BioSystems xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons

Christopher L. Buckley*, Seth Bullock

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK

ARTICLE INFO

Article history:
Received 23 April 2007
Received in revised form 2 November 2007
Accepted 23 May 2008

Keywords:
Signal propagation
Centre crossing networks
Linear stability analysis
May-Wigner threshold
Network dynamics

ABSTRACT

In this paper we demonstrate that signal propagation across a laminar sheet of recurrent neurons is maximised when two conditions are met. First, neurons must be in the so-called centre crossing configuration. Second, the network's topology and weights must be such that the network comprises strongly coupled nodes, yet lies within the weakly coupled regime. We develop tools from linear stability analysis with which to describe this regime in terms of the connectivity and weight strengths of a network. We use these results to examine the apparent tension between the sensitivity and instability of centre crossing networks.

1. Signal Propagation

Signal propagation is central to the control systems of all adaptive agents in that it is crucial for the effective transduction of sensory input into motor output. Biological systems seem to achieve successful signal propagation over extended networks of neurons with relative ease. Feedforward neural architectures have been employed to investigate how signals propagate across networks and can construct complex mappings between input and output (Litvak et al., 2003). However, in general, biological neural networks are recurrent, even in systems that have previously been idealised as feedforward in nature, e.g., the columns within the visual cortex have recurrent connections within and between layers (Carlson, 1991). Signal propagation across such recurrent networks is likely to be more complex than in feedforward networks, where it is taken for granted.

There has been a deal of speculation in neuroscience concerning mechanisms that could promote signal propagation across a sequence of neurons (Turrigiano, 1999). One set of ideas involves the behaviour of nodes that tend to interact at the centre of their operating ranges. In general, networks of such neurons are thought to be computationally rich. More specifically, in this regime, nodes are maximally sensitive to input, potentially facilitating more efficient signal propagation across extended networks. Moreover, Turrigiano (1999) describes how *homeostatic processes* (HPs) might actively "keep neurons at the centre of their operating ranges" (Turrigiano, 1999).

Inspired by this work, Williams (2006) studied how an abstraction of these HPs affected the ability of a continuous time recurrent neural network (CTRNN) to propagate signals. In this work, HP provided a simple feedback mechanism that altered the gain and bias of a node such that its input tended to lie at the centre of its transfer function. He hypothesized that networks composed of such nodes would be better able propagate signals, because local HP at the level of individual nodes would drive networks into the most sensitive region of their dynamics.

Williams found that HP drove systems toward a configuration that has been identified as significant within the CTRNN literature. In this so-called "centre crossing configuration" all nodes in a CTRNN interact at the centre of their sigmoid transfer functions (Mathayomchan and Beer, 2002), a mathematical property that bears close resemblance to the biological ideas highlighted by Turrigiano, amongst others. Williams also demonstrated that signal propagation was improved within such centre crossing networks.

However, this signal propagation was impoverished within larger networks, and did not approach the performance achieved by an equivalent feedforward architecture even for small networks (personal communication). One possible reason for these results can be induced from the original work on centre crossing CTRNNs (Mathayomchan and Beer, 2002). Here, it was demonstrated that the generation of rhythmic patterns evolved more readily in such networks. This is due to the fact that centre crossing networks are likely to produce oscillatory dynamics. Such oscillatory behaviour is likely to corrupt the transmission of signals across extended net-

* Corresponding author.
E-mail address: C.L.Buckley@soton.ac.uk (C.L. Buckley).

0303-2647/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2008.05.026

works and explain why such networks would be outperformed by feedforward networks that do not exhibit such autonomous oscillations.

There seems to be a tension between these two accounts of the utility of centre crossing networks. On the one hand, their sensitivity helps to prevent signal loss while, on the other, they are prone to interfering reverberant activity. Here, we attempt to resolve this conflict through the use of linear stability analysis and random matrix theory. We start by briefly outlining the theoretical tools that we employ before applying them to a simple model of signal propagation across a laminar CTRNN. First, we show that the May-Wigner stability theorem, originally derived for linear systems, constitutes a bound on the stability of nonlinear CTRNNs. Then we go on to demonstrate that signal propagation is maximised just below the May-Wigner threshold for a number of CTRNN topologies. We consider these results in relation to the centre crossing ideas developed within research on both natural and artificial neural networks.

2. Linear Stability Analysis

In this section, we show that it is possible to obtain insights into the properties of an ensemble of nonlinear systems using linear stability analysis (LSA). Specifically, we claim that this type of analysis allows us to demarcate stable and unstable regions in the CTRNN parameter space. We start by demonstrating that the absolute strength of the coupling around all CTRNN equilibria is bounded by their weights. We then present results that allow us to describe a stable region in the parameter space of linear networks with weights that are normally distributed. Finally, we combine these insights and argue that they allow us to numerically and analytically calculate a bound for the stability of CTRNNs.

2.1. Coupling in a CTRNN at Equilibrium

The CTRNN (Beer, 1995) or leaky integrator, equation for N nodes is given by

$$\dot{y}_i = -y_i + \tanh\left(\sum_{j=1}^{N} \omega_{ij} y_j + \theta_i + I_i\right) \tag{1}$$

where y_i represents activation at the ith neuron, ω_{ij} is the weight on the connection between neurons i and j, θ_i is the bias value at the ith neuron, and I_i is the sensory input at node i and is assumed to be zero unless specified. Note: in this work, for simplicity, we neglect the usual timescale parameter τ_i. Hence each CTRNN is fully parameterized by the set of weights and biases $[\Omega, \Theta]$.

