Reservoir computing simulation. Review on best results .

1 Model

The model is constructed in the same way as before. It is an opto-electronic delay system with a multiple delays feedback loop. The system can be modeled by an integro-differential equation with delays,

$$\tau \frac{\mathrm{d}x}{\mathrm{d}t} + x(t) + \frac{1}{\theta}\int x(s)\mathrm{d}s = \beta \sin^2\left[\sum_{i=1}^{N_{neur}} w_i x(t - i\cdot\delta\tau_D) + \gamma v(t)\varepsilon_n(t) + \phi_0\right]. \tag{1}$$

For simplicity of the numerical analysis the system can be rewritten in the form of two differential equations with delays and normalized to the response time of the low-pass filter, $\tau = 1$,

$$\begin{cases} \dot{x} = -x - \dfrac{1}{\theta}y + \beta\sin^2\left[\displaystyle\sum_{i=1}^{N_{neur}} w_i x(t - i\cdot\delta\tau_D) + \gamma v(t)\varepsilon_n(t) + \phi_0\right], \\ \dot{y} = x \end{cases} \tag{2}$$

In real experiments with FPGA setup, the number of neurons is $N_{neur} = 150$ and the time parameters are as follows,

τ	$132.629\,\mu s$	1
$\delta\tau_D$	$26.526\,\mu s$	0.2
τ_D	$3.98\,ms$	$= N_{neur}\delta\tau_D = 30$
θ	$80\,ms$	603.2

2 *Weights*

Weights w_i are random numbers normalized in such way that $\sum_{i=1}^{N_{neur}} w_i = 1$. Because simulation results very sensitive to the choice of the weights, w_i are generated once and stored in Matlab data file for common use in simulation and real experiments. Matlab code for generation of weights is very simple:

```
%% Connection vector & matrix
%load('common/w');
neur=150;
w=rand(neur,1)-0.5;
w=w/sum(abs(w));
```

3 Input signal design: NARMA(10) and mask

The input signal $\gamma v(t)\varepsilon_n(t)$ is generated from NARMA(10) test to perform first and simplest benchmark of developed reservoir. NARMA test is discrete real value data points generated by certain rule using previous point (history). The order of NARMA system (10 in our case) is equal to the number of point taking from history of generated sequence to calculated next point.

$$\omega_{n+1} = 0.3\omega_n + 0.05\omega_n \sum_{i=1}^{10} \omega_{n-i+1} + 1.5\varepsilon_{n-9}\varepsilon_n + 0.1 \tag{3}$$

Generally ε_n is a random normally distributed sequence. Then ω_n is generated by Eq.3 dynamical variable. We will call ω_n 'target' as it is the desired output of the reservoir. The examples of inputs and targets are shown in Fig. 1. Matlab code for generation of NARMA points is quite straightforward:

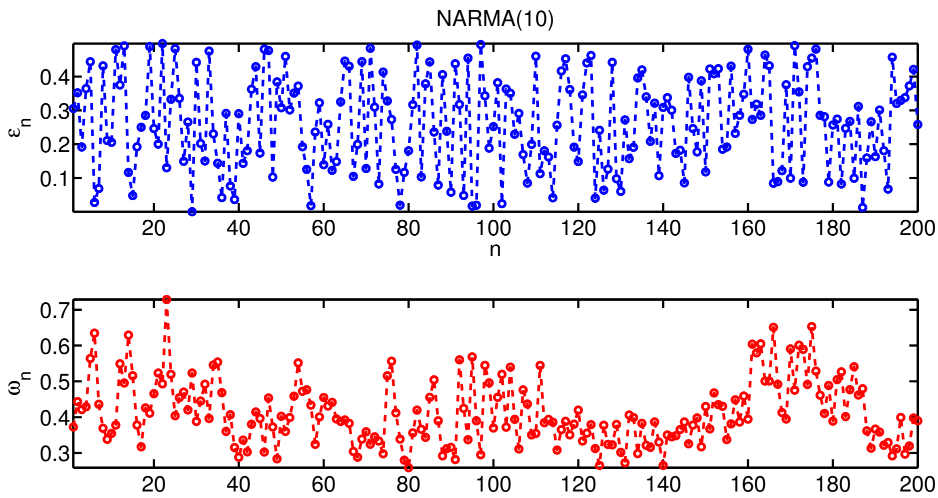

Fig. 1: Examples of input ε_n and target ω_n sequences for NARMA(10)

```
% order 10
a=.3;
b=.05;
c=1.5;
d=.1; % 0.1 -- 0.001

inputs = cell(1,nr_samples);
outputs = cell(1,nr_samples);

for i = 1:nr_samples
    %Create random input sequence
    inputs{i} = rand(1,len)*0.5;
    outputs{i} = 0.1*ones(1,len);
    for n = systemorder+1:len-1
        outputs{i}(n+1) = a*outputs{i}(n) + ...
        b*outputs{i}(n)*sum(outputs{i}(n-systemorder+1:n)) +...
        c*inputs{i}(n-systemorder+1) * inputs{i}(n) + d;
    end
    inputs{i}=inputs{i}(warm+1:end);
    outputs{i}=outputs{i}(warm+1:end);
end
```

Input signal $In(t)$ for a sample with length N_{sample} is produced by the information input $\varepsilon_n(t)$ and mask $v(t)$: $In(t) = \gamma v(t)\varepsilon_n(t)$, where γ is the final amplification of the signal before input to reservoir. Information input is step-wise function produced by input NARMA sequence ε_n,

$$\varepsilon_n(t) = \sum_{n=1}^{N_{sample}} \varepsilon_n \chi_{(n\tau_D, n\tau_D+\tau_D]}(t).$$

The function $I(t)$ is equal to 0 inside the interval $(0, N_{skip}\tau_D]$ for transition of the reservoir to a steady state. Finally one can write $I(t)$ as follows,

$$I(t) = 0\chi_{(0, n_{nrm}N_{skip}\tau_D]}(t) + \sum_{n=1}^{N_{sample}} \varepsilon_n \chi_{(n\tau_D, n\tau_D+\tau_D]}(t).$$

The mask $v(t)$ is a periodic function with period τ_D. The period is formed by step-wise function which has N_{neur} steps with values -1 or 1 for simplest way. In more common case mask can be formed by random numbers between -1 and 1. Thus the expression for $v(t)$ for one period is as follows,

$$v(t)_{(0,T]} = \sum_{n=1}^{N_{nod}} u(n)\chi_{((n-1)\delta T, n\delta T]}(t)$$

where $u(n)$ is randomly generated sequence of -1's and 1's or random number from $(-1, 1)$.

