Plasma Phys. Control. Fusion **40** (1998) 295–318. Printed in the UK PII: S0741-3335(98)83179-5

Stabilization of burn conditions in a thermonuclear reactor using artificial neural networks

Javier E Vitela and Julio J Martinell

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, 04510, México, DF

Received 8 April 1997, in final form 28 August 1997

Abstract. In this work we develop an artificial neural network (ANN) for the feedback stabilization of a thermonuclear reactor at nearly ignited burn conditions. A volume-averaged zero-dimensional nonlinear model is used to represent the time evolution of the electron density, the relative density of alpha particles and the temperature of the plasma, where a particular scaling law for the energy confinement time previously used by other authors, was adopted. The control actions include the concurrent modulation of the D–T refuelling rate, the injection of a neutral He-4 beam and an auxiliary heating power modulation, which are constrained to take values within a maximum and minimum levels. For this purpose a feedforward multilayer artificial neural network with sigmoidal activation function is trained using a back-propagation through-time technique. Numerical examples are used to illustrate the behaviour of the resulting ANN–dynamical system configuration. It is concluded that the resulting ANN can successfully stabilize the nonlinear model of the thermonuclear reactor at nearly ignited conditions for temperature and density departures significantly far from their nominal operating values. The NN–dynamical system configuration is shown to be robust with respect to the thermalization time of the alpha particles for perturbations within the region used to train the NN.

1. Introduction

It is expected that future practical thermonuclear reactors will be of tokamak type operating under ignited or near-ignited conditions. In the first case, the energy provided by the alpha particles produced by the D–T fusion will suffice to keep a steady thermonuclear reaction level for long periods of time, assuming an appropriate refuelling rate is provided. Near-ignition states on the other hand, require some kind of external heating which must be small compared to the alpha particle heating in order to be of interest. However, nominal operating points of most ignited or nearly ignited tokamak reactors will be inherently unstable if they are designed to operate in the low-temperature region, i.e. any perturbation will drive the system away from the operating point; this is due to the fact that in this region the energy produced by fusion dominates over the energy losses when the temperature increases and *vice versa*. It is desired to control these thermal instabilities not only to keep the plasma under appropriate burn conditions for long periods of time in order to effectively burn the D–T plasma fuel, but also to prevent wall damage in the reactor chamber due to plasma disruptions.

Several studies have been made in the past in this regard using volume-averaged zero-dimensional particle and energy balance equations [1–8]. In particular Haney *et al* [7], suggested that ITER (International Thermonuclear Experimental Reactor) should operate in the low-temperature region and studied the feasibility of the active control of burn

0741-3335/98/020295+24$19.50

conditions for a tentative model of ITER by analysing the individual effects of modulated neutral beam heating, refuelling rate and the controlled injection of high-Z impurities. More recently Hui *et al* [8], using the characteristics of the tokamak reactor proposed by the ITER-CDA (conceptual design activity) group, studied the design and performance of a refuelling-rate-based robust control for the burn conditions of a fusion reactor using the H_∞ algorithm [9]. The performance of the constant-gain proportional feedback controller synthesized by these means was evaluated under several model uncertainties such as helium ash confinement time and controller time delay. The refuelling-rate-based controller proved to be able to control both density and temperature excursions. However, because of the constant-gain characteristic, this refuelling rate control law yields extremely large positive or negative values, except for very small deviations from the nominal operating point. This may constitute a major drawback in the practical implementation of such a controller device, since real control actions are always constrained to lie between some maximum and minimum levels.

In spite of the many advances in nonlinear control techniques, the great diversity of nonlinear systems has prevented the development of a systematic and generally applicable theory for nonlinear control design [10, 11]. In the last decade, artificial neural networks (NN) have received considerable attention by the dynamical systems community for the identification and control of nonlinear systems due to their ability to approximate arbitrary nonlinear mappings [12]. Hence, they provide us with the capability of dealing with nonlinear systems and with the possibility of nonlinear control designs that cannot be synthesized with traditional control techniques [13].

The purpose of this paper is to demonstrate the capabilities of neural networks, by means of a multilayer feedforward NN, for the stabilization of a thermonuclear tokamak reactor operating in the low-temperature region at nearly ignited burn conditions. The control actions considered here include the simultaneous modulation of the D–T refuelling rate, the injection of a neutral He-4 beam and an auxiliary heating power, provided for instance by appropriate radiofrequency electromagnetic waves. All these techniques have been previously explored by different authors, but to our knowledge, neither the use of neural networks for this purpose, nor the concurrent action of the three methods with imposed maximum and minimum limits that any feasible control action is constrained to take, have been considered.

The volume-averaged zero-dimensional model adopted in this work is similar to that used by Hui *et al* [8], in which for completeness purposes, we have also included the effect of ohmic heating. In this model, ions and electrons are assumed to have Maxwellian distributions and share the same temperature at all times. The plasma is assumed to be solely composed of D–T in equal proportions, n_{DT}; fully ionized helium ions n_α, also called helium ash, and electrons; with the electron density determined by the quasineutrality condition, $n_{\mathrm{e}} = n_{\mathrm{DT}} + 2n_\alpha$. No high-$Z$ impurities are considered. The model assumes instantaneous thermalization of the alpha particles produced by the fusion reactions, a common assumption made by most authors in previous works but the effect of a finite thermalization time is also studied. In addition, both energy and particle transport are taken into account through a scaling law for the energy confinement time τ_{E}, a helium ash confinement time τ_α and a D–T fuel confinement time τ_P. Assuming that n_{DT} is kept constant, Taylor *et al* [14] showed that for a D–T plasma no steady state burn conditions can be achieved when $\tau_\alpha > 15\tau_{\mathrm{E}}$; and that the presence of high-Z, impurities lowers this limit. Here we will assume $\tau_\alpha = 7\tau_{\mathrm{E}}$ and $\tau_P = 3\tau_{\mathrm{E}}$ [8]. Bremsstrahlung radiation losses are included as an energy loss mechanism, while synchrotron radiation is neglected since, for a wide range of reactor parameters, it becomes comparable to the former only above 30 keV [14].

The rest of the paper is organized as follows: in section 2, the zero-dimensional dynamical equations are stated, a nominal operating steady state is determined and, a normalized set of dynamical equations is derived. Section 3 is concerned with the NN-system configuration and the algorithm used to train the network. Section 4 discusses the results of the training process and section 5 presents some illustrative simulation results of the NN–thermonuclear system behaviour for some particular state perturbations, showing the capabilities of the NN to deal with the nonlinear aspects of the thermonuclear systems. The effect of the thermalization time of the alpha particles in the stability of the NN–dynamical system configuration is studied in section 6, where several constant delay times, together with a density and temperature dependent thermalization time, are considered. Finally, section 7 contains some conclusions concerning this work and further extensions and research suggested in this direction.

2. The dynamical system model

As described above, the zero-dimensional model of the ignited tokamak reactor adopted in this work is similar to the one used by Hui *et al* [8] with the inclusion of ohmic heating and the concurrent modulation of the refuelling rate, the neutral He-4 injection, and the auxiliary heating power as control variables in the dynamical equations. Thus, with these considerations we have the following coupled set of evolution equations

$$\frac{\mathrm{d}}{\mathrm{d}t}n_{\mathrm{DT}} = S_{\mathrm{f}} - 2\left(\frac{n_{\mathrm{DT}}}{2}\right)^2\langle\sigma v\rangle - \frac{n_{\mathrm{DT}}}{\tau_P} \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}n_{\alpha} = S_{\alpha} + \left(\frac{n_{\mathrm{DT}}}{2}\right)^2\langle\sigma v\rangle - \frac{n_{\alpha}}{\tau_{\alpha}} \tag{2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{3}{2}(n_{\mathrm{e}} + n_{\mathrm{DT}} + n_{\alpha})T\right] = P_{\mathrm{aux}} + Q_{\alpha}\left(\frac{n_{\mathrm{DT}}}{2}\right)^2\langle\sigma v\rangle + \eta j^2 - A_{\mathrm{b}}Z_{\mathrm{eff}}^2 n_{\mathrm{e}}^2 T^{1/2} - \frac{3}{2}(n_{\mathrm{e}} + n_{\mathrm{DT}} + n_{\alpha})\frac{T}{\tau_{\mathrm{E}}} \tag{3}$$

which corresponds to the balance equations for the densities of D–T, the helium ash and the thermal energy, respectively. In these equations, S_{f} represent the refuelling rate, S_{α} the neutral He-4 injection rate, and P_{aux} the auxiliary heating power density; j is the average plasma current density, $Q_{\alpha} = 3.5$ Mev is the energy carried by the fusion alpha particles, $\langle\sigma v\rangle$ is the D–T reactivity [15, 16], A_{b} is the coefficient associated with the bremsstrahlung radiation losses [15] and η is the neoclassical parallel resistivity [17]. Finally τ_{E}, τ_P and τ_{α} represent the energy, the D–T fuel and the helium ash confinement times, respectively.