Analysis of the dynamics of these systems is difficult even in the two-node case, see Beer (1995). However progress can be made by examining the dynamics around such a system's equilibrium positions (Strogatz, 1994). For example consider a two-node CTRNN given by

$$\begin{aligned} \dot{y}_1 &= -y_1 + \tanh(\omega_{11} y_1 + \omega_{12} y_2 + \theta_1) \\ \dot{y}_2 &= -y_2 + \tanh(\omega_{21} y_1 + \omega_{22} y_2 + \theta_2) \end{aligned} \tag{2}$$

Now the equilibrium positions of this system correspond to points in phase space where both of the derivatives with respect to time of the system are equal to zero. Setting the LHS of each of Eq. (2) to zero gives the equations

$$\begin{aligned} y_2 &= \frac{\operatorname{atanh}(y_1) - \omega_{11} y_1}{\omega_{12}} - \frac{\theta_1}{\omega_{12}} \\ y_1 &= \frac{\operatorname{atanh}(y_2) - \omega_{22} y_2}{\omega_{21}} - \frac{\theta_2}{\omega_{21}} \end{aligned} \tag{3}$$

where $\operatorname{atanh}(x)$ is just the inverse of the hyperbolic tangent function, $\tanh(x)$. Plotting the resulting curves, which are known as the *nullclines* of the system, yields Fig. 1. The equilibrium positions of the system are given by the intersection of the curves. In general there may be multiple equilibria but let us inspect the dynamics around one particular equilibrium $(y_1 = y_1^*, y_2 = y_2^*)$.

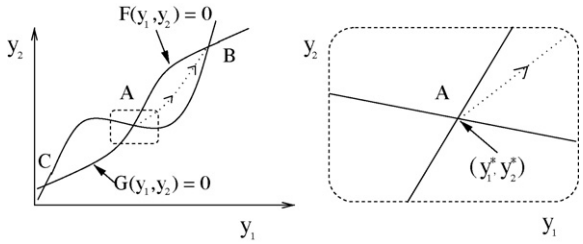

Fig. 1. A schematic of the nullclines of Eq. (2) plotted in phase space. The left-hand panel shows both nullclines intersecting at three points. (A) is an unstable equilibrium and (B) and (C) are stable equilibria. The dotted line marks the system trajectory through the phase space which starts from an initial condition very close to (A) and then diverges toward (B). The right-hand panel shows an enlargement of the region marked by the dotted box in the left-hand panel. In this region the nullclines are approximately linear.

In order to determine the dynamics of the system let us continue and linearize the system around a general equilibrium (y_1^*, y_2^*). To do this we construct the Jacobian (Strogatz, 1994; Murray, 1989) of the system as

$$J = \begin{pmatrix} \omega_{11}^{\text{eff}} - 1 & \omega_{12}^{\text{eff}} \\ \omega_{21}^{\text{eff}} & \omega_{22}^{\text{eff}} - 1 \end{pmatrix}_{y_1^*, y_2^*} \tag{4}$$

where for notational ease we have made the following substitutions

$$\omega_{ij}^{\text{eff}} \equiv \omega_{ij} \frac{\mathrm{d}[\tanh(U_i)]}{\mathrm{d}U_i} \tag{5}$$

and

$$U_1 = \omega_{11} y_1 + \omega_{12} y_2 + \theta_1 U_2 = \omega_{21} y_1 + \omega_{22} y_2 + \theta_2 \tag{6}$$

The Jacobian consists of a set of *effective weights* $(\omega_{ij}^{\text{eff}})$ that comprise a matrix that constitutes the operator of a linearised system describing the dynamics in a local region around the equilibrium (y_1^*, y_2^*); see right-hand panel of Fig. 1. These effective weights not only depend on the *actual weights* but are also modified by the parameter θ_i and, more generally, by the equilibrium position (y_1^*, y_2^*) through Eqs. (5) and (6). Intuitively one can think of this as a modification of the linearised interaction of the variables that depends on the slope of their transfer functions around the equilibrium position. For example if the equilibrium of a system lies at the extremities of two units' transfer functions (e.g. $y_1^* = 0.9$ and $y_2^* = 0.9$) then they would interact in a much weaker way than if the equilibrium were at the centres of their transfer functions (e.g. $y_1^* = 0$ and $y_2^* = 0$). Consequently the former would have low effective weights while the latter would have high effective weights.

This process is easily generalised to an N-node CTRNN. Specifically its possible to linearise the dynamics of an N-node CTRNN around the general equilibrium $\mathbf{y}^* \equiv (y_1^*, \ldots, y_N^*)$. This yields a Jacobian of the form

$$J = \begin{pmatrix} \omega_{11}^{\text{eff}} - 1 & \cdots & \omega_{1N}^{\text{eff}} \\ \vdots & & \vdots \\ \omega_{N1}^{\text{eff}} & \cdots & \omega_{NN}^{\text{eff}} - 1 \end{pmatrix}_{\mathbf{y}^*} \tag{7}$$

where we have made the following substitutions

$$\omega_{ij}^{\text{eff}} \equiv \omega_{ij} \frac{\mathrm{d}[\tanh(U_i)]}{\mathrm{d}U_i} \tag{8}$$

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026

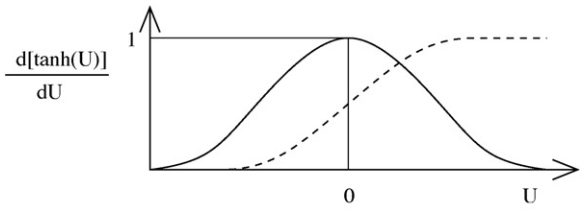

Fig. 2. A typical sigmoidal transfer function, here, the hyperbolic tangent (the dashed line) and its first derivative, in this case $\text{sech}^2(x)$ (the solid line). By definition, this derivative reaches a maximum coincident with the maximum gradient of the sigmoidal transfer function which, for any canonical transfer function, will occur for $U = 0$.

and

$$U_i = \sum_{j=1}^{N} \omega_{ij} y_j + \theta_i \tag{9}$$

It can be shown that the stability of the system around $\mathbf{y}^*$ is completely determined by the eigenvalues of the Jacobian (Strogatz, 1994). Specifically, a system is said be stable (i.e., it will quickly return after small perturbations) if *all* the real parts of the eigenvalues of the Jacobian are negative. Conversely, it will be unstable (i.e., small perturbations from equilibrium will diverge away) if *any* of the real parts of the eigenvalues of the Jacobian are positive (Strogatz, 1994). Consequently, for any given equilibrium in any given network it is possible to numerically calculate its stability by calculating, and then inspecting, its eigenvalues. We shall use this technique extensively in later sections.

