Available online at www.sciencedirect.com
ScienceDirect
Chaos, Solitons and Fractals 39 (2009) 717–730
CHAOS SOLITONS & FRACTALS
www.elsevier.com/locate/chaos

Transiently chaotic neural networks with piecewise linear output functions

Shyan-Shiou Chen [a], Chih-Wen Shih [b,*,1]

[a] Department of Mathematics, National Taiwan Normal University, Taipei, Taiwan, ROC
[b] Department of Applied Mathematics, National Chiao Tung University, 1001 Ta-Hsueh Road, Hsinchu, Taiwan, ROC

Accepted 2 January 2007

Communicated by Professor Ji-Huan He

Abstract

Admitting both transient chaotic phase and convergent phase, the transiently chaotic neural network (TCNN) provides superior performance than the classical networks in solving combinatorial optimization problems. We derive concrete parameter conditions for these two essential dynamic phases of the TCNN with piecewise linear output function. The confirmation for chaotic dynamics of the system results from a successful application of the Marotto theorem which was recently clarified. Numerical simulation on applying the TCNN with piecewise linear output function is carried out to find the optimal solution of a travelling salesman problem. It is demonstrated that the performance is even better than the previous TCNN model with logistic output function.

1. Introduction

Solving combinatorial optimization problems such as the travelling salesman problem (TSP) has been one of the main motifs for the development of neural networks [1]. While iterations in the classical Hopfield model may be trapped at local minimum and fail to reach global minimum of the objective functions, the transiently chaotic neural network (TCNN) was developed to provide a global searching ability and thus achieved better performance in solving combinatorial optimization problems [2–8]. The TCNN model employed in [2–4] is

$$x_i(t+1) = \mu x_i(t) - \omega_{ii}(t)[y_i(t) - a_{0i}] + \alpha\left[\sum_{j=1, j\neq i}^{n} \omega_{ij} y_j + v_i\right], \tag{1}$$

$$|\omega_{ii}(t+1)| = |(1-\beta)\omega_{ii}(t)| \tag{2}$$

for $i = 1, \ldots, n, t \in \mathbb{N}$ (positive integers), where x_i is the internal state of neuron i; y_i is the output of neuron i, which corresponds to x_i through an output function; the output function adopted therein is the logistic function given by

* Corresponding author. Tel.: +886 3 5131209; fax: +886 3 5724679.
E-mail address: cwshih@math.nctu.edu.tw (C.-W. Shih).
[1] Partially supported by The National Science Council, and the MOEATU program, of ROC on Taiwan.

0960-0779/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chaos.2007.01.103

$y_i(t) = 1/[1 + \exp(-x_i(t)/\varepsilon)]$; μ is the damping factor of nerve membrane; ω_{ii} is the self-feedback connection weight; a_{0i} is the self-recurrent bias of neuron i; ω_{ij} is the connection weight from neuron j to neuron i; v_i is the input bias of neuron i; β with $0 < \beta < 1$, is the damping factor for ω_{ii}. Eq. (2) represents an exponential cooling schedule in the annealing procedure.

There are chaotic phase and convergent phase for the TCNN with different parameters. The purpose of this presentation is to provide in-depth understanding on the dynamics of the TCNN. In particular, we present detailed analysis and concrete parameter conditions for the chaotic and convergent phases of the TCNN with piecewise linear output functions:

$$y_i(t) = g_\varepsilon(x_i(t)) := \left[2 + \left|\frac{x_i(t)}{\varepsilon} + 1\right| - \left|\frac{x_i(t)}{\varepsilon} - 1\right|\right] \Big/ 4, \quad \varepsilon > 0. \tag{3}$$

Notably, piecewise linear output functions have been employed in various neural networks, cf. [9–13]. There are several advantages for the use of piecewise linear output function in the TCNN. In this presentation, it will be demonstrated that the mathematical description on the chaotic phase for the TCNN with output function (3) is more succinct than with the logistic one, cf. [14]. Subsequently, the parameter conditions derived are more concise. In addition, the model considered herein is a successful application of the Marotto's theorem [15,16] in concluding chaotic dynamics for multi-dimensional and high dimensional maps. Moreover, there also exists a Lyapunov function which is Lipschitz though not differentiable, for such a TCNN model. Theoretical confirmations for convergence of iterations or evolutions of the network system to a steady state can then be established. Moreover, the parameters can be controlled so that the output component of the network tends to either exact one or zero. This consequence, for example, provides a lucid signal to further tune the parameters if the outcome is not feasible in the computation task of solving combinatorial optimization problems.

As an extension of the Li and York theorem on the one-dimensional maps, Marotto [15] established a significant theorem in confirming the chaotic dynamics for multi-dimensional systems. With presence of the so-called "snap-back repeller", the phase space possesses a topological structure which includes infinitely many periodic points and a scrambled set. Very erratic behaviors of the system then occur, including the lack of global stability for solutions, and the existence of an uncountable collection of orbits which do not eventually approach any periodic points. The definition of snap-back repellers is further clarified for rigorousness recently in [16]. It will be demonstrated that the formulations and analysis in the considered model, i.e., the TCNN with piecewise linear output function, fit into the new definition of snap-back repellers in [16] pertinently.

We discuss the chaotic dynamics in Section 2, and the convergent dynamics in Section 3, for the TCNN. We will then make use of the TCNN properties to select suitable parameters for the computation of solving a travelling salesman problem. The task of solving the TSP via neural networks is to search for the minimizer of an objective function through iterations or evolutions of the network system. We arrange the setting for the application of the TCNN to the TSP in Section 4. Some numerical simulations are performed in Section 5.

2. Chaotic dynamics for the TCNN with piecewise linear output function

In this section, we plan to investigate the transiently chaotic behaviors for the TCNN with piecewise linear output functions. We consider the following system:

$$x_i(t+1) = \mu x_i(t) + \omega[y_i(t) - a_{0i}] + \sum_{j=1}^{n} \omega_{ij} y_j(t) + v_i, \tag{4}$$

where $t \in \mathbb{N}$, and $y_i(t) = g_\varepsilon(x_i(t)),\ i = 1, \ldots, n$, is the piecewise linear output function defined in (3), and illustrated in Fig. 1. Eqs. (1) and (2) are in fact a skew system of Eq. (4) over Eq. (2), cf. [17]. We denote by $\mathbf{x}(t+1) = \mathbf{F}(\mathbf{x}(t))$ the iterations generated by Eq. (4) with $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{F} = (F_1, \ldots, F_n)$. We shall analyze the chaotic dynamics of the network from a geometrical observation on Eq. (4). These chaotic dynamics correspond to the transiently chaotic behaviors in the annealing process of the TCNN.

2.1. Single neuron maps

Let us consider the following family of single neuron maps. For a fixed $h > 0$ and $a_0 \in \mathbb{R}$, let $f_\varsigma : \mathbb{R} \to \mathbb{R}$ be a function defined by

$$f_\varsigma(\xi) = \mu\xi + \omega[g_\varepsilon(\xi) - a_0] + \varsigma, \tag{5}$$

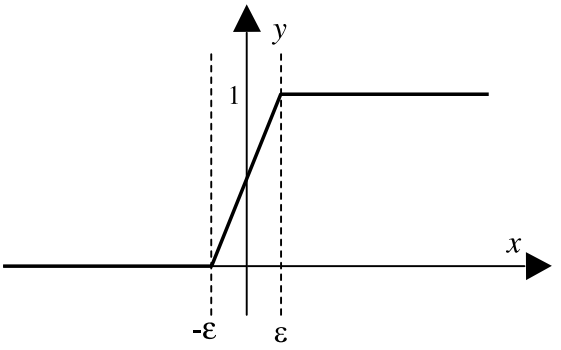

Fig. 1. The piecewise linear output function g_ε with $\varepsilon > 0$.

where $-h \leqslant \varsigma \leqslant h$. We set $h_i = \sum_{j=1}^{n} |\omega_{ij}| + |v_i|$, for each $i = 1, \ldots, n$, and define the *upper map* $\hat{f}_i$ and the *lower map* $\check{f}_i$ as

$$\hat{f}_i(\xi) = \mu\xi + \omega[g_\varepsilon(\xi) - a_{0i}] + h_i, \quad \check{f}_i(\xi) = \mu\xi + \omega[g_\varepsilon(\xi) - a_{0i}] - h_i. \tag{6}$$

Then for $i = 1, \ldots, n$,

$$\check{f}_i(x_i) \leqslant F_i(\mathbf{x}) \leqslant \hat{f}_i(x_i), \quad \text{for all } \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n. \tag{7}$$

In the following discussions, we shall propose a sequence of parameter conditions which guarantee the existence of repelling fixed points and chaotic behaviors for each f_ς in Eq. (5). These conditions are based on geometrical observations on the configurations of the maps f_ς. In our considerations, there are two sets of parameter conditions labelled as (P_1-) and (P_2-), which correspond to two types of configurations for single neuron maps f_ς, as depicted in Figs. 2a and b and 3, respectively.

For a fixed $\varepsilon > 0$, we partition the real line into the left (ℓ), middle (m), right (r) parts; namely,

$$\Omega_\ell := (-\infty, -\varepsilon), \quad \Omega_{\mathrm{m}} := [-\varepsilon, \varepsilon], \quad \Omega_{\mathrm{r}} := (\varepsilon, \infty). \tag{8}$$

The following parameter conditions are concerned with the existence of fixed point for every member in the family of single neuron maps (5).

(P_1-i) $\mu < \min\{-1 - [\omega a_0 + h]/\varepsilon, -1 - [\omega(1 - a_0) + h]/\varepsilon\}$.
(P_2-i)
(a) $\omega > 0$, $0 < \frac{1-\mu}{\omega} < \frac{1}{2\varepsilon}$, $-\frac{1-\mu}{\omega}\varepsilon - \frac{h}{\omega} + a_0 > 0$, $\frac{1-\mu}{\omega}\varepsilon + \frac{h}{\omega} + a_0 < 1$,
(b) $\omega < 0$, $\frac{1-\mu}{\omega} < \frac{1}{2\varepsilon}$, $-\frac{1-\mu}{\omega}\varepsilon + \frac{h}{\omega} + a_0 > 0$, $\frac{1-\mu}{\omega}\varepsilon - \frac{h}{\omega} + a_0 < 1$.

