_r^0\right),\ \ x_2\left(t,c_r^0\right),\ \ \ldots,\ \ x_n\left(t,c_r^0\right)\right)$$

of the generating problem (2.4), (2.5), which may correspond to the solution $x(t,w,\varepsilon)$ of the problem (2.2), (2.3). If the equation (3.5) has no solutions, then problem (2.2), (2.3) also does not have the desired solution. Note that since we are considering the original problem in real form, we are only talking about real solutions of equation (3.5).

Thus, the following statement is true.

Theorem 3.2 (Necessary condition)**.** *If the boundary-value problem* (2.2), (2.3) *has a solution, which for* $\varepsilon = 0$ *turns into one of the solutions* $x\left(t,c_r^0\right)$ *generating boundary-value problem* (2.4), (2.5), *then the vector* $c_r^0 \in \mathbb{R}^r$ *must be a real solution to the equation for the generating vectors* (3.5).

4 Optimization of the objective function

One of the important research questions of exploring neural networks is finding the extremum (minimum) or the objective function by solving the considered model. In particular, the such problems, arising in medicine, neurobiology, machine learning, were studied in [11–13, 36, 42, 44]. In this paper, we consider the problem of finding the minimum of the objective function $L(x(t, w, \varepsilon), w)$:

$$L(x(t, w, \varepsilon), w) \to \min_{w \in \mathbb{R}^M},$$

$$L(\cdot, w) \in \mathbf{C}\left[\|x - x_0\| \le q\right], \quad L(x, \cdot) \in \mathbf{C}\left(\mathbb{R}^M\right)$$

on the solutions of boundary-value problem (2.2), (2.3), which at $\varepsilon = 0$ turn into generating solution of (2.4), (2.5). Here, x_0 is the generating solution and q is a small parameter.

Suppose that when ε tends to 0 function $L(x(t, w, \varepsilon), w)$ takes the form of quadratic functional by vector of parameters $w \in \mathbb{R}^M$, that is

$$L\left(x\left(t, c_r^0\right), w\right) = \Phi(w) = (Sw, w) - 2(f, w) \to \min_{w \in \mathbb{R}^M}, \tag{4.1}$$

where $x\left(t, c_r^0\right)$ is a solution of the generating problem (2.4), (2.5) and $S : \mathbb{R}^M \to \mathbb{R}^M$ is the linear self-adjoint bounded positive operator (positive definite quadratic form), that is

$$\gamma_1 \|u\|^2 \le (Su, u) \le \gamma_2 \|u\|^2, \quad \gamma_2 > \gamma_1 > 0, \quad \forall u \in \mathbb{R}^M, \tag{4.2}$$

$f \in \mathbb{R}^M$. Restriction (4.1), as known from [30], is equivalent to finding solutions w of the following equation

$$Sw = f. \tag{4.3}$$

To minimize functional (4.1) we use the accelerated method of conjugate gradients, which, as known from [31, 32], improves the convergence of the method of steepest descent and the conjugate gradient method, expands their scope and is more robust to rounding errors. Since S satisfies condition (4.2), the functional (4.1) has a unique minimum w^* (equation (4.3) has a unique solution for any f) (see [30]).

Let us take a closer look at the accelerated method of conjugate gradients. Its essence for the minimization of functional (4.1) is, that based on some initial value approximation $w = w^0$, the following approximate solutions are determined according to the formulas

$$w^{k+1} = w^k + \alpha_k r_k + \beta_k \delta_k + \sigma_k, \tag{4.4}$$

$$r_k = f - Sw^k, \quad \delta_k = w^k - w^{k-1}, \quad \sigma_k = \sum_{i=1}^{n_0} a_i^k \varphi_i, \tag{4.5}$$

where φ_i, $i = \overline{1, n_0}$, $n_0 \le M$ is a system of linearly independent elements and the unknown parameters α_k, β_k and a_i^k we will determine from the system of linear algebraic equations

$$\frac{\partial \Phi\left(w^{k+1}\right)}{\partial \alpha_k} = 0, \quad \frac{\partial \Phi\left(w^{k+1}\right)}{\partial \beta_k} = 0, \quad \frac{\partial \Phi\left(w^{k+1}\right)}{\partial a_i^k} = 0. \tag{4.6}$$

Note that in [31], using the form of functional (4.1) and the rule of differentiation of scalar product, a convenient for practical application computational scheme of the method (4.4)–(4.6) is given.

Remark 4.1. As known (see [31]), the use of the method (4.4)–(4.6) in the space $\mathbb{R}^M$ allows to obtain an exact solution to equation (4.3) in $k \leq M$ iterations.

Remark 4.2. For $\sigma_k = 0$, the accelerated method of conjugate gradients (4.4)–(4.6) transforms into the conjugate gradient method, and for $\delta_k = 0$, $\sigma_k = 0$ – into the method of steepest descent (see [31,32]).

Let us formulate, using the results of [31], an estimate for the rate of convergence of the corresponding conjugate gradient method (4.4)–(4.6) for our optimization problem. Let H_{n_0}, $n_0 \leq M$ be the subspace spanned on a system of linearly independent elements $\{\varphi_i\}_{i=1}^{n_0}$. We introduce into consideration a self-adjoint mapping in the space V_{n_0},

$$\mathbb{R}^M = H_{n_0} \oplus V_{n_0},$$

operator $K = SZ$ which satisfies the condition

$$\eta_1 \|v\|^2 \leq (Kv, v) \leq \eta_2 \|v\|^2, \quad \gamma_1 \leq \eta_1 \leq \eta_2 \leq \gamma_2, \quad \forall v \in V_{n_0}.$$

Here the operator Z is linear and is defined by the formula

$$Zg = g + h,$$

where $g \in \mathbb{R}^M$ is an arbitrary element and $h \in H_{n_0}$ is a solution of equation

$$PS(g + h) = 0,$$

where P is the operator of orthogonal projection $\mathbb{R}^M$ onto H_{n_0}.

The following statement is true (see [31]).

Theorem 4.3. *Let the operator S satisfy condition* (4.2)*. Then, the accelerated method of conjugate gradients* (4.4)–(4.6) *converges and the rate of its convergence is characterized by estimate*

$$\left\| w^* - w^k \right\| \leq \frac{q_k}{\sqrt{\gamma_1 \eta_1}} \left\| f - Sw^0 \right\|,$$

where

$$q_k = \frac{2\rho^k}{1 + \rho^{2k}}, \quad \rho = \frac{\sqrt{\eta_2} - \sqrt{\eta_1}}{\sqrt{\eta_2} + \sqrt{\eta_1}}.$$

5 A sufficient condition for the solvability of problem (2.2), (2.3)

For the further investigation of the problem (2.2), (2.3), let us fix the value of the vector of parameters $w = w^*$, which is found using the accelerated method of conjugate gradients (4.4)–(4.6). To obtain a sufficient condition for the existence of a solution, we make the following change in variables in the boundary-value problem (2.2), (2.3):

$$x(t, w^*, \varepsilon) = x\left(t, c_r^0\right) + y(t, w^*, \varepsilon), \tag{5.1}$$

where $x\left(t, c_r^0\right)$ is a solution of the generating boundary-value problem (2.4), (2.5),

$$y(t, w, \varepsilon) = \operatorname{col}\left(y_1(t, w, \varepsilon), \quad y_2(t, w, \varepsilon), \quad \ldots, \quad y_n(t, w, \varepsilon)\right)$$

and $c_r^0 \in \mathbb{R}^r$ is a solution to the equation for the generating vectors (3.5). By replacing the variables in (5.1), the study of the existence of a solution to problem (2.2), (2.3) is reduced to the corresponding question for the boundary-value problem

$$y'(t, w^*, \varepsilon) = Ay(t, w^*, \varepsilon) + \varepsilon \left(\hat{I}(t) + W^* Z\left(x\left(t, c_r^0\right) + y(t, w^*, \varepsilon)\right)\right), \tag{5.2}$$
$$ly(\cdot, w^*, \varepsilon) = 0. \tag{5.3}$$