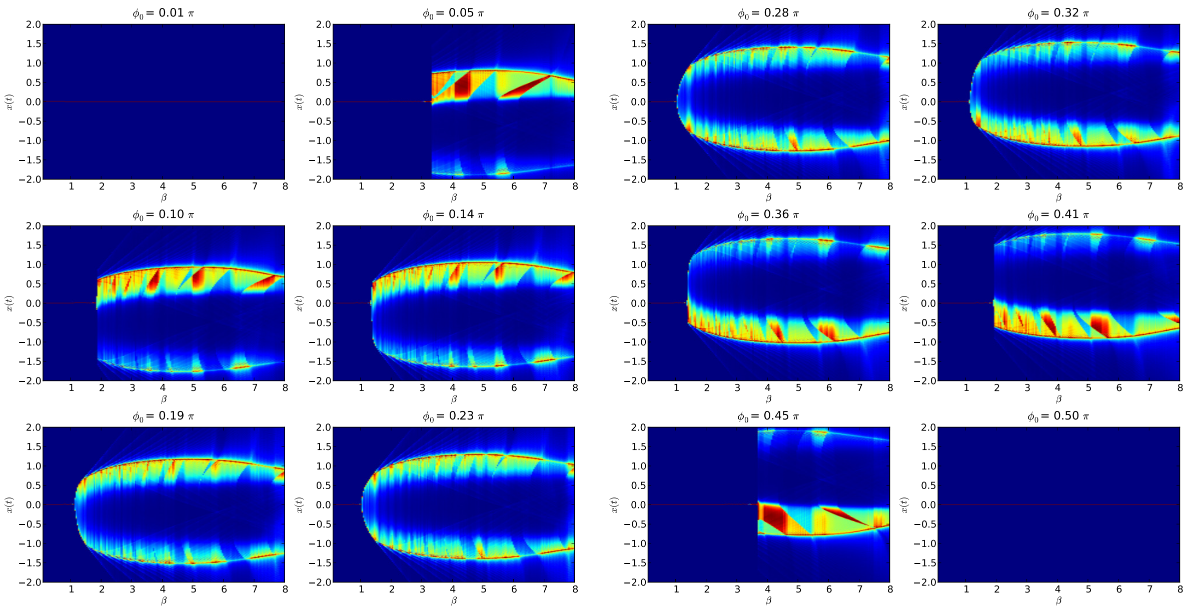

Fig. 2: Bifurcation diagrams for the system with multiple delays

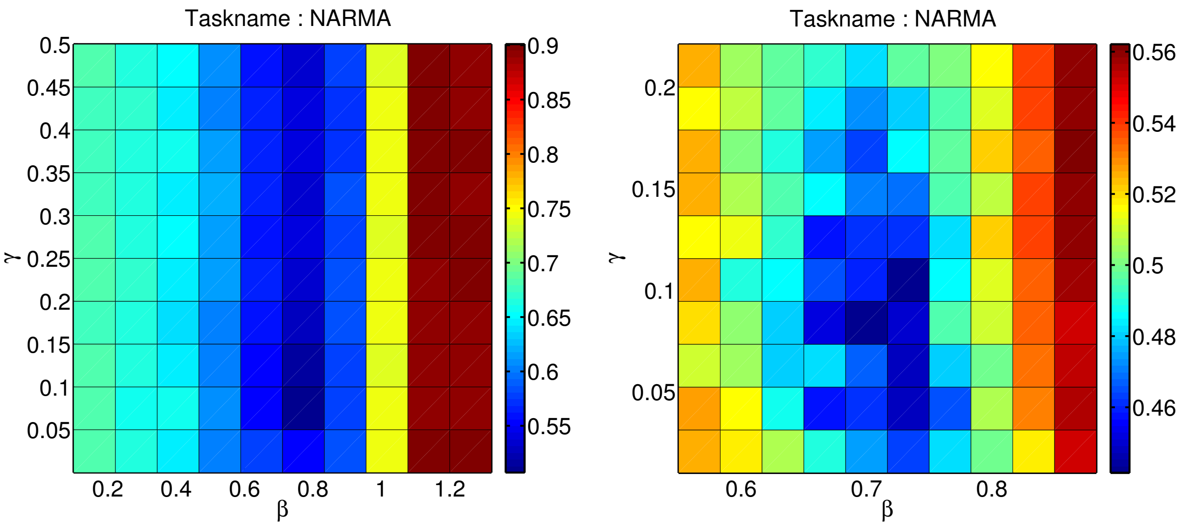

Fig. 3: Scanning results for multiple delays system

4 Simulation experiments: scanning on γ and β

Parameter scanning on γ and β were made during simulation experiments. The first bifurcation point depends on operation point ϕ_0. Thus to determine the range for the parameters β and to fix ϕ_0 the bifurcation diagrams for several different values of ϕ_0 in range $[0; \pi/2]$ were calculated. All of them are presented in Fig.2 The bifurcation point β_c for $\phi_0 = \pi/4$ is close to 1. This point was selected for first parameters scans.

Parameter scanning results for the multiple delays system are shown in Fig.3 The right picture is the zoom of best region in the left picture.

Best result for the system with multiple delays: $NRMSE_{multi} = 0.44154$, $\beta = 0.68333$, $\gamma = 0.073333$.

For comparison the parameter scanning for the system with single delay line were made in same parameters ranges. The Fig.4 demonstrate the result.

Best result for the system with single delay: $NRMSE_{single} = 0.29503$, $\beta = 0.58333$, $\gamma = 0.17889$

5 Several NARMA points on total time delay: scanning on γ and β

The idea behind this setup is to take into account the fact that NARMA(10) system has limited memory (delay) equal to 10 in discrete steps. On the other hand the opto-electrical system has total delay time

