Stability regions of recurrent type neural networks

Article *in* Electronics Letters · June 1992

DOI: 10.1049/el:19920649 · Source: IEEE Xplore

2 authors, including:

M.K. Ciliz
Bogazici University
42 PUBLICATIONS **1,140** CITATIONS

SEE PROFILE

by butting dielectric mirrors to the cleaved fibre ends and launching the pump light through one mirror.

Table 1

	Alumino-silicate	ZBLAN	GPBZK
Maximum phonon energy (cm^{-1})	1150 [1]	590 [2]	920
$^3F_4$ lifetime measured (μs)	20 [10]	1100 [3]	250
$^3H_4$ lifetime measured (μs)	500 [6]	6400 [3]	3000
Refractive index	~1·5	1·52 [11]	1·812

Lasing was achieved with only 3 cm of 7 μm core fibre pumped at 790 nm with a GaAs diode and using mirrors with 100 and 98% reflectivities from 1·8 to 2·2 μm. The laser threshold was 9 mW incident on the launch objective and the lasing wavelength was 1·905 μm. This incident power is equivalent to 3·6 mW launched, assuming a combined launch efficiency and objective transmission of 40%. With an initial fibre length of only 3 cm it was not possible to perform a fibre cutback to more accurately assess the launch efficiency. This low threshold, for a nonoptimised system, is already superior to the best thresholds achieved for Tm^{3+} doped aluminosilicate fibre, reported threshold 4·4 mW [6], and Tm^{3+} doped ZBLAN, reported threshold 50 mW [5, 4], and reflects the achievement of our aim of a low threshold fibre laser in the 2 μm region. The short length of this diode-pumped device also shows considerable promise for practical devices.

To make an initial assessment of the slope efficiency achievable from this fibre a further laser cavity was constructed with 27·3 cm of 7 μm core fibre using Ti : sapphire pumping at 794 nm. A 100% reflecting input mirror was used with lasing achieved with just Fresnel reflection feedback from the bare fibre end. The high refractive index of the GPBZK core provides a Fresnel reflection from the fibre end of ~10%, giving approximately 90% output coupling. Lasing was at a wavelength of 1·88 μm with a threshold of 43 mW of launched pump power. Slope efficiency of output power against launched pump power was measured to be 13%, equivalent to 30% quantum efficiency. This quantum efficiency can be compared to a maximum pump quantum efficiency available for this system of ~60%, an upper estimate calculated as the ratio of nonradiative to radiative emission probabilities from the 795 nm pumped $^3F_4$ level. Further improvement in the observed quantum efficiency should be achieved by optimising the fibre length. However, even without optimisation this GPBZK system, with an achieved slope efficiency of 13%, compares well to the Tm : alumina-silica system, where 17% was achieved in the optimised diode pumped system [6] and where the pump efficiency is ~100%. This 13% slope efficiency is also considerably superior to that achieved in Tm^{3+} : ZBLAN where 5·3% has been reported [5], improving to 8·3% with simultaneous lasing at 2·3 μm [3].

In conclusion, we have demonstrated that it is possible to fabricate a low-loss, high strength, intermediate phonon energy optical glass fibre doped with a rare earth. Furthermore, as anticipated from the phonon energy of the glass, we have shown that Tm^{3+} : GPBZK fibre lases in the 2 μm region with thresholds lower than silica-based fibres. The relatively high rare-earth concentration incorporated into this fibre and the high emission cross-section has resulted in very short fibre lengths for the lasers, a convenient feature for many applications such as Q switching. The combination of all these properties, together with the high mechanical strength and excellent optical properties yield a fibre which appears ideal for commercial applications.