As follows from the vector-function $Z(x(t, w^*, \varepsilon))$, it is differentiable in the neighbourhood of the generating solution $x\left(t, c_r^0\right)$, therefore, the following representation holds:

$$Z\left(x\left(t, c_r^0\right) + y(t, w^*, \varepsilon)\right) = Z\left(x\left(t, c_r^0\right)\right) + A_1(t)y(t, w^*, \varepsilon) + \mathcal{R}(y(t, w^*, \varepsilon)),$$

where

$$Z\left(x\left(t, c_r^0\right)\right) = \operatorname{col}\left(\tanh\left(\frac{a_1 x_1\left(t, c_r^0\right)}{2}\right), \quad \tanh\left(\frac{a_2 x_2\left(t, c_r^0\right)}{2}\right), \quad \ldots, \quad \tanh\left(\frac{a_n x_n\left(t, c_r^0\right)}{2}\right)\right)$$

is a limit to which the function $Z(x(t, w^*, \varepsilon))$ tends under ε tends towards 0 and $c = c_r^0$,

$$\begin{aligned} A_1(t) &= Z'_x(v)|_{v = x(t, c_r^0)} \\ &= \frac{1}{2} \operatorname{diag}\left\{\frac{a_1}{\cosh^2\left(\frac{a_1 x_1(t, c_r^0)}{2}\right)}, \quad \frac{a_2}{\cosh^2\left(\frac{a_2 x_2(t, c_r^0)}{2}\right)}, \quad \ldots, \quad \frac{a_n}{\cosh^2\left(\frac{a_n x_n(t, c_r^0)}{2}\right)}\right\} \end{aligned}$$

is derivative in the sense of Fréchet, and $\mathcal{R}(y(t, w^*, \varepsilon))$ are higher-order members

$$\begin{aligned} \mathcal{R}(y(t, w^*, \varepsilon)) &= Z\left(x\left(t, c_r^0\right) + y(t, w^*, \varepsilon)\right) - Z\left(x\left(t, c_r^0\right)\right) - A_1(t)y(t, w^*, \varepsilon) \\ &= \begin{pmatrix} \tanh\left(\frac{a_1\left(x_1\left(t, c_r^0\right) + y_1(t, w^*, \varepsilon)\right)}{2}\right) - \tanh\left(\frac{a_1 x_1\left(t, c_r^0\right)}{2}\right) - \frac{a_1 y_1(t, w^*, \varepsilon)}{2\cosh^2\left(\frac{a_1 x_1\left(t, c_r^0\right)}{2}\right)} \\ \tanh\left(\frac{a_2\left(x_2\left(t, c_r^0\right) + y_2(t, w^*, \varepsilon)\right)}{2}\right) - \tanh\left(\frac{a_2 x_2\left(t, c_r^0\right)}{2}\right) - \frac{a_2 y_2(t, w^*, \varepsilon)}{2\cosh^2\left(\frac{a_2 x_2\left(t, c_r^0\right)}{2}\right)} \\ \vdots \\ \tanh\left(\frac{a_n\left(x_n\left(t, c_r^0\right) + y_n(t, w^*, \varepsilon)\right)}{2}\right) - \tanh\left(\frac{a_n x_n\left(t, c_r^0\right)}{2}\right) - \frac{a_n y_n(t, w^*, \varepsilon)}{2\cosh^2\left(\frac{a_n x_n\left(t, c_r^0\right)}{2}\right)} \end{pmatrix}. \end{aligned}$$

Thus, the boundary-value problem (5.2), (5.3) takes the following form

$$\begin{aligned} y'(t, w^*, \varepsilon) &= Ay(t, w^*, \varepsilon) \\ &\quad + \varepsilon\left(\hat{I}(t) + W^*\left(Z\left(x\left(t, c_r^0\right)\right) + A_1(t)y(t, w^*, \varepsilon) + \mathcal{R}(y(t, w^*, \varepsilon))\right)\right), \end{aligned} \tag{5.4}$$
$$ly(\cdot, w^*, \varepsilon) = 0. \tag{5.5}$$

According to the Theorem 3.1, under d conditions

$$\begin{aligned} &P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*\left(Z\left(x\left(\tau, c_r^0\right)\right) + A_1(\tau)y(\tau, w^*, \varepsilon) + \mathcal{R}(y(\tau, w^*, \varepsilon))\right) d\tau \\ &\quad = -P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)\hat{I}(\tau) d\tau, \end{aligned} \tag{5.6}$$

boundary-value problem (5.4), (5.5) has an r-parametric family of solutions of the following form

$$\begin{aligned} y(t,w^*,\varepsilon) &= U(t)P_{Q_r}c_r + \overline{y}(t,w^*,\varepsilon) \quad \forall c_r \in \mathbb{R}^r, \\ \overline{y}(t,w^*,\varepsilon) &= \varepsilon\left(G\left[\hat{I}(t) + W^*Z\left(x\left(t,c_r^0\right) + y(t,w^*,\varepsilon)\right),0\right]\right)(t,\varepsilon). \end{aligned} \tag{5.7}$$

Using condition (3.5), relation (5.6) can be rewritten as

$$P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*(A_1(\tau)y(\tau,w^*,\varepsilon) + \mathcal{R}(y(\tau,w^*,\varepsilon)))d\tau = 0. \tag{5.8}$$

Substituting (5.7) into (5.8), we obtain the following equation for c_r:

$$Bc_r = -P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*H\left(\tau,y(\tau,w^*,\varepsilon),\overline{y}(\tau,w^*,\varepsilon)\right)d\tau, \tag{5.9}$$

$$H\left(t,y(t,w^*,\varepsilon),\overline{y}(t,w^*,\varepsilon)\right) = A_1(t)\overline{y}(t,w^*,\varepsilon) + \mathcal{R}\left(y(t,w^*,\varepsilon)\right),$$

where matrix B $(d \times r)$ has the following form

$$B = P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*A_1(\tau)U(\tau)P_{Q_r}d\tau. \tag{5.10}$$

The algebraic system (5.9) is solvable if and only if d_1 conditions hold

$$P_{B_{d_1}^*}P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*H\left(\tau,y(\tau,w^*,\varepsilon),\overline{y}(\tau,w^*,\varepsilon)\right)d\tau = 0. \tag{5.11}$$

If, for example,

$$P_{B_{d_1}^*}P_{Q_d^*} = 0, \tag{5.12}$$

then the condition (5.11) is always valid, and system (5.9) has r_1-parametric solution

$$\begin{aligned} c_r &= P_{B_{r_1}}\hat{c}_{r_1} + \bar{c}_r \quad \forall \hat{c}_{r_1} \in \mathbb{R}^{r_1}, \\ \bar{c}_r &= -B^+P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*H\left(\tau,y(\tau,w^*,\varepsilon),\overline{y}(\tau,w^*,\varepsilon)\right)d\tau. \end{aligned}$$

Here $P_{B_{r_1}}\left(P_{B_{d_1}^*}\right)$ is a matrix that consists of the complete system r_1 (d_1) of linearly independent columns (rows) of the projector matrix $P_B\left(P_{B^*}\right)$, where $P_B\left(P_{B^*}\right)$ is the projector on kernel (cokernel) of the matrix B, B^+ is the Moore–Penrose pseudo-inverse to the matrix B.