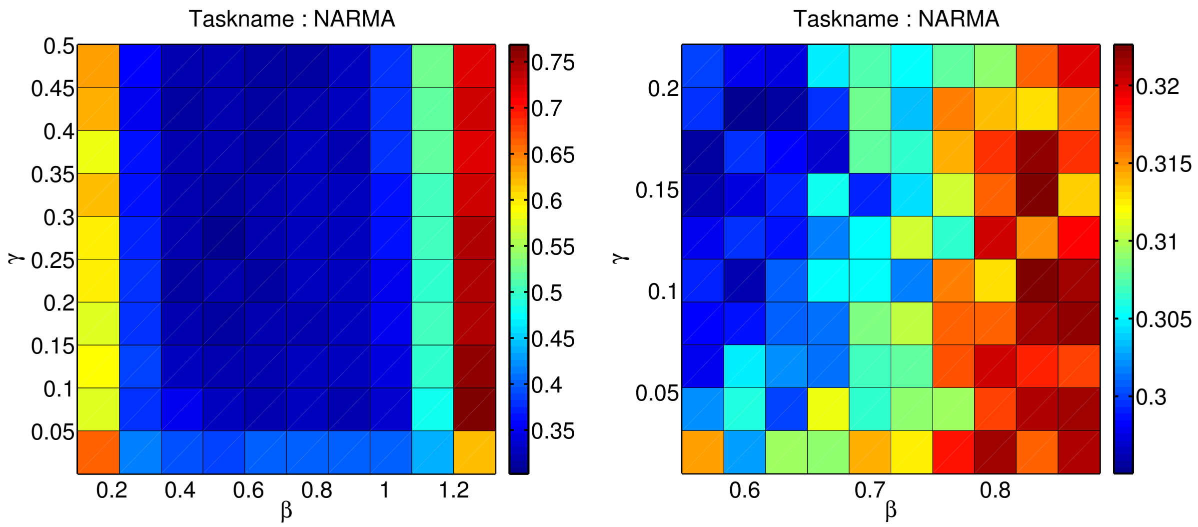

Fig. 4: Scanning results for the single delay system

equal to τ_D. One should expect that if the reservoir system is getting several NARMA points during the total delay time τ_D it will keep the history of input signal in its memory (delayed signal) and will be trained more effectively in comparison with the case when the only one data point are fed during total delay. In this section this concept is investigated.

From mathematical point of view the system looks the same,

$$\begin{cases} \dot{x} = -x - \dfrac{1}{\theta} y + \beta \sin^2 \left[\displaystyle\sum_{i=1}^{N_{neur}} w_i x(t - i \cdot \delta\tau_D) + \gamma v(t) \varepsilon_n(t) + \phi_0 \right], \\ \dot{y} = x \end{cases} \tag{4}$$

The only difference is going to the input signal design $\gamma v(t)\varepsilon_n(t)$. Now it is designed in such way that during one time delay τ_D several (n_{mult}) NARMA points ε_n inserted in the reservoir. First simulation was done with $n_{mult} = 5$. Note, that the the mask is designed in the same way as before. It means that every NARMA data point is extended to N_{neur} inputs points by the mask. It implies that time interval between successive input data point (let's define it as $\delta\tau_{inp}$) is n_{mult} time less then so-called 'neuron width' $\delta\tau_D$, $\delta\tau_D = n_{mult}\delta\tau_{inp}$. To keep the reservoir excited during the input signal it is straightforward to set $\delta\tau_{inp} = 0.2\tau$. It gives new value to the 'neuron width' $\delta\tau_D = 0.2 n_{mult}\tau$. Finally for the system with $N_{neur} = 150$ and $n_{mult} = 5$ the parameters can be summarized in the table,

τ	$132.629\,\mu s$	1
$\delta\tau_{inp}$	$26.526\,\mu s$	0.2
$\delta\tau_D$	$132.629\,\mu s$	1
τ_D	$19.89\,ms$	$N_{neur}\delta\tau_D = 150$
θ	$80\,ms$	603.2

The first two obtained diagrams are very promising. We did two simulation with $n_{mult} = 5$ and $n_{mult} = 10$. They are presented in Fig.5 Best result for the system with multiple delays and 5 NARMA point over total delay: $NRMSE_{multi} = 0.22034$, $\beta = 0.4636$, $\gamma = 0.4181$. Best result for the system with 10 NARMA points on delay time: $NRMSE = 0.23981$, $\beta = 0.20526$, $\gamma = 0.41842$. The result is better than for the system with single delay line and 400 neurons! The phase offset is the same as before, $\phi_0 = \pi/4$. It should be directly compared with Fig.3(left) where the scan for the same system were presented in the same ranges of parameters.

6 Simulation experiments: scanning on ϕ_0 and β

Scanning on ϕ_0 and β were made around best values from the previous diagrams. The parameter β runs over range $[0.35; 1.5]$ and the parameter ϕ_0 runs over range $[0.05\pi; 0.5\pi]$. The result is depicted in Fig6. Left side diagram is for multiple delays and right side diagram is for single delay. The pictures confirm that best result is located around point $\phi_0 = \pi/4$ and β around the bifurcation point $\beta = 1$, which were selected for all previous simulations.

