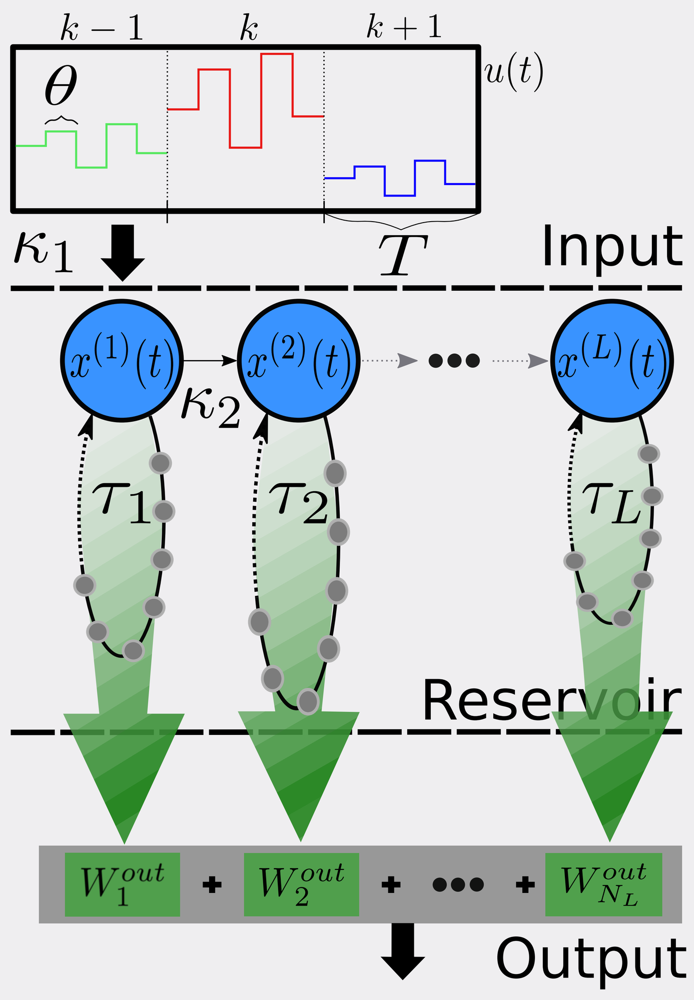
Deep Time-Delay Reservoir Computing: Dynamics and Memory Capacity

Mirko Goldmann[1], Felix Köster[2], Kathy Lüdge[2], Serhiy Yanchuk[3]

1. Instituto de Fisica Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus Universitat Illes Balears, E-07122 Palma de Mallorca, Spain
2. Institute of Theoretical Physics, Technische Universität Berlin, D-10623, Germany
3. Institute of Mathematics, Technische Universität Berlin, D-10623, Germany

mirko@ifisc.uib-csic.es

1. Model



Goal:

- supervised learning of temporal sequences $(s(k), \hat{o}(k))$

Input Preprocessing

- time-multiplexing with T - periodic mask m_j

$$u(t) = u_{k,j} = s(k)m_j \text{ for } t \in [kT + (j-1)\theta, kT + j\theta]$$

- $N_V = \frac{T}{\theta}$ virtual nodes

Reservoir

- L unidirectional instantaneous coupled layers
- input only to first layer
- dynamics given by DDEs with different delays

$$\dot{x}^{(l)}(t) = F^{(l)}(x^{(l)}(t), x^{(l)}(t-\tau_l), J^{(l)}(t))$$

$$J^{(l)}(t) = \begin{cases} u(t) \text{ for } l = 1 \\ x^{(l-1)}(t) \text{ else} \end{cases}$$

Output

- Linear Combination of all node states $x_g(k)$

$$o(k) = W^{out} x_g(k)$$

- $W^{out} \in \mathbb{R}^{LN_V}$ via simple linear regression

Memory Capacity (MC): [2]

MC Degree	sum over recalls of	Examples
Linear MC	previous inputs (linear)	$P_1(k-1)$, $P_1(k-2)$, $P_1(k-3)$
Quadratic MC	products of up to two inputs (nonlinear)	$P_2(k-1)$, $P_1(k-1)P_1(k-2)$, $P_1(k-1)P_1(k-3)$
Cubic MC	products of up to three inputs (nonlinear)	$P_3(k-1)$, $P_2(k-1)P_1(k-2)$, $P_1(k-1)P_2(k-2)$, $P_1(k-1)P_1(k-2)P_1(k-3)$