We remark that Fig. 2a and b correspond to the configurations of f_ς satisfying (P_1-i), while Fig. 3 corresponds to the configuration of f_ς satisfying (P_2-i). The following proposition relies upon observations on the graphs of f_ς and g_ε. Its proof is sketched in Appendix A.

Proposition 1. *Consider the single neuron maps f_ς with $-h \leqslant \varsigma \leqslant h$ defined in Eq. (5).* (i) *If the parameters $(\mu, \omega, \varepsilon, a_0, h)$ satisfy (P_1-i), then there exists a fixed point $\bar{x}_\varsigma^{\mathrm{m}} \in \Omega_{\mathrm{m}}$ for f_ς.* (ii) *If the parameters $(\mu, \omega, \varepsilon, a_0, h)$ satisfy (P_2-i) (a) or (P_2-i) (b), then there exist three fixed points $\bar{x}_\varsigma^{\mathrm{r}} \in \Omega_{\mathrm{r}}$, $\bar{x}_\varsigma^{\mathrm{m}} \in \Omega_{\mathrm{m}}$ and $\bar{x}_\varsigma^{\ell} \in \Omega_\ell$ for f_ς.*

Let us recall the notion of snap-back repeller [15,16]. Consider a map $\mathbf{x} \mapsto \mathbf{F}(\mathbf{x})$ where $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{F}$ is C^1 or piecewise C^1. Suppose $\bar{\mathbf{x}}$ is a fixed point of $\mathbf{F}$ with all eigenvalues of $D\mathbf{F}(\bar{\mathbf{x}})$ exceeding 1 in magnitude, and there exists a point $\mathbf{x}_0 \neq \bar{\mathbf{x}}$ in a repelling neighborhood of $\bar{\mathbf{x}}$, such that $\mathbf{F}^m(\mathbf{x}_0) = \bar{\mathbf{x}}$ and $\det(D\mathbf{F}^m(\mathbf{x}_0)) \neq 0$, for some positive integer m. Then $\bar{\mathbf{x}}$ is called a *snap-back repeller* of $\mathbf{F}$. A scenario for such a snap-back repeller is depicted in Fig. 4. The chaotic dynamics induced by the presence of a snap-back repeller have been established in Marotto's theorem quoted in Appendix B.

The following parameter conditions are concerned with the existence of snap-back repellers for every member in the family of single neuron maps (5).

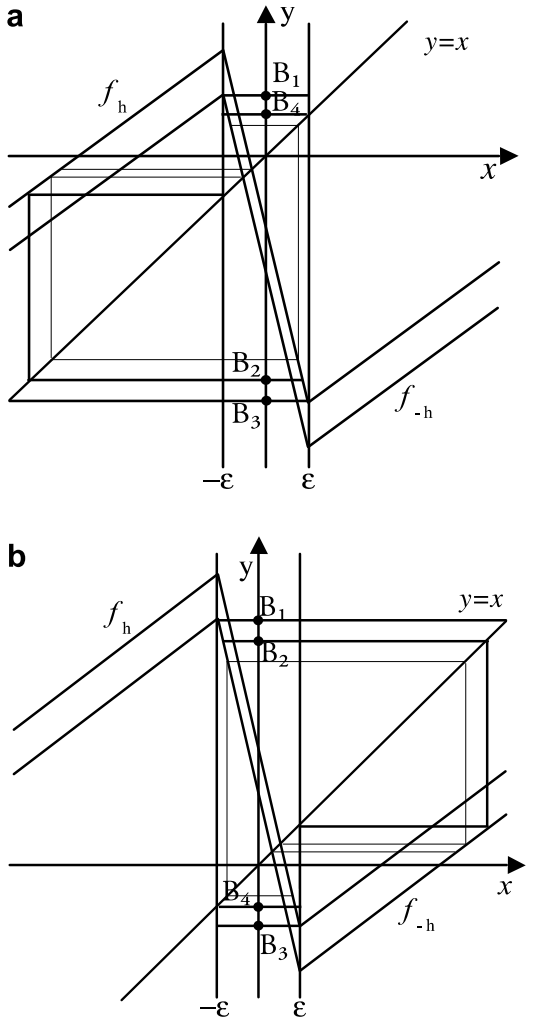

Fig. 2. The configurations for the graphs of f_h, f_{-h} and the relative positions of certain points which correspond to the inequality in condition (P_1-i). (a) $B_1:(0,-\mu\varepsilon-\omega a_0-h)$, $B_2:(0,(-\varepsilon+\omega a_0-h)/\mu)$, $B_3:(0,\mu\varepsilon+\omega(1-a_0)+h)$, $B_4:(0,\varepsilon)$. (b) $B_1:(0,-\mu\varepsilon-\omega a_0+h)$, $B_2:(0,(\varepsilon-\omega(1-a_0)+h)/\mu)$, $B_3:(0,\mu\varepsilon+\omega(1-a_0)-h)$, $B_4:(0,-\varepsilon)$.

(P_1-ii) $\mu>0, \frac{1}{\mu}(\varepsilon-\omega a_0+h)<-\mu\varepsilon-\omega(1-a_0)-h.$
(P_1-iii) $\mu>0, \frac{1}{\mu}(\varepsilon-\omega(1-a_0)+h)<-\mu\varepsilon-\omega a_0+h.$

Notably, (P_1-ii) (resp. (P_1-iii)) is formulated for finding preimages of fixed points from the left (resp. right) part of the graph of f_ς; Fig. 2a (resp. Fig. 2b) corresponds to condition (P_1-ii) (resp. (P_1-iii)).

Proposition 2. *If the parameters* $(\mu,\omega,\varepsilon,a_0,h)$ *satisfy* (P_1-i) *and* (P_1-ii) *or* (P_1-iii), *then the fixed point* $\bar{x}^{\mathrm{m}}_\varsigma\in\Omega_{\mathrm{m}}$ *is a snap-back repeller for* f_ς *with* $-h\leqslant\varsigma\leqslant h$.

Proof. We develop a scheme for constructing preimages of a repelling fixed point, and confirm that this fixed point is a snap-back repeller. We only explain the case satisfying (P_1-i) and (P_1-ii), with configuration illustrated in Fig. 2a. Notably, the following quantities "b_i" correspond to the vertical coordinates of points "B_i" in Fig. 2a. Since (P_1-i) holds, there exists a fixed point $\bar{x}^{\mathrm{m}}_\varsigma$ lying in Ω_{m}, for each f_ς with $-h\leqslant\varsigma\leqslant h$. In addition, ($P_1$-i) also implies $\mu+\frac{\omega}{2\varepsilon}<-1$; subsequently, there exists a point $x^{(-1)}_\varsigma\in\Omega_\ell$ such that $f_\varsigma(x^{(-1)}_\varsigma)=\bar{x}^{\mathrm{m}}_\varsigma$, if $\mu>0$. Let $b_2=(-\varepsilon+\omega a_0-h)/\mu$ with $\mu>0$, then

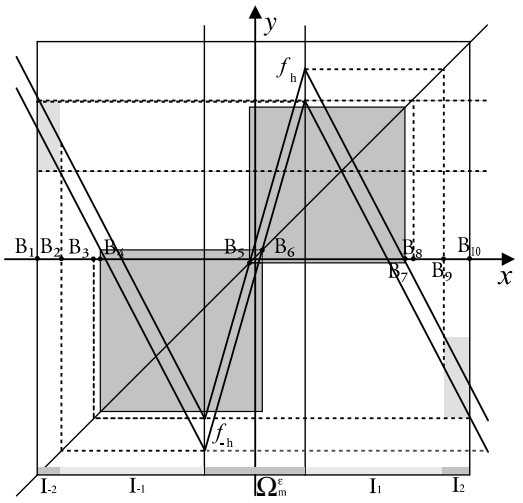

Fig. 3. The configurations for the graphs of f_h, f_{-h} and the relative positions of certain points which correspond to the inequality in condition (P_2-i). $B_1 : (\varepsilon + (\omega - 2h)/\mu, 0)$, $B_2 : (-\mu\varepsilon - \omega a_{0i} - h, 0)$, $B_3 : (-\mu\varepsilon - \omega a_{0i} + h, 0)$, $B_4 : (((\frac{1}{2}\omega - \omega a_{0i} - h)/(1 - \mu - \frac{\omega}{2\varepsilon}) + \omega a_{0i} + h)/\mu, 0)$, $B_5 : ((\frac{1}{2}\omega - \omega a_{0i} + h)/(1 - \mu - \frac{\omega}{2\varepsilon}), (\frac{1}{2}\omega - \omega a_{0i} + h)/(1 - \mu - \frac{\omega}{2\varepsilon}))$, $B_6 : ((\frac{1}{2}\omega - \omega a_{0i} - h)/(1 - \mu - \frac{\omega}{2\varepsilon}), (\frac{1}{2}\omega - \omega a_{0i} - h)/(1 - \mu - \frac{\omega}{2\varepsilon}))$, $B_7 : ((\frac{1}{2}\omega - \omega a_{0i} + h)/(1 - \mu - \frac{\omega}{2\varepsilon}) - \omega(1 - a_{0i}) - h)/\mu, 0)$, $B_8 : (\mu\varepsilon + \omega(1 - a_{0i}) - h, 0)$, $B_9 : (\mu\varepsilon + \omega(1 - a_{0i}) + h, 0)$, $B_{10} : (-\varepsilon - (\omega - 2h)/\mu, 0)$.