Using the quasineutrality condition $n_{\mathrm{e}} = n_{\mathrm{DT}} + 2n_{\alpha}$, the above set of equations can be transformed into the equation

$$\frac{\mathrm{d}n_{\mathrm{e}}}{\mathrm{d}t} = S_{\mathrm{f}} - \left(\frac{2f_{\alpha}}{\tau_{\alpha}} + \frac{1-2f_{\alpha}}{\tau_P}\right)n_{\mathrm{e}} + 2S_{\alpha} \tag{4}$$

for the electron density, the equation

$$\frac{\mathrm{d}f_{\alpha}}{\mathrm{d}t} = \frac{1}{4}n_{\mathrm{e}}(1-2f_{\alpha})^2\langle\sigma v\rangle - \frac{f_{\alpha}}{n_{\mathrm{e}}}S_{\mathrm{f}} + f_{\alpha}(1-2f_{\alpha})\left(\frac{1}{\tau_P} - \frac{1}{\tau_{\alpha}}\right) + S_{\alpha}\frac{1}{n_{\mathrm{e}}}(1-2f_{\alpha}) \tag{5}$$

for the relative fraction of helium ions defined by $f_\alpha = n_\alpha/n_\mathrm{e}$, and the following equation for the plasma temperature

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{2}{3}\frac{P_\mathrm{aux}}{n_\mathrm{e}(2-f_\alpha)} + \left(\frac{1}{6}Q_\alpha + \frac{1}{4}T\right) n_\mathrm{e}\langle\sigma v\rangle \frac{(1-2f_\alpha)^2}{2-f_\alpha} + \frac{2(1-2f_\alpha)}{2-f_\alpha}\frac{T}{\tau_P}$$
$$-\frac{2}{3}A_\mathrm{b}\frac{(1+2f_\alpha)}{2-f_\alpha}n_\mathrm{e}T^{1/2} - \frac{T}{\tau_\mathrm{E}} + \frac{3f_\alpha}{2-f_\alpha}\frac{T}{\tau_\alpha} + \frac{2}{3}A_\mathrm{h}\frac{(1+2f_\alpha)^{0.5}}{n_\mathrm{e}(2-f_\alpha)T^{3/2}}$$
$$\times\frac{1+1.198(1+2f_\alpha)^{0.5}+0.222(1+2f_\alpha)}{1+2.966(1+2f_\alpha)^{0.5}+0.75(1+2f_\alpha)} - \frac{2T}{n_\mathrm{e}(2-f_\alpha)}S_\mathrm{f} - \frac{3T}{n_\mathrm{e}(2-f_\alpha)}S_\alpha \tag{6}$$

in which the dependence of the neoclassical parallel resistivity η on the particle densities and temperature is explicitly written [17], and where the following expression for the D–T reactivity $\langle\sigma v\rangle$ is used [16],

$$\langle\sigma v\rangle = 10^{-2}\exp\left(\frac{a_1}{T^{0.2935}} + a_2 + a_3T + a_4T^2 + a_5T^3 + a_6T^4\right)\ \mathrm{m}^3\ \mathrm{s}^{-1} \tag{7}$$

where the constants $a_1, \ldots, a_6$ are given in table 1.

Table 1. Parameters not specified in the text regarding $\langle\sigma v\rangle$ and the dynamical equations [16].

a_1	$-21.377\,692\ \mathrm{keV}^{0.2935}$
a_2	$-25.204\,054$
a_3	$-7.101\,3427\times10^{-2}\ \mathrm{keV}^{-1}$
a_4	$1.937\,5451\times10^{-4}\ \mathrm{keV}^{-2}$
a_5	$4.924\,6592\times10^{-6}\ \mathrm{keV}^{-3}$
a_6	$-3.983\,6572\times10^{-8}\ \mathrm{keV}^{-4}$
A_b	$3.325\,071\times10^{-21}\ \mathrm{keV}^{1/2}\ \mathrm{m}^3\ \mathrm{s}^{-1}$
A_h	$3.401\,153\times10^{20}\ \mathrm{keV}^{5/2}\ \mathrm{m}^{-3}\ \mathrm{s}^{-1}$
A_e	$3.181\,82\times10^{10}\ \mathrm{keV}^{0.47}\ \mathrm{m}^{-1.41}\ \mathrm{s}^{0.53}$

Haney *et al* [7] used the following offset-linear scaling for the L-mode energy confinement time of ITER

$$\tau_\mathrm{E}(L) = 0.064A_\mathrm{i}^{0.2}I_P^{0.8}R^{1.6}a^{0.6}\kappa_\mathrm{s}^{0.5}\tilde{n}_\mathrm{e}^{0.6}B_0^{0.35}/P_\mathrm{net} + 0.04A_\mathrm{i}^{0.5}I_P^{0.5}R^{0.3}a^{0.8}\kappa_\mathrm{s}^{0.6} \tag{8}$$

and assumed an H-mode operation by taking $\tau_\mathrm{E} = H\ \tau_\mathrm{E}(L)$, where H is the enhancement factor. However, here we adopt the same scaling law used by Hui *et al* [8]

$$\tau_\mathrm{E} = 0.082I_P^{1.02}R^{1.6}B_0^{0.15}A_\mathrm{i}^{0.5}\kappa_x^{-0.19}P_\mathrm{net}^{-0.47} \tag{9}$$

where P_net represents the net plasma heating given by $V_\mathrm{core}\ (P_\mathrm{aux} + P_\alpha + P_\mathrm{oh} - P_\mathrm{br})$; with V_core the volume of the plasma and P_aux, P_α, P_oh and P_br are the auxiliary, alpha-particle heating, ohmic heating and the bremsstrahlung energy radiation densities, respectively. The reactor operating parameters are also assumed to be similar to those used by Hui *et al* [8] and a small list of relevant parameters is shown in table 2. For the parameters specified in this table, the ignited steady-state condition, i.e. where $P_\mathrm{aux} = 0$ and $S_\alpha = 0$, is obtained for $Sf_0 = 4.16\times10^{18}\ \mathrm{m}^{-3}\ \mathrm{s}^{-1}$ when $n_0 = 9.8\times10^{19}\ \mathrm{m}^{-3}$, $T_0 = 8.28$ keV and $f_0 = 0.0624$, values that will be assumed to constitute the desired operating point for the ignited tokamak reactor of this work.

For reasons that will become clear in the next sections, we define the following state variables normalized with respect to the ideal ignited operating point

$$z_1 = n_\mathrm{e}/n_0 \qquad z_2 = f_\alpha/f_0 \qquad \text{and} \qquad z_3 = T/T_0. \tag{10}$$

Table 2. Some relevant parameters of the ITER-CDA tokamak model used in the numerical simulations.

Electron density	$n_e = 9.8 \times 10^{19}$ m^{-3}
Alpha-particle fraction	$f_\alpha \equiv n_\alpha / n_e = 0.0624$
Temperature	$T = 8.28$ keV
50:50 D–T refuelling rate	$S = 4.16 \times 10^{18}$ m^{-3} s^{-1}
Plasma current	$I = 21.58$ MA
Magnetic field	$B_0 = 4.85$ T
Major radius	$R = 6$ m
Minor radius	$a = 2.15$ m
Elongation at x	$\kappa_x = 2.2$
Isotopic number	$A_i = 1.5$
Volume of plasma	$V_{\rm core} = 892$ m^3
Cross sectional area	$A_{\rm sec} = 23.7$ m^2

Using these normalized variables the dynamical equations (4)–(6) can now be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}z_1 = \hat{S}_{\rm f} - \frac{1}{\tau_{\rm E}}\left(\frac{1}{3} - \frac{8}{21}f_0 z_2\right) z_1 + 2 f_0 \hat{S}_\alpha \tag{11}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}z_2 = {} & 2.5 \times 10^{-7}\frac{n_0}{f_0} z_1 (1 - 2 f_0 z_2)^2 \\ & \times \exp\left(\frac{a_1}{T_0^{0.2935} z_3^{0.2935}} + a_2 + a_3 T_0 z_3 + a_4 T_0^2 z_3^2 + a_5 T_0^3 z_3^3 + a_6 T_0^4 z_3^4\right) \\ & - \frac{1}{z_1} z_2 \hat{S}_{\rm f} + \frac{4}{21} z_2 (1 - 2 f_0 z_2)\frac{1}{\tau_{\rm E}} + \hat{S}_\alpha \frac{1}{z_1}(1 - 2 f_0 z_2)\end{aligned} \tag{12}$$

and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}z_3 = {} & \frac{\hat{S}_{\rm aux}}{z_1(2 - f_0 z_2)} - \frac{2 z_3}{z_1(2 - f_0 z_2)}\hat{S}_{\rm f} - \frac{3 z_3 f_0}{z_1(2 - f_0 z_2)}\hat{S}_\alpha - \frac{1}{21}\left(2 + \frac{24}{2 - f_0 z_2}\right)\frac{z_3}{\tau_{\rm E}} \\ & + \left(\frac{1}{6}\frac{Q_\alpha}{T_0} + \frac{1}{4} z_3\right) \times 10^{-6} \times n_0 z_1 \frac{(1 - 2 f_0 z_2)^2}{2 - f_0 z_2} \\ & \times \exp\left(\frac{a_1}{T_0^{0.2935} z_3^{0.2935}} + a_2 + a_3 T_0 z_3 + a_4 T_0^2 z_3^2 + a_5 T_0^3 z_3^3 + a_6 T_0^4 z_3^4\right) \\ & - \frac{2}{3} A_{\rm b} \frac{n_0}{T_0^{1/2}} \frac{1 + 2 f_0 z_2}{2 - f_0 z_2} z_1 z_3^{1/2} + \frac{2}{3}\frac{A_{\rm h}}{n_0 T_0^{5/2}}\frac{1}{z_3^{3/2}} \\ & \times \frac{(1 + 2 f_0 z_2)^{0.5}}{z_1(2 - f_0 z_2)} \frac{1 + 1.198(1 + 2 f_0 z_2)^{0.5} + 0.222(1 + 2 f_0 z_2)}{1 + 2.966(1 + 2 f_0 z_2)^{0.5} + 0.75(1 + 2 f_0 z_2)}\end{aligned} \tag{13}$$

where we have assumed $\tau_P = 3\tau_{\rm E}$ and $\tau_\alpha = 7\tau_{\rm E}$. The energy confinement time $\tau_{\rm E}$, is given explicitly by the following expression

$$\begin{aligned}\tau_{\rm E} = A_{\rm e}\Bigg[& \frac{3}{2} n_0 T_0 \hat{S}_{\rm aux} + A_{\rm h}\frac{(1 + 2 f_0 z_2)^{0.5}}{T_0^{3/2}} \frac{1 + 1.198(1 + 2 f_0 z_2)^{0.5} + 0.222(1 + 2 f_0 z_2)}{1 + 2.966(1 + 2 f_0 z_2)^{0.5} + 0.75(1 + 2 f_0 z_2)} \frac{1}{z_3^{3/2}} \\ & + A_{\rm a} n_0^2 z_1^2 (1 - 2 f_0 z_2)^2 \\ & \times \exp\left(\frac{a_1}{T_0^{0.2935} z_3^{0.2935}} + a_2 + a_3 T_0 z_3 + a_4 T_0^2 z_3^2 + a_5 T_0^3 z_3^3 + a_6 T_0^4 z_3^4\right) \\ & - A_{\rm b} n_0^2 z_1^2 (1 + 2 f_0 z_2) T_0^{1/2} z_3^{1/2}\Bigg]^{-0.47}\end{aligned} \tag{14}$$

where the coefficient A_e, takes into account all the parameters in equation (9) that do not depend on P_{net}. Table 1 contains the numerical values of parameters in the above equations not specified in the text. In (11)–(13) we have used the following definitions:

$$\hat{S}_f \equiv S_f/n_0 \tag{15}$$

for the normalized refuelling rate,

$$\hat{S}_\alpha \equiv S_\alpha/(f_0 n_0) \tag{16}$$

for the normalized source of neutral helium atoms, and

$$\hat{S}_{aux} \equiv P_{aux}/(\tfrac{3}{2} n_0 T_0) \tag{17}$$

for the normalized auxiliary heating power density.

