# Increased robustness and intermittent dynamics in structured Reservoir Networks with feedback

Sarah Jarvis<sup>1,2</sup>, Stefan Rotter<sup>1,3</sup> and Ulrich Egert<sup>1,2</sup> \*

 Bernstein Center Freiburg Freiburg im Breisgau - Germany
Biomicrotechnology - Department of Microsystems Engineering

Albert-Ludwig University, Freiburg im Breisgau - Germany
3- Computational Neuroscience - Faculty of Biology
Albert-Ludwig University, Freiburg im Breisgau - Germany

#### Abstract.

Recent studies using feedforward Echo State Networks (ESN) demonstrate that reservoir stability can be strongly affected by reservoir substructures, such as clusters. Here, we evaluate the impact of including feedback on clustered ESNs and assert that certain cluster configurations extend the permissible range of spectral radius values. We also report a new class of reservoir activity: intermittent dynamics, characterized by variable periods of chaotic activity before returning to quiescent behaviour. Using a non-linear benchmark data set, we establish clustered ESNs have comparable performance against conventional ESNs, but importantly display an increased tolerance and robustness to spectral radius and input choice.

## 1 Background

Echo State Networks (ESN) are recurrent neural networks, characterized by a reservoir of sparse connections that remain fixed for the life time of the network while only the output weights are trained [1]. Demonstrated to have excellent performance, especially for prediction of non-linear signals, attention has been recently focused on defining conditions for their stability and dynamical range of reservoir activity.

While stringent criteria for stability have been identified for ESNs with feedforward architecture, the introduction of feedback from output to the reservoir can greatly alter network behaviour [2] to the point where ESNs easily enter chaotic regimes.

Here, we investigate how introducing clustered reservoirs can increase robustness of reservoirs in feedback ESNs, prompted by our previous work which established that clusters can improve the robustness of a ESN for feedforward architectures [3]. We observe a novel class of reservoir activity we term intermitent that occurs for a large range of clustered reservoir configurations and exhibits both stable and chaotic dynamics. After evaluating the performance of all networks on a NARMA benchmark test, the advantages of ESNs displaying these properties are then discussed.

<sup>\*</sup>This work was supported by the German BMBF (FKZ 01GQ0420, FKZ 01GQ0830) and by the EC (NEURO, No. 12788).

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

## 2 Clustered ESNs

ESNs consist of an input population u, reservoir population x and output population y. These populations are coupled according to a connection weight matrix  $W^{in}$  for input to reservoir units, a sparse directed connection matrix  $W^{res}$  for connections within the reservoir,  $W^{out}$  for connections from reservoir to output units and  $W^{fb}$  from output to reservoir units. The dynamic equations, modified from [4], are given in Eqn. 1-2, with activation functions f and g chosen as tanh.

$$x(n) = f\left(\mathbf{W}^{in}u(n) + \mathbf{W}^{res}x(n-1) + \mathbf{W}^{fb}y(n-1)\right)$$
(1)

$$y(n) = g\left(\mathbf{W}^{out}x(n)\right) \tag{2}$$

#### 2.1 Setting up clustered reservoirs

In our study, the input is a 1 dimensional vector, which is scaled and shifted to place it in operating range of the network. The reservoir populations examined ranged from 50 to 1000 units (see following subsection for more details), chosen for computational tractability.  $W^{in}$  and  $W^{fb}$  are both fully connected matrices with weights drawn from [-1, 1] with uniform probability. The construction of  $W_{res}$  is outlined below. As for all ESNs,  $W^{in}$ ,  $W^{fb}$  and  $W^{res}$  remain fixed for the duration of the network simulation, ensuring that training did not affect the structure of the reservoir. Only  $W^{out}$  was modified during training, which was performed by presenting a target signal to an initialized network and calculating  $W^{out}$  as the pseudo-inverse of the product of the observed reservoir state matrix and target output signal y [4], after discarding some washout period  $T_0 = 100$  (dynamics) or 200 (NARMA task).

Clustered reservoirs were generated for a given total reservoir size R, the number of clusters n and the number of intercluster units per cluster b. Each cluster had size R/n and sparse connectivity  $conn_{intra}$ . Intercluster connectivity was obtained by coupling the first b units of each cluster to each other by specifying a connectivity matrix of size bR/n and sparse connectivity  $conn_{inter}$ . All weights were initially drawn randomly from a uniform probability on [-1, 1] and rescaled to the defined spectral radius. Here,  $conn_{inter} = conn_{inter} = 0.7$  and b = 1 unless otherwise stated.

