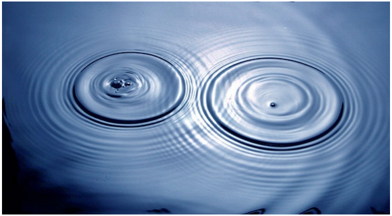
Analogue hardware for energy efficient Artificial Intelligence



Laurent Larger

FEMTO-ST institute / Optics Dept.
CNRS / University Bourgogne-Franche-Comté
Besançon, France

December 3, 2020 / ICE virtual seminar
Information, Communication, Engineering
Institut Polytechnique de Paris, France

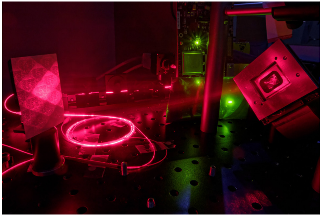
Analogue hardware for energy efficient Artificial Intelligence

Laurent Larger

FEMTO-ST institute / Optics Dept.
CNRS / University Bourgogne-Franche-Comté
Besançon, France

December 3, 2020 / ICE virtual seminar
Information, Communication, Engineering
Institut Polytechnique de Paris, France

Analogue hardware for energy efficient Artificial Intelligence



Laurent Larger

FEMTO-ST institute / Optics Dept.
CNRS / University Bourgogne-Franche-Comté
Besançon, France

December 3, 2020 / ICE virtual seminar
Information, Communication, Engineering
Institut Polytechnique de Paris, France

Outline

Background and motivations

Basics in Reservoir Computing (RC)

Delay dynamics as an emulated network of neurons

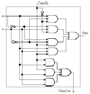
Reservoir Computing with Photonic DDE

Conclusion & perspectives

Why would we need alternative hardwares?

Adapting technology to breakthrough, solving bottlenecks

- AI is a new concept, implemented with unmatched and decades old hardware computing principles



- Turing – von Neumann machines are matched to arithmetic computing, not neural network processing

- Digital computers are inefficiently simulating NN models (multi-task 30W brain vs. MW single task AlphaGo)

- Energy efficiency is a major drawback of currently available processors (autonomous car driving bottleneck)

Change of viewpoints?

Possibly requirred change of paradigm or framework

- Digital vs. Analog
  (Boolean vs. Continuous functions)
- Clocked discrete time vs. Continuous time varying processing
  (e.g. breakthrough from sampling theory to compressive sensing)
- Error-free & repeatable vs. robust & fault tolerant
  (fully deterministic rules vs. freedom & randomness)
- Finite (though high-) vs. Infinite dimensional

Scientific challenges

Examples of a few

- RNN chips: hardware, devices & technologies, are to be discovered
- New paradigms are explored for groundbreaking architectures to be determined (high risk & unconventional approaches requirred)
- Theory & understanding: The magic “black box” of RNN needs to be turned into grey, or even white (processing, learning, operation boundaries)
- Extreme cross-disciplinary research is needed to succeed (far beyond the nowadays dominating realm of “current” Computer Science)
- Totally new and disruptive approaches are needed (concepts, principles, design, programming, etc. . .)

In this framework, many paths are to be explored,

photonic Reservoir Computing is one of them

Outline

Background and motivations

Basics in Reservoir Computing (RC)

Delay dynamics as an emulated network of neurons

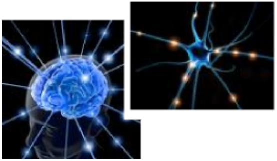
Reservoir Computing with Photonic DDE

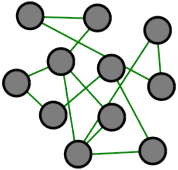
Conclusion & perspectives

A few RC concepts

Concepts

- Novel paradigm refered as to Echo State Network (ESM), Liquid State Machine (LSM) and also Reservoir Computing (RC)



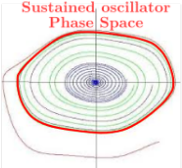


A few RC concepts

Concepts

- Novel paradigm refered as to Echo State Network (ESM), Liquid State Machine (LSM) and also Reservoir Computing (RC)
- Processing of time varying information through nonlinear transients observed in a high-dimensional phase space (could be named Nonlinear Transient Computing)

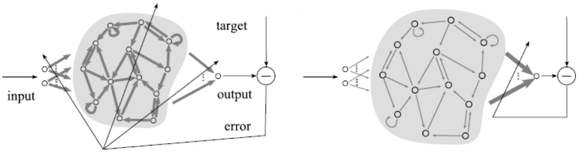
Asymptotic vs. Transient dynamics (huge space for transients out of the stable solution)



A few RC concepts

Concepts

- Novel paradigm refered as to Echo State Network (ESM), Liquid State Machine (LSM) and also Reservoir Computing (RC)
- Processing of time varying information through nonlinear transients observed in a high-dimensional phase space (could be named Nonlinear Transient Computing)
- Derived from RNN, however learning is simplified, limited to the output layer only (other weights, input and internal, chosen at random)



A few RC concepts

Concepts

- Novel paradigm refered as to Echo State Network (ESM), Liquid State Machine (LSM) and also Reservoir Computing (RC)
- Processing of time varying information through nonlinear transients observed in a high-dimensional phase space (could be named Nonlinear Transient Computing)
- Derived from RNN, however learning is simplified, limited to the output layer only (other weights, input and internal, chosen at random)
- Physical implementation: Eventually escape from the technologically unmatched details of the original neural network idea...
  The structure of a Network of Neurons is not necessarily the optimal technological solution

A few RC concepts

Concepts

- Novel paradigm refered as to Echo State Network (ESM), Liquid State Machine (LSM) and also Reservoir Computing (RC)
- Processing of time varying information through nonlinear transients observed in a high-dimensional phase space (could be named Nonlinear Transient Computing)
- Derived from RNN, however learning is simplified, limited to the output layer only (other weights, input and internal, chosen at random)
- Physical implementation: Eventually escape from the technologically unmatched details of the original neural network idea...
  The structure of a Network of Neurons is not necessarily the optimal technological solution

Dimensionality, complexity, degrees of freedom might be the correct ingredients, not the RNN architecture itself

Outline

Outline

Background and motivations

Basics in Reservoir Computing (RC)

Delay dynamics as an emulated network of neurons
- Basics about delay differential dynamics
- Network complexity found in DDE: Chimera States

Reservoir Computing with Photonic DDE

Conclusion & perspectives

Introduction to delay dynamics

Introduction to delay dynamics

Actually not that an unsual dynamics

- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing,. . .)

Introduction to delay dynamics

Actually not that an unsual dynamics


- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing,. . .)
- Traffic jam, accordeon car flow


Introduction to delay dynamics

Actually not that an unsual dynamics

- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing,. . .)
- Traffic jam, accordeon car flow
- Distant control of satelites or rockets in space

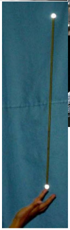
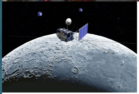
Introduction to delay dynamics

Actually not that an unsual dynamics

- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing,. . .)
- Traffic jam, accordeon car flow
- Distant control of satelites or rockets in space
- Game of vertical stick control at the tip of a finger





Introduction to delay dynamics

Actually not that an unsual dynamics

- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing,. . .)
- Traffic jam, accordeon car flow
- Distant control of satelites or rockets in space
- Game of vertical stick control at the tip of a finger
- Human stand-up position control (and effects of increased perception delay after some kinds of drinks)

Introduction to delay dynamics

Actually not that an unsual dynamics

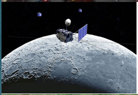
- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing,. . .)
- Traffic jam, accordeon car flow
- Distant control of satelites or rockets in space
- Game of vertical stick control at the tip of a finger
- Human stand-up position control (and effects of increased perception delay after some kinds of drinks)

Introduction to delay dynamics

Actually not that an unsual dynamics





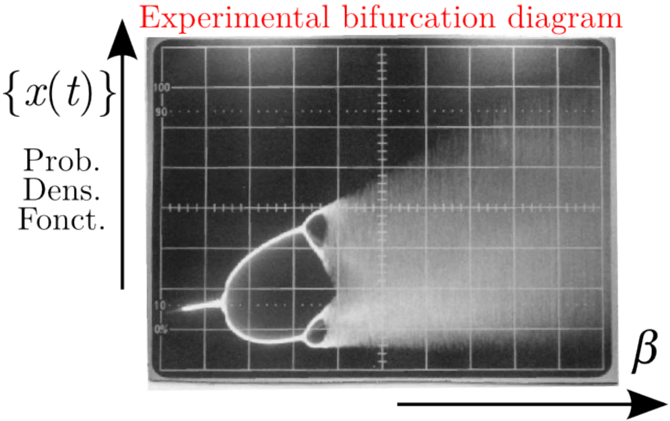



- Living systems (population dynamics, blood cell regulation mechanisms, people reaction after perception and neural system processing, . . .)
- Traffic jam, accordeon car flow
- Distant control of satelites or rockets in space
- Game of vertical stick control at the tip of a finger
- Human stand-up position control (and effects of increased perception delay after some kinds of drinks)
- Hot and cold oscillations at shower start