In order to evaluate the effect of the different control variables on the temporal behaviour of the state variables, we take partial derivatives of the right-hand side of (11)–(13) with respect to the different control variables, (15)–(17). Since both state and control variables are normalized, these quantities can be used as a measure of the relative importance of the control variables. Thus, with $\dot{z}_j \equiv \mathrm{d}z_j/\mathrm{d}t$, we have

$$\frac{\partial \dot{z}_1}{\partial \hat{S}_f} = 1 \tag{18}$$

$$\frac{\partial \dot{z}_2}{\partial \hat{S}_f} = z_2/z_1 \tag{19}$$

$$\frac{\partial \dot{z}_3}{\partial \hat{S}_f} = -z_3/z_1 \tag{20}$$

for the relative importance of $\hat{S}_f$;

$$\frac{\partial \dot{z}_1}{\partial \hat{S}_\alpha} = 2 f_0 \tag{21}$$

$$\frac{\partial \dot{z}_2}{\partial \hat{S}_\alpha} = (1 - 2 f_0 z_2)/z_1 \tag{22}$$

$$\frac{\partial \dot{z}_3}{\partial \hat{S}_\alpha} = 3 f_0 z_3/z_1 (2 - f_0 z_2) \tag{23}$$

for the importance of $\hat{S}_\alpha$; and finally

$$\frac{\partial \dot{z}_1}{\partial \hat{S}_{aux}} = \frac{1}{\tau_E^2}\left(\frac{1}{3} - \frac{8}{21} f_0 z_2\right) z_1 \frac{\partial \tau_E}{\partial \hat{S}_{aux}} \tag{24}$$

$$\frac{\partial \dot{z}_2}{\partial \hat{S}_{aux}} = -\frac{4}{21} z_2 (1 - 2 f_0 z_2) \frac{1}{\tau_E^2} \frac{\partial \tau_E}{\partial \hat{S}_{aux}} \tag{25}$$

and

$$\frac{\partial \dot{z}_3}{\partial \hat{S}_{aux}} = \frac{1}{2 - f_0 z_2} + \frac{1}{21} z_3 \left(2 - \frac{24}{2 - f_0 z_2}\right) \frac{1}{\tau_E^2} \frac{\partial \tau_E}{\partial \hat{S}_{aux}} \tag{26}$$

for the relative importance of the normalized auxiliary heating. It is apparent that the importance of the different control variables is a function of the current state of the system. Nevertheless, for small perturbations around the nominal operating point, we can approximate their values by those at the steady state. Thus, in this region: the importance of the refuelling rate modulation on the z_1, z_2 and z_3 is $+1.0$, -1.0 and -1.03, respectively; the importance of the injection rate modulation of neutral helium is approximately, $+0.12$,

$+0.88$, and -0.1, respectively; finally, the importance of the auxiliary power heating on the corresponding state variables is -0.08, $+0.04$, and $+0.35$, respectively. Hence, we can conclude that around the nominal operating point, the refuelling rate is highly relevant for all three state variables; while auxiliary power heating modulation and helium injection are significant almost exclusively for the plasma temperature and the helium fraction, respectively.

In the next section we describe the feedforward NN proposed to stabilize the thermonuclear system at a nominal operating point close to the ideal ignited state (n_0, f_0, T_0), review some of its characteristics and discuss the algorithm implemented to train the neural network for this purpose.

3. Artificial neural network and the training algorithm

The purpose of the NN is to provide the sequence of control actions that drives the system from an arbitrarily selected initial state z_1 within a given region in phase space, toward a target state z_t, without violating the constraints imposed on the state and/or the control variables. The type of dynamical systems we are concerned with here are time-discrete dynamical systems of the form $z_{k+1} = F(z_k, u_k)$, where z_k and u_k represent the state and control variables respectively, at time step k ($k = 1, 2, \ldots$). It is desired to stabilize the dynamical system, through the use of a feedforward NN to provide a control law of the form $u = u(z)$, such that any initial perturbation in the state variables is suppressed by the neural network, returning to a nominal operating point in the vicinity of the target state specified by $z_T = (1, 1, 1)$. The joint NN–dynamical system configuration is illustrated in figure 1, where the output of the neural network are the normalized control variables u as will be explained later in this section.

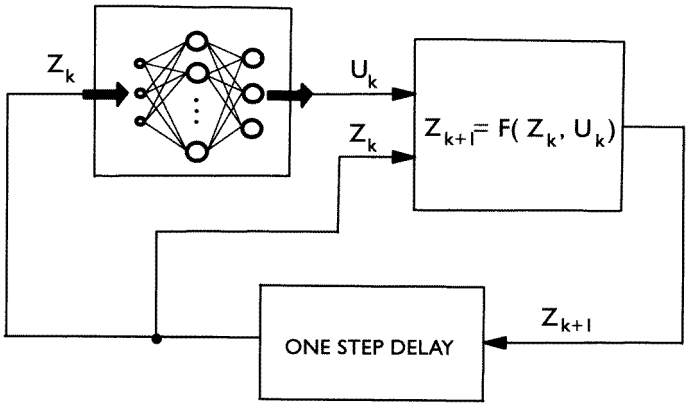

Figure 1. Artificial neural network–dynamical system configuration used for the feedback stabilization of the thermonuclear reactor model.

Feedforward neural networks can be considered as nonlinear continuous multivariate mappings. In general these NNs are composed of nonlinear computational elements or nodes arranged in L layers with $L \geqslant 3$. (See figure 2.) The output of the network is obtained by processing the values of a multidimensional input variable from the first layer or input to the Lth layer or output, passing through the $L-2$ remaining hidden layers. The

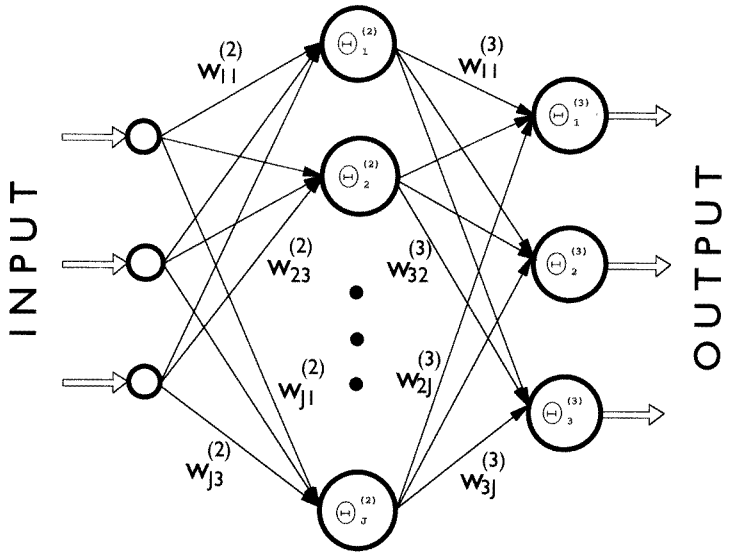

Figure 2. Three layered feedforward artificial neural network.

name *hidden* is used because such layers do not communicate directly with the external environment. The nodes in the first and last layers of the network are fixed by the nature of the problem itself; however the number of hidden layers and the number of nodes in each of these layers are not known *a priori*, and in general are determined after some trial and error process. The number of nodes in each layer will be denoted here by $J(l)$ with $l = 1, \ldots, L$. Each one of the $J(1)$ nodes in the first layer are mapped through the identity function, i.e. these nodes receive as input one of the $J(1)$ components of the input vector and transmit them unaltered to serve as the input values to the nodes of the second layer. Communication exists only between neurons in adjacent layers and is implemented through unidirectional feedforward connections known as weights. The nodes in the hidden and output layers are constituted of nonlinear functions, mapping a multidimensional input received from the nodes of the immediately preceding layer into a one-dimensional output. In this work each element of the input vector corresponds to the normalized values of the electron density, the relative fraction of helium ash and the plasma temperature at a given time step. The elements of the output vector on the other hand, are associated with the normalized D–T refuelling rate, neutral helium injection rate, and auxiliary power heating, as will be further discussed later on in this section. The activation of the jth node in the lth layer, $\mathcal{O}_j^{(l)}$, is given by

$$\mathcal{O}_j^{(l)} = f_j(\text{input}_j^{(l)}) \tag{27}$$

where $\text{input}_j^{(l)}$, the effective input to node j, is the weighted sum of the outputs $\mathcal{O}_i^{(l-1)}$ of the immediately preceding layer, i.e.

$$\text{input}_j^{(l)} = \sum_{i=1}^{J(l-1)} \omega_{ji}^{(l)} \mathcal{O}_i^{(l-1)} + \theta_j^{(l)} \tag{28}$$

where $\omega_{ji}^{(l)}$ denotes the weight that connects the output of node i in layer $l-1$ with the node j in layer l; $\theta_j^{(l)}$ and f_j correspond to the threshold value and the nonlinear activation

function of this node, respectively. The most widely used activation function in feedforward NN is the sigmoid function [18],

$$f = \frac{1}{1 + \mathrm{e}^{-\mathrm{input}_j^{(l)}}}. \tag{29}$$

Since the weights and thresholds connecting the nodes can take any real value, the output of this activation function is bounded between 0 and 1. For short, in general, we will refer as weights $\boldsymbol{w}$ to both, the parameters $\omega_{ij}^{(l)}$ and the thresholds $\theta_j^{(l)}$.