From now on we will restrict ourselves to the particular solution $c_r = \bar{c}$ of the system (5.9). So, for defining the solution of the problem (5.2), (5.3) we come to the system of equations

$$\begin{aligned} y(t,w^*,\varepsilon) &= U(t)P_{Q_r}c_r + \overline{y}(t,w^*,\varepsilon), \\ c_r &= -B^+P_{Q_d^*}l\int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^*H\left(\tau,y(\tau,w^*,\varepsilon),\overline{y}(\tau,w^*,\varepsilon)\right)d\tau, \\ \overline{y}(t,w^*,\varepsilon) &= \varepsilon\left(G\left[\hat{I}(t) + W^*Z\left(x\left(t,c_r^0\right) + y(t,w^*,\varepsilon)\right),0\right]\right)(t,\varepsilon), \end{aligned}$$

which can be solved using a convergent iterative process explained in detail in [7]. The following statement is true.

Theorem 5.1 (Sufficient condition). *Let the generating problem for* (2.2), (2.3) *problem* (2.4), (2.5), *subject to the conditions of d linearly independent conditions* (3.1), *have an r-parametric family of solutions $x\left(t,c_r^0\right)$* (3.2) *and the operator S satisfy condition* (4.2). *Then for every real value of the* *vector $c_r^0 \in \mathbb{R}^r$, which satisfies the equation for the generating vectors (3.5), for the value of the parameter vector $w^* \in \mathbb{R}^M$ minimizing the quadratic functional (4.1), and when conditions (5.12) hold, the boundary-value problem (2.2), (2.3) has a solution that can be found using the following iterative process*

$$c_r^k = -B^+ P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^* H\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right) d\tau,$$

$$\overline{y}^{k+1}(t, w^*, \varepsilon) = \varepsilon\Big(G\Big[\hat{I}(t) + W^*\Big(Z\left(x\left(t, c_r^0\right)\right) + A_1(t)U(t)P_{Q_r}c_r^k \\ + H\left(t, y^k(t, w^*, \varepsilon), \overline{y}^k(t, w^*, \varepsilon)\right)\Big), 0\Big]\Big)(t, \varepsilon),$$

$$y^{k+1}(t, w^*, \varepsilon) = U(t)P_{Q_r}c_r^k + \overline{y}^{k+1}(t, w^*, \varepsilon),$$

$$x^k(t, w^*, \varepsilon) = y^k(t, w^*, \varepsilon) + x\left(t, c_r^0\right), \quad x(t, w^*, \varepsilon) = \lim_{k\to\infty} x^k(t, w^*, \varepsilon),$$

$$y^0(t, w^*, \varepsilon) = \overline{y}^0(t, w^*, \varepsilon) = 0.$$

Corollary 5.2. *Let $r = d$ and nonlinearity $F(c_r)$ has the inverse to $F'\left(c_r^0\right)$ for the vector c_r^0, that satisfies the equation (3.5). Then $F'\left(c_r^0\right) = B$, and for such each c_r^0, the boundary-value problem (2.2), (2.3) has a unique solution.*

Proof. Consider the difference

$$F(c_r + h) - F(c_r) = P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^* Z(x(\tau, c_r + h))d\tau \\ - P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^* Z(x(\tau, c_r))d\tau.$$

Based on the representation (3.2), that is $x(\tau, c_r) = U(\tau)P_{Q_r}c_r + (G[I, \alpha])(\tau)$, we obtain that

$$Z(x(\tau, c_r + h)) = Z\left(x(\tau, c_r) + U(t)P_{Q_r}h\right) = Z(x(\tau, c_r)) + A_1(\tau)U(\tau)P_{Q_r}h + \mathcal{R}\left(U(\tau)P_{Q_r}h\right),$$

where $\mathcal{R}\left(U(\tau)P_{Q_r}h\right)$ contains terms higher than the first order in h. Substituting the received equality in the difference $F(c_r + h) - F(c_r)$, we get the following:

$$F(c_r + h) - F(c_r) = P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^* \left(Z(x(\tau, c_r)) + A_1(\tau)U(\tau)P_{Q_r}h + \mathcal{R}\left(U(\tau)P_{Q_r}h\right)\right) d\tau \\ - P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^* Z(x(\tau, c_r))d\tau \\ = Bh + P_{Q_d^*} l \int_0^{\cdot} U(\cdot)U^{-1}(\tau)W^* \mathcal{R}\left(U(t)P_{Q_r}h\right) d\tau.$$

From the equation above we obtain that $F'\left(c_r^0\right) = B$. Thus, reversibility of $F'\left(c_r^0\right)$ implies the invertibility of the matrix B. From this follows, that $r_1 = d_1 = 0$, $P_{B_{r_1}} = P_{B_{d_1}^*} = 0$, condition (5.12) is satisfied, and non-linear boundary-value problem (2.2), (2.3) has a unique solution for each such c_r^0. □

Remark 5.3. To calculate the projectors and the Moore–Penrose pseudo-inverse matrices, one can use the well-known formulas (see [7, p. 48], [28, p. 454]).

6 Examples

Example 6.1. Consider the underdetermined boundary-value problem (2.2), (2.3) for three equations in case when boundary condition (2.3) is T-periodic in part of the coordinates and has the form

$$l\begin{pmatrix} x_1(\cdot,w,\varepsilon) \\ x_2(\cdot,w,\varepsilon) \\ x_3(\cdot,w,\varepsilon) \end{pmatrix} = \begin{pmatrix} x_1(T,w,\varepsilon)-x_1(0,w,\varepsilon) \\ x_2(T,w,\varepsilon)-x_2(0,w,\varepsilon) \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In this case, the matrix Q is defined by the equality

$$Q = \begin{pmatrix} e^{-\frac{T}{R_1}}-1 & 0 & 0 \\ 0 & e^{-\frac{T}{R_2}}-1 & 0 \end{pmatrix}.$$

The Moore–Penrose pseudo-inverse matrix Q^+ and vector g have the following form:

$$Q^+ = \begin{pmatrix} \frac{e^{\frac{T}{R_1}}}{1-e^{\frac{T}{R_1}}} & 0 \\ 0 & \frac{e^{\frac{T}{R_2}}}{1-e^{\frac{T}{R_2}}} \\ 0 & 0 \end{pmatrix}, \quad g = -\int_0^T \begin{pmatrix} e^{\frac{\tau-T}{R_1}} I_1(\tau) \\ e^{\frac{\tau-T}{R_2}} I_2(\tau) \end{pmatrix} d\tau,$$

and orthoprojectors P_Q, P_{Q^*} look as follows

$$P_Q = E_3 - Q^+Q = \operatorname{diag}\{0, \quad 0, \quad 1\}, \quad P_{Q^*} = E_2 - QQ^+ = O_2,$$

where E_2, E_3 identity matrices of dimensions 2 and 3, respectively, O_2 is zero matrix of dimension 2. That is, $r = 1$, $d = 0$ and the condition of solvability (3.1) holds, and the linear boundary-value problem (2.4), (2.5) has a one-parameter set of solutions of the following form:

$$x(t,c_r) = \begin{pmatrix} \frac{-1}{1-e^{\frac{T}{R_1}}}\int_0^T e^{\frac{\tau-t}{R_1}} I_1(\tau)d\tau + \int_0^t e^{\frac{\tau-t}{R_1}} I_1(\tau)d\tau \\ \frac{-1}{1-e^{\frac{T}{R_2}}}\int_0^T e^{\frac{\tau-t}{R_2}} I_2(\tau)d\tau + \int_0^t e^{\frac{\tau-t}{R_2}} I_2(\tau)d\tau \\ e^{-\frac{t}{R_3}}c_3 + \int_0^t e^{\frac{\tau-t}{R_3}} I_3(\tau)d\tau \end{pmatrix}.$$

The equation for the generating vectors (3.5) takes the form of the identity $F(c_r) \equiv 0$ for any vector c_r.