The Jacobian for this system, and hence its stability, is dependent on $\mathbf{y}^*$, Θ and Ω; see Eqs. (8) and (9). However, it is possible to determine an upper bound on the absolute values of the entries of the Jacobian in terms of only the weight values, Ω. Specifically, the contribution of $\mathbf{y}^*$ and Θ to the Jacobian is constrained by the maximum value of the first derivative of the transfer function (in this case the hyperbolic tangent function). Now Fig. 2 shows how a hyperbolic tangent function, and its first derivative, $\text{sech}^2(x)$, vary with their arguments. The latter reaches a maximum value of unity when $x = 0$ and tends towards zero on either side.

We can therefore deduce that the maximum absolute values of the effective weights will be achieved when this function evaluates to unity in which case they will equal the original network weights, i.e.,

$$\text{Max}\,[|\omega_{ij}^{\text{eff}}|] = |\omega_{ij}| \tag{10}$$

As such, the maximum possible absolute magnitude of the coupling around *any* equilibrium in a nonlinear system with weights Ω is just determined by the weights themselves. Variation in the equilibrium position $\mathbf{y}^*$ and the biases Θ may only reduce the coupling from this value. Furthermore, it is centre crossing networks that achieve this maximum possible absolute coupling, since, for such networks, nodes interact at the centre of their transfer functions, where the derivative is maximal. We present a more detailed version of this argument elsewhere (Buckley, 2007).

2.2. The May-Wigner Threshold

In order to characterise general conditions for stability in CTRNNs, we will make use of work originally carried out by Gardner and Ashby (1970) and later formalized by May (1972). In a now classic study, Gardner and Ashby (1970)[1] investigated stability criteria for large complex systems in terms of the effect of a network's size, connectivity and weight strength on its tendency to exhibit a stable point attractor. Gardner and Ashby considered linear networks of the following form

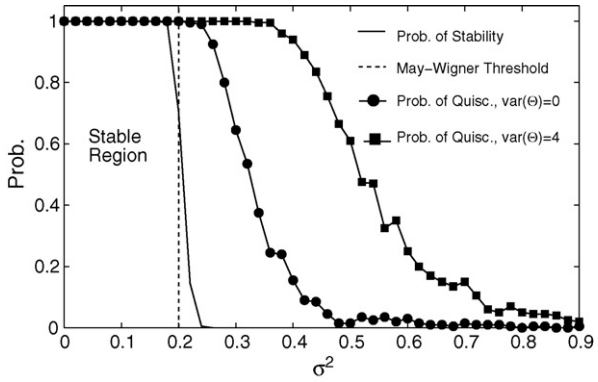

Fig. 3. Curves represent the impact of variance in a network's weight strengths, σ^2, on the probability of network stability for three different classes of network, with the vertical line indicating the prediction made by the May-Wigner threshold. First (solid line), probability of stability in linear networks comprising $N = 100$ nodes of the kind studied by Gardner and Ashby (1970). Each network is fully connected and is defined by a normally distributed weight matrix with variance σ^2 and zero mean, normally distributed biases with variance var(Θ) and zero mean, and self-weights $\omega_{ii} + \beta$, where $\beta = -1$ (see text for details). The probability of stability is determined by inspection of the numerically calculated eigenvalues of the Jacobian. Second (circles), fully connected networks of 100 unbiased CTRNN nodes (i.e., var(Θ) = 0) with connection weights specified by the same matrices. Third (squares), the same sample of CTRNNs with node biases drawn from a normal distribution with zero mean and variance, var(Θ) = 4. The proportion of networks that do not exhibit oscillatory behaviour is employed as a proxy for stability. Each data point is the average of 50 network realisations, and the error for all data points is less than 2%.

$$\dot{y}_i = -y_i + \sum_{j=1}^{N} \omega_{ij} y_j \tag{11}$$

They employed a numerical method of determining stability in networks of varying size, N, and connectivity, C. For example consider a network where the entries of Ω are drawn from a normal distribution with zero mean, have a variance, σ^2, and where a small negative value β is added to each self-weight to stabilise the system. Gardner and Ashby (1970) found that the probability of network stability, p, falls with increasing network size. Furthermore, it was observed that networks have a high probability of stability if either σ^2 or C are low, and that this probability decreases with increasing C or σ^2 (see Fig. 3, solid line).

Later, May (1972) was able formalize these findings using analytical results from random matrix theory (Wigner, 1959; Mehta, 1967). He was able to derive a critical threshold below which any network has a high probability of stability. Consequently it is possible to show that in the limit of large system size ($N \gg 1$) a system is almost certainly stable if

$$\sqrt{NC}\sigma < 1 - \beta \tag{12}$$

This result, often referred to as the May-Wigner stability theorem, corresponds well with Gardner and Ashby's original findings and still holds as a very important threshold (Sinha and Sinha, 2005). Note: here we have made a small amendment to May's original formulation by incorporating the term β to account for the inclusion

[1] Solow et al. (1999) point out an error in this paper. However, this error only constitutes a quantitative correction to the paper's numerical results and does not impact on the overall message of the paper.

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026

of an additional self-weight term, see Buckley (2007). From Eq. (12) we can then determine critical thresholds for both the variance and connectivity. Specifically, a system is almost certainly stable if the variance, σ^2, conforms to the following inequality

$$\sigma^2 < \sigma^2_{MW} \equiv \frac{(1-\beta)^2}{NC}. \quad (13)$$

Or, equivalently, a system is almost certainly stable if the connectivity, C, conforms to

$$C < C_{MW} \equiv \frac{(1-\beta)^2}{N\sigma^2}. \quad (14)$$

Fig. 3 reconfirms the canonical result that the impact of variance in a network's weights on its probability of stability (as determined numerically for a sample of networks) agrees well with the analytically derived May-Wigner threshold. In order to make use of this powerful result, we will need to consider how it might be applied to the nonlinear CTRNN.

Note: several papers have claimed that some of the conclusions in May (1972) are incorrect. These concern the observations about the stability of modular systems (Solow et al., 1999) or the fact that there are exceptions to the prediction of instability in the limit of large system size (Cohen and Newman, 1985). These criticisms do not alter the overall message of the paper nor the derivation of the May-Wigner threshold and consequently are not considered in the work presented here

2.3. The Stability of CTRNNs

Given the results of the last two sections, it is now possible to determine general conditions for the stability of CTRNNs with normally distributed weights. We will achieve this by using linear analysis tools to consider their (nonlinear) behaviour around an arbitrary equilibrium, $\mathbf{y}^*$. Consider a fully connected CTRNN with N nodes and weights normally distributed with variance σ^2 and $\beta = -1$. Here, the network biases, Θ, are similarly distributed normally with zero mean and variance var(Θ).