The spectral radius  $\rho$  is traditionally used as an indicator of stability of network dynamics, with  $\rho$  less than unity being sufficient to ensure stable dynamics. We examined the effect of  $\rho$  on stability of clustered ESNs by increasing  $\rho$  from 0.5 to 2 in increments of 0.05.

#### 2.2 Quantifying robustness

Several different methods of quantifying reservoir performance exist [5]. Here, we experimentally estimated the Lyapunov exponent  $\hat{\lambda}$  as described previously in our work on feedforward ESNs [3]. Briefly, positive Lyapunov exponents indicate

a system where small perturbations lead to exponentially diverging trajectories, suggestive of a chaotic system.

To relate this to reservoir activity, we also considered the distribution of dwell times, determined by a sequence consisting of a repeated input impulse presented with a constant interstimulus interval of 50 timesteps. The up states were defined as periods during which the activity exceeded a threshold that was 50% of the maximum normalized activity value. This was repeated for 100 independent realizations of the the same network configuration with zero noise. The resulting mean and standard deviations for the collected dwell times in the up state were then assessed.

We observed that three types of response exist: (i) the network showed some activation that quickly died away (Fig. 1A), having dwell times that had both low mean and low standard deviation; (ii) the network either never stabilized to begin with or was unstable after the presentation of the first stimulus (Fig. 1C), with large mean dwell times and low standard deviation; or (iii) intermittent behaviour, characterized by highly variable dwell times (Fig. 1B). Intermittent responses were characterized by fast to instantaneous ignition of clusters of reservoir units with maximal reservoir activity.

Importantly, this was in complete contrast to the transient responses observed for feedforward ESNs, where reservoir activity settled at some non-zero level of background activity that increased with increasing  $\rho$  values until saturation of reservoir activity [3], thus characterizing a new class of ESN dynamics.



Fig. 1: Normalized total reservoir activity in response to a identical stimulus with fixed frequency for (a) stable, (b) intermittent and (c) chaotic activity. Dwell time is determined as the duration for which the response was larger than threshold  $\theta_{50}$ , set to be 50% of the maximum for total reservoir activity.

#### 3 Results

#### 3.1 Reservoir dynamics

To test how the responses change as the spectral radius is increased, we analysed a reservoir of 200 units, increasing  $\rho$  from 0.5 to 2 in 0.05 increments. Clusters were then introduced for  $n = \{1, 2, 4, 5, 8, 10, 20, 25, 50, 100\}$  and we estimated the dwell times and the Lyapunov exponents  $\hat{\lambda}$  from 100 realizations of each reservoir configuration. (Fig. 2A-C). All reservoir configurations with  $\rho \leq 1$  had stable responses, characterized by negative  $\hat{\lambda}$  and dwell time distributions with low mean and low variability. As for all ESNs, reservoir dynamics become chaotic with increasing  $\rho$ . However, the presence of clusters increased the spectral radius value for which this transition occurred (indicated by dotted line in Fig. 2C) ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.



Fig. 2: Reservoir activity for mean (left columns) and standard deviation (middle) of dwell times, and Lyapunov exponent (right). (A-C) Distributions of normalized reservoir activity values and the pseudo-Lyapunov exponent  $\hat{\lambda}$  as  $\rho$ increased for R=200. Dashed line in (C) indicates the  $\hat{\lambda}=0$  contour. Values in (C) have been clipped to minimum value. (D-F) Dwell times and Lyapunov exponent estimate for collection of reservoir size and cluster configuration for fixed  $\rho=1.2$ . R taken for {50, 100, 200, 500, 750, 1000}. White lines indicate fixed reservoir size and dots indicate network configuration tested, interpolation is used to visualize trend.

resulting in reservoirs  $20 \le n \le 40$  clusters having the largest spectral radius range for stable dynamics.

Intermittent dynamics were observed during the transition from stable to chaotic dynamics as  $\rho$  was increased, but only for clustered reservoirs. Interestingly, although the variability of dwell times to a repeated stimulus was high, the Lyapunov exponent was negative, indicating that intermittent dynamics are not classically chaotic.

The impact of total reservoir size on dynamics was investigated by considering reservoirs of 50 to 1000 units using a fixed spectral radius value  $\rho = 1.2$  (Fig 2D-F). We observed that the resulting dynamic regime was determined by a combination of the absolute reservoir size, absolute cluster size and the number of clusters. As observed for R=200, networks with fewer but larger clusters proved to have the strongest tendency to enter chaotic regimes when  $\rho > 1$ , while the configurations that transitioned most slowly through the different responses had relative cluster sizes that were 1-5% of the total reservoir size.