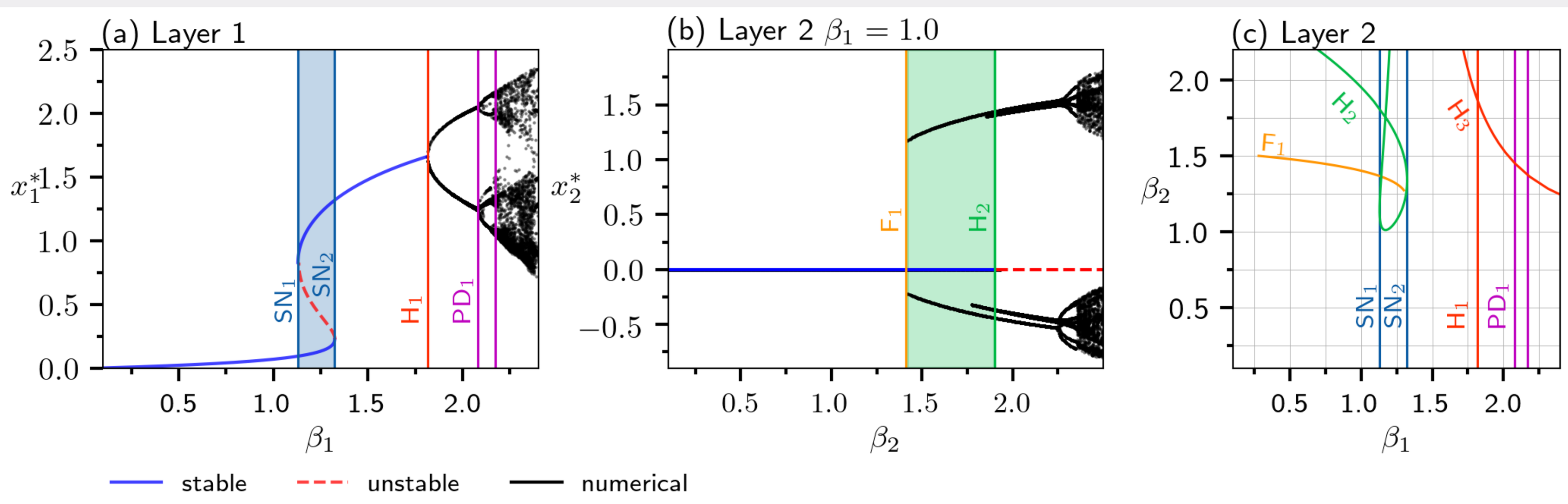
$P_d(\dots)$ Legendre Polynomial of degree d
Total MC $\hat{=}$ sum over all recall-abilities and bounded by readout dimension LN_V

2. Ikeda System

$$\dot{x}^{(l)}(t) = -x^{(l)}(t) - \delta_l y^{(l)}(t) + \beta_l \sin^2(x^{(l)}(t-\tau_l) + \kappa_l J^{(l)}(t) + b_l)$$
$$\dot{y}^{(l)}(t) = x^{(l)}(t)$$

Parameter	Description	Value
δ_1	damping	0.01
β_l	feedback gain	$[0, 2.5]$
b	bias	0.2
$\kappa_{l>1}$	coupling strength	1
$\kappa_{l=1}$	input gain	0.01
τ_l	delay time	30
T	clock cycle	25
N_V	virtual nodes	25

- coupling in nonlinear regime of $\sin^2(\dots)$
- consistency condition of Reservoir Computing: autonomous system is asymptotically stable



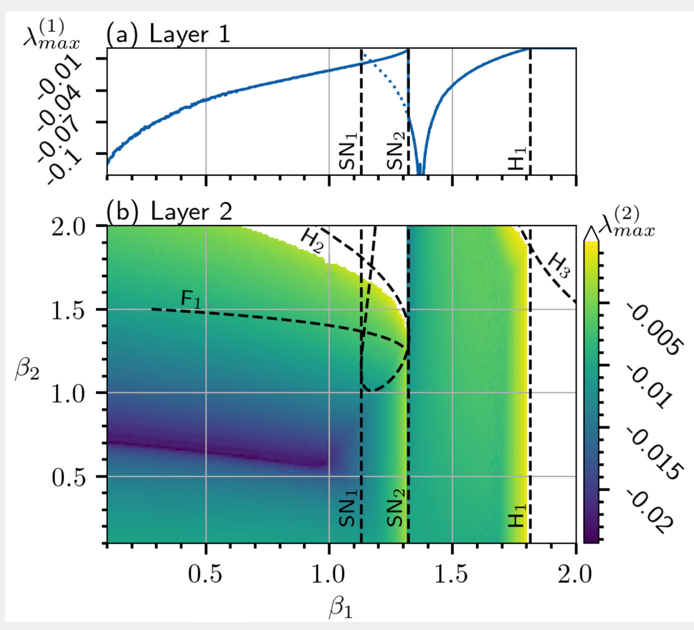
SN - saddle-node bifurcation, H (red) - supercrit. Hopf bifurcation,
H (green) - subcrit. Hopf bifurcation, F - Fold Bifurcation, PD - Period Doubling