While the May-Wigner threshold applies directly to systems of linear elements, the nonlinear transfer functions of a CTRNN require us to employ the reasoning introduced in Section 2.1. Since the absolute values of the effective weights associated with a particular equilibrium in a network's dynamics cannot be greater than that of their corresponding actual weights ($|\omega^{eff}_{ij}| \leq |\omega_{ij}|$), and these weights are normally distributed with zero mean, the variance of the effective weights cannot be greater than that of the actual weights ($\sigma^2_{eff} \leq \sigma^2$; see Appendix A). Hence, if the variance of a network's actual weights lies below the May-Wigner threshold ($\sigma^2_{act} < \sigma^2_{MW}$), so will the variance of any set of effective weights ($\sigma^2_{eff} < \sigma^2_{MW}$). There is a slight complication here, since the May-Wigner threshold does not speak to weight distributions with non-zero mean, or weight distributions that are non-normal. However, the crucial consideration is that each eigenvalue associated with any network equilibrium is never more likely to have a positive real part than the equivalent eigenvalue associated with the centre crossing equilibrium. Consequently, for any CTRNN where $\sigma^2_{act} < \sigma^2_{MW}$, each and all of its equilibria will have a high probability of stability. Furthermore, the maximum absolute coupling and hence the least stable network equilibrium will occur for an equilibrium at the centre crossing point, i.e., when all nodes intersect at the centre of their sigmoidal functions such that $\omega^{eff}_{ij} = \omega_{ij}, \forall i, j$.

The network simulation results presented in Fig. 3 are consistent with this line of reasoning, suggesting that it is valid to co-opt these linear stability results in order to characterise the onset of instability in nonlinear CTRNN systems. Here, networks of 100 nodes with weights and biases distributed as described above

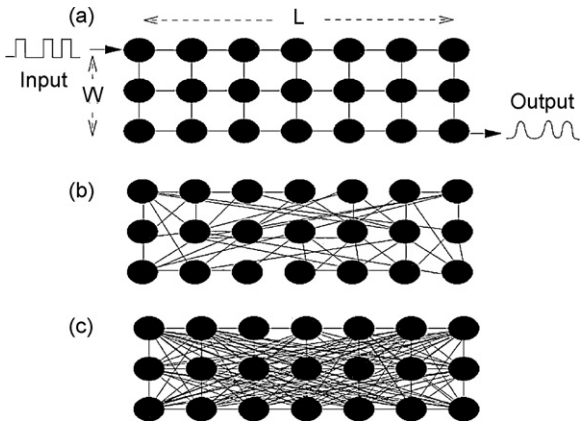

Fig. 4. A laminar sheet of N CTRNN nodes arranged in an array with width, W, and length, L, is driven by a square wave input signal at one corner node. The correlation between this input signal and the output taken from the diametrically opposed node is measured for three different topologies: (a) a rectangular lattice, (b) the same lattice randomly rewired such that every node is assigned $k = 4$ incoming edges at random, but out degree is free to vary, (c) a fully connected network.

were forward integrated for 1000 time steps with a Euler step of $\delta t = 0.05$. As an indication of instability in these networks, oscillatory behaviour is tested for by measuring the average deviation of each variable from the mean after a transient period of 500 time steps. The May-Wigner threshold and a numerical calculation of stability around the centre crossing point (i.e., the point where the coupling around equilibrium is maximal) are also presented. Both for cases where nodes are unbiased (var(Θ) = 0, circles) and where biases vary (var(Θ) = 0, squares), no simulated network below the May-Wigner threshold exhibits oscillatory behaviour. Clearly, networks may be unstable in other ways, but the fact that we see no oscillatory behaviour below the threshold can be taken as reasonable evidence for the stability in that region. CTRNNs in this stable region bounded by the May-Wigner threshold can be considered to be examples of *weakly coupled systems* which are studied throughout neuroscience and are considered to be good models of the dynamics of networks of neurons in many parts of the nervous system (Hoppensteadt and Izhikevich, 1997).

3. Signal Propagation in a Recurrent Sheet of CTRNN Nodes

Here we examine signal propagation across laminar sheets of CTRNN nodes utilising the tools developed within the previous section. Each sheet consists of $N = 60$ nodes arranged in a $L \times W$ rectangular array. The networks are connected according to various topologies, see Fig. 4. Each connection within the network (i.e., the value of each entry in the weight matrix, Ω) is drawn from a normal distribution with zero mean and variance σ^2. Similarly, the biases of each network are drawn from a normal distribution with zero mean and variance var(Θ). The self-weight parameter, β, is set to negative one for all networks such that each node is intrinsically stable. All networks are forward integrated with an Euler step of $\delta = 0.05$.

A square wave signal is applied to the input node $i = 1$. This comprises intervals of low stimulation, $I_1 = 0$, for periods uniformly distributed over the interval [50, 400] time steps, and high stimulation, $I_1 = 1$, with length uniformly distributed over the interval [50, 200] time steps, see the top two panels of Fig. 5. We measure the correlation between the input signal and the output signal from the diametrically opposite node, see Fig. 4. Note: calculating correlation involves scaling each signal by its variance and is therefore insensitive to the absolute magnitude of the signal. However, here we apply a small magnitude noise signal to each node ($\approx 10^{-6}$) at