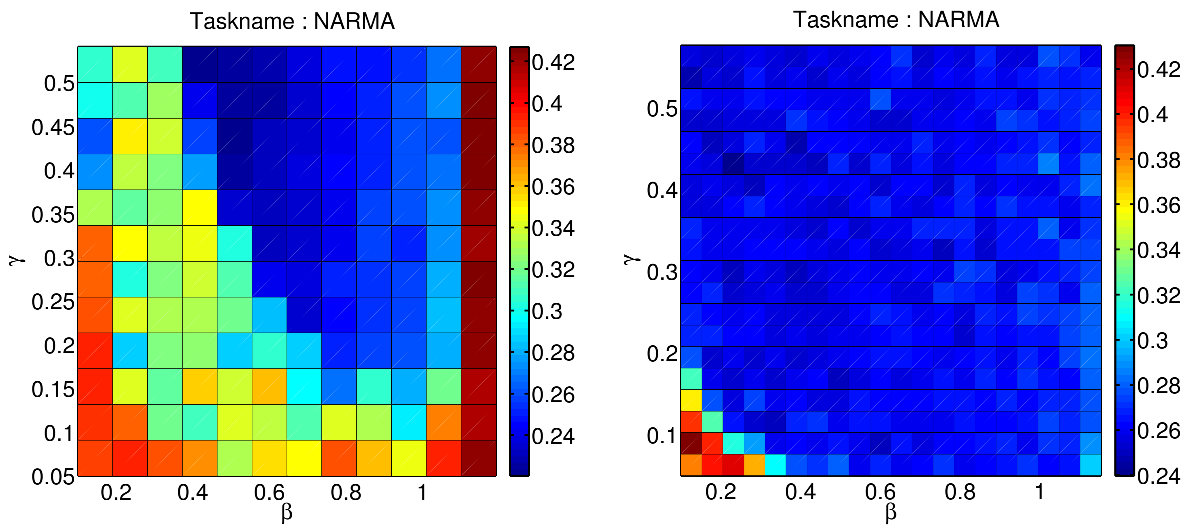

Fig. 5: Scanning results for multiple delays system with 5 NARMA points (left) and 10 NARMA points (right) over total delay