1st April 1992

J. R. Lincoln and W. S. Brocklesby (*Department of Physics, University of Southampton, Southampton SO9 5NH, United Kingdom*)

C. J. Mackechnie, J. Wang, R. S. Deol, D. C. Hanna and D. N. Payne (*Optoelectronics Research Centre, University of Southampton, Southampton SO9 5NH, United Kingdom*)

References

1 MATSON, D. W., SHARMA, S. K., and PHILPOTTS, J. A.: 'The structure of high-silica alkali-silicate glasses: a Raman spectroscopic investigation', *J. Non-Crystalline Solids*, 1983, **58**, pp. 323–352
2 BENDOW, B., BANERJEE, P. K., DREXHAGE, M. G., GOLTMAN, J., MITRA, S. S., and MOYNIHAN, C. T.: 'Comparative study of vibrational characteristics of fluorozirconate and fluorohafnate glasses', *J. Am. Ceram. Soc.*, 1982, **65**, (1), pp. C8–C9
3 CARTER, J. N., SMART, R. G., TROPPER, A. C., and HANNA, D. C.: 'Thulium doped fluorozirconate fibre lasers', *J. Non-Crystalline Solids*, 1992, **140**, pp. 10–15
4 ALLAIN, J. Y., MONERIE, M., and POIGNANT, H.: 'Tunable CW lasing around 0·82, 1·48, 1·88 and 2·35 μm in thulium-doped fluorozirconate fibre', *Electron. Lett.*, 1989, **25**, pp. 1660–1662
5 SMART, R. G., CARTER, J. N., TROPPER, A. C., and HANNA, D. C.: 'Continuous-wave oscillation of Tm^{3+}-doped fluorozirconate fibre lasers around 1·47 μm, 1·9 μm and 2·3 μm when pumped at 790 nm', *Opt. Comm.*, 1991, **82**, pp. 563–570
6 BARNES, W. L., and TOWNSEND, J. E.: 'Highly tunable and efficient diode pumped operation of Tm^{3+} doped fibre lasers', *Electron. Lett.*, 1990, **26**, (11), pp. 746–747
7 LAYNE, C. B., LOWDERMILK, W. H., and WEBER, M. J.: 'Multiphonon relaxation of rare-earth ions in oxide glasses', *Phys. Rev. B.*, 1977, **16**, (1), pp. 10–20
8 TAYLOR, E. R., TAYLOR, D. T., LI, L., TACHIBANA, M., TOWNSEND, J. E., WANG, J., WELLS, P. J., REEKIE, L., MORKEL, P. R., and PAYNE, D. N.: 'Application-specific optical fibres manufactured from multicomponent glasses'. Optical Fiber Materials and Processing Materials Research Conf., Symp. Proc., Boston, 1989, pp. 321–327
9 IMBUSCH, G. F., and KOPELMEN, R.: *In* YEN, W. M., and SELZER, P. M.: 'Laser excitation fluorescence spectroscopy in glass' (Springer Verlag, New York, 1981), p. 3
10 LINCOLN, J. R., BROCKLESBY, W. S., CUSSO, F., TOWNSEND, J. E., TROPPER, A. C., and PEARSON, A.: 'Time resolved and site selective spectroscopy of thulium doped into germano- and alumino-silicate optical fibres and preforms', *J. Luminescence*, 1991, **50**, pp. 297–308
11 FRANCE, P. W., CARTER, S. F., MOORE, M. W., and DAY, C. R.: 'Progress in fluoride fibres for optical communications', *Br. Telecom. Technol. J.*, 1987, **5**, (2), pp. 28–43

STABILITY REGIONS OF RECURRENT TYPE NEURAL NETWORKS

K. Ciliz and A. Harova

Indexing terms: Stability, Neural networks

The stability regions of the stable equilibria of the Hopfield type recurrent neural networks are investigated. The dimensions of these regions directly determine how much partial information is needed for an imperfect pattern to approach a desired pattern. Therefore estimation of such regions is instrumental in the analysis of analogue recurrent neural networks.

Introduction: Artificial neural networks (ANNs) have emerged as viable alternatives to conventional computational systems. Specifically, recurrent type neural networks which were introduced by Hopfield [1] were shown to be used for various optimisation problems [2] and also as effective associative memory elements [1].