. . . Any time when information transport occurs (at finite speed), thus resulting in longer propagation time compared to intrinsic dynamical time scales

Optoelectronic delay oscillators (& Apps)

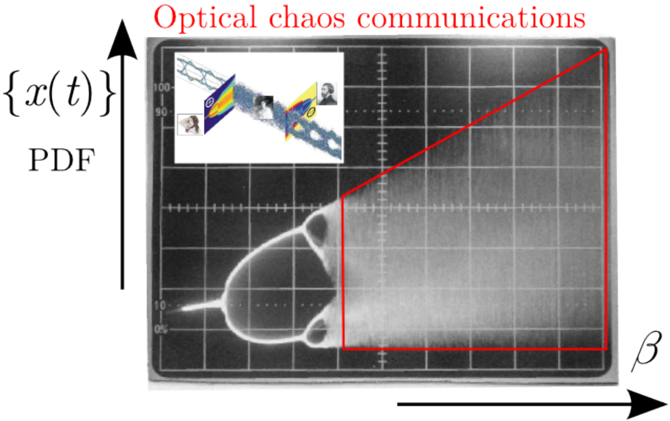
$$\varepsilon\dot{x}(t) = -x(t) + \beta\,\sin^2[x(t-1) + \Phi_0]$$



Y.K. Chembo *et al.*, *Review of Modern Physics*, 2019

Optoelectronic delay oscillators (& Apps)

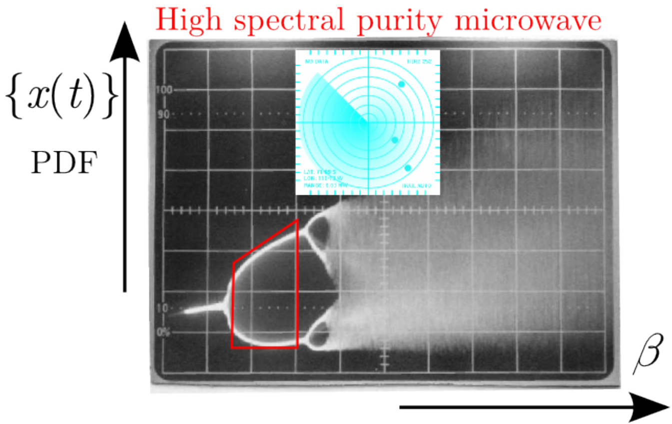
$$\varepsilon \dot{x}(t) = -x(t) + \beta \sin^2[x(t-1) + \Phi_0]$$



Y.K. Chembo *et al.*, *Review of Modern Physics*, 2019

Optoelectronic delay oscillators (& Apps)

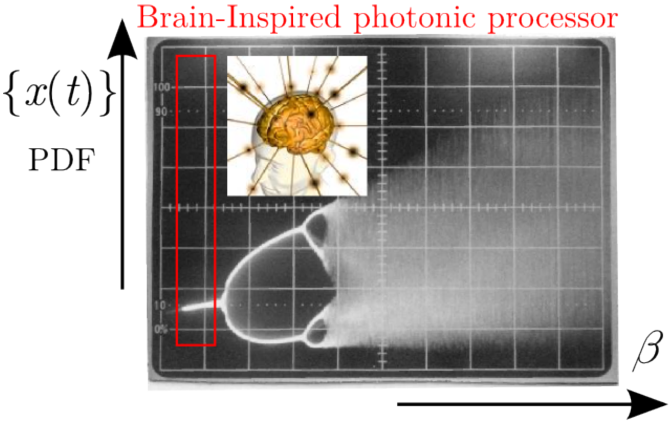
$$\varepsilon\dot{x}(t) = -x(t) + \beta \sin^2[x(t-1) + \Phi_0]$$



Y.K. Chembo *et al.*, *Review of Modern Physics*, 2019

Optoelectronic delay oscillators (& Apps)

$$\varepsilon\dot{x}(t) = -x(t) + \beta \sin^2[x(t-1) + \Phi_0]$$



Y.K. Chembo *et al.*, *Review of Modern Physics*, 2019

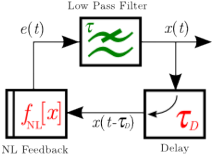
A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE

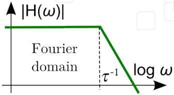
$$\tau \cdot \frac{\mathrm{d}x}{\mathrm{d}t}(t) = -x(t) + f_{\mathrm{NL}}[x(t-\tau_D)]$$

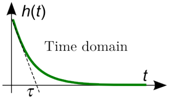
A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$\tau \cdot \frac{\mathrm{d}x}{\mathrm{d}t}(t) = -x(t) + f_{\mathrm{NL}}[x(t-\tau_D)]$$



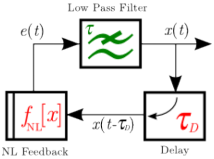


Non-delayed (instantaneous) terms:

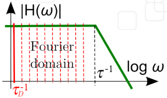
- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

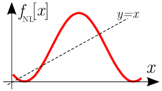
A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$\tau \cdot \frac{dx}{dt}(t) = -x(t) + f_{NL}[x(t-\tau_D)]$$





Non-delayed (instantaneous) terms:

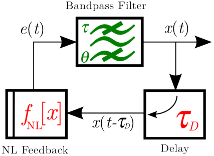
- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

Delayed (feedback) term:

- Non-linearity (slope sign, # extrema, multi-stability),
- Delay (infinite degrees of freedom, stability)
- Large delay case, $\tau_D \gg \tau$

A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$\tau \cdot \frac{dx}{dt}(t) + \frac{1}{\theta} \int_{t_0}^{t} x(\xi)\, d\xi = -x(t) + f_{NL}[x(t-\tau_D)]$$

Non-delayed (instantaneous) terms:

- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

Delayed (feedback) term:

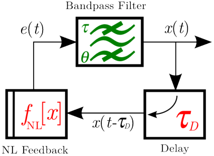
- Non-linearity (slope sign, # extrema, multi-stability),
- Delay (infinite degrees of freedom, stability)
- Large delay case, $\tau_D \gg \tau$

Unusual features for delay dynamics

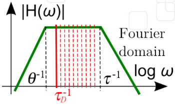
- Bandpass Fourier filter, or integro-differential delay equation

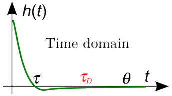
A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$\tau \cdot \frac{\mathrm{d}x}{\mathrm{d}t}(t) = -x(t) - y(t) + f_{\mathrm{NL}}[x(t-\tau_D)]$$
$$\theta \cdot \frac{\mathrm{d}y}{\mathrm{d}t}(t) = x(t)$$





Non-delayed (instantaneous) terms:

- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

Delayed (feedback) term:

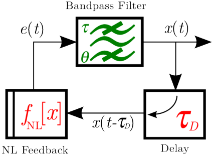
- Non-linearity (slope sign, # extrema, multi-stability),
- Delay (infinite degrees of freedom, stability)
- Large delay case, $\tau_D \gg \tau$

Unusual features for delay dynamics

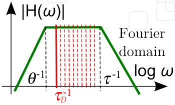
- Bandpass Fourier filter, or integro-differential delay equation

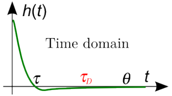
A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$x(t) = \int_{-\infty}^{t} h(t-\xi) \cdot f_{\mathrm{NL}}[x(\xi - \tau_D)]\, \mathrm{d}\xi$$





Non-delayed (instantaneous) terms:

- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

Delayed (feedback) term:

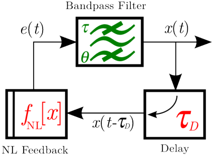
- Non-linearity (slope sign, # extrema, multi-stability),
- Delay (infinite degrees of freedom, stability)
- Large delay case, $\tau_D \gg \tau$

Unusual features for delay dynamics

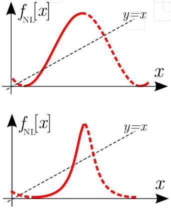
- Bandpass Fourier filter, or integro-differential delay equation
- Dynamical model with a convolution product

A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$x(t) = \int_{-\infty}^{t} h(t-\xi) \cdot f_{\mathrm{NL}}[x(\xi - \tau_D)]\,\mathrm{d}\xi$$