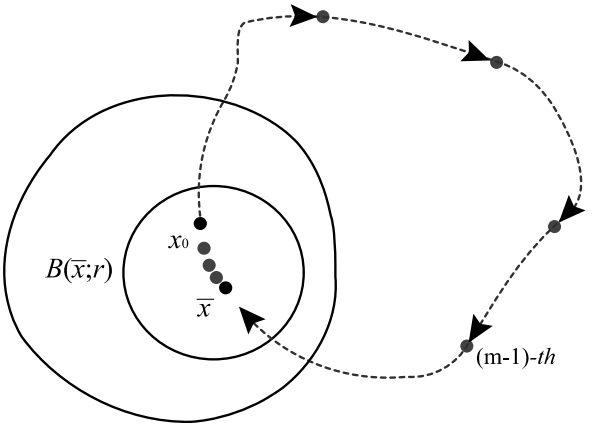

Fig. 4. Illustration of a snap-back repeller $\bar{x}$ with a snapback point x_0.

$f_h(b_2) = -\varepsilon$. It follows from (P_1-ii) that $b_2 > b_3 = f_h(\varepsilon) = \mu\varepsilon + \omega(1 - a_0) + h$. Thus there exists $x_\varsigma^{(-2)} \in \Omega_m$ such that $f_\varsigma(x_\varsigma^{(-2)}) = x_\varsigma^{(-1)}$. Let $b_1 = -\mu\varepsilon - \omega a_0 - h$, then $b_1 = f_h(-\varepsilon) > b_4 = \varepsilon$, due to ($P_1$-i). Hence, there exists a sequence of points $\{x_\varsigma^{(-i)} | i = 3, 4, 5 \ldots\}$ lying in region Ω_m, which are successive preimages of $\bar{x}_\varsigma^m$ under f_ς. Consequently, $\bar{x}_\varsigma^m$ is a snap-back repeller lying in Ω_m and $\{x_\varsigma^{(-i)} | i = 1, 2, \ldots\}$ is a homoclinic orbit for f_ς. □

There are also conditions for the fixed points in Proposition 1(ii) to be snap-back repellers; namely

(P_2-ii) (a) $\mu\varepsilon + \omega(1 - a_0) - h > \frac{1}{\mu}[(\frac{1}{2}\omega - \omega a_0 + h)/(1 - \mu - \frac{\omega}{2\varepsilon}) - \omega + \omega a_0 - h]$,
(b) $\mu\varepsilon + \omega a_0 - h > -\frac{1}{\mu}[(\frac{1}{2}\omega - \omega a_0 - h)/(1 - \mu - \frac{\omega}{2\varepsilon}) + \omega a_0 + h]$.

Proposition 3. *If the parameters* $(\mu, \omega, \varepsilon, a_0, h)$ *satisfy either* (P_2-i)(a) *and* (P_2-ii)(a) *or* (P_2-i)(b) *and* (P_2-ii)(b), *then the fixed points in Proposition* 1(ii) *are snap-back repellers for the single neuron maps* f_ς *with* $-h \leqslant \varsigma \leqslant h$.

Boundedness of iterations are necessary in the applications of TCNN. For the case $0 < \mu < 1$, i.e., the situations of Fig. 2a and b, interval $[-M, M]$ is a trapping region for map (5), provided $M > 0$ is large enough. For the case $\mu < -1$, the following condition is needed:

(P_2-iii) $\varepsilon + \frac{1}{\mu}(\omega - 2h) < \min\{-\mu\varepsilon - \omega a_0 - h, -\mu\varepsilon - \omega(1 - a_0) - h\}$.

Proposition 4. *Assume that the parameters* $(\mu, \omega, \varepsilon, a_0, h)$ *satisfy* $\mu < -1$, (P_2-i)(a) *and* (P_2-iii), *then interval* $[\varepsilon + \frac{\omega}{\mu} - \frac{2h}{\mu}, -\varepsilon - \frac{\omega}{\mu} + \frac{2h}{\mu}]$ *is a trapping region for the single neuron maps* f_ς *with* $-h \leqslant \varsigma \leqslant h$.

2.2. Multi-dimensional chaotic neural networks

We shall apply Propositions 2 and 3 for single neuron maps to establish the chaotic behaviors for the n-dimensional neural networks (4). Note that $\mathbb{R}^n$ can be written as the direct sum of the following subsets:

$$\mathbf{\Omega}_{q_1\cdots q_n} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_i \in \Omega_{q_i}; q_i = \text{“r”}, \text{“m”}, \text{“}\ell\text{”}; i = 1, \ldots, n\}, \tag{9}$$

as illustrated in Fig. 5 for $n = 2$. $\mathbf{\Omega}_{\mathrm{m}\cdots\mathrm{m}}$ is called the *interior region*; each $\mathbf{\Omega}_{q_1\cdots q_n}$, with $q_i =$ “ℓ”, “r”, for all i, is called a *saturated region*; each $\mathbf{\Omega}_{q_1\cdots q_n}$, with $q_i =$ “ℓ”, or “r”, for some i, and $q_j =$ “m” for some j, is called a *mixed region*.

We derive the following theorem on the existence of fixed point for system (4). Recall $h_i = \sum_{j=1}^n |\omega_{ij}| + |v_i|$.

Theorem 1. *If the parameters* $(\mu, \omega, \varepsilon, a_{0i}, h_i)$, $i = 1, \ldots, n$, *satisfy* (P_1-i) (*resp.* (P_2-i)(a)), *then there exists one fixed point* (*resp.* 3^n *fixed points*) *in* $\mathbf{\Omega}_{\mathrm{m}\cdots\mathrm{m}}$ (*resp. the* 3^n *regions* $\mathbf{\Omega}_{q_1\cdots q_n}$ *defined in* (9)) *for system* (4).

Proof. We prove the case that (P_2-i)(a) holds for parameters $(\mu, \omega, \varepsilon, a_{0i}, h_i)$, $i = 1, \ldots, n$. For a given $(\tilde{\xi}_1, \ldots, \tilde{\xi}_n) \in \mathbb{R}^n$, by Proposition 1(ii), there exist points $\xi_i^\ell \in \Omega_\ell$, $\xi_i^{\mathrm{m}} \in \Omega_{\mathrm{m}}$, $\xi_i^{\mathrm{r}} \in \Omega_{\mathrm{r}}$ such that

$$\xi_i^* = \mu\xi_i^* + \omega[g_\varepsilon(\xi_i^*) - a_{0i}] + \sum_{j=1}^n \omega_{ij} g_\varepsilon(\tilde{\xi}_j) + v_i, \tag{10}$$

where $* =$ “ℓ”, “m”, “r” for each i. Restated, each ξ_i^* is a fixed point of the one-dimensional map $\xi_i \mapsto \mu\xi_i + \omega[g_\varepsilon(\xi_i) - a_{0i}] + \sum_{j=1}^n \omega_{ij} g_\varepsilon(\tilde{\xi}_j) + v_i$. In fact, each of these $(\xi_1^*, \ldots, \xi_n^*)$ lies in a compact proper subset $\mathbf{\Omega}^*_{q_1\cdots q_n}$ of $\mathbf{\Omega}_{q_1\cdots q_n}$, according to our formation in Section 2.1. Consider a fixed region $\Omega^*_{q_1\cdots q_n}$ for certain $q_i =$ “ℓ”, “m”, “r” ($i = 1, \ldots, n$). Let $\mathbf{H} : \mathbf{\Omega}^*_{q_1\cdots q_n} \mapsto \mathbf{\Omega}^*_{q_1\cdots q_n}$ be defined by $\mathbf{H}(\tilde{\xi}_1, \ldots, \tilde{\xi}_n) = (\xi_1^{q_1}, \ldots, \xi_n^{q_n})$. We shall show that there exists a fixed point for $\mathbf{H}$. Define $\mathbf{G} : \mathbf{\Omega}^*_{q_1\cdots q_n} \times \mathbf{\Omega}^*_{q_1\cdots q_n} \to \mathbf{\Omega}^*_{q_1\cdots q_n}$, $\mathbf{G} = (G_1, \ldots, G_n)$, by

$$G_i(\tilde{\mathbf{x}}, \mathbf{x}) = x_i - \mu x_i - \omega[g_\varepsilon(x_i) - a_{0i}] - \sum_{j=1}^n \omega_{ij} g_\varepsilon(\tilde{x}_j) - v_i,$$

where $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)$, $i = 1, \ldots, n$. Notably, $\mathbf{G}(\tilde{\mathbf{x}}, \mathbf{H}(\tilde{\mathbf{x}})) = 0$ for every $\tilde{\mathbf{x}} \in \mathbf{\Omega}_{q_1\cdots q_n}$, by Eq. (10). Now,

$$\frac{\partial \mathbf{G}}{\partial \mathbf{x}} = \mathrm{diag}[\chi_1, \ldots, \chi_n],$$

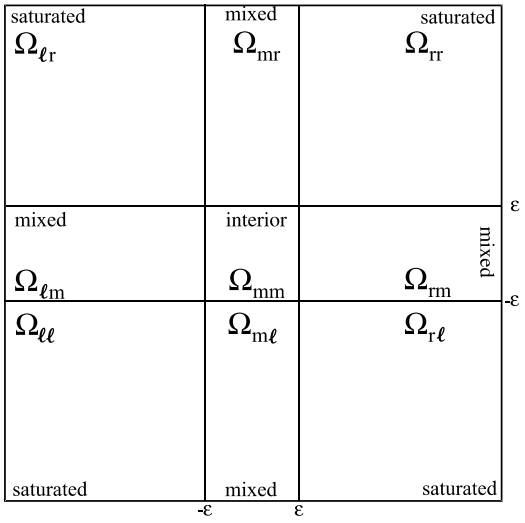

Fig. 5. Illustration of $\Omega_{q_1 q_2}$ in $\mathbb{R}^2$, where q_1 and q_2 are “ℓ” or “m” or “r”.

where $\chi_i = 1 - \mu - \omega g'_\varepsilon(x_i)$, $i = 1, \ldots, n$. Note that $g'_\varepsilon = 0$ in regions Ω_r, Ω_ℓ, and $g'_\varepsilon = \frac{1}{2\varepsilon}$ in the interior region Ω_m. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbf{\Omega}^*_{q_1 \cdots q_n}$, we have $\chi_i = 1 - \mu - \frac{\omega}{2\varepsilon}$ or $1 - \mu$ for each i, which are nonzero due to $0 < (1-\mu)/\omega < 1/(2\varepsilon)$. It follows that $\mathbf{H}$ is a C^1 function on region $\mathbf{\Omega}^*_{q_1 \cdots q_n}$, by the implicit function theorem. Thus, there exists one fixed point $\bar{\mathbf{x}}$ of $\mathbf{H}$ in $\mathbf{\Omega}^*_{q_1 \cdots q_n}$, which is also a fixed point of system (4), by the Brouwer's fixed point theorem. Consequently, there are 3^n fixed points of system (4) in $\mathbb{R}^n$. □

Theorem 2. *If the parameters* $(\mu, \omega, \varepsilon, a_{0i}, h_i)$, $i = 1, \ldots, n$, *satisfy* (P_1-i), *and* (P_1-ii) *or* (P_1-iii), *then there exists a snap-back repeller in the interior region* $\mathbf{\Omega}_{m \cdots m}$ *for system* (4).