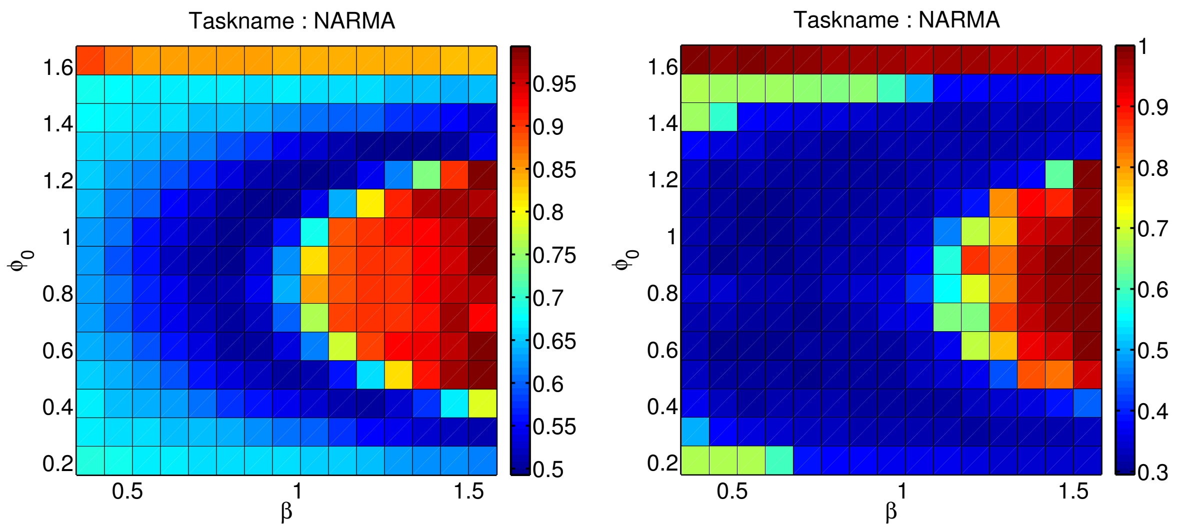

Fig. 6: Scanning on ϕ_0 and β results for multiple and single delay systems

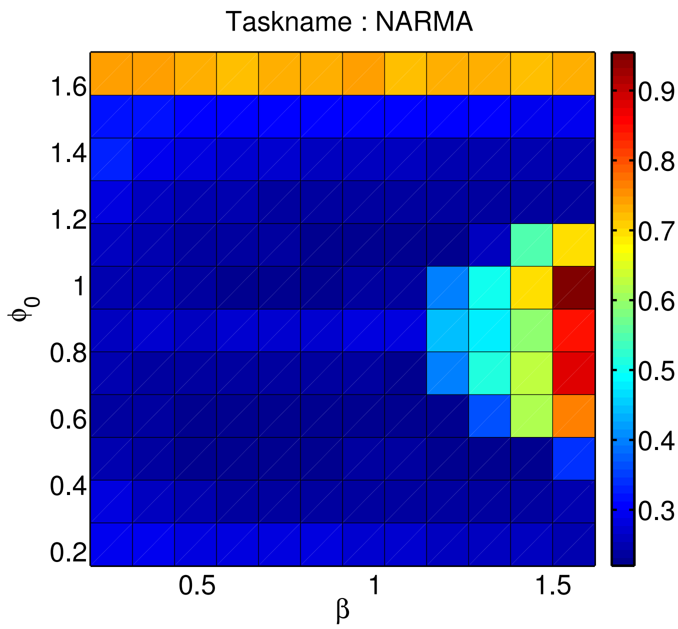

Fig. 7: Scanning on ϕ_0 and β results for multiple delays systems with 5 NARMA data points over total delay time τ_D

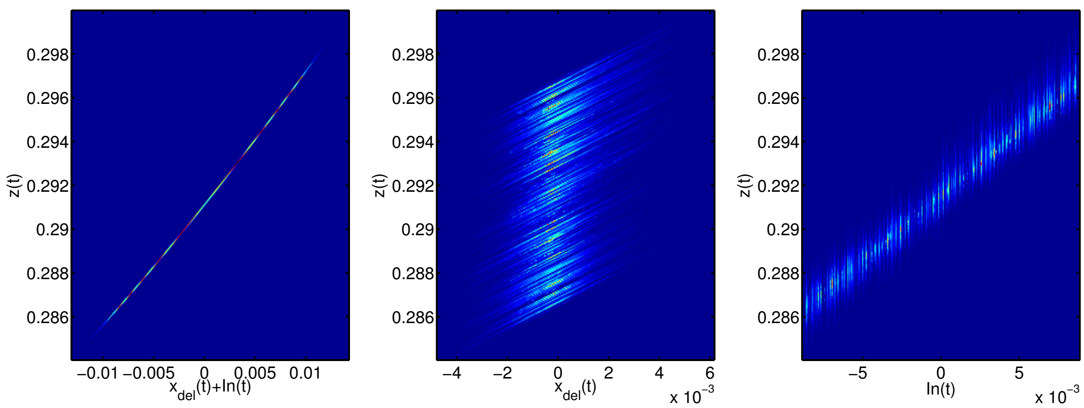

Fig. 8: Histograms for single delay system

The same scan was made for the system with 5 NARMA points over total delay time. The result is presented in Fig.7 It is clear seen that blue area with small values of the errors wide-spread in same way as in Fig.6(right) for the system with multiple delay lines and one NARMA point on the delay time. The difference is that the absolute values for the system with several data points on the delay is much lower then for the setup with one data point. The best parameters for the diagram in Fig.7 is $NRMSE = 0.21999$ at $\beta = 1.0273$ and $\phi_0 = 0.54264$.

7 Operating range

To verify the range in winch system operates the histograms of several derived variables were constructed for the best values of parameters from Sec.4. The following variables were taken in to account,

$$\begin{cases} x_{del}(t) & = \sum_{i=1}^{N_{nod}} w_i x(t - i \cdot \delta\tau_D) \\ In(t) & = \gamma v(t) \varepsilon_n(t) \\ z(t) & = \beta \sin^2 \left[x_{del}(t) + In(t) + \phi_0 \right] \end{cases}$$

Results for single delay ($\beta = 0.58333$, $\gamma = 0.17889$, $\phi_0 = \pi/4$) is shown in Fig.8
Results for multiple delays ($\beta = 0.68333$, $\gamma = 0.073333$, $\phi_0 = \pi/4$) is shown in Fig.9