Non-delayed (instantaneous) terms:

- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

Delayed (feedback) term:

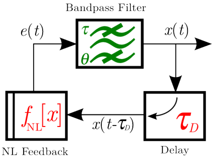
- Non-linearity (slope sign, # extrema, multi-stability),
- Delay (infinite degrees of freedom, stability)
- Large delay case, $\tau_D \gg \tau$

Unusual features for delay dynamics

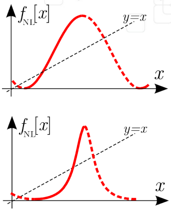
- Bandpass Fourier filter, or integro-differential delay equation
- Dynamical model with a convolution product
- Positive slope operating point, carved nonlinear function profile

A bit of DDE physics & modelling

Mackey–Glass- or Ikeda-like DDE



$$x(t) = \int_{-\infty}^{t} h(t-\xi) \cdot f_{\mathrm{NL}}[x(\xi - \tau_D)]\, \mathrm{d}\xi$$



Non-delayed (instantaneous) terms:

- Linear differential equation, rate of change $\gamma = 1/\tau$
- Stable linear Fourier filter, frequency cut-off $(2\pi\tau)^{-1}$
- A few degrees of freedom $\equiv$ filter or diff.eq. order

Delayed (feedback) term:

- Non-linearity (slope sign, # extrema, multi-stability),
- Delay (infinite degrees of freedom, stability)
- Large delay case, $\tau_D \gg \tau$

Unusual features for delay dynamics

- Bandpass Fourier filter, or integro-differential delay equation
- Dynamical model with a convolution product
- Positive slope operating point, carved nonlinear function profile
- Multiple delay architectures, coupled delay dynamics: many possibilities

Outline

Chimera states. . .

Y. Kuramoto and D. Battogtokh, *Nonlinear Phenom. Complex Syst.* **5**, 380 (2002); D. M. Abrams and S. H. Strogatz, *Phys. Rev. Lett.* **93**, 174102 (2004); I. Omelchenko *et al. Phys. Rev. Lett.* **106** 234102 (2011); A. M. Hagerstrom *et al.* & M. Tinsley *et al.*, *Nat. Phys.* 8, 658 & 662 (2012)

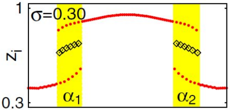
Chimera states. . .

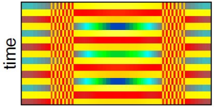
What is a Chimera state?

- Network of coupled oscillators with clusters of incongruent motions
- Predicted numerically in 2002, derived for a particular case in 2004, and 1st observed experimentally in 2012
- Not observed (initially) with local coupling, neither with global one

Y. Kuramoto and D. Battogtokh, *Nonlinear Phenom. Complex Syst.* **5**, 380 (2002); D. M. Abrams and S. H. Strogatz, *Phys. Rev. Lett.* **93**, 174102 (2004); I. Omelchenko *et al. Phys. Rev. Lett.* **106** 234102 (2011); A. M. Hagerstrom *et al.* & M. Tinsley *et al.*, *Nat. Phys.* 8, 658 & 662 (2012)

Chimera states. . .





What is a Chimera state?

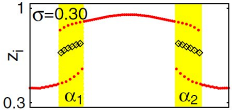
- Network of coupled oscillators with clusters of incongruent motions
- Predicted numerically in 2002, derived for a particular case in 2004, and 1st observed experimentally in 2012
- Not observed (initially) with local coupling, neither with global one

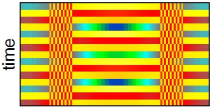
Features allowing for Chimera states?

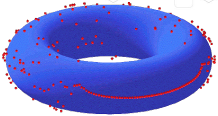
- Network of coupled <u>identical</u> oscillators, spatio-temporal dynamics
- Requires <u>non-local</u> nonlinear coupling between oscillator nodes
- Important parameters: <u>coupling strength</u>, and <u>coupling distance</u>

Y. Kuramoto and D. Battogtokh, *Nonlinear Phenom. Complex Syst.* **5**, 380 (2002); D. M. Abrams and S. H. Strogatz, *Phys. Rev. Lett.* **93**, 174102 (2004); I. Omelchenko *et al. Phys. Rev. Lett.* **106** 234102 (2011); A. M. Hagerstrom *et al.* & M. Tinsley *et al.*, *Nat. Phys.* 8, 658 & 662 (2012)

Chimera states...







What is a Chimera state?

- Network of coupled oscillators with clusters of incongruent motions
- Predicted numerically in 2002, derived for a particular case in 2004, and 1st observed experimentally in 2012
- Not observed (initially) with local coupling, neither with global one

Features allowing for Chimera states?

- Network of coupled identical oscillators, spatio-temporal dynamics
- Requires non-local nonlinear coupling between oscillator nodes
- Important parameters: coupling strength, and coupling distance

Y. Kuramoto and D. Battogtokh, *Nonlinear Phenom. Complex Syst.* **5**, 380 (2002); D. M. Abrams and S. H. Strogatz, *Phys. Rev. Lett.* **93**, 174102 (2004); I. Omelchenko *et al. Phys. Rev. Lett.* **106** 234102 (2011); A. M. Hagerstrom *et al.* & M. Tinsley *et al.*, *Nat. Phys.* 8, 658 & 662 (2012)

Spatio-temporal setup demonstating Chimera

Light controlled Belousov-Zhabotinsky chemical reaction

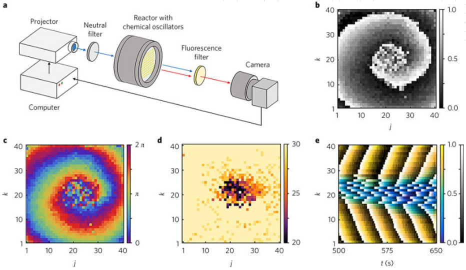
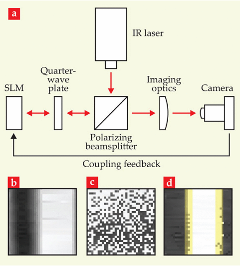


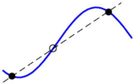
Image formation in a CCD camera – SLM feedback loop

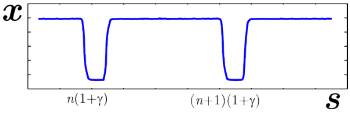


M. Tinsley *et al.*, and A. M. Hagerstrom *et al. Nat. Phys.* 2012

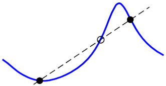
DDE recipe for chimera states

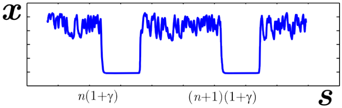
Symmetric $f_{\mathrm{NL}}[x]$: Similar $\sigma-$"clusters" for $x < 0$ and $x > 0$





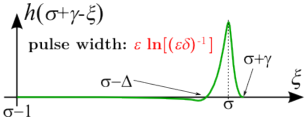
Asymmetric $f_{\mathrm{NL}}[x]$: Distinct $\sigma-$clusters for $x < 0$ and $x > 0$





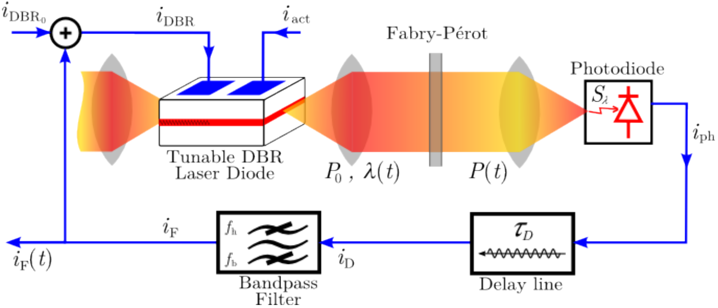
And i DDE

$$\delta \int_{s_0}^{s} x(\xi)\,\mathrm{d}\xi + x(s) + \varepsilon \frac{\mathrm{d}x}{\mathrm{d}s}(s) = f_{\mathrm{NL}}[x(s-1)]$$



Laser based delay dynamics experiment

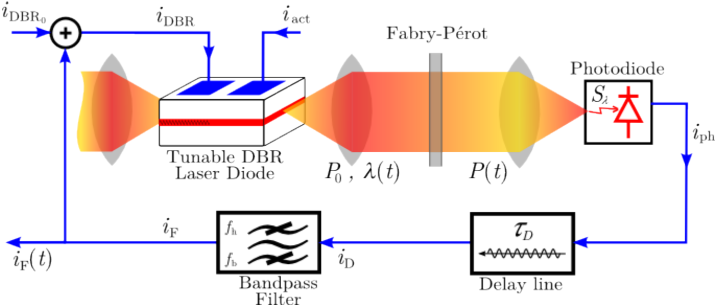
Tunable SC Laser setup, for i DDE Chimera



LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

Laser based delay dynamics experiment

Tunable SC Laser setup, for i DDE Chimera



$f_{\mathrm{NL}}[x]$: the Airy function of a Fabry-Pérot interferometer

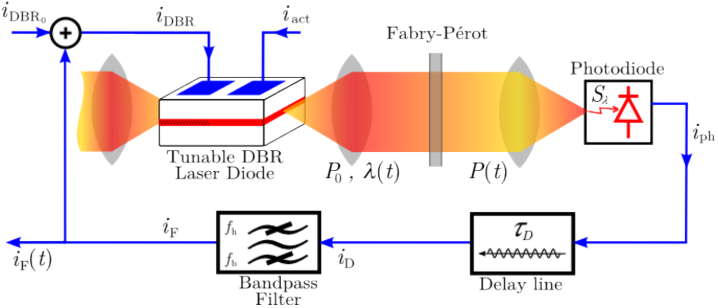
$$\Rightarrow f_{\mathrm{NL}}[\lambda] = \frac{\beta}{1+m\ \sin^2(2\pi ne/\lambda)} = \frac{\beta}{1+m\ \sin^2(x+\Phi_0)}$$

with $x = \frac{2\pi ne}{\lambda_0^2}\,\delta\lambda$ and $\Phi_0 = \frac{2\pi ne}{\lambda_0 + S_{\mathrm{tun.}}\, i_{\mathrm{DBR}_0}}$

LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

Laser based delay dynamics experiment

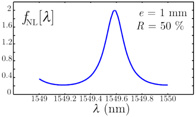
Tunable SC Laser setup, for i DDE Chimera

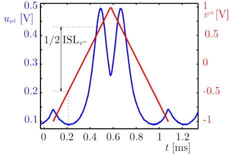


$f_{\mathrm{NL}}[x]$: the Airy function of a Fabry-Pérot interferometer

$$\Rightarrow f_{\mathrm{NL}}[\lambda] = \frac{\beta}{1+m\,\sin^2(2\pi ne/\lambda)} = \frac{\beta}{1+m\,\sin^2(x+\Phi_0)}$$

with $x = \dfrac{2\pi ne}{\lambda_0^2}\,\delta\lambda$ and $\Phi_0 = \dfrac{2\pi ne}{\lambda_0 + S_{\mathrm{tun.}}\, i_{\mathrm{DBR}_0}}$



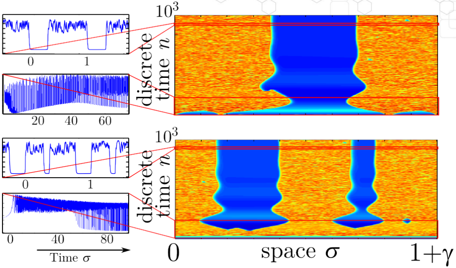


LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

1^{st} Chimera in $(\sigma, n)-$space

Numerics:

- $\beta = 0.6$, $\nu_0 = 1$, $\varepsilon = 5.10^{-3}$, $\delta = 1.6 \times 10^{-2}$
- Initial conditions: small amplitude white noise (repeated several times with different noise realizations)
- Calculated durations: Thousands of n

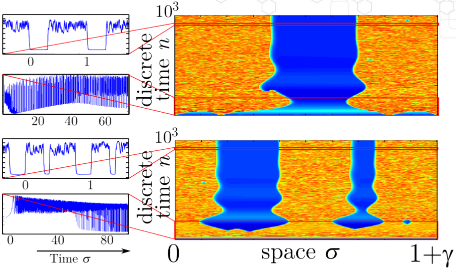


LL *et al. Phys. Rev. Lett.* 2013.

1^{st} Chimera in $(\sigma, n)-$space

Numerics:

- $\beta = 0.6$, $\nu_0 = 1$, $\varepsilon = 5.10^{-3}$, $\delta = 1.6 \times 10^{-2}$
- Initial conditions: small amplitude white noise (repeated several times with different noise realizations)
- Calculated durations: Thousands of n



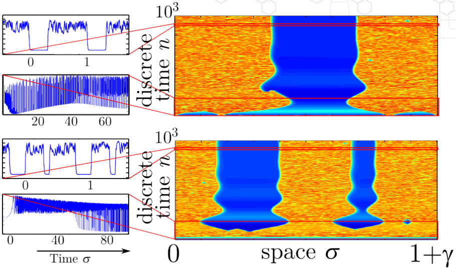
Experiment. . .

LL *et al. Phys. Rev. Lett.* 2013.

1^{st} Chimera in $(\sigma, n)-$space

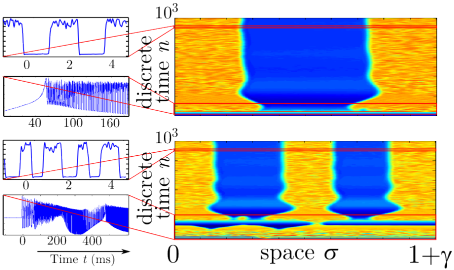
Numerics:

- $\beta = 0.6$, $\nu_0 = 1$, $\varepsilon = 5.10^{-3}$, $\delta = 1.6 \times 10^{-2}$
- Initial conditions: small amplitude white noise (repeated several times with different noise realizations)
- Calculated durations: Thousands of n



Experiment...

- Very close amplitude and time parameters, $\tau_D = 2.54$ms, $\theta = 160$ms, $\tau = 12.7\mu$s
- Initial conditions forced by background noise
- Recording of up to 16×10^6 points, allowing for a few thousands of n



LL *et al. Phys. Rev. Lett.* 2013.

How can a DDE emulate a Network?

Normalization wrt Delay τ_D: $s = t/\tau_D$, and $\varepsilon = \tau/\tau_D$

$$\varepsilon\,\dot{x}(s) = -x(s) + f_{\mathsf{NL}}[x(s-1)], \quad \text{where} \quad \dot{x} = \frac{\mathrm{d}x}{\mathrm{d}s}.$$

Large delay case: $\varepsilon \ll 1$, potentially high dimensional attractor

$\infty-$dimensional phase space, initial condition: $x(s), s \in [-1, 0]$

How can a DDE emulate a Network?

Normalization wrt Delay τ_D: $s = t/\tau_D$, and $\varepsilon = \tau/\tau_D$

$$\varepsilon\,\dot{x}(s) = -x(s) + f_{\mathrm{NL}}[x(s-1)], \quad \text{where} \quad \dot{x} = \frac{\mathrm{d}x}{\mathrm{d}s}.$$

Large delay case: $\varepsilon \ll 1$, potentially high dimensional attractor

$\infty-$dimensional phase space, initial condition: $x(s), s \in [-1, 0]$

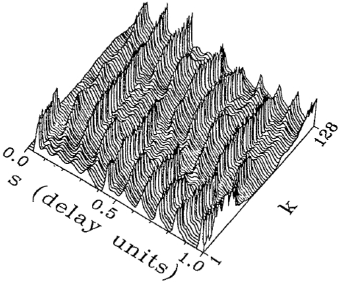
Space-time representation

- Virtual space variable σ,

  $\sigma \in [0; 1+\gamma]$ with $\gamma = O(\varepsilon)$.

- Discrete time n

$$\begin{array}{ccc} n & \rightarrow & (n+1) \\ s = n(1+\gamma)+\sigma & \rightarrow & s = (n+1)(1+\gamma)+\sigma \end{array}$$



F.T. Arecchi, *et al. Phys. Rev. A* 1992

How can a DDE emulate a Network?

Convolution product involving the linear impulse response,

$$h(t) = \textbf{FT}^{-1}[H(\omega)]$$

$$x(s) = \int_{-\infty}^{s} h(s-\xi) \cdot f_{\mathrm{NL}}[x(\xi - 1)]\,\mathrm{d}\xi \quad \text{with} \quad s = n(1+\gamma) + \sigma$$

LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

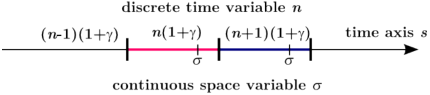
How can a DDE emulate a Network?

Convolution product involving the linear impulse response,

$$h(t) = \mathbf{FT}^{-1}[H(\omega)]$$

$x(s) = \int_{-\infty}^{s} h(s-\xi)\cdot f_{\mathrm{NL}}[x(\xi-1)]\,\mathrm{d}\xi$ with $s = n(1+\gamma)+\sigma$

...partitioning the time domain:



$$]-\infty; s] =]-\infty\,;\, n(1+\gamma)+\sigma\,] \quad\cup\quad]n(1+\gamma)+\sigma\,;\,(n+1)(1+\gamma)+\sigma\,]$$

LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

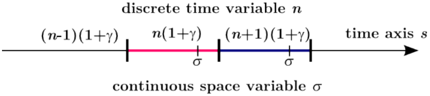
How can a DDE emulate a Network?