In order to stabilize the system around a given state, the neural network must provide appropriate values for the control variables, according to the current state of the system. In contrast with the training process of NNs for pattern recognition tasks [18–20], in which a set of input–output teaching patterns must be provided, the correct control values are not known during the training process: only the deviations of the state variables of the reactor with respect to their target values are available. By considering the NN–dynamical system joint configuration as a single unit, a set of input–output teaching patterns for training can be generated. The error between the target state $\boldsymbol{z}_{\mathrm{t}}$ and the actual final state $\boldsymbol{z}_{pk_{\mathrm{f}}}$ reached by the phase-space trajectory generated by the joint NN–dynamical system configuration for a given initial condition p, will be denoted by $\mathcal{E}_p$. For a particular set of weights specifying the neural network, and a subset of P initial conditions that the system can take, the total error $\mathcal{E}$ is defined then as the sum of these individual errors after the generation of the set of P trajectories, i.e.

$$\mathcal{E} = \sum_{p=1}^{P} \mathcal{E}_p = \frac{1}{2} \sum_{p=1}^{P} |\boldsymbol{z}_{pk_{\mathrm{f}}} - \boldsymbol{z}_{\mathrm{t}}|^2. \tag{30}$$

Thus, given a topology of the feedforward NN a successful training consists of determining a set of weights $\boldsymbol{w}$ for which the total error $\mathcal{E}$ defined above, is a global minimum. For this purpose a formal methodology known as back-propagation through-time (BPTT) [21, 22] was used to calculate the gradient of the error $\mathcal{E}$ in weight-space, $\nabla\mathcal{E}$, such that the error $\mathcal{E}$ can be reduced by updating the weights iteratively using a steepest descent method, or some other optimization technique. The systematic procedure of reducing the error $\mathcal{E}$ is known as *training of the neural network*, regardless of the mathematical technique chosen to adjust the weights.

In contrast to other training procedures which make use of a learning parameter and a momentum term, the code written for this purpose speeds up the training process making use of the method of conjugate gradients, which is a quadratically convergent gradient method that locates an unconstrained local minimum of a multivariate function [19, 23, 24]. Apart from round-off errors, this technique finds the minimum of a quadratic function in a finite number of steps. However, similarly to all gradient-based optimization methods, for non-quadratic functions this procedure is iterative, and although it is not guaranteed to find the global minimum, the algorithm will lead to the bottom of whatever valley it starts in. The conjugate gradient technique has the additional advantage that it avoids the undesirable phenomenon of premature saturation of the network output units, which precludes significant improvements in the network weights causing an unnecessary increase in the number of iterations required to train the network [25].

The following steps summarize the iterative algorithm used to train a feedforward neural network as a dynamic system controller:

- Step 1. An admissible volume of the phase space (which in our case is a region around the steady-state ignited conditions) is divided in a number P of cells.
- Step 2. Randomly select the initial values of the set of weights that specify the NN.

• Step 3. Generate a set of P trajectories by randomly selecting an initial state from each one of the cells in phase space, and then allowing the NN–dynamical system to evolve in time until a pre-specified maximum number of time steps is reached, the state of the system has moved out of a certain region, or satisfied some other criteria.

• Step 4. Convergence is achieved when the entire set of P trajectories satisfies the convergence criterion $|z_i - z_{it}| \leqslant \epsilon$, for all i, where ϵ is a pre-specified error range and the subindex i corresponds to the components of the state-space vector $\boldsymbol{z}$. Otherwise proceed to the next step.

• Step 5. Using BPTT, calculate the gradient of the error, $\nabla\mathcal{E}$, produced by all the different trajectories and update the weights using the method of conjugate gradients.

• Step 6. Repeat steps 3–5 until the training of the NN is successfully completed, i.e. when the entire set of P trajectories, each of which starts from a different cell, reaches the target $\boldsymbol{z}_{\mathrm{t}}$ within the error range ϵ.

Basically the above training algorithm involves two phases. In the first phase, the joint NN–dynamical system configuration evolves in time from an initial state $\boldsymbol{z}(1)$, until a final state $\boldsymbol{z}(N)$. This final state is compared with its target value, resulting in an error signal for each of the state variables of the system. In the second phase, the error signals are propagated recursively backward in time from the final time step N until the initial time step 1, and are used to calculate the adjustments needed in the weights composing the NN using the components of $\nabla\mathcal{E}$.

A systematic way to obtain the components of $\nabla\mathcal{E}$ makes use of the concept of ordered partial derivatives, which are derivatives whose constant and varying terms are determined by means of an ordered set of equations [21, 22]. A set of coupled algebraic equations determining the values of n variables, is said to be ordered when the values of variables, for example $\{y_1, y_2, \ldots, y_n\}$, are determined by means of a set of recursive equations of the form

$$y_i = f_i(y_1, y_2, \ldots, y_{i-1}) \qquad i = 2, \ldots, n. \tag{31}$$

Let us assume we are interested in a quantity which is a function of the above variables, e.g. $\mathcal{F} = \mathcal{F}(y_1, y_2, \ldots, y_n)$; the ordered partial derivative of $\mathcal{F}$, denoted by a $\dagger$ superscript in the partial derivative symbol, with respect to one of these variables is defined as

$$\frac{\partial^{\dagger}\mathcal{F}}{\partial y_i} = \left(\frac{\partial \mathcal{F}}{\partial y_i}\right)_{y_1,\ldots,y_{i-1}}. \tag{32}$$

The value of this ordered derivative can then be obtained by either one of the following two chain rule expansions [21, 22],

$$\frac{\partial^{\dagger}\mathcal{F}}{\partial y_i} = \frac{\partial \mathcal{F}}{\partial y_i} + \sum_{k=i+1}^{n} \frac{\partial^{\dagger}\mathcal{F}}{\partial y_k}\frac{\partial y_k}{\partial y_i} \tag{33}$$

or

$$\frac{\partial^{\dagger}\mathcal{F}}{\partial y_i} = \frac{\partial \mathcal{F}}{\partial y_i} + \sum_{k=i+1}^{n} \frac{\partial \mathcal{F}}{\partial y_k}\frac{\partial^{\dagger} y_k}{\partial y_i}. \tag{34}$$

Thus, given a particular trajectory produced by the joint NN–dynamical system configuration in figure 1, $\{\boldsymbol{z}(n),\ n = 1, \ldots, N\}$, in which the system evolved from the initial state $\boldsymbol{z}(1)$, at time step 1, until $\boldsymbol{z}(N)$ at the final time step N, the error produced by this particular trajectory will be measured by an expression of the form

$$\mathcal{E} = \mathcal{E}[\boldsymbol{z}(N), \boldsymbol{z}_{\mathrm{T}}]. \tag{35}$$

Identifying the set of equations produced by the time evolution of the NN–discrete dynamical system, with the set in (31) and using the chain rule in (33), it follows that the components of the gradient of the error with respect to the weights, $\partial\mathcal{E}/\partial\omega_{pq}^{(l)}$, are obtained from the expression

$$\frac{\partial\mathcal{E}}{\partial\omega_{pq}^{(l)}} = \sum_{n=1}^{N-1}\sum_{j=1}^{J(L)} \frac{\partial^{\dagger}\mathcal{E}}{\partial u_j(N-n)} \frac{\partial u_j(N-n)}{\partial\omega_{pj}^{(l)}} \tag{36}$$

where the symbol † in the term $\partial^{\dagger}\mathcal{E}/u_j(N-n)$, denotes ordered derivative, as discussed above. The first factor on the right-hand side of (36) is obtained using the following recursive set of equations

$$\frac{\partial^{\dagger}\mathcal{E}}{\partial u_j(N-n)} = \sum_{i=1}^{J(1)} \frac{\partial^{\dagger}\mathcal{E}}{\partial z_i(N-n+1)} \frac{\partial z_i(N-n+1)}{\partial u_j(N-n)} \qquad \text{for } n = 1,\ldots,N-1 \tag{37}$$

where

$$\frac{\partial^{\dagger}\mathcal{E}}{\partial z_i(N-n+1)} = \sum_{j=1}^{J(L)} \frac{\partial^{\dagger}\mathcal{E}}{\partial u_j(N-n+1)} \frac{\partial u_j(N-n+1)}{\partial z_i(N-n+1)} + \sum_{k=1}^{J(1)} \frac{\partial^{\dagger}\mathcal{E}}{\partial z_k(N-n+2)} \frac{\partial z_k(N-n+2)}{\partial z_i(N-n+1)} \qquad \text{for } n = 2,3,\ldots,N-1 \tag{38}$$

is used with the following initialization

$$\frac{\partial^{\dagger}\mathcal{E}}{\partial z_i(N)} = \frac{\partial\mathcal{E}}{\partial z_i(N)}. \tag{39}$$

In the above equations $u_j(N-n)$, i.e. $\mathcal{O}_j^{(L)}(N-n)$, the jth output of the NN, is the value of the jth control variable at time step $N-n$; and $z_i(N-n+1)$ corresponds to the value of the ith state variable at time step $N-n+1$. The name *back-propagation through-time* is due to the fact that the recursive set of equations initialize at the last time step and then proceeds backwards in time until the first time step.

It is necessary to point out that the algorithm described in this section is not strictly a conjugate gradient method, because at each successive iteration an entire new set of randomly chosen initial conditions is used to calculate the total error $\mathcal{E}$ and its gradient. Nevertheless, this technique allows us to sample uniformly the region of interest in the phase space, avoiding the problem of memorizing or over-fitting, which is common in pattern recognition tasks using NNs [20]. As with many optimization techniques, global minimization cannot be guaranteed and often it is necessary to try several NN topologies as well as a different starting set of weights before the global or a satisfactory local minimum is found.

Actual control actions are always bounded between a minimum and a maximum level, and the type of artificial neural network we propose to use incorporates in a natural way these constraints. The output of the sigmoidal nodes in this network is bounded between 0 and 1 as can be deduced from (29), and in order to take into account this effect into the dynamics of the system, we use the following transformation equations relating the output of the NN with the normalized control variables defined in (15)–(17)

$$\hat{S}_{\mathrm{f}} = \frac{S_0}{n_0} k_1 u_1 \tag{40}$$

$$\hat{S}_{\alpha} = k_2(2u_2 - 1)^2 \tag{41}$$

and

$$\hat{S}_{\text{aux}} = k_3(2u_3 - 1)^2 \tag{42}$$

where u_1, u_2 and u_3 are the bounded outputs of the neural network and the parameters k_1, k_2 and k_3 are prespecified positive constants, defining the maximum values that the refuelling rate, the neutral helium injection and the auxiliary power heating can take, respectively. Because $u_j \in (0, 1)\,\forall j$, the control variables in the dynamical system are constrained to take values within the following sets, $0 \leqslant \hat{S}_\alpha < k_2$, $0 < \hat{S}_{\text{f}} < k_1 S_0/n_0$ and $0 \leqslant \hat{S}_{\text{aux}} < k_3$.