Example 6.2. Consider the original problem for three equations with the boundary condition (2.3) of this form

$$l\begin{pmatrix} x_1(\cdot,w,\varepsilon) \\ x_2(\cdot,w,\varepsilon) \\ x_3(\cdot,w,\varepsilon) \end{pmatrix} = \begin{pmatrix} x_1(T,w,\varepsilon)-x_1(0,w,\varepsilon) \\ x_2(T,w,\varepsilon)-x_2(0,w,\varepsilon) \\ x_3(T,w,\varepsilon)-x_3(0,w,\varepsilon) \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}. \tag{6.1}$$

In this case

$$Q = \operatorname{diag}\left\{e^{-\frac{T}{R_1}}-1, \quad e^{-\frac{T}{R_2}}-1, \quad e^{-\frac{T}{R_3}}-1\right\}$$

and the Moore–Penrose pseudo-inverse matrix Q^+ coincides with matrix Q^{-1}

$$Q^+ = Q^{-1} = \operatorname{diag}\left\{ \frac{e^{\frac{T}{R_1}}}{1-e^{\frac{T}{R_1}}}, \quad \frac{e^{\frac{T}{R_2}}}{1-e^{\frac{T}{R_2}}}, \quad \frac{e^{\frac{T}{R_3}}}{1-e^{\frac{T}{R_3}}} \right\},$$

the vector g has the form

$$g = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} - \int_0^T \begin{pmatrix} e^{\frac{\tau-T}{R_1}} I_1(\tau) \\ e^{\frac{\tau-T}{R_2}} I_2(\tau) \\ e^{\frac{\tau-T}{R_3}} I_2(\tau) \end{pmatrix} d\tau.$$

The orthoprojectors P_Q, P_{Q^*} are given by the relations

$$P_Q = E_3 - Q^+Q = P_{Q^*} = E_3 - QQ^+ = O_3,$$

where O_3 is the zero matrix of dimension 3. The condition of solvability (3.1) is fulfilled automatically, and the linear boundary-value problem (2.4), (2.5) has only one solution of the following form:

$$x(t, c_r) = x(t) = \begin{pmatrix} \dfrac{e^{\frac{T-t}{R_1}}}{1-e^{\frac{T}{R_1}}}\alpha_1 - \dfrac{1}{1-e^{\frac{T}{R_1}}} \displaystyle\int_0^T e^{\frac{\tau-t}{R_1}} I_1(\tau) d\tau + \int_0^t e^{\frac{\tau-t}{R_1}} I_1(\tau) d\tau \\ \dfrac{e^{\frac{T-t}{R_2}}}{1-e^{\frac{T}{R_2}}}\alpha_2 - \dfrac{1}{1-e^{\frac{T}{R_2}}} \displaystyle\int_0^T e^{\frac{\tau-t}{R_2}} I_2(\tau) d\tau + \int_0^t e^{\frac{\tau-t}{R_2}} I_2(\tau) d\tau \\ \dfrac{e^{\frac{T-t}{R_3}}}{1-e^{\frac{T}{R_3}}}\alpha_3 - \dfrac{1}{1-e^{\frac{T}{R_3}}} \displaystyle\int_0^T e^{\frac{\tau-t}{R_3}} I_3(\tau) d\tau + \int_0^t e^{\frac{\tau-t}{R_3}} I_3(\tau) d\tau \end{pmatrix}.$$

Any vector c_r satisfies the equation for generating vectors (3.5) since $P_{Q^*} = O_3$. From (5.10) follows that matrix $B = O_3$. Thus, the boundary-value problem (2.2), (6.1) has a solution which, according to Theorem 5.1, can be found using the iterative process

$$\begin{aligned} x^{k+1}(t, w, \varepsilon) &= y^{k+1}(t, w, \varepsilon) + x(t), \\ y^{k+1}(t, w, \varepsilon) &= \overline{y}^{k+1}(t, w, \varepsilon) \\ &= \varepsilon \left(\hat{I}(t) + G \left[W \left(Z(x(t)) + A_1(t) y^k(t, w, \varepsilon) + \mathcal{R}\left(y^k(t, w, \varepsilon) \right) \right), 0 \right] \right)(t, \varepsilon). \end{aligned}$$

If we rewrite this componentwise, then we obtain the following iterative procedure for finding the solutions of the boundary-value problem (2.2), (6.1):

$$\begin{aligned} y_1^{k+1}(t, w, \varepsilon) = \varepsilon \int_0^t e^{\frac{\tau-t}{R_1}} & \left(\hat{I}_1(\tau) + w_{12} \tanh\left(\frac{a_2 \left(x_2(\tau) + y_2^k(\tau, w, \varepsilon) \right)}{2} \right) \right) d\tau \\ & + \varepsilon w_{13} \int_0^t e^{\frac{\tau-t}{R_1}} \tanh\left(\frac{a_3 \left(x_1(\tau) + y_3^k(\tau, w, \varepsilon) \right)}{2} \right) d\tau \\ & - \varepsilon \int_0^T \frac{e^{\frac{\tau-t}{R_1}}}{1-e^{\frac{T}{R_1}}} \left(\hat{I}_1(\tau) + w_{12} \tanh\left(\frac{a_2 \left(x_2(\tau) + y_2^k(\tau, w, \varepsilon) \right)}{2} \right) \right) d\tau \\ & - \varepsilon w_{13} \int_0^T \frac{e^{\frac{\tau-t}{R_1}}}{1-e^{\frac{T}{R_1}}} \tanh\left(\frac{a_3 \left(x_1(\tau) + y_3^k(\tau, w, \varepsilon) \right)}{2} \right) d\tau, \end{aligned}$$

$$
\begin{aligned}
y_2^{k+1}(t,w,\varepsilon) = {} & \varepsilon \int_0^t e^{\frac{\tau-t}{R_2}} \left(\hat{I}_2(\tau) + w_{12} \tanh \left(\frac{a_1 \left(x_1(\tau) + y_1^k(\tau,w,\varepsilon) \right)}{2} \right) \right) d\tau \\
& + \varepsilon w_{23} \int_0^t e^{\frac{\tau-t}{R_2}} \tanh \left(\frac{a_3 \left(x_3(\tau) + y_3^k(\tau,w,\varepsilon) \right)}{2} \right) d\tau \\
& - \varepsilon \int_0^T \frac{e^{\frac{\tau-t}{R_2}}}{1-e^{\frac{T}{R_2}}} \left(\hat{I}_2(\tau) + w_{12} \tanh \left(\frac{a_1 \left(x_1(\tau) + y_1^k(\tau,w,\varepsilon) \right)}{2} \right) \right) d\tau \\
& - \varepsilon w_{23} \int_0^T \frac{e^{\frac{\tau-t}{R_2}}}{1-e^{\frac{T}{R_2}}} \tanh \left(\frac{a_3 \left(x_3(\tau) + y_3^k(\tau,w,\varepsilon) \right)}{2} \right) d\tau, \\
y_3^{k+1}(t,w,\varepsilon) = {} & \varepsilon \int_0^t e^{\frac{\tau-t}{R_3}} \left(\hat{I}_3(\tau) + w_{13} \tanh \left(\frac{a_1 \left(x_1(\tau) + y_1^k(\tau,w,\varepsilon) \right)}{2} \right) \right) d\tau \\
& + \varepsilon w_{23} \int_0^t e^{\frac{\tau-t}{R_3}} \tanh \left(\frac{a_2 \left(x_2(\tau) + y_2^k(\tau,w,\varepsilon) \right)}{2} \right) d\tau \\
& - \varepsilon \int_0^T \frac{e^{\frac{\tau-t}{R_3}}}{1-e^{\frac{T}{R_3}}} \left(\hat{I}_3(\tau) + w_{13} \tanh \left(\frac{a_1 \left(x_1(\tau) + y_1^k(\tau,w,\varepsilon) \right)}{2} \right) \right) d\tau \\
& - \varepsilon w_{23} \int_0^T \frac{e^{\frac{\tau-t}{R_3}}}{1-e^{\frac{T}{R_3}}} \tanh \left(\frac{a_2 \left(x_2(\tau) + y_2^k(\tau,w,\varepsilon) \right)}{2} \right) d\tau.
\end{aligned}
$$