Convolution product involving the linear impulse response,

$$h(t) = \mathbf{FT}^{-1}[H(\omega)]$$

$$x(s) = \int_{-\infty}^{s} h(s-\xi)\cdot f_{\mathrm{NL}}[x(\xi-1)]\,\mathrm{d}\xi \quad \text{with} \quad s = n(1+\gamma)+\sigma$$

...partitioning the time domain:



$$]-\infty;s] =]-\infty\,;\,n(1+\gamma)+\sigma\,] \quad \cup \quad]n(1+\gamma)+\sigma\,;\,(n+1)(1+\gamma)+\sigma\,]$$

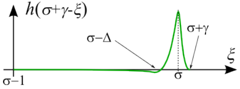
and make a change of integration variable $\xi \leftrightarrow \xi-(n+1)(1+\gamma)+\gamma$

LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

How can a DDE emulate a Network?

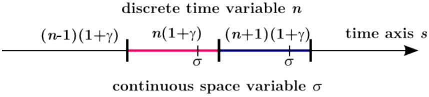
Convolution product involving the linear impulse response,

$$h(t) = \mathbf{FT}^{-1}[H(\omega)]$$



$x(s) = \int_{-\infty}^{s} h(s-\xi)\cdot f_{\mathsf{NL}}[x(\xi-1)]\,\mathrm{d}\xi$ with $s = n(1+\gamma)+\sigma$

...partitioning the time domain:



$$]-\infty; s] \;=\;]-\infty\,;\, n(1+\gamma)+\sigma\,] \quad \cup \quad]n(1+\gamma)+\sigma\,;\,(n+1)(1+\gamma)+\sigma\,]$$

and make a change of integration variable $\xi \leftrightarrow \xi-(n+1)(1+\gamma)+\gamma$

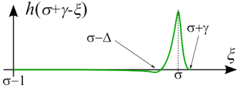
$$\Rightarrow \quad x_{n+1}(\sigma) = I_\epsilon(n,\sigma) + \int_{\sigma-1}^{\sigma+\gamma} h(\sigma+\gamma-\xi)\cdot f_{\mathsf{NL}}[x_n(\xi)]\,\mathrm{d}\xi, \quad \text{with } I_\epsilon \ll x_n(\sigma)$$

LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

How can a DDE emulate a Network?

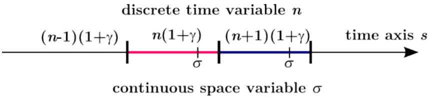
Convolution product involving the linear impulse response,

$$h(t) = \mathbf{FT}^{-1}[H(\omega)]$$



$$x(s) = \int_{-\infty}^{s} h(s-\xi)\cdot f_{\mathrm{NL}}[x(\xi-1)]\,\mathrm{d}\xi \quad \text{with} \quad s = n(1+\gamma)+\sigma$$

...partitioning the time domain:

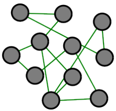


$$]-\infty; s] \;=\;]-\infty\,;\, n(1+\gamma)+\sigma\,] \quad \cup \quad]n(1+\gamma)+\sigma\,;\,(n+1)(1+\gamma)+\sigma\,]$$

and make a change of integration variable $\xi \leftrightarrow \xi-(n+1)(1+\gamma)+\gamma$

$$\Rightarrow \quad x_{n+1}(\sigma) = I_\epsilon(n,\sigma) + \int_{\sigma-1}^{\sigma+\gamma} h(\sigma+\gamma-\xi)\cdot f_{\mathrm{NL}}[x_n(\xi)]\,\mathrm{d}\xi, \quad \text{with} \;\; I_\epsilon \ll x_n(\sigma)$$

$$\left\{ \frac{\partial \phi}{\partial t} = \omega - \int_{-\pi}^{\pi} G(x-x')\cdot \sin[\phi(x,t)-\phi(x',t)+\alpha]\,\mathrm{d}x \right\}$$

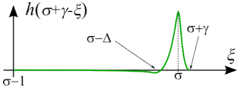


LL, Penkovsky, Maistrenko, *Nat. Commun.* 2015

How can a DDE emulate a Network?

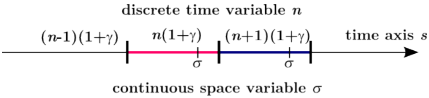
Convolution product involving the linear impulse response,

$$h(t) = \textsf{FT}^{-1}[H(\omega)]$$



$x(s) = \int_{-\infty}^{s} h(s-\xi)\cdot f_{\mathsf{NL}}[x(\xi-1)]\,\mathsf{d}\xi$ with $s = n(1+\gamma)+\sigma$

...partitioning the time domain:



$$]-\infty; s] \;=\;]-\infty\,;\, n(1+\gamma)+\sigma\,] \quad \cup \quad]n(1+\gamma)+\sigma\,;\,(n+1)(1+\gamma)+\sigma\,]$$

and make a change of integration variable $\xi \leftrightarrow \xi-(n+1)(1+\gamma)+\gamma$

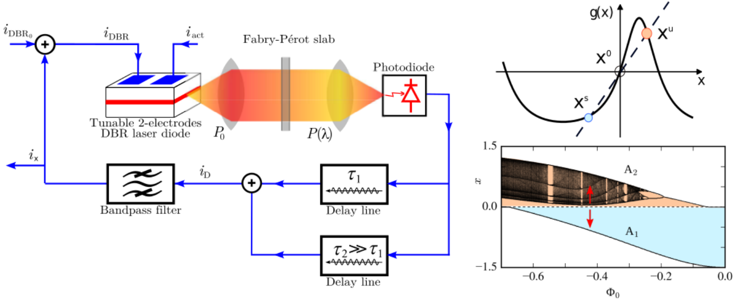
$$\Rightarrow \quad x_{n+1}(\sigma) = I_\epsilon(n,\sigma) + \int_{\sigma-1}^{\sigma+\gamma} h(\sigma+\gamma-\xi)\cdot f_{\mathsf{NL}}[x_n(\xi)]\,\mathsf{d}\xi, \quad \text{with} \;\; I_\epsilon \ll x_n(\sigma)$$

$$\left\{ \frac{\partial\phi}{\partial t} = \omega - \int_{-\pi}^{\pi} G(x-x')\cdot\sin[\phi(x,t)-\phi(x',t)+\alpha]\,\mathsf{d}x \right\}$$

Remark: the NL dynamics and coupling features of each virtual oscillator are by construction <u>identical</u> at any position σ!!!

Double delay dynamics: toward 2D chimera

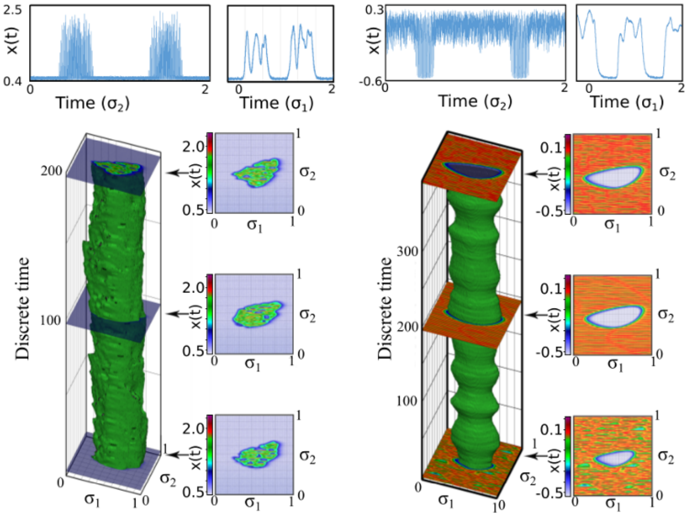
Setup and delay dynamics features



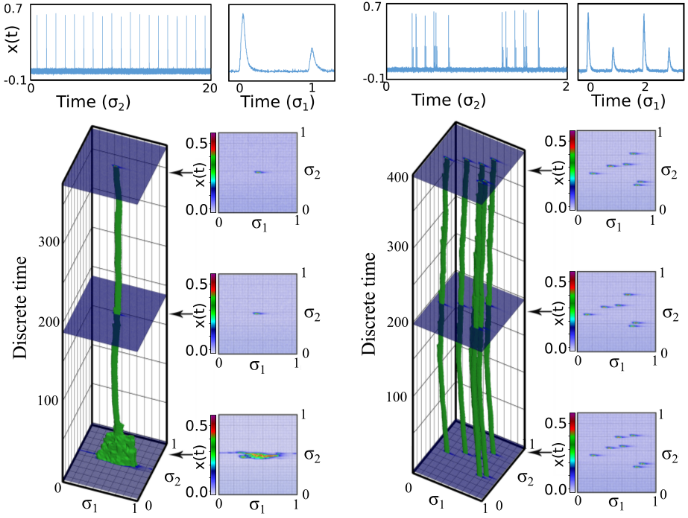
Double delay nonlinear integro-differential equation

$$\varepsilon \frac{\mathrm{d}x}{\mathrm{d}t}(t) + x(t) + \delta \int x(\xi)\mathrm{d}\xi = (1-\gamma) f_{\mathrm{NL}}[x(t-\tau_1)] + \gamma f_{\mathrm{NL}}[x(t-\tau_2)]$$