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026

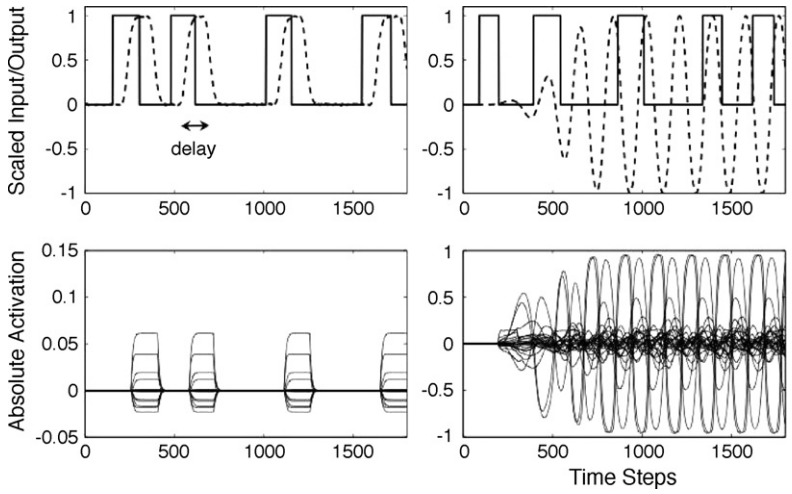

Fig. 5. Plots of network activity over time for the lattice network reported in Fig. 4 parameterised below the May-Wigner threshold (left-hand panels) and above it (right-hand panels). The two top panels show the input signal and the scaled output signal, solid and dashed lines respectively. The bottom two panels show a representative selection of the absolute activation values for all nodes. Note the difference in scale of *y*-axes on the bottom pair of graphs. The delay between the input and output signal is marked on the top-left panel.

every time step, which effectively masks any correlation between the input and extremely small output signals. Finally, the phase delay between input and output signal imposed by the shortest path length separating the input node from the output node is corrected for such that, for every measurement, the correlation is maximised, see the top left panel of Fig. 5.

First let us consider signal propagation across networks in a centre crossing configuration. Unlike the more complicated criterion for centre crossing configuration employed by (Mathayomchan and Beer, 2002), for the system described by Eq. (1) a centre crossing configuration can be straightforwardly obtained by setting $\theta_i = 0, \forall i$. This condition ensures that there is a network equilibria when all node activations are zero, $\mathbf{y}^* = [0]$. Note: for a given parameter set there is only one centre crossing configuration.

Fig. 5 shows typical traces of the input, output and inter node activations for a lattice network (see Fig. 4a). The two left-hand panels depict the dynamics of a lattice parameterised to lie within the weakly coupled region below the May-Wigner threshold. The output signal closely maps the input with some consistent delay, but the absolute magnitudes of the node activations are very small, since the signal is significantly attenuated as it traverses the lattice. As a result, signal propagation performance is critically dependent on the scale of any noise within the system. For systems with small weight values, the output signal is so small that it is washed away by the internal noise injected at each node. The two right-hand panels depict the dynamics associated with a lattice parameterised to lie above the May-Wigner threshold. Networks in this region exhibit high magnitude complex dynamics unrelated to the input signal. In general the absolute value of the propagated signal increases with weight variance. Note the difference in scale on the y-axes of the lower panels.

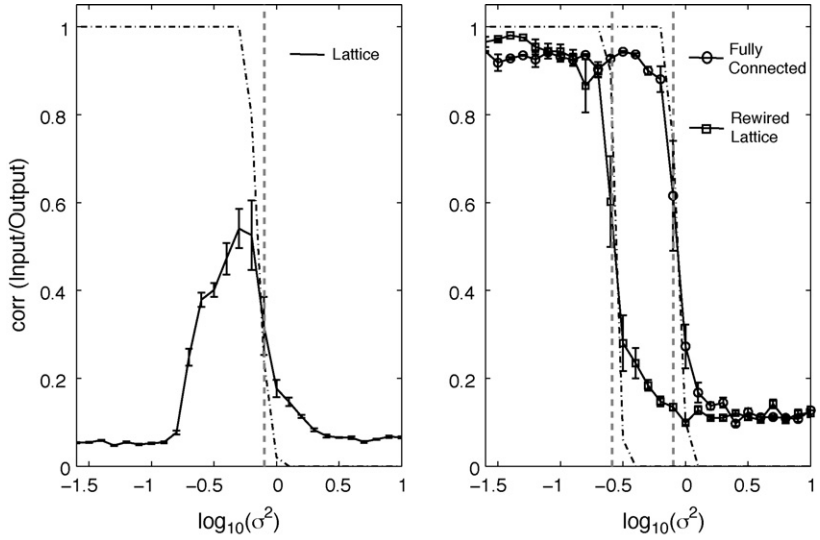

Fig. 6. The input/output correlation, corr(Input/Output), versus the log of the variance of the weights, $\log_{10}(\sigma^2)$ for rectangular laminar networks with length ($L = 15$) and width ($W = 4$) and all biases, Θ, set to zero. The solid line in the left-hand panel and the circles and squares in the right-hand panel show the correlation for a lattice network (see Fig. 4a), randomly rewired lattice network (see Fig. 4b) and fully connected network (see Fig. 4c), respectively. The dot-dashed lines are the respective numerically calculated probabilities of stability, and the vertical lines represent the analytically derived May-Wigner thresholds. Each data point is calculated as the average if 50 network realisations. Error-bars are given for each measurement of correlation.

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026

Fig. 6 shows how the input/output correlation, corr(Input/Output), varies with the log of the variance of the weights, $\log_{10}(\sigma^2)$ in a lattice with $L \times W = 15 \times 4$ for the three different network connection topologies given in Fig. 4. Note: the connectivity, C, in each network is dictated by their structure. Specifically for both the lattice and the rewired lattice each node has $k = 4$ incoming connections which is equivalent to a connectivity of $C = Nk/(N^2 - N)$ while the fully connected network has connectivity of $C = 1$.

The left-hand panel presents results for a lattice network (see Fig. 4a), and shows that the input/output correlation rises and then falls with the variance of the weights. More specifically, there is an intermediate region where the coupling between nodes is high enough to resist signal attenuation, but low enough to avoid instability. This "sweet spot" is located just below the May-Wigner threshold (the vertical dashed line given by Eq. (13)).

The right-hand panel of Fig. 6 presents results for a rewired lattice (Fig. 4b) and a fully connected network (Fig. 4b). For these topologies, the short path length between input and output nodes ensures that the signal attenuation problem suffered by the lattice is not as significant. As a result, high input/output correlation can be achieved with low weight variance and furthermore the peak correlation achieved is appreciably higher than in the lattice. However, the figure confirms that signal propagation still falls sharply above the May-Wigner threshold for these networks, despite the potential advantage conferred by their short minimum path lengths. Interestingly the input/output correlation can still be impoverished even if a direct connection between input and output nodes exists. This is because in an unstable network connections incident on the output node will act as a source of noise.