Example 6.3. Let us consider the Hopfield model for three neurons described by the boundary-value problem (2.2), (2.3) of the form

$$
\begin{aligned}
x_1'(t,w,\varepsilon) &= \varepsilon \left(w_{12} \tanh \left(x_2(t,w,\varepsilon) \right) + w_{13} \tanh \left(x_3(t,w,\varepsilon) \right) \right), \\
x_2'(t,w,\varepsilon) &= \varepsilon \left(w_{12} \tanh \left(x_1(t,w,\varepsilon) \right) + w_{23} \tanh \left(x_3(t,w,\varepsilon) \right) \right), \\
x_3'(t,w,\varepsilon) &= \varepsilon \left(2 + w_{13} \tanh \left(x_1(t,w,\varepsilon) \right) + w_{23} \tanh \left(x_2(t,w,\varepsilon) \right) \right),
\end{aligned} \tag{6.2}
$$

$$
\begin{aligned}
x_1(1,w,\varepsilon) - x_1(0,w,\varepsilon) &= 0, \\
\int_0^1 x_1(t,w,\varepsilon) dt &= 1
\end{aligned} \tag{6.3}
$$

and the generating problem for it

$$
x_j'(t) = 0, \quad j = 1,2,3, \tag{6.4}
$$

$$
\begin{aligned}
x_1(1) - x_1(0) &= 0, \\
\int_0^1 x_1(t) dt &= 1.
\end{aligned} \tag{6.5}
$$

That is, in our case $R_1 = R_2 = R_3 = \infty$, $a_1 = a_2 = a_3 = 2$, $I_1(t) = I_2(t) = I_3(t) = 0$, $\hat{I}_1(t) = \hat{I}_2(t) = 0$, $\hat{I}_3(t) = 2$, $l = \operatorname{col}\left(l_1, \quad l_2\right)$, $\alpha = \operatorname{col}\left(0, \quad 1\right)$.

Let us investigate the problem of finding the extremum (minimum) of the loss function

$$
\begin{aligned}
L(x(t,w,\varepsilon),w) = {} & 2 \left(2w_{12}^2 + w_{13}^2 + 2w_{23}^2 - 2w_{12}w_{13} + w_{13}w_{23} \right) x_1(0,w,\varepsilon) \\
& - 2 \coth(1)(6w_{12} - 4w_{13} - w_{23}) x_1(1,w,\varepsilon) \to \min_{w \in \mathbb{R}^3},
\end{aligned} \tag{6.6}
$$

on the solutions of the boundary-value problem (6.2), (6.3), which at $\varepsilon = 0$ turn into solutions of the generating problem (6.4), (6.5), by the vector of parameters $w = \operatorname{col}\left(w_{12}, \quad w_{13}, \quad w_{23}\right) \in \mathbb{R}^3$.

Using the well-known formulas (see [7, p. 48], [28, p. 501]), we get $r = 2$, $d = 1$,

$$Q = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad Q^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad P_{Q_2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_{Q_1^*} = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{6.7}$$

and the vector g has the following form

$$g = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The solvability condition (3.1), in our case, is satisfied and due to Theorem 3.1, the solution of the generating problem (6.4), (6.5) takes the form

$$x(t, c_r) = \begin{pmatrix} 1 \\ c_2 \\ c_3 \end{pmatrix}. \tag{6.8}$$

The necessary condition for the existence of a solution $x(t, w, \varepsilon)$ of the problem (6.2), (6.3), which by $\varepsilon = 0$ turns into one of the solutions $x(t, c_r)$ (6.8) of the generating problem (6.4), (6.5), in our case has the following representation:

$$F(c_r) = \int_0^1 (w_{12} \tanh(c_2) + w_{13} \tanh(c_3))\, dt = 0 \tag{6.9}$$

or

$$c_2 = -\tanh^{-1}\left(\frac{w_{13}}{w_{12}} \tanh(c_3)\right).$$

The values of the parameters $c_2 = c_3 = 0$, which are the solution of the system of equations (6.9), determine the generating solution

$$x\left(t, c_r^0\right) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{6.10}$$

to which corresponds the solution $x(t, w, \varepsilon)$ of the problem (6.2), (6.3).

Let us return to the problem of finding the minimum of functional (6.6). When ε tends to 0, taking into consideration $x(t, w, \varepsilon) \to x\left(t, c_r^0\right)$, where $x\left(t, c_r^0\right)$ taking into consideration (6.10), we obtain the quadratic functional for the vector of parameters w

$$\begin{aligned} \Phi(w) = 4w_{12}^2 + 2w_{13}^2 + 4w_{23}^2 - 4w_{12}w_{13} + 2w_{13}w_{23} \\ - \coth(1)(12w_{12} - 8w_{13} - 2w_{23}) \to \min_{w \in \mathbb{R}^3}. \end{aligned} \tag{6.11}$$

The problem of finding the minimum of the quadratic functional (6.11) is equivalent to the solution of the following equation

$$Sw = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 2 & 1 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} w_{12} \\ w_{13} \\ w_{23} \end{pmatrix} = \coth(1) \begin{pmatrix} 6 \\ -4 \\ -1 \end{pmatrix} = f. \tag{6.12}$$

To find the solution of the equation (6.12), or, which is the same, find the minimum of the quadratic functional (6.11), we use the accelerated method of conjugate gradients (4.4)–(4.6) and compare it with the method of steepest descent and the conjugate gradient method. Now let us consider the case where

$$w^0 = \frac{3}{2}\coth(1)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \sigma_k = a_k\varphi, \quad \varphi = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

If the system φ_i, $i = \overline{1,n}$, $n \leq 3$ consists of more than one linearly independent element, then the rate of convergence of the accelerated method of conjugate gradients increases. Without cluttering the example above with calculations that can be made following the computational scheme from [31], we present successive approximations to the minimum of functional (6.11) obtained by the accelerated method of conjugate gradients

$$w^1 = \frac{\coth(1)}{14}\begin{pmatrix} 19 \\ -4 \\ -4 \end{pmatrix}, \quad w^2 = \coth(1)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix},$$

and the conjugate gradient method

$$w^1 = \frac{\coth(1)}{4}\begin{pmatrix} 6 \\ -1 \\ -1 \end{pmatrix}, \quad w^2 = \frac{\coth(1)}{10}\begin{pmatrix} 11 \\ -6 \\ -2 \end{pmatrix}, \quad w^3 = \coth(1)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Therefore, as can be proved by substituting the obtained approximations into the equation (6.12), the minimum of functional (6.11) is achieved by the accelerated method of conjugate gradients in the second approximation w^2, and by the conjugate gradient method in the third approximation w^3, and is equal to

$$w^*_{12} = \coth(1), \quad w^*_{13} = -\coth(1), \quad w^*_{23} = 0. \tag{6.13}$$

Note that the rate of convergence of the method of steepest descent, in our case, is considerably slower and even the approximation w^5 is far away from the value (6.13)

$$w^1 = \frac{\coth(1)}{4}\begin{pmatrix} 6 \\ -1 \\ -1 \end{pmatrix}, \quad w^3 = \frac{\coth(1)}{32}\begin{pmatrix} 42 \\ -17 \\ -5 \end{pmatrix}, \quad w^5 \approx \coth(1)\begin{pmatrix} 1,206881 \\ -0,683004 \\ -0,103441 \end{pmatrix}.$$