#### 3.2 Performance on NARMA prediction

The performance of clustered ESNs was tested using the 10th order NARMA system, given in Eqn. 3, a non-linear prediction task commonly used to benchmark ESNs [4]. Driving the system with u(t) randomly drawn from [-0.1,0.5], we evaluated network performance using normalized root mean squared error (NRMSE). Our results for reservoirs with 200 units (Fig 3A-B) confirm that the best performance is achieved when  $\rho \approx 1$  with NRMSE lowest for reservoirs when  $n \leq 10$ , although these networks displayed the highest NRMSE values as  $\rho$  was increased to 2.

Trends for performance and stability as cluster configurations were varied are simultaneously plotted in Fig 3C. Both plots demonstrate inflections for clustered configurations and that these occur for the same number of clusters. Furthermore, the contours reveal that while the trends for stability and performance are first inversely correlated for clustered networks ( $10 \le n \le 50$ ) as  $\rho$ is increased from 1, this is reversed as the  $\rho \simeq 1.5$ . This clarifies that clustered reservoirs perform similarly to homogeneous reservoirs but for larger values of  $\rho$ ; importantly, the degradation of performance and stability also occurs more gradually. Thus, while traditional ESNs may achieve the lowest error rates, clustered ESNs perform comparably and across a broader range of spectral radius values, resulting in a slower transition to a chaotic regime.

$$d(n+1) = 0.3d(n) + 0.05d(n) \left[\sum_{i=0}^{9} d(n-i)\right] + 1.5u(n-9)u(n) + 0.1$$
 (3)



Fig. 3: Average *NRMSE* for 100 instances of each cluster configuration, calculated during (A) training and (B) testing, obtained for reservoir R=200. Plots have been clipped to maximum value. (C) Contour plots for both  $\hat{\lambda}$  (dashed) and *NRMSE*<sub>test</sub> (solid). Thicker lines indicate  $\hat{\lambda}=0$  and *NRMSE*<sub>test</sub>=1,2.

# 4 Conclusion and future work

Using clustered ESNs we established that, when feedback is present, network stability and dynamics also depend on the relative cluster size as the spectral radius is increased, similar to feedforward ESNs [3]. We demonstrate that intermediate cluster sizes display the most robust performance as spectral radius is increased. However, although the spectral radius has previously been used to characterize stability, it is not an absolute predictor [5], and is to some extent input-dependent. Therefore, as clustered reservoirs demonstrate more tolerance for large values of the spectral radius, they allow greater flexibility in parameter choice than homogeneous reservoirs (equivalent here to n=1).

Importantly, our results also display a novel type of reservoir activity: intermittent responses, which were characterized by highly dynamic responses to a fixed stimulus. Occurring during the transition of stable to chaotic dynamics for cluster configurations with intermediate cluster sizes and a high spectral radius values ( $\rho > 1$ ), their activity suggests that they may lie at the edge of chaotic dynamics, supported by their negative pseudo-Lyapunov exponents and highly variable dwell times. The presence of intermittent activity suggests the existence of architectures that allow both stable (input-driven) and chaotic (self-sustained) activity. As we observed that reservoirs with identical cluster configurations but different weight values displayed a different distribution of dwell times, this suggests that further factors contribute. Nevertheless, we conclude that both clusters and feedback are necessary, since intermittent activity is absent for all spectral radius values when the reservoir is homogeneous and was never observed when feedforward clustered ESNs were similarly investigated.

The exploitation of chaotic activity has recently been suggested by devising a learning rule that exploits self-sustained activity in feedback networks to generate coherent patterns of activity from chaotic ESNs [6]. However, such architectures offer no opportunity to return to an input driven regime without retraining; thus it would be advantageous to have an architecture that supports both input-driven and self-sustained regimes during runtime. This could be easily achieved by gating clusters to provide either type of activity as required. Here, we have made the first steps in the identification of specific factors of a reservoir that facilitate intermittent activity and thereby shift between stable and chaotic regimes.

## References

- H. Jaeger. The "echo state" approach to analysing and training recurrent neural networks. Technical Report 148, German National Research Center for Information Technology, 2001.
- [2] M. Lukosevicius. Echo State Networks with trained feedbacks. Technical Report 4, International University Bremen, 2007.
- [3] S. Jarvis, S. Rotter, and U. Egert. Extending stability through hierarchical clusters in echo state networks. *Front Neuroinformatics*, 4, 2010.
- [4] H. Jaeger. Advances in neural information processing systems, chapter Adaptive nonlinear system identification with echo state networks. 2003.
- [5] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt. An experimental unification of reservoir computing methods. *Neural Netw*, 20:391–403, Apr 2007.
- [6] D. Sussillo and L. F. Abbott. Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63:544–557, Aug 2009.