Proof. By Theorem 1, there exists a fixed point $\bar{\mathbf{x}} = (\bar{x}_1, \ldots, \bar{x}_n)$ in $\mathbf{\Omega}_{m \cdots m}$ such that $F_i(\bar{\mathbf{x}}) = \bar{x}_i$, $i = 1, \ldots, n$. To find preimages of $\bar{\mathbf{x}}$ under $\mathbf{F}$ for the system, we need to solve $F_i(\mathbf{x}) = \bar{x}_i, i = 1, \ldots, n$. To achieve this, recall Eq. (7) that $\check{f}_i(x_i) \leqslant F_i(\mathbf{x}) \leqslant \hat{f}_i(x_i), i = 1, \ldots, n$. We observe that for a given $(\tilde{x}_1, \ldots, \tilde{x}_n) \in \mathbf{\Omega}_{\ell \cdots \ell}$, one can always find $(x_1, \ldots, x_n) \in \mathbf{\Omega}_{\ell \cdots \ell}$ satisfying

$$\bar{x}_i = \mu x_i + \omega[g_\varepsilon(x_i) - a_{0i}] + \sum_{j=1}^{n} \omega_{ij} g_\varepsilon(\tilde{x}_j) + v_i, \tag{11}$$

if the parameters $(\mu, \omega, \varepsilon, a_{0i}, h_i), i = 1, \ldots, n$, satisfy conditions ($P_1$-i), and ($P_1$-ii), according to our previous formulations and Proposition 2. Define a map $\mathbf{G} : \mathbf{\Omega}_{\ell \cdots \ell} \times \mathbf{\Omega}_{\ell \cdots \ell} \to \mathbb{R}^n$ by

$$G_i(\tilde{\mathbf{x}}, \mathbf{x}) = \bar{x}_i - \mu x_i - \omega[g_\varepsilon(x_i) - a_{0i}] - \sum_{j=1}^{n} \omega_{ij} g_\varepsilon(\tilde{x}_j) - v_i,$$

where $\mathbf{G} = (G_1, \ldots, G_n)$. Then $(\partial \mathbf{G}/\partial \mathbf{x})(\tilde{\mathbf{x}}, \mathbf{x}) = \mathrm{diag}[\chi_1, \ldots, \chi_n]$, where $\chi_i = -\mu - \omega g'_\varepsilon(x_i)$, $i = 1, \ldots, n$. Thus, $\det(\partial \mathbf{G}/\partial \mathbf{x})(\tilde{\mathbf{x}}, \mathbf{x}) \neq 0$. By similar arguments as the ones in the proof of Theorem 1, there exists a point $\underline{\mathbf{x}} \in \mathbf{\Omega}_{\ell \cdots \ell}$ such that $\mathbf{G}(\underline{\mathbf{x}}, \underline{\mathbf{x}}) = 0$. Denoting $\bar{\mathbf{x}}^{(-1)} = \underline{\mathbf{x}}$, it follows that $\mathbf{F}(\bar{\mathbf{x}}^{(-1)}) = \bar{\mathbf{x}}$. Successively, by similar arguments, we also obtain $\bar{\mathbf{x}}^{(-2)}$ in $\mathbf{\Omega}_{m \cdots m}$ with $F_i(\bar{\mathbf{x}}^{(-2)}) = \bar{x}_i^{(-1)}, i = 1, \ldots, n$, and then a sequence $\{\bar{\mathbf{x}}^{(-k)} | k \geqslant 3, k \in \mathbb{N}\}$ in $\mathbf{\Omega}_{m \cdots m}$ with $F_i(\bar{\mathbf{x}}^{-(k+1)}) = \bar{x}_i^{(-k)}, i = 1, \ldots, n$. Furthermore, the entries for the Jacobian matrix of $\mathbf{F}$ in $\mathbf{\Omega}_{m \cdots m}$ are

$$[D\mathbf{F}]_{ij} = \mu + \frac{\omega + \omega_{ii}}{2\varepsilon}, \quad \text{if } i = j, \qquad = \frac{\omega_{ij}}{2\varepsilon}, \quad \text{if } i \neq j.$$

Notably, $\mu + \frac{\omega}{2\varepsilon} + \frac{1}{2\varepsilon} \max_{i=1,\ldots,n} (\sum_{j=1}^{n} |\omega_{ij}|) < -1$, due to conditions ($P_1$-i), and ($P_1$-ii). Thus, the absolute values of all eigenvalues of $D\mathbf{F}$ are greater than one, by the Gerschgorin's theorem. Therefore, $\mathbf{F}$ is expanding in $\mathbf{\Omega}_{m \cdots m}$. Hence, the sequence $\{\bar{\mathbf{x}}^{(-k)} | k = 1, 2, \ldots\}$ approaches the fixed point in $\Omega_{m \cdots m}$ and thus the fixed point $\bar{\mathbf{x}}$ is a snap-back repeller. □

Theorem 3. *If the parameters* $(\mu, \omega, \varepsilon, a_{0i}, h_i)$, $i = 1, \ldots, n$, *satisfy* (P_2-i)(a), *and* (P_2-ii)(a) *or* (P_2-ii)(b), *then there exists a snap-back repeller for the multi-dimensional neural network* (4).

The proof employs Proposition 3 and resembles the one of Theorem 2.

3. Stability analysis for the TCNN

We plan to study the asymptotic behaviors for the TCNN in this section. Analogous to the TCNN with logistic output function, there also exists a time-dependent Lyapunov function for the TCNN with piecewise linear output function (3). We shall derive this Lyapunov function in Section 3.1 and employ the non-autonomous discrete-time LaSalle's invariant principle to analyze the convergence of the TCNN. In Section 3.2, we analyze the instability of the fixed points in the interior region and the mixed regions so that the evolutions of TCNN, with suitably chosen parameters, settle at saturated regions from almost all initial points.

3.1. Lyapunov function for the TCNN

We consider the TCNN with more general cooling schedule, namely

$$x_i(t+1) = \mu x_i(t) + (1-\beta)^{q(t)} \omega[g_\varepsilon(x_i(t)) - a_{0i}] + \sum_{j=1}^{n} \omega_{ij} g_\varepsilon(x_j(t)) + v_i, \tag{12}$$

where $i = 1, \ldots, n$, $0 < \beta < 1$; g_ε is defined in (3); $q(t)$ satisfies the condition that there exists an $n_1 \in \mathbb{N}$ such that $q(t) - t \geqslant 0$ for all $t > n_1$. The standard cooling process simply takes $q(t) = t$. Correspondingly, we define

$$F_i(t, \mathbf{x}) = \mu x_i + (1-\beta)^{q(t)} \omega [g_\varepsilon(x_i) - a_{0i}] + \sum_{j=1}^{n} \omega_{ij} g_\varepsilon(x_j) + v_i. \tag{13}$$

Let us recall the notion of Lyapunov function for discrete-time non-autonomous system, cf. [18]. Let $\mathbb{N}$ be the set of positive integers. For a given continuous function $\mathbf{F} : \mathbb{N} \times \mathbb{R}^n \to \mathbb{R}^n$, we consider the non-autonomous dynamical equation