Let us fix the value of the vector of parameters $w = w^*$ (6.13). Let us now find the sufficient condition for the existence of solutions of the problem (6.2), (6.3). For this we make the substitution

$$x(t, w^*, \varepsilon) = x\left(t, c_r^0\right) + y(t, w^*, \varepsilon),$$

where $x\left(t, c_r^0\right)$ has the form (6.10). After such a substitution we obtain the following boundary-value problem

$$\begin{aligned}
y_1'(t, w^*, \varepsilon) &= \varepsilon\coth(1)\left(\tanh\left(y_2(t, w^*, \varepsilon)\right) - \tanh\left(y_3(t, w^*, \varepsilon)\right)\right), \\
y_2'(t, w^*, \varepsilon) &= \varepsilon\coth(1)\tanh\left(1 + y_1(t, w^*, \varepsilon)\right), \\
y_3'(t, w^*, \varepsilon) &= \varepsilon\left(2 - \coth(1)\tanh\left(1 + y_1(t, w^*, \varepsilon)\right)\right), \\
y_1(1, w^*, \varepsilon) - y_1(0, w^*, \varepsilon) &= 0, \\
\int_0^1 y_1(t, w^*, \varepsilon)dt &= 0.
\end{aligned}$$

For the vector-function $Z\left(x\left(t, c_r^0\right) + y(t, w^*, \varepsilon)\right)$, in the neighbourhood of the generating solution $x\left(t, c_r^0\right)$ (6.10), the following representation holds

$$Z(x(t, c_r^0) + y(t, w^*, \varepsilon)) = Z(x(t, c_r^0)) + A_1(t) y(t, w^*, \varepsilon) + \mathcal{R}(y(t, w^*, \varepsilon)),$$

where

$$\begin{aligned}
Z\left(x\left(t, c_r^0\right)\right) &= \operatorname{col}\left(\tanh(1), \quad 0, \quad 0\right), \\
A_1(t) &= \operatorname{diag}\left\{\cosh^{-2}(1), \quad 1, \quad 1\right\}, \\
\mathcal{R}(y(t, w^*, \varepsilon)) &= \begin{pmatrix} \mathcal{R}_1(y(t, w^*, \varepsilon)) \\ \mathcal{R}_2(y(t, w^*, \varepsilon)) \\ \mathcal{R}_3(y(t, w^*, \varepsilon)) \end{pmatrix} \\
&= \begin{pmatrix} \tanh\left(1 + y_1(t, w^*, \varepsilon)\right) - \tanh(1) - \dfrac{y_1(t, w^*, \varepsilon)}{\cosh^2(1)} \\ \tanh\left(y_2(t, w^*, \varepsilon)\right) - y_2(t, w^*, \varepsilon) \\ \tanh\left(y_3(t, w^*, \varepsilon)\right) - y_3(t, w^*, \varepsilon) \end{pmatrix}.
\end{aligned}$$

The function $H\left(t, y(t, w^*, \varepsilon), \overline{y}(t, w^*, \varepsilon)\right)$ has the form:

$$\begin{aligned}
H\left(t, y(t, w^*, \varepsilon), \overline{y}(t, w^*, \varepsilon)\right) &= \begin{pmatrix} H_1\left(t, y(t, w^*, \varepsilon), \overline{y}(t, w^*, \varepsilon)\right) \\ H_2\left(t, y(t, w^*, \varepsilon), \overline{y}(t, w^*, \varepsilon)\right) \\ H_3\left(t, y(t, w^*, \varepsilon), \overline{y}(t, w^*, \varepsilon)\right) \end{pmatrix} \\
&= \begin{pmatrix} \dfrac{\overline{y}_1(t, w^*, \varepsilon)}{\cosh^2(1)} + \mathcal{R}_1(y(t, w^*, \varepsilon)) \\ \overline{y}_2(t, w^*, \varepsilon) + \mathcal{R}_2(y(t, w^*, \varepsilon)) \\ \overline{y}_3(t, w^*, \varepsilon) + \mathcal{R}_3(y(t, w^*, \varepsilon)) \end{pmatrix}.
\end{aligned}$$

Matrices B, B^+, $P_{B_{r_1}}$, $P_{B^*_{d_1}}$ in our case takes the following view

$$B = \coth(1)\begin{pmatrix} 1 & -1 \end{pmatrix}, \quad B^+ = \frac{\tanh(1)}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad P_{B_1} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad P_{B^*} = 0. \tag{6.14}$$

Using (6.7), (6.14), we verify the validity of condition (5.12). Following Theorem 5.1, we can find an approximate solution of (6.2), (6.3), which under $\varepsilon = 0$ turns into $x\left(t, c^0\right)$ (6.10) of the generating problem (6.4), (6.5), following this algorithm

$$\begin{aligned}
c_3^k(w^*, \varepsilon) = -c_2^k(w^*, \varepsilon) = \frac{1}{2}\int_0^1 & H_2\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right) d\tau \\
-\frac{1}{2}\int_0^1 & H_3\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right) d\tau,
\end{aligned} \tag{6.15}$$

$$\begin{aligned}
\overline{y}_1^{k+1}(t, w^*, \varepsilon) = \varepsilon \coth(1) &\int_0^t \left(2c_2^k(w^*, \varepsilon) + H_2\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right)\right) d\tau \\
&- \varepsilon \coth(1) \int_0^t H_3\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right) d\tau \\
&- \varepsilon \coth(1) \int_0^1 \int_0^t \left(2c_2^k(w^*, \varepsilon) + H_2\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right)\right) d\tau dt \\
&+ \varepsilon \coth(1) \int_0^1 \int_0^t H_3\left(\tau, y^k(\tau, w^*, \varepsilon), \overline{y}^k(\tau, w^*, \varepsilon)\right) d\tau dt,
\end{aligned} \tag{6.16}$$

$$\overline{y}_2^{k+1}(t,w^*,\varepsilon) = \varepsilon \coth(1) \int_0^t \tanh\left(1 + y_1^k(\tau,w^*,\varepsilon)\right) d\tau \\ + \frac{2\varepsilon}{\sinh(2)} \int_0^t \left(\overline{y}_1^k(\tau,w^*,\varepsilon) - y_1^k(\tau,w^*,\varepsilon)\right) d\tau, \tag{6.17}$$

$$\overline{y}_3^{k+1}(t,w^*,\varepsilon) = 2\varepsilon t - \overline{y}_2^{k+1}(t,w^*,\varepsilon), \tag{6.18}$$

$$y_1^{k+1}(t,w^*,\varepsilon) = \overline{y}_1^{k+1}(t,w^*,\varepsilon), \quad y_2^{k+1}(t,w^*,\varepsilon) = c_2^k(w^*,\varepsilon) + \overline{y}_2^{k+1}(t,w^*,\varepsilon), \tag{6.19}$$

$$y_3^{k+1}(t,w^*,\varepsilon) = c_3^k(w^*,\varepsilon) + \overline{y}_3^{k+1}(t,w^*,\varepsilon), \tag{6.20}$$

$$x_1^k(t,w^*,\varepsilon) = 1 + y_1^k(t,w^*,\varepsilon), \quad x_2^k(t,w^*,\varepsilon) = y_2^k(t,w^*,\varepsilon), \quad x_3^k(t,w^*,\varepsilon) = y_3^k(t,w^*,\varepsilon), \tag{6.21}$$

$$x_i(t,w^*,\varepsilon) = \lim_{k\to\infty} x_i^k(t,w^*,\varepsilon), \quad y_i^0(t,w^*,\varepsilon) = \overline{y}_i^0(t,w^*,\varepsilon) = 0, \quad i = \overline{1,3}. \tag{6.22}$$

Let us construct the first approximation $x^1(t,w^*,\varepsilon)$. Since $y^0(t,w^*,\varepsilon) = \overline{y}^0(t,w^*,\varepsilon) = 0$, then the constants $c_2^0(w^*,\varepsilon)$, $c_3^0(w^*,\varepsilon)$ defined by formula (6.15) take the form $c_2^0(w^*,\varepsilon) = c_3^0(w^*,\varepsilon) = 0$. From (6.16)–(6.20) we obtain

$$y^1(t,w^*,\varepsilon) = \overline{y}^1(t,w^*,\varepsilon) = \varepsilon \begin{pmatrix} 0 \\ t \\ t \end{pmatrix}$$

and, following (6.21), we get

$$x^1(t,w^*,\varepsilon) = \begin{pmatrix} 1 \\ \varepsilon t \\ \varepsilon t \end{pmatrix}. \tag{6.23}$$

Continuing calculations according to (6.15)–(6.22), we see that

$$c_2^k(w^*,\varepsilon) = c_3^k(w^*,\varepsilon) = 0, \quad \forall k \geq 1$$

and all subsequent approximations $x^k(t,w^*,\varepsilon)$, $k \geq 2$ are equal to the first approximation, that is vector-function (6.23), as can be seen by a simple substitution, is the solution of (6.2), (6.3), which at $\varepsilon = 0$ turns into the generating solution (6.10) for the values of parameters (6.13), minimizing functional (6.11).