The reason for using the particular transformations in (41) and (42) is the following: the output level of sigmoidal activation functions yields the values 0 and 1 only asymptotically, when the net input to the unit attains very large absolute values. On the other hand, the steady state ignited operating point requires $\hat{S}_\alpha = 0$, and $\hat{S}_{\text{aux}} = 0$, something that could not be satisfied by expressions similar to (40), unless the output of the corresponding unit is saturated to zero. In order to avoid this undesirable situation, the quadratic form $(2u - 1)^2$ is used, where the zeros of $\hat{S}_{\text{aux}}$ and $\hat{S}_\alpha$ are achieved when the activation of the output unit is 0.5; a value that is obtained when the net input to the unit is zero.

Since the purpose of this work is to demonstrate the capabilities of an NN for the stabilization at nearly ignited conditions in a tokamak reactor, we choose arbitrarily the following values for the constants k_i:

$$k_1 = 4.0 \qquad k_2 = 0.1\ \text{s}^{-1} \qquad \text{and} \qquad k_3 = 0.1\ \text{s}^{-1}. \tag{43}$$

4. Neural network training process

With the algorithm described in the previous section, we trained a NN to suppress density and temperature excursions in the nonlinear ITER-type thermonuclear reactor model described in section 2. The continuous time dynamical equations given in (11)–(13) in the form

$$\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = \boldsymbol{g}(\boldsymbol{z}, \boldsymbol{u}) \tag{44}$$

were discretized in the simplest form following the Euler's method as

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \boldsymbol{g}(\boldsymbol{z}_k, \boldsymbol{u}_k)\Delta t \tag{45}$$

where Δt is a fixed time-step length, and k ($k = 1, 2, \ldots$), denotes the time step number as discussed in section 3. After some preliminary study we chose to use $\Delta t = 0.1$ s, in order to avoid an excessively large number of time steps during the transient simulations required by the training process. Given an initial perturbation, it is not possible to know *a priori*, the number of time steps required by the NN–dynamical system configuration to return to the operating state; hence, it is not uncommon in these situations, that convergence, as established in Step No 4 of the training algorithm in section 3, is not satisfied for an arbitrarily small convergence parameter during the training process when the maximum number of time steps specified by the algorithm is smaller than the actual number required. Nevertheless the resulting NN can show a satisfactory behaviour, returning the system not to the ideal ignited state but to a nearly ignited state, that will constitute the nominal operating point for our joint NN–dynamical system configuration.

After some trial and error, the multilayer feedforward NN selected was a three layered network of $3 \times 16 \times 3$ architecture, i.e. three input nodes, corresponding to the normalized electron density, the relative helium ash fraction and the plasma temperature; one hidden layer with sixteen nodes; and three output nodes, whose activation levels u_1, u_2 and u_3 determine the values of the normalized refuelling rate, He-4 neutral beam injection and

the auxiliary heating power density by means of equations (40)–(42). The initial values of the weights composing the network were randomly selected between -0.05 and $+0.05$. A cubic region centred at the target state $z_T = (1, 1, 1)$ with lateral size 0.20, was constructed and 27 initial states were chosen randomly at each iteration from within the 27 cubic cells of equal size in which the above region was divided. Therefore the initial values for any normalized state variable were restricted to lie between 0.9 and 1.1.

The training process was divided into four stages each consisting of 1150 iterations of the algorithm discussed in section 3. In the first stage the maximum number of time steps was set to 300 while in the second stage it was set to 600. The reason for this is that during the earliest part of the training the weights are far from their optimal values yielding a bad performance; by restricting the maximum number of time steps to a relatively small number, computational time is not spent unnecessarily at the beginning of the training, when the NN is merely learning to control the fluctuations within a reasonable range. During the transients, in these first two stages of the training, the state variables were restricted to remain within the interval 1.0 ± 0.35, stopping the transient if the phase-space trajectory leaves this region. In the third and fourth stages the maximum number of time steps in a transient was set to 700 and the state variables restricted to remain in the interval 1.0 ± 0.4, except when the initial values of the normalized temperature and electron density were both greater than 1.08, in which case the maximum number of time steps was set to 900 and the variables in the phase-space trajectory constrained to the interval 1.0 ± 0.6. During the first three stages an additional stopping criterion was set: the phase-space trajectory was stopped whenever the normalized electron density crossed the target value of one. During the fourth and last stage this criterion was removed and the trajectories were allowed to expand the entire maximum number of time steps before stopping.

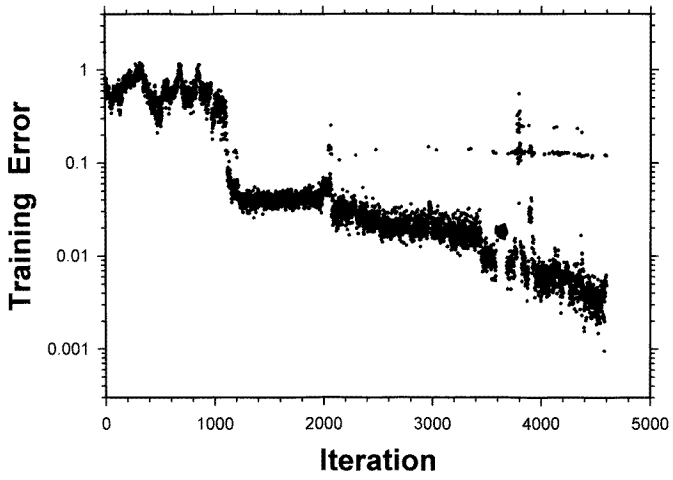

Figure 3. Training error as a function of the iteration number showing a stochastic-like behaviour as a consequence of the random choice in the initial states of the system at each iteration as discussed in the text.

In figure 3, the behaviour of the total error $\mathcal{E}$ is shown, (30), as a function of the number of iterations in the training process. It is observed that the training error does not monotonically decrease as function of the iteration number of the training process. This is

expected because at each iteration the initial conditions are randomly chosen as explained above, hence the weight update that reduces the error at one particular iteration, i.e. for a given set of initial conditions, does not necessarily reduce the error for a different set of initial conditions. Nevertheless the total error shows a tendency to be confined within a region bounded by a minimum and a maximum value that decrease monotonically; the points lying outside this region are due to those trajectories that leave the constrained phase space and are then forced to stop, as discussed above. The training process is thus slowly adjusting the weights so that the error is eventually reduced for a wide set of different initial conditions. In tables 3 and 4 we show the weight values of the NN resulting from the training session just described.

Table 3. Weights connecting the ith node in the input layer with the jth node in the hidden layer.

	i			
j	1	2	3	θ
1	−1.144 10	0.625 92	−2.776 08	1.472 44
2	−1.379 67	0.645 95	−2.554 00	1.577 85
3	−1.638 10	0.692 57	−2.243 72	1.754 29
4	−1.436 91	0.677 95	−2.579 38	1.677 74
5	−1.175 95	0.312 33	−2.298 03	1.223 85
6	0.244 81	1.915 63	−4.442 59	2.236 38
7	−1.066 74	0.876 29	−3.122 99	1.686 17
8	−1.506 79	0.433 48	−1.745 03	1.351 76
9	−1.512 08	0.629 87	−2.163 22	1.485 75
10	−0.988 73	0.818 94	−3.104 87	1.632 86
11	−1.434 35	0.582 12	−2.423 71	1.618 00
12	−1.216 02	1.061 06	−3.185 16	1.798 01
13	−0.438 56	1.366 54	−3.561 60	1.905 59
14	−1.441 30	0.564 16	−2.281 54	1.513 75
15	−1.000 83	0.946 89	−3.278 29	1.787 92
16	−1.039 53	0.860 75	−3.157 65	1.712 48

Extensive testing of the NN–dynamical system configuration showed that the resulting controller is able to return the state of the system to a neighbourhood within a few per cent of the ideal target operating point, following perturbations in any state variable or combination of state variables within the training set used, i.e. 1.0 ± 0.1. Even more, as is shown in the examples in the next section, the NN is able to generate appropriate control actions that suppress excursions with initial conditions outside the region used in the training process, showing the intrinsic generalization capabilities of the feedforward NN. The nearly ignited nominal operating state $\boldsymbol{z}^*$ resulting from the above training process is $z_1^* = 1.015$, $z_2^* = 0.995$ and $z_3^* = 0.990$, which is obtained with an auxiliary heating power density of $\sim$3% its maximum allowable value.

In the next sections we present three typical cases that demonstrate the capabilities of the resulting NN, and a robustness study of the resulting NN to the thermalization time of the alpha particles. In addition to showing the behaviour of the NN–dynamical system for the different perturbations, we will also show the time behaviour of the energy confinement time, (14), and the gain factor defined by the ratio between the instantaneous rate of energy deposited by the alpha particles produced by the D–T fusion and the auxiliary heating power

Table 4. Weights connecting the ith node in the hidden layer with the jth node in the output layer.

	i		
j	1	2	3
1	−2.806 69	−1.274 37	−1.515 46
2	−2.705 81	−1.560 53	−1.532 41
3	−2.573 60	−1.911 74	−1.532 30
4	−2.719 36	−1.620 07	−1.590 68
5	−2.381 98	−1.415 99	−1.337 77
6	−4.273 16	0.497 68	−2.361 07
7	−3.124 05	−1.119 98	−1.704 14
8	−2.082 81	−1.857 34	−1.353 71
9	−2.408 42	−1.737 87	−1.529 50
10	−3.106 49	−1.059 11	−1.645 65
11	−2.602 35	−1.662 04	−1.527 72
12	−3.209 40	−1.214 80	−1.843 17
13	−3.560 97	−0.443 32	−1.944 76
14	−2.523 62	−1.692 90	−1.431 78
15	−3.263 42	−1.050 17	−1.720 25
16	−3.134 93	−1.098 50	−1.695 74
θ	8.583 33	3.162 70	5.096 59

density, i.e. $P_\alpha/(P_{\text{aux}} + P_{\text{ohmic}})$, which is given by

$$Q = \frac{A_\alpha n_0^2 z_1^2 (1 - 2 f_0 z_2)^2 \exp(a_1/T_0^{0.2935} z_3^{0.2935} + a_2 + a_3 T_0 z_3 + \cdots)}{P_{\text{aux}} + \eta j^2}. \quad (46)$$

This factor is a measure of how close the thermonuclear system is to true ignition conditions. In a practical thermonuclear system it is expected that $Q > 25$, in order to be economically viable [26]. The gain value associated with the nearly ignited nominal operating point obtained with the resulting NN–dynamical system configuration here, is $Q \approx 180$.