Across the different network topologies explored here, the fall in performance associated with high weight variance is well predicted both by the numerically calculated probability of stability and the analytically calculated May-Wigner threshold, further supporting the arguments made in Section 2.3. Specifically, as the weight variance exceeds this threshold, reverberant oscillation and node saturation associated with the unstable regime destructively interferes with the transmission of information.

It is clear from Fig. 6 that path length as well as network stability impacts on the signal propagation in CTRNNs. Consequently, here we explore this relationship in more detail. To do this we start by constructing a 2-D lattice of size $L \times W$ (see Fig. 4a) where each node has $k = 4$ incoming connections (in degree). Each connection then has a probability, p_r, of being randomly rewired to another node while preserving the in degree at each node. This rewiring process alters the network from a lattice ($p_r = 0$), with a long minimum path between input and output, to a random graph ($p_r = 1$) with a much shorter minimum path. This method is similar to that employed by Strogatz (1994) in investigating the small-world phenomenon. Note: the rewiring process preserves both the connectivity and the variance of the weights of each network and hence does not effect their stability (see Sinha, 2005 for a detailed discussion of this).

Fig. 7 shows how the correlation between input and output varies with log of the rewiring probability for rewired lattices of size $L \times W = 15 \times 4$ (solid line), 30×2 (circles), and 60×1 (squares). The variance of the weights, σ^2, is set according to Fig. 6 such that it maximises signal propagation.

The dotted line gives an indication of the path length between input and the output. This is calculated as the ratio between actual path length, λ, and the path length in the original lattice, $\lambda_0 \equiv \lambda(p_r = 0)$.

The poorest signal propagation occurs in the longest lattice (i.e., $L = \times W = 60 \times 1$ and $p_r \to 0$) because this has the largest path length between input and output and consequently the signal is highly attenuated. Both random rewiring and shorter lattice length have a beneficial effect on signal propagation because they reduce the path length between input and output and consequently reducing signal attenuation.

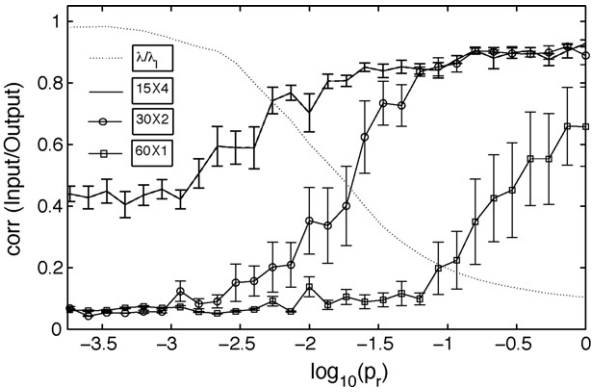

Fig. 7. The input/output correlation, corr(Input/Output), versus the log of the rewiring probability, $\log_{10}(p_r)$, for randomly rewired lattices of size $L \times W = 15 \times 4$ (solid line), 30×2 (circles) and 60×1 (squares). All biases, Θ, are set to zero. The dotted line gives an indication of the path length between input and output. This is calculated as the ratio between shortest path length between input and output, λ, and the path length between the input and output in a lattice, λ_0 (i.e., when $p_r = 0$). Each data point is calculated as the average over 50 network realisations. Error-bars are given for each measurement of correlation.

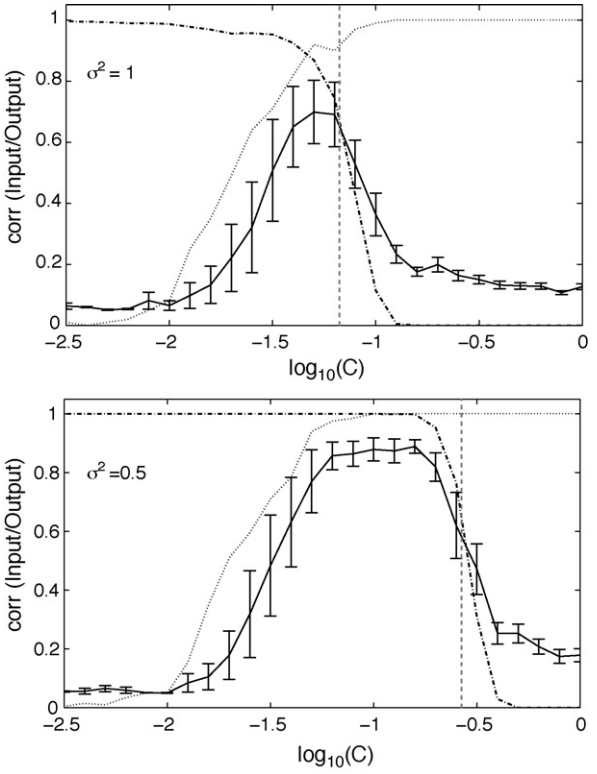

Fig. 8. The input/output correlation, corr(Input/Output), versus the log of the connectivity, $\log_{10}(C)$, for two random graphs. The top and bottom panels present the results for networks with variance $\sigma^2 = 1$ and $\sigma^2 = 0.5$, respectively. All biases, Θ, are set to zero. The dot-dashed lines are the respective numerically calculated probabilities of stability, and the vertical lines represent the analytically derived May-Wigner thresholds. The dotted line gives the probability that a route between input and output exists. This is calculated as the ratio between the number of network with and without routes between input and output for a given connectivity, C. Each data point is calculated as the average over 50 network realisations. Error-bars are given for each measurement of correlation.