Note that one of the important concepts in the study of the problem for finding the extremum of a function on solutions of an equation, including problem (2.2), (2.3), (4.1), is the concept of solution sensitivity with respect to the parameters

$$s(t,\varepsilon) = \frac{\partial x(t,w,\varepsilon)}{\partial w}.$$

In the literature [9, 10, 12, 16, 18, 24, 42] there are two approaches to find $s(t,\varepsilon)$: the direct method, which uses the chain rule to find the complete derivative, and adjoint sensitivity method from Pontryagin papers [39]. The use of the conjugate sensitivity method reduces the computational costs when finding the gradient by parameters when the number of parameters is much greater than the dimension of the set of required functions. When the number of parameters is much less than the number of desired functions, the advantages of this method are lost due to the complexity of solving the auxiliary coupled system. In Example 6.3 we consider a boundary-value problem for systems of three differential equations with three parameters. Therefore, the use of the direct method and adjoint sensitivity method for finding the gradient by parameters w_{12}, w_{13}, w_{23} are equivalent. However, when investigating the problem of optimization of function on solutions of the Hopfield network for n ($n \gg 3$) neurons, in which the number of weights by far exceeds the number of potentials ($M \gg n$), the adjoint sensitivity method has advantages over the direct method. The study of relationships of the direct method, adjoint sensitivity method and the accelerated method of conjugate gradients for solving the presented paper tasks will be devoted to our future research.

7 Conclusions

Necessary and sufficient conditions for the solvability were established, as well as a constructive algorithm for finding solutions to a boundary-value problem for a system of weakly non-linear differential equations describing Hopfield network for n neurons is presented. The problem of minimizing a functional on the solutions of the given problem was investigated and the application of the accelerated method of conjugate gradients to its solutions was explored. The results are demonstrated by examples of problems for the case of three neurons.

Acknowledgements

The work was supported by the National Research Foundation of Ukraine (Project number 2020.20/0089).

References

[1] E. Agliari, A. Barra, M. Notarnicola, The relativistic Hopfield network: rigorous results, *J. Math. Phys.* **60**(2019), No. 3, Paper No. 033302, 11 pp. https://doi.org/10.1063/1.5077060; MR3927839; Zbl 1421.82027

[2] M. Akhmet, E. M. Alejaily, Domain-structured chaos in a Hopfield neural network, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **29**(2019), No. 14, Paper No. 1950205, 7 pp. https://doi.org/10.1142/S0218127419502055; MR4045912; Zbl 1429.37021

[3] A. Boichuk, J. Diblík, D. Khusainov, M. Růžičková, Boundary-value problems for weakly nonlinear delay differential systems, *Abstr. Appl. Anal.* **2011**, Paper No. 631412, 19 pp. https://doi.org/10.1155/2011/631412; MR2802836; Zbl 1222.34075

[4] A. A. Boichuk, I. A. Korostil, M. Fečkan, Bifurcation conditions for a solution of an abstract wave equation, *Differ. Equ.* **43**(2007), No. 4, 495–502. https://doi.org/10.1134/S0012266107040076; MR2358690; Zbl 1146.35010

[5] A. A. Boichuk, M. Medved, V. P. Zhuravliov, Fredholm boundary-value problems for linear delay systems defined by pairwise permutable matrices, *Electron. J. Qual. Theory Differ. Equ.* **2015**, No. 23, 1–9. https://doi.org/10.14232/ejqtde.2015.1.23; MR3346934; Zbl 1349.34254

[6] A. A. Boichuk, A. A. Pokutnyi, Perturbation theory of operator equations in the Fréchet and Hilbert spaces, *Ukrainian Math. J.* **67**(2016), No. 9, 1327–1335. https://doi.org/10.1007/s11253-016-1156-y; MR3473723; Zbl 06817053

[7] A. A. Boichuk, A. M. Samoilenko, *Generalized inverse operators and Fredholm boundary-value problems*, 2nd ed., Inverse and Ill-Posed Problems Series, Vol. 59, De Gruyter, Berlin, 2016. https://doi.org/10.1515/9783110378443; MR3585692; Zbl 1394.47001

[8] A. A. Boichuk, V. F. Zhuravlev, Solvability criterion for integro-differential equations with degenerate kernel in Banach spaces, *Nonlinear Dyn. Syst. Theory* **18**(2018), No. 4, 331–341. MR3891950; Zbl 1411.34105

[9] J. Calver, W. Enright, Numerical methods for computing sensitivities for ODEs and DDEs, *Numer. Algorithms* **74**(2017), No. 4, 1101–1117. https://doi.org/10.1007/s11075-016-0188-6; MRMR3626330; Zbl 1362.65070

[10] Y. Cao, S. Li, L. Petzold, R. Serban, Adjoint sensitivity analysis for differential-algebraic equations: the adjoint DAE system and its numerical solution, *SIAM J. Sci. Comput.* **24**(2003), No. 3, 1076–1089. https://doi.org/10.1137/S1064827501380630; MR1950525; Zbl 1034.65066

[11] Z. Che, S. Purushotham, K. Cho, D. Sontag, Y. Liu, Recurrent neural networks for multivariate time series with missing values, *Sci. Rep.* **8**(2018), Paper No. 6085. https://doi.org/10.1038/s41598-018-24271-9

[12] R. T. Q. Chen, Yu. Rubanova, J. Bettencourt, D. Duvenaud, Neural ordinary differential equations, in: *NeurIPS 2018 (Proc. of the International Conference on Neural Information Processing Systems, 2018)*, Curran Associates Inc., New York, 2018, pp. 6571–6583.

[13] E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, J. Sun, Doctor AI: predicting clinical events via recurrent neural networks, in: *Machine Learning for Healthcare Conference (Proceedings of the 1st Machine Learning for Healthcare Conference, Vol. 56 of Proceedings of Machine Learning Research)*, PMLR, 2016, pp. 301–318.