5. Simulation results

In order to illustrate the capabilities of the NN trained according to the algorithm described earlier, we present in this section three cases. In the first case, the values of the perturbations in the plasma density, the helium fraction and plasma temperature lie in the border of the initial states region used for the NN training; this initial state was chosen because it manifests an interesting characteristic of the resulting NN controller: the control of temperature fluctuations dominates over the control of the variations in the plasma density; in addition, according to the simple control importance analysis made at the end of section 2, the modulation of the refuelling rate is the most influential of the three control techniques used. As a consequence, an increase in temperature is neutralized by temporally increasing the refuelling rate and *vice versa*. The level of difficulty in the control actions is therefore higher when the values of the density and the temperature of the plasma are both either smaller or larger than their nominal operating values.

In the other cases the initial values of the perturbations lie outside the range used during the training session, and are aimed at demonstrating the capability of the resulting NN to generalize. Thus, in the second case the initial perturbations of the electron density and the helium ash are smaller than their nominal operating values while the plasma temperature

is larger; and finally we show a case in which the electron density and the helium ash fraction have both initial values larger than their nominal operating values but the plasma temperature is smaller.

As discussed in the last section the simulation results show that the normalized state variables do not return exactly to the ideal ignited operating point $\boldsymbol{z}_{\mathrm{T}} = (1, 1, 1)$ but to a slightly different one, $\boldsymbol{z}^* = (1.015, 0.995, 0.990)$. This later state corresponds to a nearly ignited steady state associated with a refuelling rate value $S_0^* \cong 1.015 \times S_0$, an auxiliary heating power density $P_0^* = 2.85 \times 10^{-3} \times \frac{3}{2} n_0 T_0$ s^{-1}, and a rate of neutral He-4 injection $S_{\alpha_0}^* \cong 0$, yielding a Q factor of $\sim$180, with a value of the energy confinement time of 7.27 s. These values will constitute our nominal operating conditions for the NN–thermonuclear system configuration.

We shall recall that the output of the NN, u_1, u_2 and u_3, are always bounded between 0 and 1 as discussed in section 3 and it follows from (15)–(17) and (40)–(43), that the refuelling rate S_{f}, the neutral He-4 injection rate S_α, and the auxiliary heating power density P_{aux}, are constrained to take values within the following control sets: $S_{\mathrm{f}} \in (0, 4S_0)$, $S_\alpha \in [0, 0.1 f_0 n_0$ s$^{-1})$, $P_{\mathrm{aux}} \in [0, 0.15 n_0 T_0$ s$^{-1})$, where $S_0 = 4.16 \times 10^{18}$ m^{-3} s^{-1}, $n_0 = 9.8 \times 10^{19}$ m^{-3}, $T_0 = 8.28$ keV and $f_0 = 0.0624$ as discussed in earlier sections.

Case (a). Here we present the resulting transient behaviour associated with the following initial conditions: the electron density is $1.1 \times n_0$, the relative fraction of helium ions is $0.9 \times f_0$, and the plasma temperature is $1.1 \times T_0$. In figure 4 the first 100 s of the resulting time behaviour of the state and control variables are shown. It is observed that initially there is an excess of plasma density in the system and the simultaneous injection of the 'cold' D–T fuel and atomic helium brings down the plasma temperature, at the expense of temporarily increasing even further the plasma density. After the first 25 s into the transient helium injection ceases and the behaviour of the helium fraction is a consequence of the alpha particles produced by the D–T fusion, its large confinement time $7\tau_{\mathrm{E}}$ and the electron density. Eventually the entire system relaxes to a state with a D–T refuelling rate $S = 1.015 S_0$, very close to the desired value S_0, and an auxiliary heating power density of $\sim$3% its maximum allowable value. We have to point out that although increasing the electron density reduces τ_{E}, while reducing the plasma temperature increases it, the combination of the two, in this case, increases the energy confinement time as observed in figure 5 (top). The energy confinement time reaches a maximum and starts to decrease when the temperature is low enough and the auxiliary heating is turned on again. Afterwards the energy confinement time remains practically constant around its nominal operating value of 7.27 s. In figure 5 (bottom) we plot the inverse of the gain, Q^{-1}, as a function of time for the entire transient period, showing the asymptotic approach to the nominal operating value $Q \approx 180$.

Case (b). In this case we choose the following initial conditions of the system in the normalized phase space: the plasma temperature value is 1.2, the relative fraction of helium ions is taken again as 0.9 while the electron density is now 0.8, i.e. z_1 and z_2 are 20% and 10% below their desired target values and z_3 is 20% above. Figure 6 shows the behaviour of the state variables as function of time resulting from the control actions of the NN for the first hundred seconds. As we can observe, the D–T refuelling rate is almost saturated to its maximum allowable value during the first seconds into the transient in order to increase the values of the electron density. This produces a steep decrease in the relative fraction of helium ions as well as in the plasma temperature. After 10 s the refuelling rate falls below its nominal operating value, the electron density starts to decrease and the temperature and the alpha particle fraction increase again. On the other hand the auxiliary heating power decreases from a moderately high value to zero, increasing again

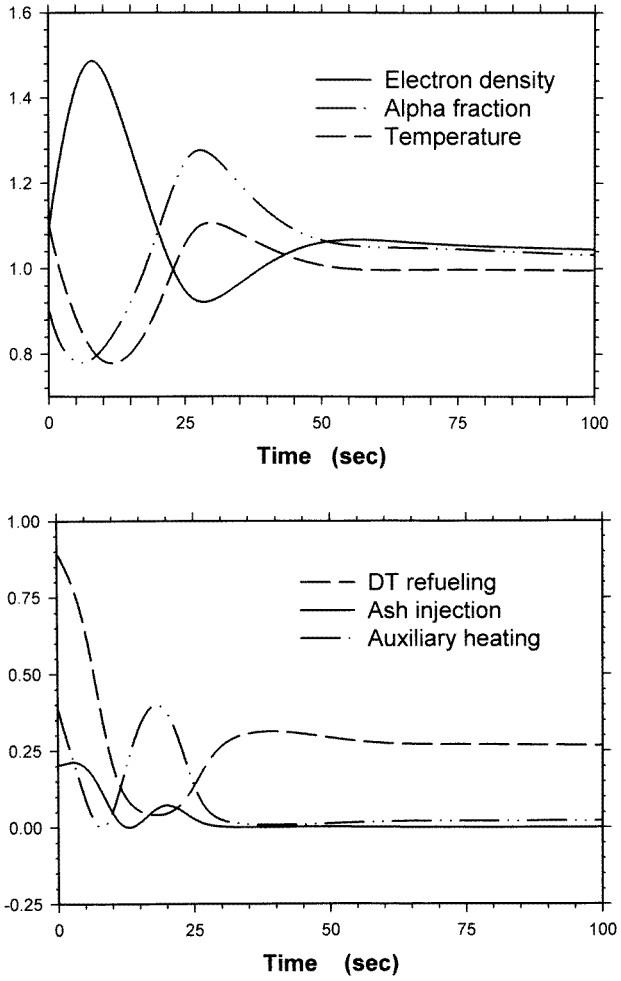

Figure 4. Time behaviour of the normalized state and control variables resulting from the NN–dynamical system configuration in figure 1. The initial values of the electron density, the alpha particle fraction and the plasma temperature are 1.1, 0.9 and 1.1, respectively. Top: behaviour of the normalized state variables defined in (10) as a function of time. Bottom: associated sequence of control actions u_1, $(2u_2 - 1)^2$ and $(2u_3 - 1)^2$ produced by normalized state variables; see (40)–(42).

to almost half its initial value at around 15 s and decreases quickly afterwards to its normal operating value. As observed in this figure the atomic helium injection rate is always negligible except during the early stage of the transient. In figure 5 the resulting time behaviours of the energy confinement time and the Q-factor are shown. Initially the energy confinement time increases because the plasma temperature is diminishing although the electron density increases. Around 10 s into the transient the electron temperature reaches a maximum above its normal operating value and starts to decrease; at the same time the auxiliary heating increases again, and since the gain Q is large, the energy confinement time is dominated by the fusion heating energy and the net result is that the plasma energy confinement time decreases asymptotically to its normal operating value.

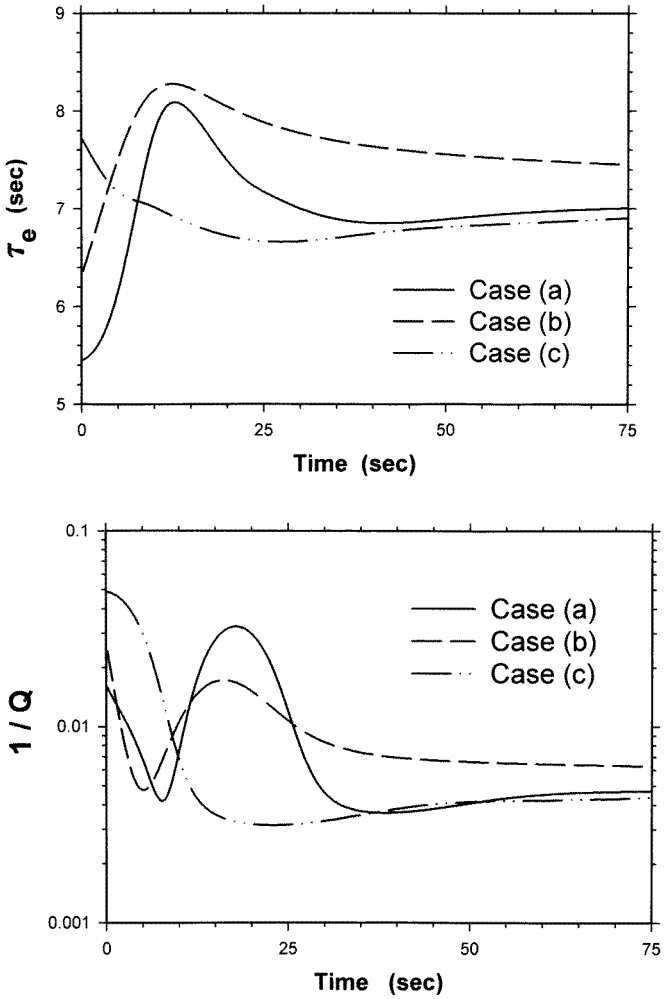

Figure 5. Behaviour of the energy confinement time (top) and inverse of the gain Q (bottom) as a function of time for the cases (a), (b) and (c) discussed in the text.