$$\mathbf{x}(t+1) = \mathbf{F}(t, \mathbf{x}(t)). \tag{14}$$

A sequence of points $\{\mathbf{x}(t) | t = 1, 2, \ldots\}$ in $\mathbb{R}^n$ is a solution of (14) if $\mathbf{x}(t+1) = \mathbf{F}(t, \mathbf{x}(t))$, for all $t \in \mathbb{N}$. Let $O_{\mathbf{x}} = \{\mathbf{x}(t) | t \in \mathbb{N}, \mathbf{x}(1) = \mathbf{x}\}$ be the orbit of $\mathbf{x}$. $\mathbf{p}$ is an ω-limit point of $O_{\mathbf{x}}$ if there exists a sequence of positive integers $\{t_k\}$ with $t_k \to \infty$ as $k \to \infty$, such that $\mathbf{p} = \lim_{k\to\infty} \mathbf{x}(t_k)$. Denote by $\omega(\mathbf{x})$ the set of all ω-limit points of $O_{\mathbf{x}}$. Let n_i be a positive integer, and let $\mathbb{N}_i$ denote the set of all positive integers larger than n_i. Denote by $\overline{\Omega}$ the closure of $\Omega \subseteq \mathbb{R}^n$. For a function $V : \mathbb{N}_0 \times \Omega \to \mathbb{R}$, we define $\dot{V}(t, \mathbf{x}) = V(t+1, \mathbf{F}(t, \mathbf{x})) - V(t, \mathbf{x})$ so that if $\{\mathbf{x}(t)\}$ is a solution of Eq. (14), then $\dot{V}(t, \mathbf{x}(t)) = V(t+1, \mathbf{x}(t+1)) - V(t, \mathbf{x}(t))$. V is said to be a *Lyapunov function* for (14) if (i) each $V(t, \cdot)$ is continuous, and (ii) for each $\mathbf{p} \in \overline{\Omega}$, there exists a neighborhood U of $\mathbf{p}$ such that $V(t, \mathbf{x})$ is bounded below for $\mathbf{x} \in U \cap \Omega$ and $t \in \mathbb{N}_1$, $n_1 \geqslant n_0$, and (iii) there exists a non-degenerate continuous function $Q_0 : \overline{\Omega} \to \mathbb{R}$ such that $\dot{V}(t, \mathbf{x}) \leqslant -Q_0(\mathbf{x}) \leqslant 0$ for all $\mathbf{x} \in \Omega$ and for all $t \in \mathbb{N}_2$, $n_2 \geqslant n_1$, or (iii)′ there exist a non-degenerate continuous function $Q_0 : \overline{\Omega} \to \mathbb{R}$ and an equi-continuous family of functions $Q(t, \cdot) : \overline{\Omega} \to \mathbb{R}$ such that $\lim_{t\to\infty} |Q(t, \mathbf{x}) - Q_0(\mathbf{x})| = 0$ for all $\mathbf{x} \in \Omega$ and $\dot{V}(t, \mathbf{x}) \leqslant -Q(t, \mathbf{x}) \leqslant 0$ for all $(t, \mathbf{x}) \in \mathbb{N}_2 \times \Omega$, $n_2 \geqslant n_1$. Define

$$S_0 = \{\mathbf{x} \in \overline{\Omega} : Q_0(\mathbf{x}) = 0\}. \tag{15}$$

We recall the following theorem for the asymptotic behaviors of solutions to Eq. (14).

Theorem 4 [19]. *Let $n_0 \in \mathbb{N}$, $V : \mathbb{N}_0 \times \Omega \to \mathbb{R}$ be a Lyapunov function for Eq. (14) and $O_{\mathbf{x}}$ be an orbit of Eq. (14) lying in Ω for all $t \in \mathbb{N}_1$. Then $\lim_{t\to\infty} Q(t, \mathbf{x}(t)) = 0$, and $\omega(\mathbf{x}) \subset S_0$.*

Further properties for $\omega(\mathbf{x})$ can be derived if there exists a limiting map $\overline{\mathbf{F}}(\mathbf{x})$ for $\mathbf{F}(t, \mathbf{x})$, i.e., there exists a continuous map $\overline{\mathbf{F}} : \mathbb{R}^{\mathbf{n}} \to \mathbb{R}^{\mathbf{n}}$ such that $\lim_{t\to\infty} |\mathbf{F}(t, \mathbf{x}) - \overline{\mathbf{F}}(\mathbf{x})| = \mathbf{0}$, for all $\mathbf{x} \in \mathbb{R}^n$. The following theorem generalizes Theorem 2.2 in [19].

Theorem 5. *Assume that $\overline{\mathbf{F}}(\mathbf{x})$ is the limiting map of $\mathbf{F}(t, \mathbf{x})$, and the orbit $O_{\mathbf{x}}$ is bounded, then $\omega(\mathbf{x})$ is positively invariant under $\overline{\mathbf{F}}$; moreover, if the ω-limit set $\omega(\mathbf{x})$ of $O_{\mathbf{x}}$ is contained in the set of fixed points of $\overline{\mathbf{F}}$, then $\omega(\mathbf{x})$ is connected. Under this circumstances, if $\overline{\mathbf{F}}$ has only finitely many fixed points, then the orbit $O_{\mathbf{x}}$ approaches some single fixed point of $\overline{\mathbf{F}}$, as time tends to infinity.*

Let us define the limiting mapping $\overline{\mathbf{F}} = (\overline{F}_1, \ldots, \overline{F}_n)$ of $\mathbf{F}$ in (13) by

$$\overline{F}_i(\mathbf{x}) = \mu x_i + \sum_{j=1}^{n} \omega_{ij} g_\varepsilon(x_j) + v_i, \quad i = 1, \ldots, n. \tag{16}$$

It follows that $|\overline{\mathbf{F}}(\mathbf{x}) - \mathbf{F}(t, \mathbf{x})|$ converges uniformly to zero in $\mathbf{x} \in \mathbb{R}^n$ as $t \to \infty$, since $|1 - \beta| < 1$, $|\overline{\mathbf{F}}(\mathbf{x}) - \mathbf{F}(t, \mathbf{x})|^2 = \sum_{i=1}^{n} |(1-\beta)^{q(t)} \omega(y_i - a_{0i})|^2$ and $\sum_{i=1}^{n} |\omega(y_i - a_{0i})|^2$ is bounded. Let us consider the function

$$V(t, \mathbf{x}) = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} y_i y_j - \sum_{i=1}^{n} v_i y_i - (\mu - 1)\varepsilon \sum_{i=1}^{n} (y_i^2 - y_i) + c^t, \tag{17}$$

where $y_i = g_\varepsilon(x_i)$, for $i = 1, \ldots, n$, $0 < c < 1$. Note that V is globally Lipschitz, but not C^1. Let $W = [\omega_{ij}]_{n\times n}$.

Theorem 6. *If $0 \leqslant \mu \leqslant 1$, $\varepsilon > 0$, $|\frac{1-\beta}{c}| < 1$ and the matrix $W + 2\varepsilon(1-\mu)I$ is positive-definite, then there exists $n_2 \in \mathbb{N}, n_2 > n_1$ so that $V(t+1, \mathbf{x}(t+1)) \leqslant V(t, \mathbf{x}(t))$ for $t \geqslant n_2$ and V is a Lyapunov function for system (12) on $\mathbb{N}_2 \times \mathbb{R}^n$.*

Proof.

$$\begin{aligned}V(t+1,\mathbf{x}(t+1))-V(t,\mathbf{x}(t)) = & -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\Delta_t y_i\Delta_t y_j-(\mu-1)\varepsilon\sum_{i=1}^{n}[y_i^2(t+1)-y_i^2(t)-y_i(t+1)+y_i(t)]\\ &+\sum_{i=1}^{n}[\mu x_i(t)-x_i(t+1)]\Delta_t y_i+\sum_{i=1}^{n}(1-\beta)^{q(t)}\omega[y_i(t)-a_{0i}]\Delta_t x_i+c^{t+1}-c^t\\ =&-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\Delta_t y_i\Delta_t y_j-(\mu-1)\varepsilon\sum_{i=1}^{n}[y_i^2(t+1)-y_i^2(t)-y_i(t+1)+y_i(t)]\\ &+\mu\sum_{i=1}^{n}[x_i(t)-x_i(t+1)]\Delta_t y_i+(\mu-1)\sum_{i=1}^{n}x_i(t+1)\Delta_t y_i+\sum_{i=1}^{n}(1-\beta)^{q(t)}\omega[y_i(t)\\ &-a_{0i}]\Delta_t y_i+c^{t+1}-c^t=-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\Delta_t y_i\Delta_t y_j-\mu\sum_{i=1}^{n}[x_i(t+1)-x_i(t)]\Delta_t y_i\\ &+(\mu-1)\sum_{i=1}^{n}\{x_i(t+1)\Delta_t y_i-\varepsilon[y_i^2(t+1)-y_i^2(t)]+\varepsilon\Delta_t y_i\}+\sum_{i=1}^{n}(1-\beta)^{q(t)}\omega[y_i(t)\\ &-a_{0i}]\Delta_t y_i+c^{t+1}-c^t. \end{aligned}\tag{18}$$

where $\Delta_t\mathbf{y}=(\Delta_t y_1,\ldots,\Delta_t y_n)$ with $\Delta_t y_i=g_\varepsilon(x_i(t+1))-g_\varepsilon(x_i(t))$. Let us compute the term in the bracket of (18). Set

$$\begin{aligned}\Gamma_i=\Gamma_i(x_i(t),x_i(t+1)) &:= x_i(t+1)\Delta_t y_i-\varepsilon[y_i^2(t+1)-y_i^2(t)]+\varepsilon\Delta_t y_i\\ &= F_i(\mathbf{x}(t))\Delta_t y_i-\varepsilon\Delta_t y_i[g_\varepsilon(F_i(\mathbf{x}(t)))+g_\varepsilon(x_i(t))-1].\end{aligned}$$

We shall justify that $\Gamma_i\geqslant\varepsilon(\Delta_t y_i)^2$ in Appendix C. On the other hand, $[x_i(t+1)-x_i(t)][g_\varepsilon(x_i(t+1))-g_\varepsilon(x_i(t))]\geqslant 0$, due to that g_ε is non-decreasing for $\varepsilon>0$. Let $\widetilde{W}=W+2\varepsilon(1-\mu)I$. Then,

$$V(t+1,\mathbf{x}(t+1))-V(t,\mathbf{x}(t))\leqslant-\frac{1}{2}(\Delta_t\mathbf{y})^T\widetilde{W}\Delta_t\mathbf{y}+M(1-\beta)^{q(t)}+c^{t+1}-c^t,$$

where $M=\frac{1}{n}\max_{i=1}^{n}\{\sup_{t\in\mathbb{N},\mathbf{y}\in\mathbb{R}^n}|\omega(y_i-a_{0i})\Delta_t y_i|\}$. Since $|\frac{1-\beta}{c}|<1$ and $q(t)-t\geqslant 0$ for all $t>n_1$, there exists $n_2\in\mathbb{N}$ such that $M(1-\beta)^{q(t)}+c^{t+1}-c^t=c^t[M(1-\beta)^{q(t)-t}(\frac{1-\beta}{c})^t+c-1]<c^t[M(\frac{1-\beta}{c})^t+c-1]<0$ if $t\geqslant n_2$. Hence,

$$V(t+1,\mathbf{x}(t+1))-V(t,\mathbf{x}(t))\leqslant-\frac{1}{2}(\Delta_t\mathbf{y})^T\widetilde{W}\Delta_t\mathbf{y}\leqslant 0.$$