[14] M.-F. Danca, N. Kuznetsov, Hidden chaotic sets in a Hopfield neural system, *Chaos Solitons Fractals* **103**(2017), 144–150. https://doi.org/10.1016/j.chaos.2017.06.002; MR3689806; Zbl 1380.65423

[15] E. E. Dudnikov, M. V. Rybashov, On the absolute stability of a class of neural networks with feedback, *Automat. Remote Control* **60**(1999), No. 12, part 1, 1700–1706. MR1813073; Zbl 1088.34524

[16] R. M. Errico, What is an adjoint model?, *Bull. Amer. Meteorol. Soc.* **78**(1997), No. 11, 2577–2591. https://doi.org/10.1175/1520-0477(1997)078<2577:WIAAM>2.0.CO;2

[17] R. Fuentes-García, R. H. Mena, S. G. Walker, Modal posterior clustering motivated by Hopfield's network, *Comput. Statist. Data Anal.* **137**(2019), 92–100. https://doi.org/10.1016/j.csda.2019.02.008; MR3921062; Zbl 07058808

[18] M. B. Giles, N. A. Pierce, An introduction to the adjoint approach to design, *Flow Turbul. Combust.* **65**(2000), No. 3–4, 393–415. https://doi.org/10.1023/A:1011430410075; Zbl 0996.76023

[19] S. D. Glyzin, A. Yu. Kolesov, N. Kh. Rozov, Relaxation self-oscillations in Hopfield networks with delay, *Izv. Math.* **77**(2013), No. 2, 271–312. https://doi.org/10.1070/IM2013v077n02ABEH002636; MR3097567; Zbl 1271.34074

[20] S. D. Glyzin, A. Yu. Kolesov, N. Kh. Rozov, Self-excited relaxation oscillations in networks of impulse neurons, *Russian Math. Surveys* **70**(2015), No. 3, 383–452. https://doi.org/10.1070/RM2015v070n03ABEH004951; MR3400564; Zbl 1361.34093

[21] E. A. Grebenikov, Yu. A. Ryabov, *Constructive methods in the analysis of nonlinear systems*, Mir Publishers, Moscow, 1983. MR0733787; Zbl 0536.65058

[22] S. Haykin, *Neural networks and learning machines*, 3rd ed., Pearson Education, New York, 2009.

[23] C. J. Hillar, N. M. Tran, Robust exponential memory in Hopfield networks, *J. Math. Neurosci.* **8**(2018), Paper No. 1, 20 pp. https://doi.org/10.1186/s13408-017-0056-2; MR3747011; Zbl 1395.92010

[24] A. C. Hindmarsh, R. Serban, *User documentation for CVODES, an ODE solver with sensitivity analysis capabilities*, Lawrence Livermore National Laboratory, Livermore, 2002.

[25] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Nat. Acad. Sci. U.S.A.* **79**(1982), No. 8, 2554–2558. https://doi.org/10.1073/pnas.79.8.2554; MR0652033; Zbl 1369.92007

[26] J. J. Hopfield, D. W. Tank, Neural computation of decisions in optimization problems, *Biol. Cybernet.* **52**(1985), No. 3, 141–152. MR0824597; Zbl 0572.68041

[27] F. C. Hoppensteadt, E. M. Izhikevich, *Weakly connected neural networks*, Applied Mathematical Sciences, Vol. 126, Springer, New York, 1997. https://doi.org/10.1007/978-1-4612-1828-9; MR1458890; Zbl 0887.92003

[28] R. A. Horn, C. R. Johnson, *Matrix analysis*, 2nd ed., Cambridge University Press, Cambridge, 2013. MR2978290; Zbl 1267.15001

[29] E. M. Izhikevich, Neural excitability, spiking and bursting, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **10**(2000), No. 6, 1171–1266. https://doi.org/10.1142/S0218127400000840; MR1779667; Zbl 1090.92505

[30] L. V. Kantorovič, Functional analysis and applied mathematics (in Russian), *Uspehi Matem. Nauk* **3**(1948), No. 6(28), 89–185. MR0027947; Zbl 0034.21203

[31] A. Yu. Luchka, O. E. Noshchenko, I. V. Sergiyenko, N. I. Tukalevskaya, Accelerated method of conjugate gradients, *U.S.S.R. Comput. Math. and Math. Phys.* **27**(1987), No. 3, 6–12. https://doi.org/10.1016/0041-5553(87)90073-5; Zbl 0655.65075

[32] A. Yu. Luchka, O. E. Noshchenko, N. I. Tukalevskaya, The variational gradient method, *U.S.S.R. Comput. Math. Math. Phys.* **24**(1984), No. 4, 1–6. https://doi.org/10.1016/0041-5553(84)90222-2; Zbl 0601.47012

[33] I. G. Malkin, *Some problems of the theory of nonlinear oscillations* (in Russian), Gos. Izdt. Tekh.-Teor. Lit., Moscow, 1956. MR0081402; Zbl 0070.08703

[34] M.-H. Matcovschi, O. Pastravanu, Flow-invariance and stability analysis for a class of nonlinear systems with slope conditions, *Eur. J. Control* **10**(2004), No. 4, 352–364. https://doi.org/10.3166/ejc.10.352-364; MR2120426; Zbl 1293.93391

[35] M. Mézard, Mean-field message-passing equations in the Hopfield model and its generalizations, *Phys. Rev. E* **95**(2017), No. 2, Paper No. 022117, 15 pp. https://doi.org/10.1103/PhysRevE.95.022117; MR3778464

[36] B. A. Pearlmutter, Learning state space trajectories in recurrent neural networks, *Neural Comput.* **1**(1989), No. 2, 263–269. https://doi.org/10.1162/neco.1989.1.2.263

[37] R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.* **51**(1955), 406–413. MR0069793 ; Zbl 0065.24603

[38] M. Pinto, D. Sepulveda, R. Torres, Exponential periodic attractor of impulsive Hopfield-type neural network system with piecewise constant argument, *Electron. J. Qual. Theory Differ. Equ.* **2018**, No. 34, 1–28. https://doi.org/10.14232/ejqtde.2018.1.34; MR3811497; Zbl 1413.34224

[39] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The mathematical theory of optimal processes*, A Pergamon Press Book, The Macmillan Co., New York, 1964. MR0186436; Zbl 0117.31702

[40] P. Rotondo, M. Marcuzzi, J. P. Garrahan, I. Lesanovsky, M. Müller, Open quantum generalisation of Hopfield neural networks, *J. Phys. A* **51**(2018), No. 11, Paper No. 115301, 11 pp. https://doi.org/10.1088/1751-8121/aaabcb; MR3766232; Zbl 1394.82014

[41] A. M. Samoilenko, A. A. Boichuk, S. M. Chuiko, Hybrid difference-differential boundary-value problem, *Miskolc Math. Notes* **18**(2017), No. 2, 1015–1031. https://doi.org/10.18514/MMN.2017.2280; MR3768335; Zbl 1399.34051

[42] B. Sengupta, K. J. Friston, W. D. Penny, Efficient gradient computation for dynamical models, *NeuroImage* **98**(2014), 521–527. https://doi.org/10.1016/j.neuroimage.2014.04.040

[43] I. N. Da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, S. F. R. Alves, *Artificial neural networks*, Springer, Switzerland, 2017. https://doi.org/10.1007/978-3-319-43162-8; MR3526246

[44] P. Stapor, F. Fröhlich, J. Hasenauer, Optimization and profile calculation of ODE models using second order adjoint sensitivity analysis, *Bioinformatics* **34**(2018), No. 13, 151–159. https://doi.org/10.1093/bioinformatics/bty230

[45] A.-M. Stoica, I. Yaesh, Hopfield networks with multiplicative noise in an anisotropic norm setup, *Ann. Acad. Rom. Sci. Ser. Math. Appl.* **11**(2019), No. 1, 79–97. MR3973488; Zbl 1438.93218

[46] C. A. Tavares, T. M. R. Santos, N. H. T. Lemes, J. P. C. Dos Santos, J. C. Ferreira, J. P. Braga, Solving ill-posed problems faster using fractional-order Hopfield neural network, *J. Comput. Appl. Math.* **381**(2021), Paper No. 112984, 14 pp. https://doi.org/10.1016/j.cam.2020.112984; MR4105673; Zbl 07241269

[47] Z. Zeng, W. X. Zheng, Multistability of two kinds of recurrent neural networks with activation functions symmetrical about the origin on the phase plane, *IEEE Trans. Neural Netw. Learn. Syst.* **24**(2013), No. 11, 1749–1762. https://doi.org/10.1109/TNNLS.2013.2262638