Case (c). Finally this example shows the first 60 s of a transient for the following initial conditions: the electron density is $1.15 \times n_0$, the relative fraction of helium ions is now assumed to be $1.1 \times f_0$, while the temperature of the plasma is $0.85 \times T_0$. Figure 7 shows the time behaviour of the control and the normalized state variables. It is observed that in this case the refuelling rate is saturated to zero value while the auxiliary heating power is slightly larger than half its maximum value during the early stages of the transient; the helium injection rate has a value of around 20% its upper limit within the first 5 s decreasing quickly to negligible values afterwards and remaining there for the rest of the transient. The energy confinement time and the behaviour of $1/Q$ during the transient are also shown in figure 5, from where it is observed that although initially the temperature is low, and the density is high, the energy confinement time is dominated by the temperature

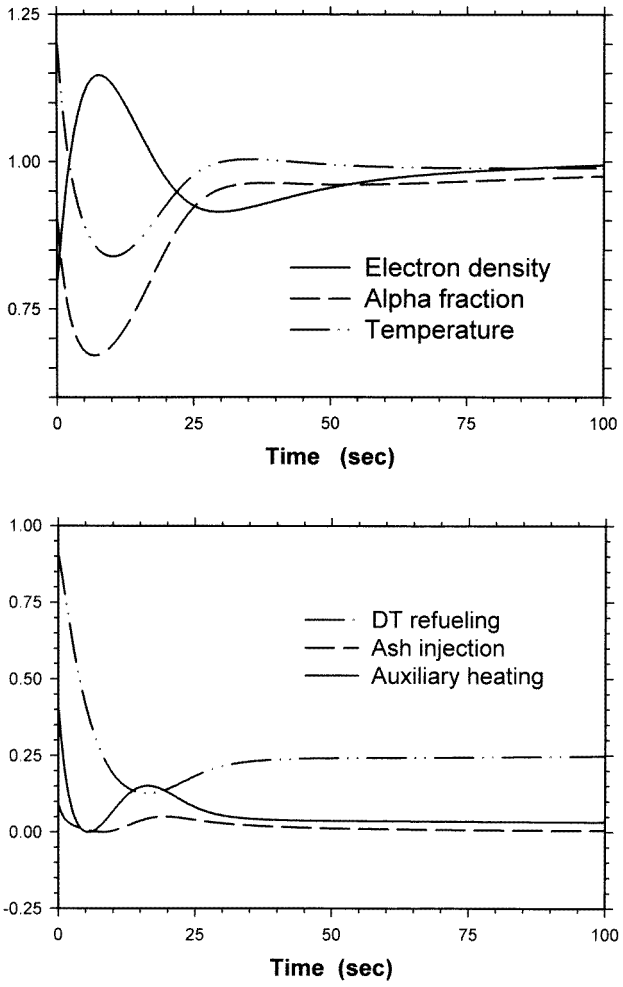

Figure 6. First hundred seconds of the transitory behaviour of the state and control variables resulting from the NN–dynamical system configuration in figure 1, for initial values of the electron density, the alpha particle fraction, and the plasma temperature of 0.8, 0.9 and 1.2, respectively. Top: time behaviour of the normalized state variables as defined in (10). Bottom: associated control actions u_1, $(2u_2 - 1)^2$ and $(2u_3 - 1)^2$; see (40)–(42).

variations. Thus, τ_E decreases because the temperature increases. As can be observed in figure 5 (bottom) the value of the gain during the first 10 s into the transient increases which means that the energy produced by the fusion reactions becomes larger compared to the total auxiliary heating power. The auxiliary heating initially is half the maximum allowable value, helping raise the plasma temperature, increasing the fusion rate and thus the Q factor, while decreasing the energy confinement time; after 15 s the auxiliary heating has already decreased to near its nominal operating value.

In the next section we present a numerical study of the effect the thermalization time of the alpha particles produced by the fusion reactions has on the stability of the joint NN–dynamical system configuration of figure 1.

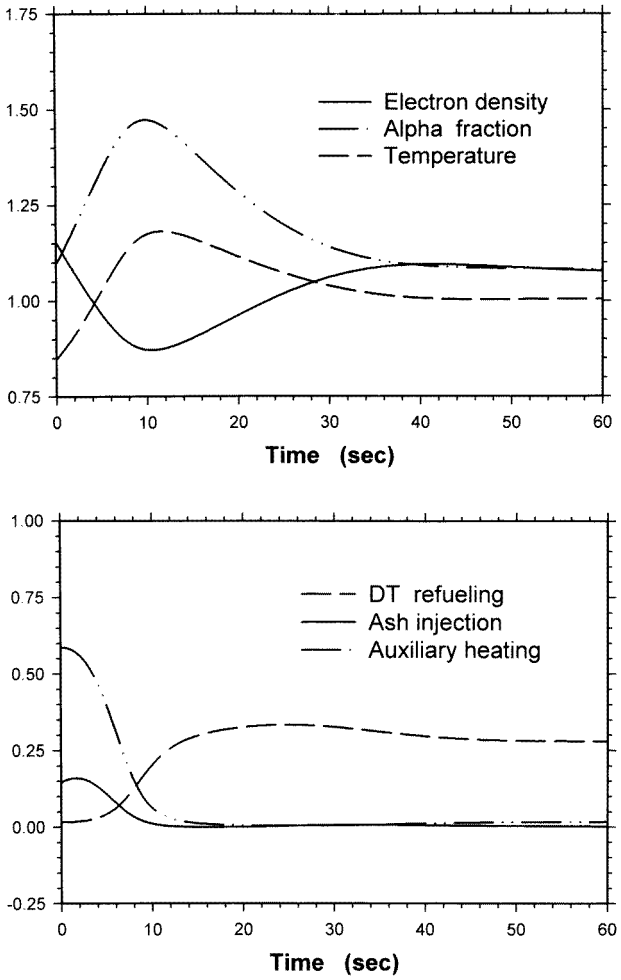

Figure 7. First 60 s of the transient corresponding to the following normalized initial conditions: the electron density, the alpha particle fraction and the plasma temperature values are 1.15, 1.1 and 0.85, respectively. Top: behaviour of the normalized state variables as defined in (10), as a function of time. Bottom: time behaviour of the associated control actions u_1, $(2u_2 - 1)^2$ and $(2u_3 - 1)^2$ in (40)–(42).

6. Effect of the thermalization time of the alpha particles

In most of the energy range between 3.5 MeV and the plasma nominal temperature of 8.28 keV, the velocity of the alpha particles is greater than the thermal velocity $\sqrt{2T/m}$ of the ions and much smaller than that corresponding to the electrons. In this range, $v_{\mathrm{th,i}} \ll v_\alpha \ll v_{\mathrm{th,e}}$, the rate of energy loss of the alpha particles is approximately given by [27–29],

$$\frac{\mathrm{d}}{\mathrm{d}t}E_\alpha = -\frac{32\sqrt{\pi}\,Z_\alpha^2 e^4 n_\mathrm{e}}{3m_\mathrm{e}m_\alpha}\frac{\ln\Lambda_\mathrm{e}}{v_{\mathrm{th,e}}^3}E_\alpha - \sum_{\mathrm{i=ions}}\frac{4\pi\,Z_\alpha^2 e^4 n_\mathrm{i}}{\sqrt{2}}\left(1+\frac{m_\alpha}{m_\mathrm{i}}\right)\frac{1}{E_\alpha^{1/2}} \qquad (47)$$

where the first term on the right-hand side is the energy absorbed by the plasma electrons while the last term is the energy lost to the ions. The alpha particles lose their energy mainly to the electrons above a certain energy value, that for the nominal operating parameters of the thermonuclear reactor of this work is 0.374 MeV. Hence, in the thermalization process, approximately 90% of the energy of the fusion alphas is absorbed by the electrons and only 10% by the ions. Using (47) it is possible to estimate the time required by the alphas to slow down from 3.5 MeV to 0.374 MeV which turns out to be 0.22 s, taking an additional 0.06 s approximately to completely thermalize to the plasma temperature of 8.28 keV. This equation can be integrated exactly to obtain a closed expression for the slowing down time of an alpha particle from 3.5 MeV to the critical energy, after which its energy is mainly absorbed by the ions, i.e.

$$\tau_{\text{crit}} = 3.33 \times 10^{17} \frac{T_0^{3/2}}{n_0} \left(\frac{z_3^{3/2}}{z_1} \right) \ln(0.5 + 340.7 T_0^{-3/2} z_3^{-3/2}). \tag{48}$$

On the other hand, the energy absorbed by the electrons is eventually shared with the plasma ions through an energy equipartition process, obeying the following law [29, 30]

$$\frac{\mathrm{d}T_{\text{e}}}{\mathrm{d}t} = \frac{T_{\text{i}} - T_{\text{e}}}{\tau_{\text{eq}}} \tag{49}$$

where T_{e} and T_{i} are the electron and ion kinetic temperatures, respectively, and τ_{eq} is the equilibration time given by

$$\tau_{\text{eq}} = \frac{3}{8\sqrt{2\pi} m_{\text{e}}^{1/2} \ln \Lambda_{\text{e}}} \frac{T_{\text{e}}^{3/2}}{\sum_{\text{ions}} n_{\text{i}}/m_{\text{i}}} \tag{50}$$

for the plasma parameters considered in this work,

$$\tau_{\text{eq}} = 1.2 \times 10^{18} \frac{T_0^{3/2}}{n_0} \left(\frac{z_3^{3/2}}{z_1} \right). \tag{51}$$

Because it is the ion kinetic temperature which directly affects the fusion rate, and because the model described in the introductory section assumes that all the components of the plasma share the same temperature at all times, it is reasonable to consider the equilibration time as a contributing factor to the effective thermalization time. Thus, we will assume here that the effective thermalization time is given by the sum of equations (48) and (51), i.e.

$$\tau_{\text{th}} = 2.1 \times 10^{18} \frac{T_0^{3/2}}{n_0} \left(\frac{z_3^{3/2}}{z_1} \right) \text{ s} \tag{52}$$

where $[T_0] =$ keV and $[n_0] = \text{m}^{-3}$. As a simplification, in this expression we have assumed that the argument of the logarithm in (48) is constant and equal to its value at the nominal operating point. At the nominal operating parameters of this work, the effective thermalization time turns out to be approximately 0.51 s.