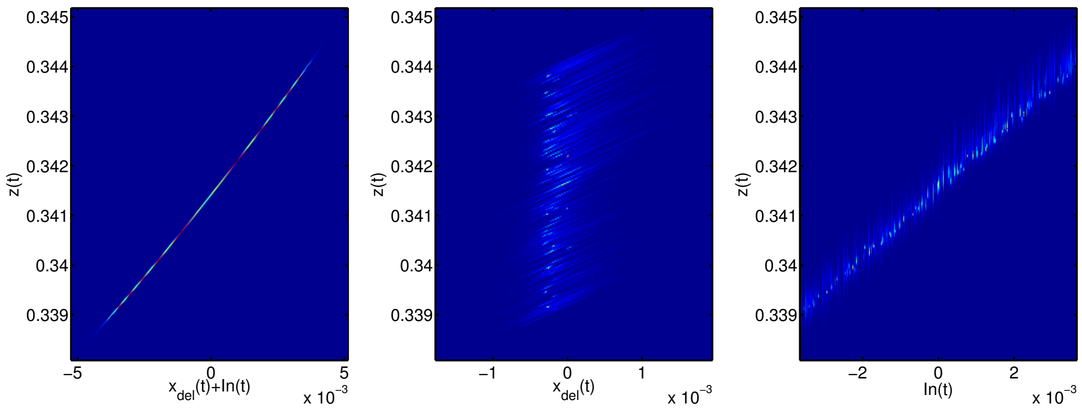

Fig. 9: Histograms for multiple delays system

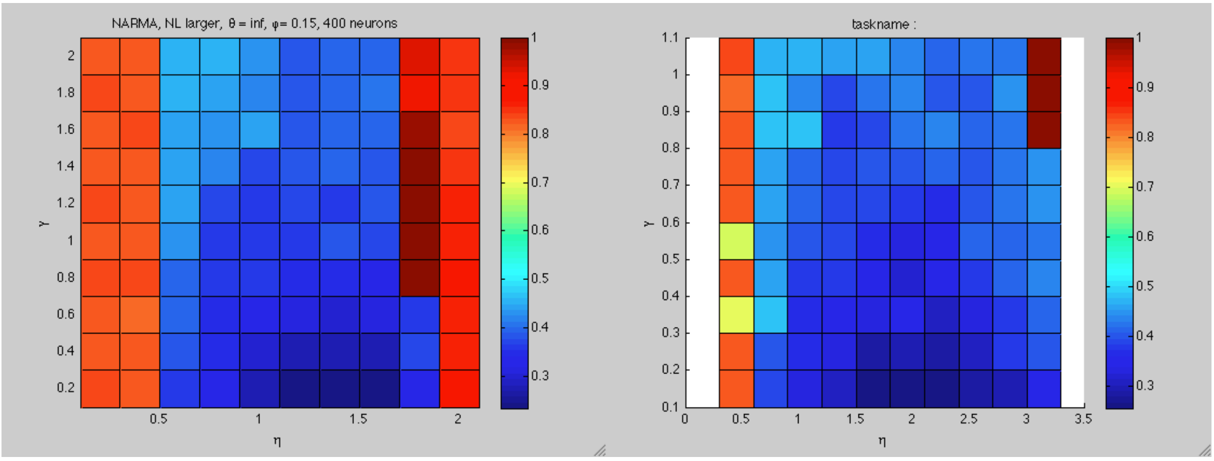

Fig. 10: Scans from Lennert presentation

8 Real experiments

We made one real experiment with multiple delays and parameters shown in Table. The result is $NRMSE_{exp} = 0.93366$

9 Best results extracted from Lennert's reports and presentations

Lennert used a normalized on $\tau = 1$ model with single delay and $N_{neur} = 400$ neurons,

$$\begin{cases} \dot{x} = -x - \dfrac{1}{\theta} y + \beta \sin^2 \left[x(t - \tau_D) + \gamma v(t) \varepsilon_n(t) + \phi_0 \right] \\ \dot{y} = x \end{cases} \tag{5}$$

Best result he got is $NRMSE = 0.24$ for $\gamma = 0.3$, $\beta = 1.58$, $\theta = \infty$, $\phi_0 = 0.15$. Some pictures with scans obtained by him are in Fig.10