In the analysis of feedback neural networks, estimation of the stability regions of equilibrium points is important, because the success of recalling a desired pattern vector from partial information is directly related to the stability boundaries of the corresponding equilibrium point. The larger the domain of attraction of an equilibrium point, the better the recalling performance for the pattern vector associated with that equilibrium point. This means that the pattern vector can be recalled with less partial information. In a recent work by Michel *et al.* [3], the dynamics of a feedback type neural network is analysed in the framework of interconnected systems and a method for the estimated of the stability region is also proposed. In this Letter a constructive and iterative procedure is used to estimate the stability region of an equi-

librium point and it is demonstrated that it gives better results than the method used in Reference 3.

Recurrent neural network dynamics: Here we consider the analogue feedback type neural networks represented by the following mathematical model [1]:

$$C_i \frac{\partial u_i}{\partial t} = \sum_{j=1}^{n} T_{ij} v_j - \frac{u_i}{r_i} + I_i \qquad i = 1, \ldots, n \tag{1}$$

In eqn. 1, u_i represents the mean soma potential of the 'i'th neuron, v_j is the output of the 'j'th neuron and is defined by the relation $v_j = g(u_j)$ where $g(\ ): \mathscr{R} \to (-\Delta, \Delta)$ is a bounded, continuous and increasing function; I_i can be considered as the external input to the 'i'th neuron and T_{ij} denotes the elements of the $n \times n$ interconnect matrix T.

The system defined in eqn. 1 can be written in state space form by assigning the cell potentials u_i to the state vector x. Let $x = [u_1, \ldots, u_n]^T \in \mathscr{R}^n$, then eqn. 1 becomes

$$\dot{x} = CT\gamma(x) - Gx + CI \tag{2}$$

where $\gamma(x) = [g(x_1), \ldots, g(x_n)]^T \in \mathscr{R}^n$ is a vector function; C is an $n \times n$ diagonal matrix with entries $C_{ii} = 1/c_i$; $I = [I_1, \ldots, I_n]^T \in \mathscr{R}^n$ is an external input vector and the diagonal matrix G can be written as $G_{ii} = 1/r_i c_i$ with $i = 1, \ldots, n$.

Without loss of generality, we can assume that the external inputs $I_i = 0$ are set to zero; then the system dynamics can be represented in a more general form as

$$\dot{x} = f(x) \tag{3}$$

where $f(x) = [CT\gamma(x) - Gx] \in \mathscr{R}^n$ is a nonlinear vector function.

An equilibrium point x_* of eqn. 3 is said to be stable if all the eigenvalues of its corresponding Jacobian matrix have negative real parts, otherwise it is unstable. For a locally stable equilibrium point x_*, the concept of stability region is defined as the set of all initial conditions which have the property that any solution trajectories starting from them eventually approach the corresponding stable equilibrium point. If $\phi(x_0, t)$ denotes a trajectory that starts at time $t = t_0$ from an initial condition x_0, then the domain of attraction or the stability region of a stable equilibrium point x_* can be written as

$$S(x_*) = \{x_0 \in \mathscr{R}^n : \phi(x_0, t) \to x_* \qquad \text{as } t \to \infty\} \tag{4}$$

Estimation of stability regions: A common approach to estimating the domain of attraction is through the use of critical level values of the associated Lyapunov functions of the system under discussion [4, 5].