A set of simulations was performed to test the NN–dynamical system stability with respect to finite thermalization times, assuming both constant time delays in the energy deposited by the alpha particles, as well as a state dependent effective thermalization time as given in (52). To illustrate the results of these simulations we calculate the time behaviour of the total error $\mathcal{E}$, as given in (30), for a set of 64 trajectories whose initial states lie inside the region used to train the NN, as described in section 4. Each of these trajectories was generated by assigning to the different state variables one of the following four initial values: 0.9, 0.9666, 1.0333 and 1.1. Figure 8 shows the time behaviour of the error $\mathcal{E}$

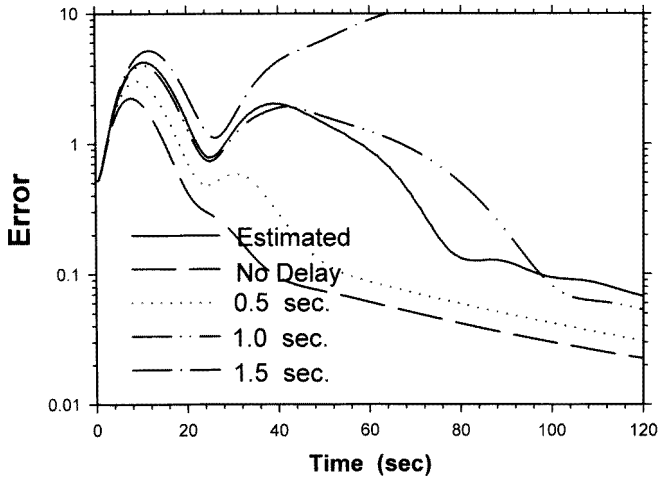

Figure 8. Time behaviour of the total error $\mathcal{E}$ corresponding to the 64 trajectories discussed in the text, for different constant thermalization time delays and for the effective thermalization time given in (52).

for three different constant thermalization time delays: 0.5 s, 1.0 s and 1.5 s; for the thermalization time in (52) and also include the case where the time delay is null for comparison purposes. It is observed that all of the 64 trajectories return to the nominal operating point for constant time delays of 0.5 s and 1.0 s, (comprising the thermalization time estimated at the nominal operating point) but when the time delay was set to 1.5 s the total error diverges in time as shown in the figure. However, the culprit of this divergent behaviour is only one single trajectory: the one originating from the initial state $z_1 = 1.1$, $z_2 = 0.9$ and $z_3 = 1.1$. Similarly, the total error diverges when the thermalization time is estimated from (52) for exactly the same reason. In figure 8 we show the total error produced, in this later case, by these 64 trajectories when this initial state was substituted by $z_1 = 1.097$, $z_2 = 0.903$ and $z_3 = 1.097$. In fact, when the thermalization time is calculated from (52), the entire cubic region of size 0.2 centred in the ideal operating point $(1, 1, 1)$ is stable except for a small region surrounding the corner point $(1.1, 0.9, 1.1)$ in the normalized phase space; this is indeed a small region, since for instance, the surrounding points $(1.1, 0.9, 1.097)$, $(1.097, 0.9, 1.1)$ and $(1.1, 0.903, 1.1)$ are all stable. Other values for the constant thermalization time were also used in the simulations, showing that this unstable region grows larger for larger thermalization times; damped oscillatory behaviour in the state variables appear when the constant thermalization time is taken to be $\sim$3.0 s; and undamped, stationary non-divergent oscillations with a period of 60 s, are found throughout the training region in phase space for $\tau_{\text{th}} = 3.5$ s. The nominal operating point of the system becomes unstable afterwards.

7. Conclusions

In this work we have shown how a feedforward multilayer NN with sigmoidal activation functions can be trained for the stabilization of a thermonuclear reactor at nearly ignited burn conditions. A volume-averaged zero-dimensional nonlinear model was used to represent the evolution of the electron density, the helium ash fraction and the plasma temperature. The

stabilization is obtained by means of the concurrent modulation of the D–T refuelling rate, neutral He-4 injection and an auxiliary heating. The stabilization task is made difficult due to the presence of constraints in the maximum and minimum levels that the control actions can take. The NN was trained using a back-propagation through-time technique in which a conjugate gradients algorithm was used to accelerate learning. Although, the resulting NN is guaranteed to stabilize the system only for initial states within the training region of the neural network, it was shown to be able to stabilize the nonlinear thermonuclear system for temperature and density departures significatively far from their nominal operating values so that the strong nonlinear behaviour of the system is manifested. From a simple sensitivity analysis, it was found that out of the three control parameters considered in this work, the D–T refuelling rate and the auxiliary heating, play a more significant role on the stabilization of the system; although, the auxiliary heating and the neutral He injection are needed mainly in the early phases of the transient. Using an approximate expression for the thermalization time, the resulting NN proved to be robust regarding delay times in the energy deposited to the plasma by the alpha particles produced by the D–T fusion, for perturbations within the training region of the phase space.

A study concerning the robustness of the NN under uncertainties in diverse parameters of the system such as energy confinement time, and the behaviour when random noise is present in the measured values of the state variables, is underway and will be reported later. Further extensions of this work improving the model by using a more recent design from the engineering design activity (EDA) for ITER, allowing different electron and ion temperatures, as well as a more general treatment of the energy confinement times will be considered in future work. In addition, load following capabilities of the NN controller, as well as optimal or near-optimal strategies, minimizing for example the transient time, the total auxiliary heating energy provided to control the transient, or the total energy generated by power overshoots also needs to be addressed.

Acknowledgments

The authors gratefully acknowledge partial financial support from DGAPA-IN101696 (UNAM) and CONACYT-3155P-A9607 (México) projects.

References

[1] Stacey W M 1973 Operating regimes of controlled thermonuclear reactors and stability against fundamental mode excursions in particle densities and temperatures *Nucl. Fusion* **13** 843

[2] Bromberg L, Fisher J L and Cohn D R 1980 Active burn control of nearly ignited plasmas *Nucl. Fusion* **20** 203

[3] Ashby D and Hughes M H 1980 Dynamic burn control of a tokamak reactor by fuel injection *Nucl. Fusion* **20** 203

[4] Harting L, Fuchs V and Bers A 1980 Creating stable tokamak reactor equilibria by suplement heating *Nucl. Fusion* **20** 803

[5] Sager G, Miley G H and Maya I 1985 Optimal control theory applied to fusion plasma thermal stabilization *Fusion Technol.* **8** 1795

[6] Ho S K and Fenstermacher M E 1989 Thermally stable operation of engineering test reactors *Fusion Technol.* **16** 185

[7] Haney S W, Perkins L J, Mandrekas J and Stacey W M 1990 Active control of burn conditions for the international thermonuclear experimental reactor *Fusion Technol.* **18** 606

[8] Hui W, Bamieh B A and Miley G H 1994 Robust burn control of a fusion reactor by modulation of the refuelling rate *Fusion Technol.* **25** 318

Hui W, Bamieh B A and Miley G H 1994 Robust burn control of a fusion reactor with modulation of refuelling rate *Fusion Technol.* **26** 1151

[9] Doyle J C, Glover K, Khargonekar P P and Francis B A 1989 State-space solutions to standard H_2 and H_∞ control problems *IEEE Trans. Automatic Control* **34** 831

[10] Vemuri V R (ed) 1992 *Artificial Neural Networks: Concepts and Control Applications* (Los Alamitos, CA: IEEE Computer Society Press)

[11] Hunt K J, Sbarbaro D, Zbikowski R and Gawthrop P J 1992 Neural networks for control systems—a survey *Automatica* **28** 1083

[12] Funahashi K 1989 On the approximate realization of continuous mappings by neural networks *Neural Networks* **2** 183

[13] Narendra K S and Parthasarathy K 1990 *Identification and Control of Dynamical Systems using Neural Networks (IEEE Trans. Neural Networks)* vol 1, no 1 p 4

[14] Taylor R J, Fried B D and Morales G J 1990 Burn threshold for fusion plasmas with helium accumulation *Internal Report* UCLA PPG # 1290 March, Institute of Plasma and Fusion Research, University of California, Los Angeles

[15] Glasstone S and Lovberg R H 1975 *Controlled Thermonuclear Reactions* (New York: Krieger)

[16] Hively L M 1977 Convenient computational forms for Maxwellian reactivities *Nucl. Fusion* **17** 873

[17] Hirshman S P and Sigmar D J 1981 Neoclassical transport of impurities in tokamak plasmas *Nucl. Fusion* **21** 1079

[18] Rumelhart D E, Hinton G E and Williams R J 1986 Learning internal representations by error propagation *Parallel Distributed Processing: Exploration in the Microstructure of Cognition* vol I, ed D E Rumelhart and J C McClelland (Massachusetts: MIT)

[19] Reifman J and Vitela J E 1994 Accelerating learning of neural networks with conjugate gradients for nuclear power plant applications *Nucl. Technol.* **106** 225

[20] Wieland A and Leighton R 1987 Geometric analysis of neural network capabilities *Proc. IEEE 1st. Int. Conf. on Neural Networks (III-385 Sheraton Harbor Island East, San Diego, CA, June 21–24 1987)*

[21] Werbos P 1990 Backpropagation through time: what it does and how to do it *Proc. IEEE* **78** 1550

[22] Piché S W 1994 Steepest descent algorithms for neural network controllers and filters *IEEE Trans. Neural Networks* **5** 198

[23] Fletcher R and Reeves C M 1964 Function minimization by conjugate gradients *Comput. J.* **7** 149

[24] Luenberger D G 1984 *Linear and Nonlinear Programming* (New York: Addison-Wesley)

[25] Vitela J E and Reifman J 1997 Premature saturation in backpropagation networks: mechanism and necessary conditions *Neural Networks* **10** 721

[26] Cohn D R 1994 High gain requirements and high field tokamak experiments *J. Fusion Energy* **13**

[27] Vitela J E and Coronado M 1990 Nonquilibrium fluctuations of a charged brownian particle in a plasma *Phys. Rev.* A **42** 2354

[28] Trubnikov B A 1965 *Review of Plasma Physics* vol 1, ed M A Leontovich (New York: Consultants Bureau)

[29] Anders A 1990 *A Formulary for Plasma Physics* (Berlin: Akademie)

[30] Spitzer L Jr 1962 *Physics of Fully Ionized Gases* 2nd edn (New York: Wiley)