2D-chimera with chaotic sea, or chaotic island



Isolated pulses



Outline

Background and motivations

Basics in Reservoir Computing (RC)

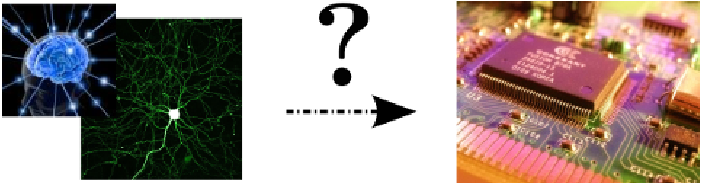
Delay dynamics as an emulated network of neurons

Reservoir Computing with Photonic DDE

Conclusion & perspectives

Emulating RNN with delay dynamics

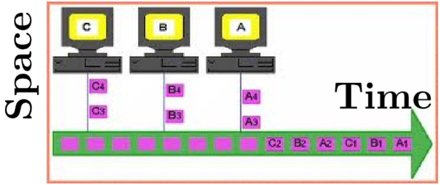
A convenient hardware solution for RC



- Designing a complex and controlled 3D network of nodes as a brain: a very difficult technological challenge

Emulating RNN with delay dynamics

A convenient hardware solution for RC



- Designing a complex and controlled 3D network of nodes as a brain: a very difficult technological challenge
- Serial processing: common in many communication systems

Emulating RNN with delay dynamics

A convenient hardware solution for RC

VOLUME 73, NUMBER 8 PHYSICAL REVIEW LETTERS 22 AUGUST 1994

Defects and Spacelike Properties of Delayed Dynamical Systems

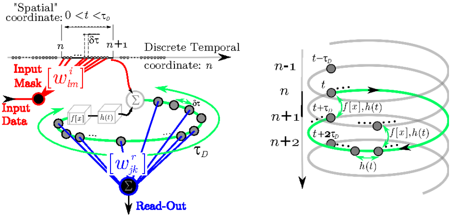
G. Giacomelli,[1,2] R. Meucci,[1] A. Politi,[1,3] and F. T. Arecchi[1,4]
[1]*Istituto Nazionale di Ottica, 50125 Firenze, Italy*
[2]*ITIS "Tullio Buzzi," Prato, Italy*
[3]*INFN, Sezione di Firenze, Firenze, Italy*
[4]*Dipartimento di Fisica, Universitá di Firenze, Firenze, Italy*
(Received 11 January 1994)

In a laser with delayed feedback operating in an oscillatory regime, phase defects appear for delays longer than the oscillation period. These defects are visualized by rearranging the data in a two-dimensional representation. Two distinct disordered phases are observed, one of weak turbulence characterized by phase fluctuations, and one of highly developed turbulence characterized by a constant density of defects. The transition between the two regimes is analyzed by studying the dependence of the defect lifetime on the delay. The experimental findings are modeled via a generalized Landau equation which includes a delayed coupling.

- Designing a complex and controlled 3D network of nodes as a brain: a very difficult technological challenge
- Serial processing: common in many communication systems
- Delay dynamics known as virtual Space-Time dynamics

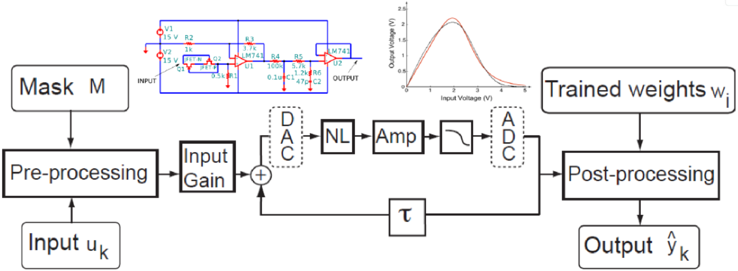
Emulating RNN with delay dynamics

A convenient hardware solution for RC



- Designing a complex and controlled 3D network of nodes as a brain: a very difficult technological challenge
- Serial processing: common in many communication systems
- Delay dynamics known as virtual Space-Time dynamics
- Schematic of RC architecture with delay dynamics

Examples of delay-based RC

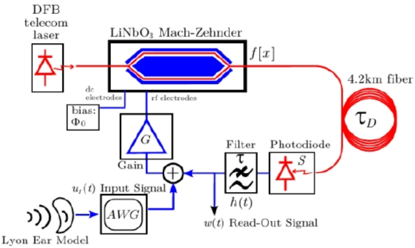


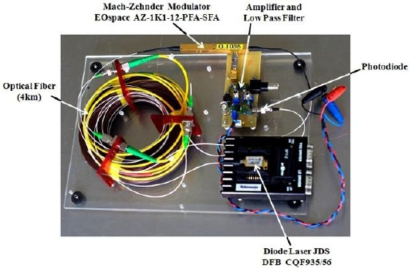
...and now even available in hardware

- Low speed analogue electronic

Appeltant *et al.*, *Nature Commun.* 2011.

Examples of delay-based RC



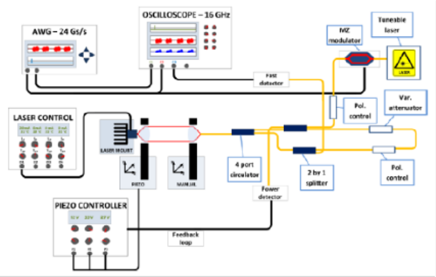


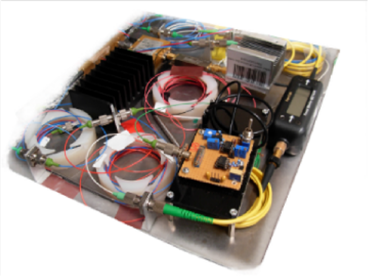
...and now even available in hardware

- Low speed analogue electronic
- Moderate speed optoelectronic

LL *et al.*, *Opt. Expr.* 2012. Paquot *et al.*, *Sci. Rep.* 2012. Martinenghi *et al.*, *Phys. Rev. Lett.* 2012.

Examples of delay-based RC





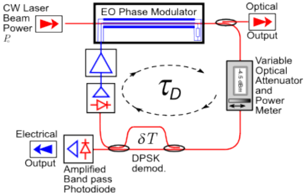
. . . and now even available in hardware

- Low speed analogue electronic
- Moderate speed optoelectronic
- High speed all-optical and optoelectronic demo

Brunner *et al.*, *Nature Comm.* 2013. LL *et al.*, *Phys. Rev. X* 2017. Fiers *et al.*, *Nature Comm.* 2014.

EO phase setup: Modeling

RF Bandpass filter, DPSK NL delayed feedback

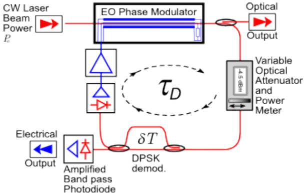


EO phase setup: Modeling

RF Bandpass filter, DPSK NL delayed feedback

- Integro-differential (linear bandpass filter) nonlinear delay equation

$$\frac{1}{\theta}\int_{t_0}^{t}\varphi(\xi)\,\mathrm{d}\xi + \varphi(t) + \tau\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t) = \beta\cdot\left[f_{(t-\tau_D)}(\varphi^\star)\right]$$



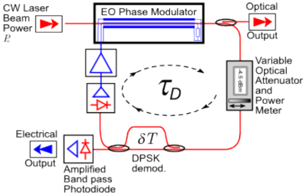
EO phase setup: Modeling

RF Bandpass filter, DPSK NL delayed feedback

- Integro-differential (linear bandpass filter) nonlinear delay equation

$$\frac{1}{\theta}\int_{t_0}^{t} \varphi(\xi)\,\mathrm{d}\xi + \varphi(t) + \tau\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t) = \beta \cdot \left[f_{(t-\tau_D)}(\varphi^\star)\right]$$

- Non linearity via imbalanced interferometer (temporal non locality)