For a local Lyapunov function $V(x)$ of a specific equilibrium point x_* of eqn. 3, let W denote an open set over which $\dot{V}(x) < 0$ except at the equilibrium point x_*. We next define a set $D_V(r)$ as the connected component, including the equilibrium point x_*, of the set $\{x \in \mathscr{R}^n : V(x) < r\}$. If the set $D_V(r)$ is contained in the open set W, then $D_V(r)$ is a subset of the complete stability region $S(x_*)$ of x_* defined in eqn. 4. That is, $D_V(r) \in W \Rightarrow D_V(r) \subset S(x_*)$. If r is the largest constant such that

$$D_V(r) \in W \tag{5}$$

holds, then the set $D_V(r)$ is the largest estimate of the stability region of the equilibrium point x_* for the chosen Lyapunov function $V(x)$. Hence the problem reduces to finding the largest possible value 'r' such that eqn. 5 holds for the local Lyapunov function $V(x)$. This value of 'r' is also known as the critical level value of $V(x)$. To solve for 'r', the problem can be formulated as follows:

$$\max_x V(x) \qquad \text{subject to} \qquad \dot{V}(x) < 0 \tag{6}$$

The idea here is to find the maximum value of $V(x)$ at a point where $\dot{V}(x)$ changes sign. To obtain a better estimate of the stability region $S(x_*)$ using a local Lyapunov function and its associated critical level value r, commonly used techniques proposed the optimisation of Lyapunov function parameters and the associated level value in such a way that the new estimates always remain in the exact stability region $S(x_*)$ [5]. In these methods, at each step a new Lyapunov function and a new level value are obtained and this leads to excessive computation.

Iterative and constructive method: First for a given stable state (equilibrium point) of the neural network architecture in eqn. 3, a local Lyapunov function is generated through the solution of the matrix Lyapunov equation, which is given as $J^TA + AJ = -Q$ with J being the Jacobian of the system defined in eqn. 3 around the stable equilibrium point x_*. Q is chosen as an $n \times n$ symmetric positive definite matrix and solution matrix A turns out to be positive definite. Based on matrix A, a local Lyapunov function can be formed as

$$V(x) = x^TAx \tag{7}$$

The critical level value r of eqn. 7 is then obtained using an optimisation technique based on the formulaton given in eqn. 6.

Once the 'r' value is obtained, the connected component $D_V(r)$ of the set $\{x \in \mathscr{R}^n : V(x) \leqslant r\}$ containing the stable equilibrium x_* gives the first estimate of the stability region $S(x_*)$. The second step is the modification of the Lyapunov function in eqn. 7 such that a better estimate would be obtained. $V(x)$ is replaced by $V_1(x) = V[x + \alpha f(x)]$ where $\alpha > 0$ and $f(x)$ is the nonlinear vector function defined in eqn. 3. The procedure cells for the computation of connected component $D_{V_1}(r) = \{x : V_1(x) \leqslant r\}$ for the new Lyapunov function $V_1(x)$, and this new estimated region is also included in the exact stability region $S(x_*)$. Next a new Lyapunov function $V_2(x) = V_1[x + \alpha f(x)]$ is generated and a new connected component $D_{V_2}(r) = \{x : V_2(x) \leqslant r\}$ is obtained using the same critical level value r.

The expansion scheme is carried on with the generation of new Lyapunov functions such as $V_1(x) = V[x + \alpha f(x)]$; $V_2(x) = V_1[x + \alpha f(x)]; \ldots; V_n(x) = V_{n-1}[x + \alpha f(x)]$ and at each step a new stability region estimate is obtained using the same critical level value of the original Lyapunov function $V(x)$. Proofs showing that, for a finite number of iterations, the above expansion procedure gives a strictly increasing sequence of estimated stability regions are given elsewhere*.

This expansion scheme is computationally simple in the sense that it only requires algebraic operations in updating the Lyapunov functions and also it does not require the recomputation of the critical level value for each newly generated Lyapunov function. We next apply this estimation scheme to a two-neuron feedback neural network architecture which is also considered in References 1 and 3.