Define

$$Q(t,\mathbf{x})=-\frac{1}{2}(\Delta_t\mathbf{y})^T\widetilde{W}\Delta_t\mathbf{y},\quad Q_0(\mathbf{x})=-\frac{1}{2}(\Delta\mathbf{y})^T\widetilde{W}\Delta\mathbf{y},$$

where $\Delta\mathbf{y}=(\Delta y_1,\ldots,\Delta y_n)$, $\Delta y_i=g_\varepsilon(\overline{F}_i(\mathbf{x}))-g_\varepsilon(x_i)$. Notice that

$$\lim_{t\to\infty}|Q(t,\mathbf{x})-Q_0(\mathbf{x})|=0$$

for all $\mathbf{x}\in\mathbb{R}^n$. Thus, $V(t+1,\mathbf{x}(t+1))\leqslant V(t,\mathbf{x}(t))$ for $t\geqslant n_2$. □

Accordingly, under the assumption of Theorem 6,

$$\omega(\mathbf{x})\subset S_0=\{\mathbf{z}:g_\varepsilon(\overline{F}_i(\mathbf{z}))-g_\varepsilon(z_i)=0,i=1,\ldots,n\}$$

for any $\mathbf{x}\in\mathbb{R}^n$, due to Theorem 4. We observe that the only points which are positively invariant under $\overline{\mathbf{F}}$ are fixed points of $\overline{\mathbf{F}}$, in respecting the definition of the limiting map $\overline{\mathbf{F}}$ defined in Eq. (16). Thus, $\omega(\mathbf{x})$ is connected and contained in the set of fixed points of $\overline{\mathbf{F}}$. If, furthermore, $\overline{\mathbf{F}}$ has only finitely many fixed points, then every orbit $O_{\mathbf{x}}$ tends to a single fixed point of $\overline{\mathbf{F}}$.

3.2. Instability for fixed points in the interior and mixed regions

For the previous TCNN models in the literatures, there is a situation that the output component tends to a value which is neither close to one nor to zero. For example, a numerical simulation for the TCNN with logistic output function exhibiting such a phenomenon is illustrated in Fig. 6.

In this section, we discuss the parameter conditions so that almost every trajectory generated by $\mathbf{F}$ tends to a stable fixed point of $\overline{\mathbf{F}}$, which lies in a saturated region. Restated, starting from almost all initial points, every entry of the

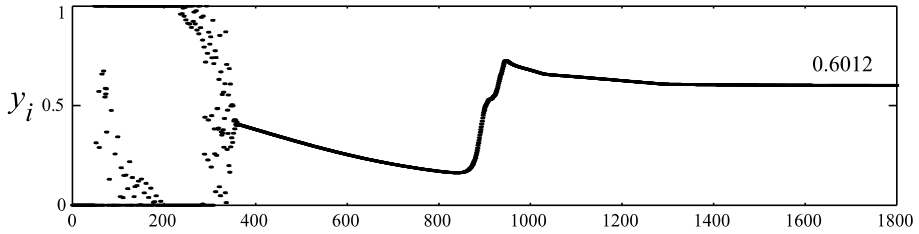

Fig. 6. An example that the TCNN with the logistic output function has an infeasible solution, i.e., there exists an output entry y_i which approaches 0.6012, neither close to 1 nor to 0, after 1400 iterations.

output matrix $[y_{ij}(t)]_{n\times n}$ tends to 0 or 1 as $t \to \infty$. This convergence then corresponds to a possible route as the TCNN is applied to solve the TSP.

For a mixed region $\mathbf{\Omega}_{q_1\cdots q_n}$, we define the index sets

$$J_0 = \{i \in \{1,\ldots,n\} || x_i| \leqslant \varepsilon\}, J_1 = \{i \in \{1,\ldots,n\} || x_i| > \varepsilon\},$$

and denote by $|J_0|$, $|J_1|$ their cardinalities. Let W_0 be the $|J_0| \times |J_0|$ submatrix of W, which is obtained by deleting the jth row and the jth column of W, for all $j \in J_1$. The linear part $D\overline{\mathbf{F}}$ of $\overline{\mathbf{F}}$ restricted to a mixed region $\mathbf{\Omega}_{q_1\cdots q_n}$, after rearranging the indices, can be represented in the matrix form:

$$D\overline{\mathbf{F}} = \begin{bmatrix} B & * \\ 0 & D \end{bmatrix}, \tag{19}$$

where $*$ is an $|J_1| \times |J_0|$ matrix, $B = \mu I_1$ and $D = \mu I_0 + \frac{1}{2\varepsilon} W_0$ with I_1 (resp. I_0) being the identity matrix of size $|J_1|$ (resp. $|J_0|$). Thus, the eigenvalues of matrix (19) are the eigenvalues of B and D. Note that W_0 is symmetric. If W_0 has a nonzero eigenvalue, which holds generically, then there exists an eigenvalue of $D\overline{\mathbf{F}}$ with moduli greater than one, provided ε is sufficiently small. By similar reasoning, it can be seen that the fixed point in the interior region $\mathbf{\Omega}_{\mathrm{m}\cdots\mathrm{m}}$ is unstable, if ε is sufficiently small. On the other hand, a fixed point in a saturated region is always stable if $0 < \mu < 1$, due to that the linear part of $\overline{\mathbf{F}}$ restricted to a saturated region is $\mu I_{n\times n}$.

Moreover, it is possible to impose a condition so that the fixed points of $\overline{\mathbf{F}}$ in certain mixed regions do not exist. For example, let us consider a mixed region $\mathbf{\Omega}_{q_1\cdots q_n}$ with $|J_0| = 1$. Without loss of generality, we suppose that $q_k =$ "m". It is computed that if $\overline{\mathbf{x}} = (\bar{x}_1,\ldots,\bar{x}_n)$ is a fixed point of $\overline{\mathbf{F}}$ in this region, then

$$\bar{x}_k = \frac{\sum_{j=1}^n \omega_{kj} y_j + v_k}{1-\mu},$$

where $y_j = 1$ or 0. Hence, if $|\sum_{j=1}^n \omega_{kj} y_j + v_k|/|1-\mu| > \varepsilon$, which is more likely to hold for smaller ε, there does not exist any fixed point of $\overline{\mathbf{F}}$ in this region. Our discussions above have demonstrated that almost every trajectory generated by $\mathbf{F}$ converges to a fixed point of $\overline{\mathbf{F}}$ in a saturated region, under conditions described in Section 3.1 and that $\varepsilon > 0$ is small enough.

4. Application to the TSP

We will make use of the properties derived in previous sections to choose suitable parameters for the TCNN to solving the TSP. The task of solving the TSP via neural networks is to search for the minimizer of an objective function through evolutions of the network system. The objective function proposed by Hopfield and Tank [1] consists of two parts which are both quadratic functions, namely,

$$E = E_1 + E_d, \tag{20}$$

where

$$E_1 = \frac{\gamma_1}{2}\left[\sum_{i=1}^n \left(\sum_{j=1}^n y_{ij} - 1\right)^2 + \sum_{j=1}^n \left(\sum_{i=1}^n y_{ij} - 1\right)^2\right],$$

$$E_d = \frac{\gamma_2}{2}\sum_{i=1}^n \sum_{k=1}^n \sum_{j=1}^n (y_{k(j+1)} + y_{k(j-1)}) y_{ij} d_{ik},$$

$y_{ij} \in [0,1], i,j = 1,\ldots,n$, is the probability for the ith city to be visited the jth time, and d_{ik} is the distance between city i and city k. We call $Y = [y_{ij}]_{n\times n}$ the output matrix. Note that E_1 attains its minimal value, i.e., zero, at a permutation matrix. E_d is related to the tour distance of the TSP. Therefore, $2E_d/\gamma_2$ gives the tour length when Y is a permutation matrix. In our setting of solving the TSP through the TCNN computations, the meaningful asymptotic output matrix should be a permutation matrix. Notably, if the global minimum of E is attained at a permutation matrix Y, then it is attained at all permutations to Y, which give the same value to E.

To apply the TCNN to the TSP, we change the setting of the TSP with two-dimensional indices into the one-dimensional form. Restated, by letting $s(i,j) = j + (i-1)n$, where n is the number of cities for the TSP, Eq. (20) becomes

$$E(\mathbf{y}) = -\frac{1}{2}\mathbf{y}\overline{W}\mathbf{y}^T - 2\gamma_1 I_{n^2\times n^2}\mathbf{y} + \gamma_1 n. \tag{21}$$

Herein, $I_{n^2\times n^2}$ is the identity matrix of size $n^2 \times n^2$, $\mathbf{y} = (y_1,\ldots,y_{s(i,j)},\ldots,y_{n^2})$ and

$$\overline{W} = -\gamma_1[I_{n\times n}\otimes 1_{n\times n} + 1_{n\times n}\otimes I_{n\times n}] - \gamma_2 D\otimes B; \tag{22}$$

$1_{n\times n}$ is the matrix whose entries are all one, $D = [d_{ij}]^T$ and $B = [b_{ij}]$ with $b_{i,j} = 0$ except that $b_{i,i+1} = b_{i,i-1} = b_{1,n} = b_{n,1} = 1$; the $n^2 \times n^2$ block matrix $A \otimes B$ is defined by the formula $[A\otimes B]_{ij} = [a_{ij}B]$, where $A = [a_{ij}]$ and $B = [b_{ij}]$. We consider the TCNN

$$x_i(t+1) = \mu x_i(t) + (1-\beta)^{q(t)}\omega[y_i(t) - a_{0i}] + \sum_{j=1}^{n^2} W_{ij}y_j(t) + 2\gamma_1, \tag{23}$$

where $W = [W_{ij}] := \overline{W} - \mathrm{diag}[\overline{W}]/2$, $i = 1,\ldots,n^2$. According to previous discussions, there is a Lyapunov function for Eq. (23):

$$V(t,\mathbf{y}) = -\frac{1}{2}\sum_{i=1}^{n^2}\sum_{j=1}^{n^2} W_{ij}y_iy_j - \sum_{i=1}^{n^2} 2\gamma_1 y_i - (\mu-1)\varepsilon\sum_{i=1}^{n^2}(y_i^2 - y_i) + c^t, \tag{24}$$

where $\mathbf{y} = (y_1,\ldots,y_{n^2})$. The conditions for chaotic and convergent dynamical phases of the TCNN are all computable. The range of the parameters satisfying these conditions can also be depicted numerically. We shall describe how we choose parameters in the numerical simulations for solving a TSP via the TCNN in the next section.