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026

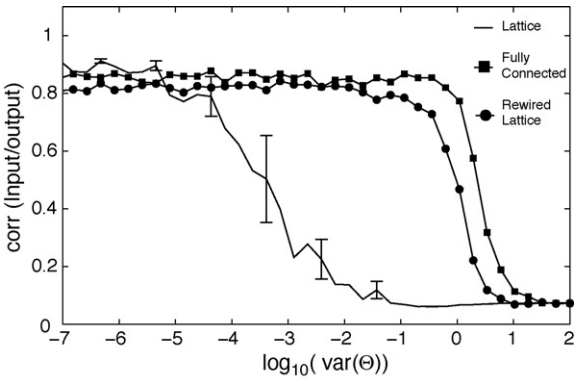

Fig. 9. The input/output correlation, corr(Input/Output), versus the log of the variance of the biases, $\log_{10}(\Theta)$, for networks of size, $L \times W = 15 \times 4$, connected as a lattice (solid line), rewired lattice (circles) and fully connected (squares). All networks have weight variance, σ^2, which maximises the signal propagation across unbiased networks. Each data point is calculated as the average of 50 network realisations and representative standard deviations are given by the error bars on the solid line.

Fig. 8 shows how the correlation between input/output varies with connectivity. Each network consists of $N = 60$ nodes where each pair of nodes is connected with probability C. The top and bottom panels show the results for network weights normally distributed with variance $\sigma^2 = 1$ and $\sigma^2 = 0.5$ respectively. In both panels the input/output correlation rises and then falls with the connectivity. Again the onset of the decline in signal propagation is well predicted by both the numerical calculation of stability (dot-dashed line) and the May-Wigner threshold (vertical dashed line given by Eq. (14)).

The dotted line in both plots gives the probability that a path exists between input and output. For an ensemble of networks with a given connectivity this is calculated as the ratio between the number of networks with and without a path between input and output. From Fig. 8 we can see that this property constitutes a lower bound on the signal propagation across the network. For example, for $\sigma^2 = 1$ (the top panel) the peak input/output correlation is less than the equivalent value when $\sigma^2 = 0.5$ (the bottom panel). This is because at the sweet spot just below the May-Wigner threshold there is a lower probability of a path existing between input and output for $\sigma^2 = 1$ (the top panel) than for $\sigma^2 = 0.5$ (the bottom panel). Again there exists an intermediate region for which signal propagation is maximised.

How do these results generalise to networks that are not in a centre crossing configuration? Fig. 9 shows how the input/output correlation varies with the log of the variance of the biases, $\log_{10}(\mathrm{var}(\Theta))$, for the three different network topologies. In each case, the variance of the weights, σ^2, is set according to Fig. 6 such that it maximises signal propagation for unbiased networks. In all cases, increasing variance damages signal propagation. Nominally, this result is in line with Williams and Noble (2007).

The effective signal propagation in both the fully connected network and the rewired lattice is more resistant to increasing variance in Θ. This is likely to stem from the involvement of fewer nodes in the path along which the signal propagates. However, the key observation here is that departure from centre crossing configurations does damage signal propagation.

4. Discussion and Conclusion

Here, we have suggested that not only is signal propagation across CTRNNs, and recurrent networks in general, maximised when they are in a centre crossing configuration, but that they must also lie within the weakly coupled regime bounded by the May-Wigner threshold. Signal propagation in networks parameterised such that they lie above this threshold will suffer from interference due to reverberative activity. However, networks that lie too far below the May-Wigner threshold are also unlikely to be effective for signal propagation. Such networks may lack a path of connected nodes linking input and output as a consequence of low connectivity, or, if such a path exists, may suffer from excessive signal attenuation along it as a consequence of low weight values. Of course, both of these problems may be minimised in networks with low diameter, where the length of the minimum path separating input and output is small. However, while such networks may achieve high performance in signal propagation, since we are interested in signal propagation as a proxy for signal transduction, a requirement for the involvement of intermediate nodes that can provide a substrate for successive computational operations is implied, ruling out short path length as a solution to signal transduction in general. These multiple considerations combine to ensure that a region j ust below the May-Wigner threshold is optimal for signal propagation in recurrent centre-crossing networks. This region combines stability (i.e., the absence of reverberant oscillation) with signal strength (i.e., the ability to resist signal attenuation over significant path lengths, see Buckley (2007)).

While this sweet spot may be associated with effective signal p ropagation, one might be concerned that the class of networks within the region are incapable of interesting signal *transduction*. However, there is some suggestion that this worry is unfounded. Echo state machines, for example, are a class of neural network that make use of neurons located within the regime that we define here, and have been shown to be capable of interesting and efficient computation (Jaeger, 2001). Such networks involve a large reservoir of recurrent neurons parameterised at random, but constrained such that they lie in the stable regime consistent with the May-Wigner threshold. While each echo state machine exhibits only a single transiently stable fixed-point attractor, nevertheless, this class of network has been shown to be a powerful computational architecture (Jaeger and Haas, 2004).

The results presented here are founded on the confirmation that the weak coupling regime is well predicted by both the numerical calculations of linear stability analysis and also the analytical derivation of the May-Wigner threshold. That linear tools such as these can be used to make headway in understanding a class of nonlinear network is encouraging, since there is real need to understand the dynamics of CTRNNs.

While some models employ bio-inspired augmentation of CTRNN-style networks (Husbands et al., 1998; Williams, 2004), there is an increasing move in CTRNN research to treat them as arbitrary dynamical systems. Indeed, this is founded on the fact that they have been proven to be universal smooth function approximators (Funahashi and Nakamura, 1993). However, it is one thing to demonstrate that a class of network is capable of arbitrary behaviour in principle, and another to characterise the type of behaviour that such networks are liable, likely, or suitable to exhibit in practice (Bullock, 2006). Consequently, one interesting task is to characterise the parts of CTRNN parameter space that readily enable adaptive behaviour of certain kinds. In particular, how might specific bio-inspired mechanisms (e.g., homeostatic plasticity, neuromodulatory gases, etc.) be associated with regions within this space that scaffold generic dynamics conducive to certain tasks. The work presented here is intended as a contribution towards answering this style of question.

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026

Appendix A

Proof of a bounded variance. In this appendix we shall prove that for a normal distribution with zero mean it is impossible to increase the variance of a data set by any reduction of the absolute magnitudes of any of the data points that comprise it.