Application to two-neuron feedback network: For a two neuron network, the interconnect matrix is given as $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and the nonlinear activation function for each neuron is $g_i(x_i) = (2/\pi) \arctan(\lambda \pi x_i/2)$ with $\lambda = 1{\cdot}4$. To be consistent with the example in Reference 3 we set the neural network dynamics as

$$\dot{x}_1 = -1{\cdot}1x_1 + g_2(x_2) \tag{8}$$

$$\dot{x}_2 = -1{\cdot}1x_2 + g_1(x_1) \tag{9}$$

The above nonlinear system has three equilibrium points at $x_{*1} = (0, 0)$, $x_{*2} = (0{\cdot}454, 0{\cdot}454)$, $x_{*3} = (-0{\cdot}454, -0{\cdot}454)$. The equilibrium points in the first and third quadrants are stable, (0, 0) is unstable. Here we investigate the stability region of (0·454, 0·454). The exact stability region of this equilibrium point is the region above the $x_2 = -x_1$ line in the phase plane [1].

Following the procedure given above, first a local quadratic Lyapunov function for the stable equilibrium (0·454, 0·454) is generated as

$$V(x_1, x_2) = \tfrac{1}{2}[(x_1 - 0{\cdot}454)^2 + (x_2 - 0{\cdot}454)^2] \tag{10}$$

* CILIZ, K., and HAROVA, A.: 'Stability region analysis of the equilibrium points of recurrent neural networks'. Under preparation

$V(x)$ is then maximised over x subject to the constraint $\dot{V}(x) < 0$ to find the critical level value. In this case, a Lagrange minimisation technique is used and the critical level value r is computed as $r = 0{\cdot}206$. The connected component of $\{x: V(x) \leqslant r\}$ gives the first estimate of the stability region. The expansion scheme given in the 'estimation of stability regions' Section is carried out for 10, 20 and 50 iterations with $\alpha = 0{\cdot}01$. The stability region estimates for these cases are shown in Fig. 1 along with the estimate obtained in Reference 3 which is given with the innermost circle in the Figure. It is clear that all the estimates are within the exact stability region and they constitute an increasing sequence.

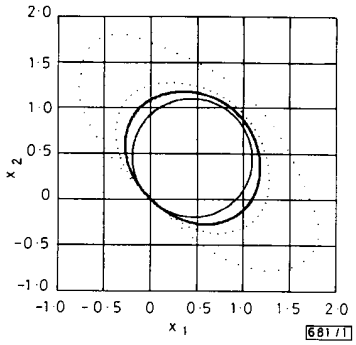

Fig. 1 *Stability region estimates of equilibrium point (0.454, 0.454) of two-neuron network for 10, 20 and 50 iterations of algorithm*

——— estimates in Reference 6
——— 10 iterations
····· 20 iterations
• • • 50 iterations

Conclusions: In this note, a constructive method is applied for the estimation of the stability regions of the stable equilibria of recurrent analogue neural networks. The estimation method gives an increasing sequence of estimated stability regions and is based on local Lyapunov functions. The scheme is demonstrated on a neural network architecture and shown to give better stability region estimates than a previously proposed method.

Acknowledgments: The first author would like to thank H. D. Chiang for introducing him to stability concepts in nonlinear systems. The work of the first author was partly sponsored by Westinghouse Education Foundation under grant No: 3597262 when he was with the Electrical Engineering Department of Syracuse University, Syracuse, NY, USA.

10th April 1992

K. Ciliz and A. Harova (*Electrical Engineering Department, Boğaziçi University, Bebek, Istanbul 80815, Turkey*)

References

1 HOPFIELD, J. J.: 'Neurons with graded response have collective computational properties like those of two-state neurons'. Proc. Natl. Acad. Sci. USA, May 1984, **81**, pp. 3088–3092
2 TANK, D. W., and HOPFIELD, J. J.: 'Simple neural optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit', *IEEE Trans.*, May 1986, **CAS-33**, pp. 533–541
3 MICHEL, A. N., FARRELL, J. A., and POROD, W.: 'Qualitative analysis of neural networks', *IEEE Trans.*, February 1989, **CAS-36**, pp. 229–243
4 CHIANG, H.-D., and THORP, J. S.: 'Stability regions of nonlinear dynamical systems: a constructive methodology', *IEEE Trans.*, December 1989, **AC-34**, pp. 1229–1241
5 MICHEL, A. N., SARABUDLA, N. R., and MILLER, R. K.: 'Analysis of complex dynamical systems: some computational methods', *Circuits, Systems and Signal Processing*, 1982, **1**, (2), pp. 171–202