5. Numerical simulations

We present numerical simulations for solving the TSP via the evolutions of the TCNN (23). We employ the ten-city TSP problem by Hopfield and Tank to illustrate the performance. The ten-city data is illustrated in Fig. 7.

Let us describe how to choose parameters in the numerical simulations. We first take an μ with $0 < \mu < 1$ for boundedness of iterations for the TCNN. Set $\omega = 0$, and take $\varepsilon, a_{0i}, h_i, i = 1,\ldots,n^2$, so that the TCNN with these parameters is

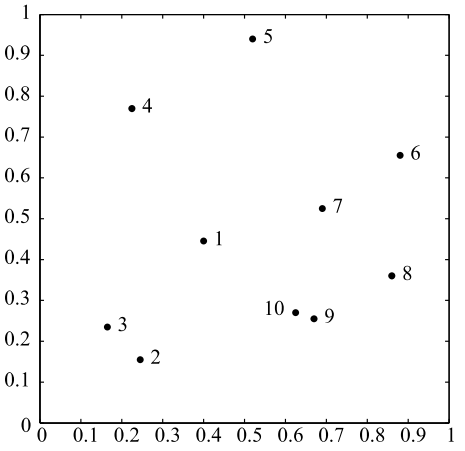

Fig. 7. The Hopfield-Tank original data.

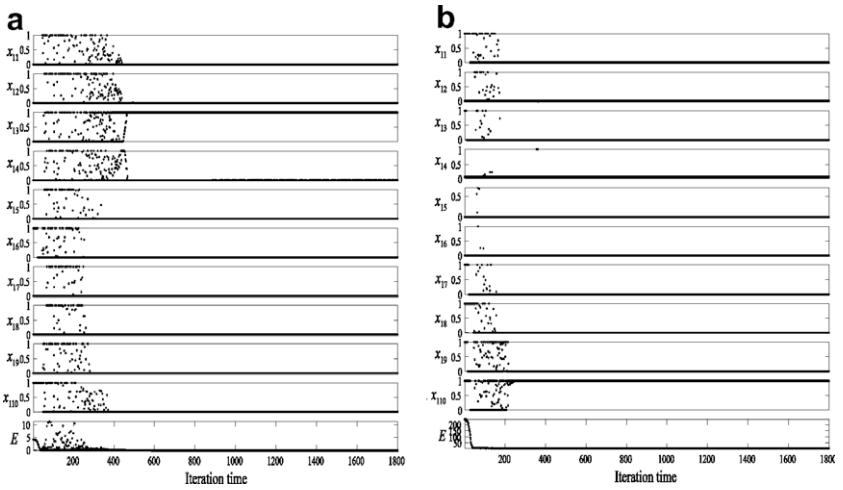

Fig. 8. The iterations of components $x_{1,1},\ldots,x_{1,10}$, and E for the TCNN in solving a 10-city TSP. The iteration is synchronous updating in (a) and cyclic updating in (b).

in convergent phase, where $h_i = \sum_{k=1}^{n^2}\{|W_{ik}| + 2|\gamma_1|\}$. We then let $|\omega|$ increase from 0 to see whether if parameters $(\mu, \varepsilon, \omega, a_{0i}, h_i)$ enter the chaotic regime. These computations can be assisted by a computer programming. In our illustration, the parameters are set as $\mu = 0.9$, $\beta = 0.005$, $\varepsilon = 0.01$, $a_{0i} = 0.65$, $\omega = -0.08$, $\gamma_1 = 0.015$, $\gamma_2 = 0.015$ and $q(t) = t$. The iterations are demonstrated in Fig. 8, with the synchronously updating mode, in accordance with our theory. We also employ the cyclic updating iterations in Fig. 8b. Our simulation indicates that the best route for this TSP is $2 \to 3 \to 1 \to 4 \to 5 \to 6 \to 7 \to 8 \to 9 \to 10 \to 2$. The other best route such as $8 \to 7 \to 6 \to 5 \to 4 \to 1 \to 3 \to 2 \to 10 \to 9 \to 8$ and $9 \to 10 \to 2 \to 3 \to 1 \to 4 \to 5 \to 6 \to 7 \to 8 \to 9$ also have the same value $E = 2.7257$. In fact, all of them represent the same loop. Under these parameters, 100% of 3000 initial conditions lead to the optimal route with $E = 2.7257$ in the evolutions of system (23).

6. Conclusion

In this presentation, we have analyzed the dynamics for the chaotic phase and the convergent phase for the TCNN with piecewise linear output function. The analysis provides concrete and computable parameter conditions for the respective dynamic phase. The investigation makes use of the structure of piecewise linear output function. The parameter conditions thus derived are much simpler than the ones for logistic output functions. We have also observed that the TCNN with piecewise linear output function has even better performance than with the logistic output function in the applications. For example, with suitably chosen parameters, the output component of the TCNN with piecewise linear output function tends to either one or zero, with the objective function decreasing, as the system evolves. If the output matrix does not tend to a permutation matrix, one can enlarge slightly the parameter γ_1 in Eq. (22). The investigation demonstrates a skillful employ of the Marotto's theorem for the n-dimensional TCNN map and is expected to contribute toward applications of the TCNN as an annealing machine in solving various combinatorial optimization problems.

Appendix A. Proof of Proposition 1

$\bar{\xi}$ is a fixed point of the single neuron map f_ς, $-h \leqslant \varsigma \leqslant h$ if and only if $\bar{\xi} = \mu\bar{\xi} + \omega[g_\varepsilon(\bar{\xi}) - a_0] + \varsigma$. To find a fixed point in Ω_{m} for the single neuron map f_ς, we seek for $\bar{\xi}$ with $-\varepsilon \leqslant \bar{\xi} \leqslant \varepsilon$ satisfying $[(1-\mu)\bar{\xi} - \varsigma]/\omega + a_0 = g_\varepsilon(\bar{\xi})$. Equivalently, we look for an intersection of equations $\eta = [(1-\mu)\xi - \varsigma]/\omega + a_0$ and $\eta = g_\varepsilon(\xi)$ in Ω_{m}. Notably, (P_1-i): $\mu < \min\{-1 - (\omega a_0 + h)/\varepsilon, -1 - [\omega(1-a_0) - h]/\varepsilon\}$ is equivalent to $f_{-h}(-\varepsilon) > \varepsilon$ and $f_h(\varepsilon) < -\varepsilon$. Thus, under condition (P_1-i), there exists a fixed point of f_ς in Ω_{m}, with reference of Fig. 2a. Let $\omega > 0$ and $L_\varsigma(x) := [(1-\mu)x - \varsigma]/\omega + a_0$. Then L_{-h} (resp. L_h) is the upper (resp. lower) line in Fig. 9. Note that $L_h(x) \leqslant L_\varsigma(x) \leqslant L_{-h}(x)$ with $-h \leqslant \varsigma \leqslant h$ for all $x \in \mathbb{R}$. If

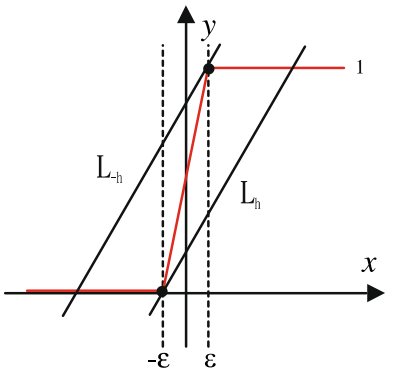

Fig. 9. The configuration explains condition (P_1-i)(a). The quantities $\frac{1-\mu}{\omega}\varepsilon+\frac{h}{\omega}+a_0$ and $\frac{1-\mu}{\omega}(-\varepsilon)-\frac{h}{\omega}+a_0$ correspond to the y-coordinates of two dotted points, respectively.