Consider a data set (S) of N points with mean and variance given by

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{15}$$

respectively. Consider a transformation of this data set (S) to another ($\hat{S}$) with mean and variance given by

$$\bar{\hat{x}} = \frac{1}{N}\sum_{i=1}^{N} t_i x_i \quad \hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i x_i - \bar{\hat{x}})^2 \tag{16}$$

respectively. Where now each data point is scaled by a value t_i which is constrained over the interval $0 \le t_i \le 1$. Consequently, this transformation can only reduce absolute values of data points. We can now restate the above conjecture as

$$\sigma^2 - \hat{\sigma}^2 \ge 0$$

Substituting in Eqs. (15) and (16) we can obtain

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 - \frac{1}{N}\sum_{i=1}^{N}(t_i x_i - \bar{\hat{x}})^2 \ge 0$$

Collecting and rearranging terms gives

$$\frac{1}{N}\sum_{i=1}^{N} x_i^2(1 - t_i^2) - 2\frac{\bar{x}}{N}\sum_{i=1}^{N} x_i + 2\frac{\bar{\hat{x}}}{N}\sum_{i=1}^{N} t_i x_i - \bar{\hat{x}}^2 + \bar{x}^2 \ge 0$$

Using Eqs. (15) and (16) this becomes

$$\frac{1}{N}\sum_{i=1}^{N} x_i^2(1 - t_i^2) - \bar{x}^2 + \bar{\hat{x}}^2 \ge 0$$

which can be rewritten as

$$\frac{1}{N}\sum_{i=1}^{N} x_i^2(1 - t_i^2) - (\bar{x} - \bar{\hat{x}})(\bar{x} + \bar{\hat{x}}) \ge 0$$

Now, again, using Eqs. (15) and (16) we can rewrite this as

$$\frac{1}{N}\sum_{i=1}^{N} x_i^2(1 - t_i^2) - \left(\frac{1}{N}\sum_{i=1}^{N} x_i(1 - t_i)\right)\left(\frac{1}{N}\sum_{i=1}^{N} x_i(1 + t_i)\right) \ge 0$$

Using the substitutions

$$a_i = x_i(1 - t_i) \quad \text{and} \quad b_i = x_i(1 + t_i)$$

and rearranging we can and obtain

$$\frac{1}{N}\sum_{i=1}^{N} a_i b_i \ge \frac{1}{N^2}\left(\sum_{i=1}^{N} a_i\right)\left(\sum_{i=1}^{N} b_i\right)$$

which is always true by the *Chebyshev sum inequality*. □

References

Beer, R.D., 1995. On the dynamics of small continuous-time recurrent neural networks. Adaptive Behavior 3 (4), 471–511.
Buckley, C.L., 2007. A systemic analysis of the ideas imminent in neuromodulation. Ph.D. Thesis. School of Electronics & Computer Science, University of Southampton.
Bullock, S., 2006. The fallacy of general purpose bio-inspired computing. In: Rocha, L.M., Yaeger, L.S., Bedau, M.A., Floreano, D., Goldstone, R.L., Vespignani, A. (Eds.), Proceedings of the Tenth International Conference on Artificial Life. MIT Press, Cambridge MA, pp. 540–545.
Carlson, N.R., 1991. Physiology of Behaviour, fourth ed. Allyn and Bacon.
Cohen, J.E., Newman, C.M., 1985. When will a large complex system be stable? Journal of Theoretical Biology 113, 153–156.
Funahashi, K., Nakamura, Y., 1993. Approximation of dynamical systems by continuous time recurrent neural networks. Neural Networks 6, 801–806.
Gardner, M.R., Ashby, W.R., 1970. Connectance of large dynamic (cybernetic) systems: critical values for stability. Nature 228, 784.
Hoppensteadt, F.C., Izhikevich, E., 1997. Weakly Connected Neural Networks. Springer-Verlag, New-York.
Husbands, P., Smith, T., N.J., O'Shea, M., 1998. Better living through chemistry: evolving GasNets for robot control, Connection Science, 10, 185–210.
Jaeger, H., 2001. The "echo state" appraoch to analysing and training recurrent neural networks. GMD-Report 148. German National Research Institute for Computer Science.
Jaeger, H., Haas, H., 2004. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science 304, 78–80.
Litvak, V., Sompolinsky, H., Segev, I., Abeles, M., 2003. On the transmission of rate code in long feedforward networks with excitatory-inhibitory balance. The Journal of Neuroscience 23 (7), 3006–3015.
Mathayomchan, B., Beer, R.D., 2002. Center-crossing recurrent neural networks for the evolution of rhythmic behavior. Neural Computation 14, 2043–2051.
May, R.M., 1972. Will a large complex system be stable. Nature 238, 413–414.
Mehta, M.L., 1967. Random Matrices. Academic Press, New York.
Murray, J.D., 1989. Mathematical Biology. Springer, Heidelberg.
Sinha, S., 2005. Complexity vs. stability in small-world networks. Physica A, 147–153.
Sinha, S., Sinha, S., 2005. Evidence of universality for the May-Wigner stability theorem for random networks with local dynamics. Physical Review Let. E 71, 1–4.
Solow, A.R., Costello, C., Beet, A., 1999. On an early result on stability and complexity. The American Naturalist 154 (5), 587–588.
Strogatz, S.H., 1994. Nonlinear Dynamics & Chaos. Addison-Wesley, Reading MA.
Turrigiano, G.G., 1999. Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. Trends in Neuroscience 22, 221–227.
Wigner, E.P., 1959. Gruppentheorie und Ihre Anwendung auf die Quantenmechanik der Atomspektren, trans. J. J. Griffin. Academic Press, New York.
Williams, H., 2004. Homeostatic plasticity in recurrent neural networks. In: Schaal, S., Ijspeert, A., Billard, A., Vijayakumar, S., Hallam, J., Meyer, J.-A. (Eds.), Proceedings of the Eighth International Conference on the Simulation of Adaptive Behavior. MIT Press, Cambridge, MA, pp. 344–353.
Williams, H.P., 2006. Homeotstatic adaptive networks. Ph.D. Thesis. Biosystems Group, School of Computing, University of Leeds.
Williams, H., Noble, J., 2007. Homeostatic plasticity improves signal propagation in continuous time recurrent neural networks. Biosystems 87 (2–3), 252–259.

Please cite this article in press as: Buckley, C.L., Bullock, S., Sensitivity and stability: A signal propagation sweet spot in a sheet of recurrent centre crossing neurons. BioSystems (2008), doi:10.1016/j.biosystems.2008.05.026