SUBTHRESHOLD AND ABOVE THRESHOLD GATE CURRENT IN HETEROSTRUCTURE INSULATED GATE FIELD-EFFECT TRANSISTORS

F. Schuermeyer, E. Martinez, M. Shur, D. Grider and J. Nohava

Indexing terms: Field-effect transistors, Transistors, Semiconductor devices and materials

Experimental data are presented which show that the gate current at zero drain bias in HIGFETs changes qualitatively when the gate voltage is varied from below to above threshold. Above threshold, the gate current flows from the entire channel, and below threshold, the gate current flows from the gate edges. A new independent method for deducing threshold voltage is introduced.

Gate leakage current presents an important limitation to the performance of heterostructure insulated gate field-effect transistors (HIGFETs) because it may restrict the maximum voltage swing. Several researchers have studied the gate current in these devices [1–3]. One of the most important results of these studies was the conclusion that the dependence of the gate current on the gate-to-source voltage changes at threshold. This change was related to the change in the band diagram at threshold. Below threshold, the gate current is limited by the Schottky barrier height, and above threshold the gate current is controlled by the conduction band discontinuity in n-channel devices and by the valence band discontinuity in p-channel devices. As a result, the gate current is considerably diminished in HIGFETs with a large mole fraction of Al in the AlGaAs layer separating the gate from the channel.

We present the results of measurements of the gate current in n-channel HIGFETs with gate lengths ranging from 0·6 to 1·0 μm. The results of this study show that the distribution of the gate current along the channel also changes dramatically at threshold. Below threshold, the gate current flows from the source and drain contacts because of the barrier separating the two ohmic contacts in the channel. Hence, the gate current is practically independent of the gate length. Above threshold, the gate current flows from the entire channel. Hence, in the above threshold regime, the gate current increases with the gate length.

The devices used in this study were fabricated using a self-aligned gate process described previously [4]. The C-HIGFET devices and circuits were fabricated employing a refractory metal (WSi) selfaligned-gate process. Projection optical lithography was used to define all of the features during the 3 inch wafer fabrication process. The first step in this C-HIGFET fabrication process is the growth of the heterostructure on a semi-insulating GaAs wafer by molecular beam epitaxy (MBE). Immediately after removing the wafer from the MBE growth chamber, WSi_x gate metal was deposited and the gates were defined using reactive ion etching (RIE). Selfaligned silicon (n^+) and beryllium (p^+) were used to form the source and drain contacts in the n- and p-HIGFETs, respectively. Following the activation of these implants using a rapid optical anneal technique, an oxygen implant/activation is used for device isolation. Source/drain ohmic contacts were formed to n-HIGFETs and p-HIGFETs using Au/Ge and Au/Zn, respectively. Finally, a gold-based two-level-metal interconnect scheme was used to connect the HIGFET devices to make complementary HIGFET circuits.

We have measured the gate current at zero drain-to-source voltage to eliminate the effect of the drain bias on the gate current. The measured dependences of the drain current on the gate voltage for transistors with gate lengths 0·6, 0·8, 1·0, and 1·2 μm are shown in Fig. 1*a* and *b* for temperatures of 298 K, and 198 and 398 K, respectively. At such low drain voltages, the drain current is proportional to the carrier concentration in the channel and the device threshold voltage can be determined using the definition of threshold current I_t, given by the unified charge control model [5]

$$I_t = \frac{qn_0 \mu V_{th} W}{L} \quad (1)$$