$L_{-h}(\varepsilon)\leqslant 1$ and $L_h(-\varepsilon)\geqslant 0$, then it is obvious that there exists a fixed point $\bar{\xi}\in\Omega_{\mathrm{m}}$ for L_ς with $-h\leqslant\varsigma\leqslant h$. Actually, under this circumstances, there exist three fixed points of f_ς for $-h\leqslant\varsigma\leqslant h$. □

Appendix B. Marotto's Theorem

Consider a dynamical system: $\mathbf{x}\mapsto\mathbf{F}(\mathbf{x})$, $\mathbf{x}\in\mathbb{R}^n$ and $\mathbf{F}$ is in $C^1(\mathbb{R}^n,\mathbb{R}^n)$ or piecewise C^1. Suppose that $\bar{\mathbf{x}}$ is a fixed point of $\mathbf{F}$ with all eigenvalues of $D\mathbf{F}(\bar{\mathbf{x}})$ exceeding 1 in magnitude, and suppose there exists a point $\mathbf{x}_0\neq\bar{\mathbf{x}}$ in a repelling neighborhood of $\bar{\mathbf{x}}$, such that $\mathbf{F}^m(\mathbf{x}_0)=\bar{\mathbf{x}}$ and $\det(D\mathbf{F}^m(\mathbf{x}_0))\neq 0$, for some $1<m\in\mathbb{N}$. Then $\bar{\mathbf{x}}$ is called a *snap-back repeller* [16] of $\mathbf{F}$, as depicted in Fig. 4. If $\mathbf{F}$ has a snap-back repeller, then the system of $\mathbf{F}$ is chaotic in the following sense: (i) There exists a positive integer m_0 such that $\mathbf{F}$ has p-periodic points for each integer $p\geqslant m_0$. (ii) There exists a scrambled set, that is, an uncountable set L containing no periodic points so that the following pertains: (a) $\mathbf{F}(L)\subset L$; (b) for every $\mathbf{y}\in L$ and any periodic point $\mathbf{x}$ of $\mathbf{F}$,

$$\limsup_{m\to\infty}\|\mathbf{F}^m(\mathbf{y})-\mathbf{F}^m(\mathbf{x})\|>0;$$

(c) for every $\mathbf{x}$, $\mathbf{y}\in L$ with $\mathbf{x}\neq\mathbf{y}$,

$$\limsup_{m\to\infty}\|\mathbf{F}^m(\mathbf{y})-\mathbf{F}^m(\mathbf{x})\|>0;$$

(iii) There exists an uncountable subset L_0 of L such that for every $\mathbf{x}$, $\mathbf{y}\in L_0$,

$$\liminf_{m\to\infty}\|\mathbf{F}^m(\mathbf{y})-\mathbf{F}^m(\mathbf{x})\|=0.$$

Marotto first reported this theorem in 1978. The theorem has been applied to justify chaotic behaviors for several dynamical systems [20]. The original definition of *snap-back repeller* employed the notion that there exists a real number $r>0$ and $\mathbf{x}_0\in B(\bar{\mathbf{x}};r)$ with $\mathbf{x}_0\neq\bar{\mathbf{x}}$, called *snap-back point*, such that all eigenvalues of $DF(\mathbf{x})$ exceed unity in norm for all $\mathbf{x}\in B(\bar{\mathbf{x}};r)$. This condition does not imply that $\mathbf{x}_0$ lies in an expanding neighborhood of $\bar{\mathbf{x}}$. Such a definition then leads to an insufficiency in the proof of the theorem. The definition of snap-back repeller is then revised recently [16]. There were other related discussions in [20,21].

Appendix C. Supplemental proof for Theorem 6

When no confusion arises, we will omit the variable t in Γ_i. We arrange the justification into three cases: $F_i(\mathbf{x})\geqslant\varepsilon$, $|F_i(\mathbf{x})|<\varepsilon$ and $F_i(\mathbf{x})\leqslant\varepsilon$.

Case I: $F_i(\mathbf{x})\geqslant\varepsilon$,

$$\begin{aligned}\Gamma_i&=F_i(\mathbf{x})\Delta_t y_i-\varepsilon\Delta_t y_i\cdot g_\varepsilon(x_i)=\Delta_t y_i[F_i(\mathbf{x})-\varepsilon g_\varepsilon(x_i)]=[1-g_\varepsilon(x_i)][F_i(\mathbf{x})-\varepsilon g_\varepsilon(x_i)]\geqslant[1-g_\varepsilon(x_i)][\varepsilon-\varepsilon g_\varepsilon(x_i)]\\&=\varepsilon[1-g_\varepsilon(x_i)]^2=\varepsilon(\Delta_t y_i)^2.\end{aligned}$$

Case II: $F_i(\mathbf{x}) \leqslant -\varepsilon$,

$$\Gamma_i = -F_i(\mathbf{x})g_\varepsilon(x_i) + \varepsilon g_\varepsilon(x_i)[g_\varepsilon(x_i) - 1] = g_\varepsilon(x_i)[-F_i(\mathbf{x}) - \varepsilon + \varepsilon g_\varepsilon(x_i)] \geqslant \varepsilon(g_\varepsilon(x_i))^2 = \varepsilon(\Delta_t y_i)^2.$$

Case III: $|F_i(\mathbf{x})| < \varepsilon$,

$$\Gamma_i = F_i(\mathbf{x})\Delta_t y_i - \varepsilon\Delta_t y_i\left[\frac{1}{2\varepsilon}F_i(\mathbf{x}) + \frac{1}{2} + g_\varepsilon(x_i) - 1\right] = \frac{1}{2}F_i(\mathbf{x})\Delta_t y_i + \frac{\varepsilon}{2}\Delta_t y_i - \varepsilon\Delta_t y_i \cdot g_\varepsilon(x_i)$$

$$= \frac{1}{2}[F_i(\mathbf{x}) + \varepsilon]\Delta_t y_i - \varepsilon\Delta_t y_i \cdot g_\varepsilon(x_i) = \left[\frac{1}{2\varepsilon}F_i(\mathbf{x}) + \frac{1}{2} - g_\varepsilon(x_i)\right]\left[\frac{1}{2}(F_i(\mathbf{x}) + \varepsilon) - \varepsilon g_\varepsilon(x_i)\right] =: *.$$

For case III, we consider three subcases: $x_i \geqslant \varepsilon$, $|x_i| < \varepsilon$ and $x_i \leqslant \varepsilon$.

Subcase III(a): $x_i \geqslant \varepsilon$,

$$* = \left[\frac{1}{2\varepsilon}F_i(\mathbf{x}) + \frac{1}{2} - 1\right]\left[\frac{1}{2}(F_i(\mathbf{x}) + \varepsilon) - \varepsilon\right] = \frac{1}{4\varepsilon}[F_i(\mathbf{x}) - \varepsilon]^2 = \varepsilon(\Delta_t y_i)^2.$$

Subcase III(b): $x_i \leqslant -\varepsilon$,

$$* = \left[\frac{1}{2\varepsilon}F_i(\mathbf{x}) + \frac{1}{2}\right]\left[\frac{1}{2}(F_i(\mathbf{x}) + \varepsilon)\right] = \frac{1}{4\varepsilon}[F_i(\mathbf{x}) + \varepsilon]^2 = \varepsilon(\Delta_t y_i)^2.$$

Subcase III(c): $|x_i| < \varepsilon$,

$$* = \left[\frac{1}{2\varepsilon}F_i(\mathbf{x}) + \frac{1}{2} - \frac{1}{2\varepsilon}x_i - \frac{1}{2}\right]\left[\frac{1}{2}(F_i(\mathbf{x}) + \varepsilon) - \varepsilon\left(\frac{1}{2\varepsilon}x_i + \frac{1}{2}\right)\right] = \frac{1}{4\varepsilon}[F_i(\mathbf{x}) - x_i]^2 = \varepsilon(\Delta_t y_i)^2.$$

References

[1] Hopfield JJ, Tank DW. Neural computation of decisions in optimization problems. Biol Cybernet 1985;52(3):141–52.
[2] Chen L, Aihara K. Chaotic simulated annealing by neural network model with transient chaos. Neural Networks 1995;8(6):915–30.
[3] Chen L, Aihara K. Chaos and asymptotical stability in discrete-time neural networks. Phys D 1997;104:286–325.
[4] Chen L, Aihara K. Global searching ability of chaotic neural networks. IEEE Trans Circuits Syst I Fund Theory Appl 1999;46(8):974–93.
[5] Kwok T, Smith KA. A unified framework for chaotic neural network approaches to combinatorial optimization. IEEE Trans Neural Networks 1999;10:978–81.
[6] Kwok T, Smith KA. Experimental analysis of chaotic neural network models for combinatorial optimization under a unifying framework. Neural Networks 2000;13:731–44.
[7] Tokuda I, Nagashima T, Aihara K. Global bifurcation structure of chaotic neural networks and its application to traveling salesman problems. Neural Networks 1997;10:1673–90.
[8] Wang L, Smith K. On chaotic simulated annealing. IEEE Trans Neural Networks 1998;9:716–8.
[9] Blondel VD, Bournez O, Koiran P, Tsitsiklis N. The stability of saturated linear dynamical systems is undecidable. J Comput Syst Sci 2001;62:442–62.
[10] Chua LO, Yang L. Cellular neural networks: theory. IEEE Trans Circuits Syst 1988;35:1257–72.
[11] Siegelmann HT, Sontage ED. Turing computability with neural nets. Appl Math Lett 1991;4:77–80.
[12] Siegelmann HT, Sontage ED. On the computational power of neural nets. J Comput Syst Sci 1995;50:132–50.
[13] Siegelmann HT, Horne BG, Lee Giles C. Computational capabilities of recurrent NARX neural networks. IEEE Trans Syst Man Cybernet B 1997;27:208–15.
[14] Chen SS, Shih CW. Transversal homoclinic orbits in a transiently chaotic neural network. Chaos 2002;12(3):654–71.
[15] Marotto FR. Snap-back repellers imply chaos in R^n. J Math Anal Appl 1978;63:199–223.
[16] Marotto FR. On redefining a snap-back repeller. Chaos, Solitons & Fractals 2005;25:25–8.
[17] Meyer KR, Zhang X. Stability of skew dynamical systems. J Differen Equat 1996;132:66–86.
[18] LaSalle JP. The stability of dynamical systems, ser. Regional conference series in applied mathematics, vol. 25. Philadelphia, PA: SIAM; 1976. p. 45–55.
[19] Chen SS, Shih CW. Asymptotic behaviors in a transiently chaotic neural network. Discr Contin Dyn Syst 2004;10(3):805–26.
[20] Chen G, Hsu SB, Zhou J. Snapback repellers as a cause of chaotic vibration of the wave equation with a van der Pol boundary condition and energy injection at the middle of the span. J Math Phys 1998;39(12):6459–89.
[21] Li CP, Chen G. An improved version of the Marotto theorem. Chaos, Solitons & Fractals 2003;18:69–77.