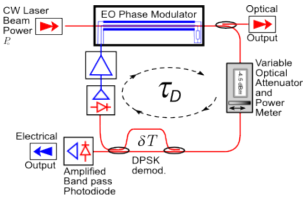
EO phase setup: Modeling

RF Bandpass filter, DPSK NL delayed feedback

- Integro-differential (linear bandpass filter) nonlinear delay equation

$$\frac{1}{\theta}\int_{t_0}^{t}\varphi(\xi)\,\mathrm{d}\xi + \varphi(t) + \tau\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t) = \beta\cdot\left[f_{(t-\tau_D)}(\varphi^\star)\right]$$

- Non linearity via imbalanced interferometer (temporal non locality)



- Standard DPSK demodulator

$$f_t(\varphi) = \{1 + \cos[\varphi(t) - \varphi(t-\delta T) + \Phi_0]\}$$

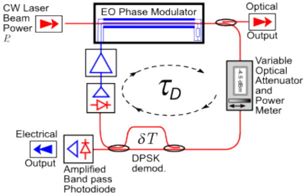
EO phase setup: Modeling

RF Bandpass filter, DPSK NL delayed feedback

- Integro-differential (linear bandpass filter) nonlinear delay equation

$$\frac{1}{\theta}\int_{t_0}^{t}\varphi(\xi)\,\mathrm{d}\xi + \varphi(t) + \tau\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t) = \beta\cdot\left[f_{(t-\tau_D)}(\varphi^\star)\right]$$

- Non linearity via imbalanced interferometer (temporal non locality)



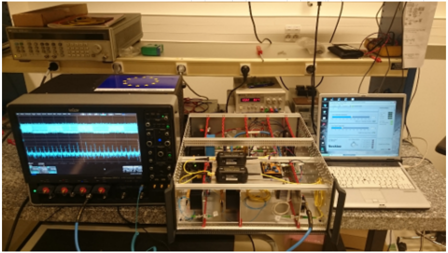
- Standard DPSK demodulator

$$f_t(\varphi) = \{1 + \cos[\varphi(t) - \varphi(t-\delta T) + \Phi_0]\}$$

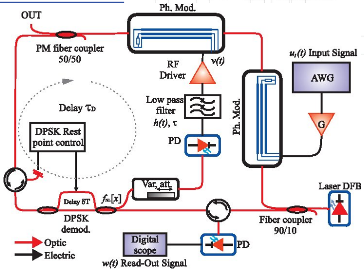
- Generalized multiple wave interferometer

$$f_t(\varphi) = F_0\left|1 + \sum_k \alpha_k\, e^{i[\varphi(t)-\varphi(t-\delta T_k)+\Phi_k]}\right|^2$$

RC operation of an EO phase setup



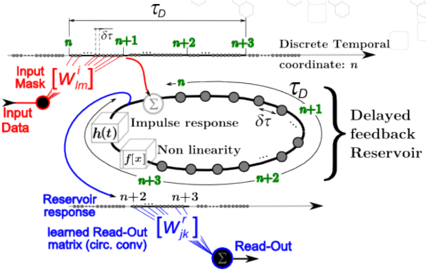
RC operation of an EO phase setup



Amplitude parameters

- Input ΦM amplitude: 1.2π
- feedback gain: $\beta \simeq 0.7$
- offset phase: $\Phi_0 \simeq 2\pi/5$ (nearly parabolic)

RC operation of an EO phase setup



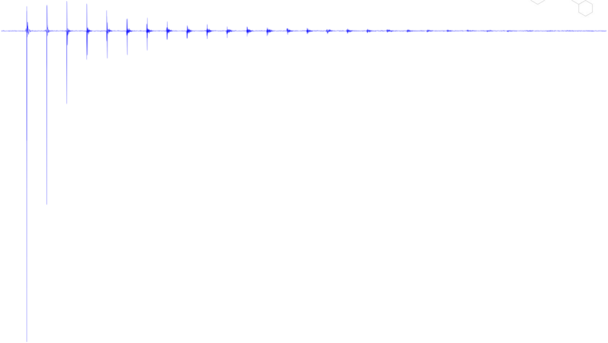
Amplitude parameters

- Input ΦM amplitude: 1.2π
- feedback gain: $\beta \simeq 0.7$
- offset phase: $\Phi_0 \simeq 2\pi/5$ (nearly parabolic)

Time parameters

- Loop filter bandwidth: 566 MHz $\Rightarrow \delta\tau \simeq 56.8$ ps (AWG limited, 17.6 GS/s)
- Time delay: $\tau_D \simeq 63.33$ ns (a few meters of fiber)
- internal input sample memory: 3 $\Rightarrow 371$ virtual nodes / input sample, or 1113 virtual nodes / time delay: **“hidden” layers within the delay**

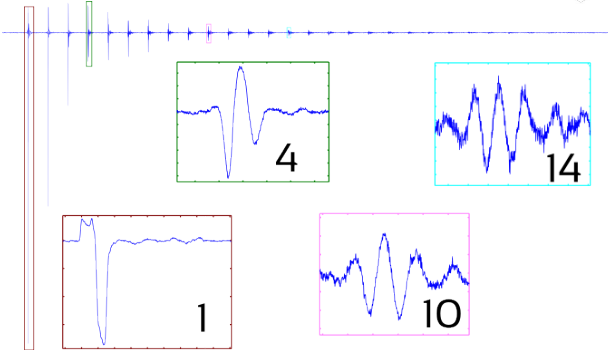
Impulse response: Delay-RC characterization



Injecting a single pulse

- Echoes (memory, feedback strength, instability neighborhood)

Impulse response: Delay-RC characterization



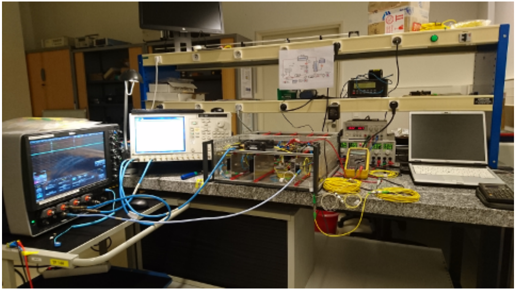
Injecting a single pulse

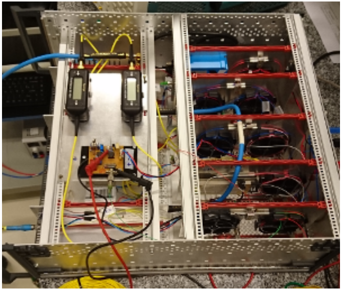
- Echoes (memory, feedback strength, instability neighborhood)
- Head: Nonlinear transformation (input amplitude, nonlinear scan)
- Tail: Linear response

Impulse response: Delay-RC characterization

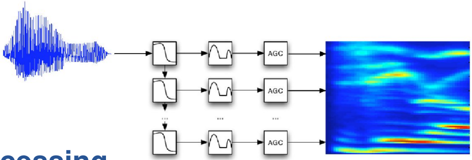
The whole setup

- Information injection by an AWG
- Dynamical processing by the EO phase DDE
- Recording of the dynamical response (oscilloscope)





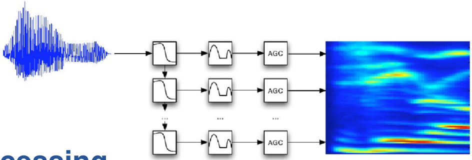
Dynamical Processing of Spoken Digits



Input pre-processing

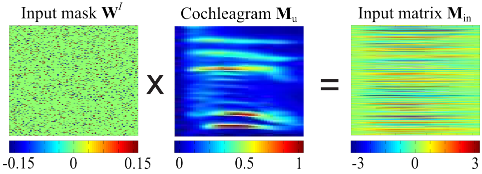
- Lyon Ear Model transformation (Time & Frequency 2D formatting, 60 Samples x 86 Freq.channel)

Dynamical Processing of Spoken Digits

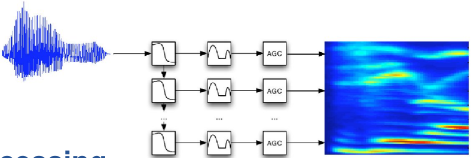


Input pre-processing

- Lyon Ear Model transformation (Time & Frequency 2D formatting, 60 Samples x 86 Freq.channel)
- Sparse “connection” of the 86 Freq. channel to the 371 neurons: random connection matrix

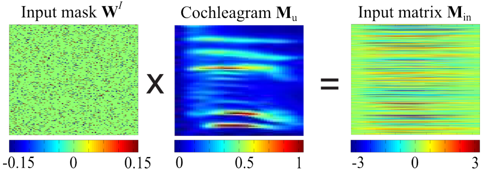


Dynamical Processing of Spoken Digits



Input pre-processing

- Lyon Ear Model transformation (Time & Frequency 2D formatting, 60 Samples x 86 Freq.channel)
- Sparse "connection" of the 86 Freq. channel to the 371 neurons: random connection matrix



Reservoir transient response:

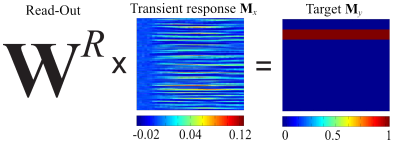
- Time series record for Read-Out post-processing

Read-Out, Training, and Testing

Training of the Read-Out with target output function

Learning: optimization of the W matrix, for each different digit

$\rightarrow$ Regression problem for $A \times W \simeq B$:
$W_{\text{opt}} = (A^T A - \lambda I)^{-1} A^T B$



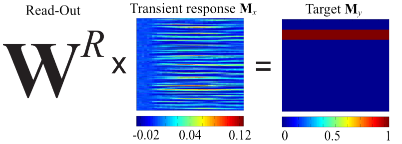
(LL *et al.*, *Phys. Rev. X* 2017)

Read-Out, Training, and Testing

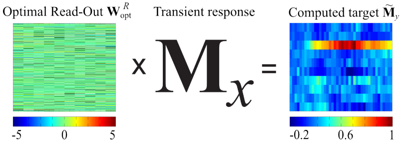
Training of the Read-Out with target output function

Learning: optimization of the W matrix, for each different digit

$\rightarrow$ Regression problem for $A \times W \simeq B$:
$W_{\text{opt}} = (A^T A - \lambda I)^{-1} A^T B$



Testing with training-defined Read-Out



(LL *et al.*, *Phys. Rev. X* 2017)

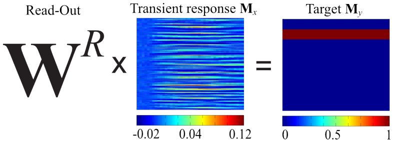
Read-Out, Training, and Testing

Training of the Read-Out with target output function

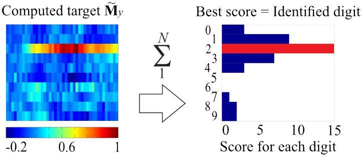
Learning: optimization of the W matrix, for each different digit

$\rightarrow$ Regression problem for $A \times W \simeq B$:

$W_{\text{opt}} = (A^T A - \lambda I)^{-1} A^T B$



Testing with training-defined Read-Out



(LL *et al.*, *Phys. Rev. X* 2017)

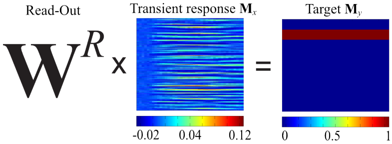
Read-Out, Training, and Testing

Training of the Read-Out with target output function

Learning: optimization of the W matrix, for each different digit

$\rightarrow$ Regression problem for $A \times W \simeq B$:

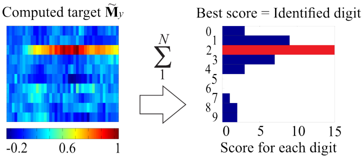
$$W_{\text{opt}} = (A^T A - \lambda I)^{-1} A^T B$$



Testing with training-defined Read-Out

Test result: State of the art
(close to 0% Word Error Rate)

With Telecom Bandwidth setup:
record speed recognition, 1M word/s



(LL *et al.*, *Phys. Rev. X* 2017)

Outline

Background and motivations

Basics in Reservoir Computing (RC)

Delay dynamics as an emulated network of neurons

Reservoir Computing with Photonic DDE

Conclusion & perspectives

Complexifying delay-based RC

Advanced architectures

- Parallel delay-RC units, with distributed filter parameters
- Cascaded delay-RC units, with different filtering features: Deep convolutional RC

Integrated optics photonic chips

- Photonic hybrid integrated technologies are mature to fabricate collective delay-RC photonic chips, with Opto-Electronic architectures
- Interfacing input and output data is still challenging

Learning: From supervised to unsupervised

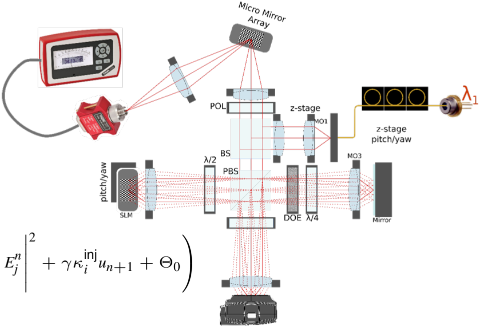
- Could learning of the Read-Out coefficients be viewed as a pattern formation triggered by some specific data feature to be filtered?
- Chimera as the spontaneous formation of a pattern allowing for feature extraction?

Grigoryeva *et al.*, *Neural Networks* 2014. Penkovsky *et al. Phys. Rev. Lett.* 2019

3D spatio-temporal photonic RC

Experimental bulk optics setup (D. Brunner, M. Jacquot)

- Nodes are spatially distributed in an image plane
- Coupling between nodes makes use of DOE
- Nonlinear is performed by SLM (polarization filtering)
- Read-Out is physically implemented (cascaded DMD and a photodiode)



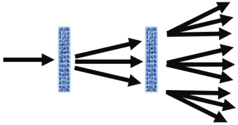
$$I_i^{n+1} = \sin^2\left(\beta\left|\sum_{j=1}^{N}\kappa_{i,j}\,E_j^n\right|^2 + \gamma\kappa_i^{\text{inj}}u_{n+1} + \Theta_0\right)$$

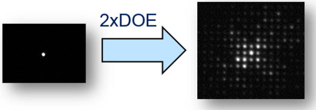
J. Bueno *et al.*, *Optica*, 2018

3D spatio-temporal photonic RC

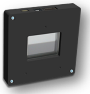
Elements characterization

- Node coupling: two cascaded DOE

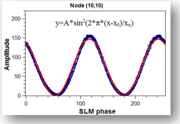




- Nonlinear transformation (SLM)



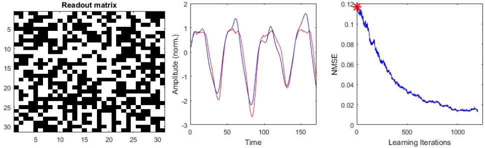
Polarization filtering



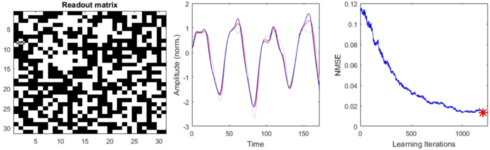
3D spatio-temporal photonic RC

Chaotic time series prediction

- Random initialization and learing



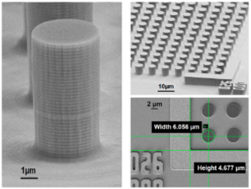
- After re-inforcement learning



Toward integrated 3D photonic RC chips

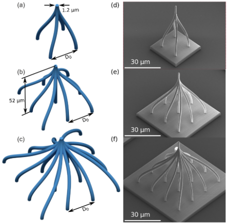
Integrated laser array

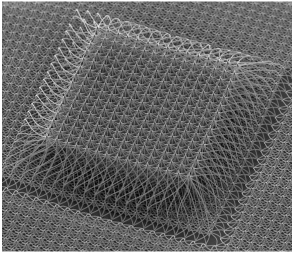
(Stefan Reizenstein *et al.*, TU Berlin)



3D-printed integrated photonic couplers

(Daniel Brunner *et al.*, FEMTO-ST)





Heuser *et al.*, *IEEE JSTQE* 2020. Moughames *et al.*, *Optical Materials Express* 2020

Thank you for attention

Thanks to many collaborators, from many disciplines:

FEMTO-ST Y.K. Chembo, M. Jacquot, D. Brunner, J.M. Dudley, M. Kadic
PhD students: R. Martinenghi, A.B. Fuentes, B. Penkovsky, B. Marquez, L. Andreoli, S. Maktoobi, N. Semenova, V. Semenov
Postdoc: X. Porte, J. Moughames

formerly @ LMB J.-P. Ortega (Sankt Gallen), L. Grigoryeva (Konstanz)

TU Berlin E. Schöll, Y. Maistrenko, R. Levchenko, S. Reizenstein

IFISC I. Fischer, P. Colet, C.R. Mirasso, M.C. Soriano
PhD students: R.M. Nguimdo, N. Oliver, J. Bueno

ULB T. Erneux
PhD student: L. Weicker

Univ. Maryland R. Roy, T.E. Murphy, Y.K. Chembo
PhD student: J.D. Hart

VUB J. Danckaert, G. Van der Sande
PhD student: L. Appeltant