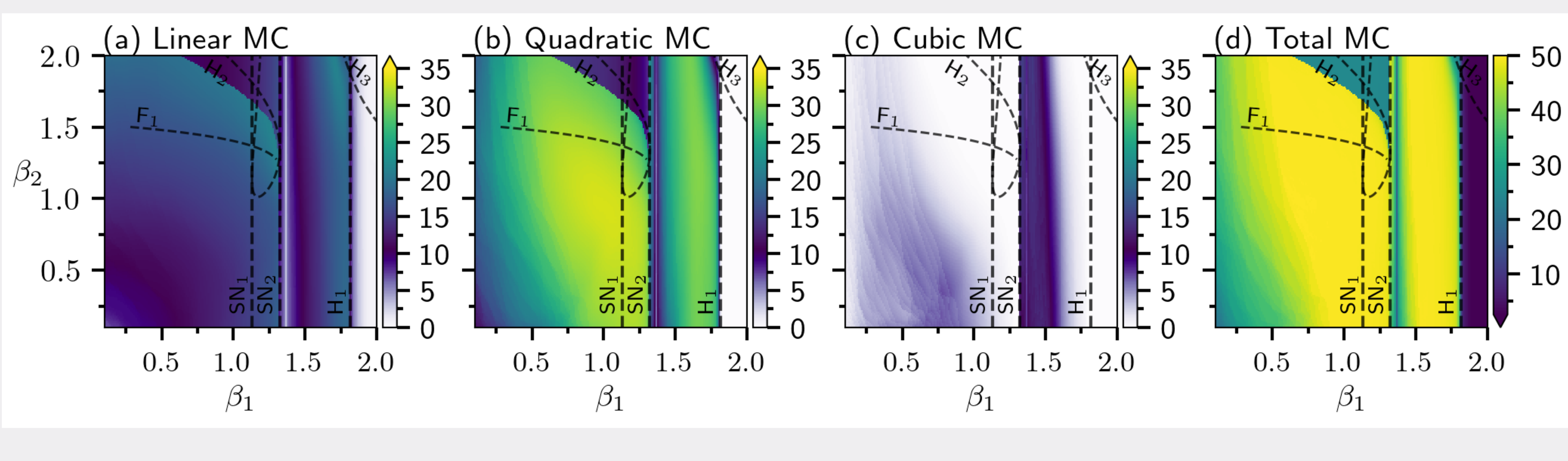
3. Conditional Lyapunov Exponent

- 2 layer system driven by input $s(k) \sim \mathcal{U}[-1, 1]$
- numerical computation of Conditional Lyapunov Exponent (CLE):
  - evaluation of states with different initial conditions $x^{(1)}(t, \phi), x^{(1)}(t, \phi')$

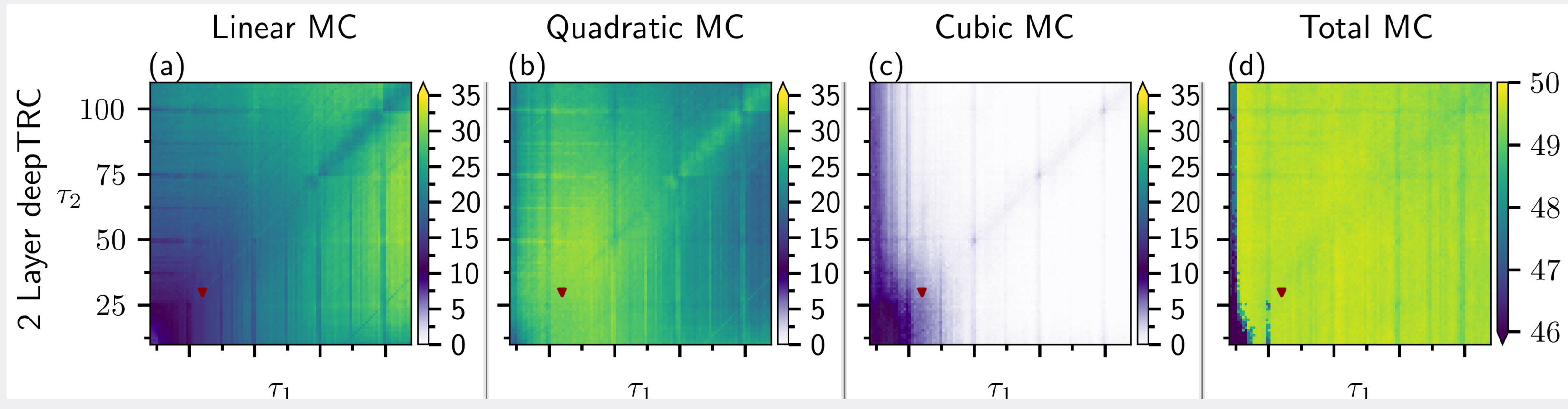
$$e^{\lambda_{max}^{(l)} t} \approx \frac{\|x^{(l)}(t) - x_p^{(l)}(t)\|}{\|x^{(1)}(0, \phi) - x^{(1)}(0, \phi')\|}$$

- negative CLE indicates generalized synchronization
  $\rightarrow$ Echo State Property
- negative near zero CLE marks high Linear MC
  $\rightarrow$ close to Hopf-Bifurcation
- Total $MC = LN_v$ is reached in a wide range

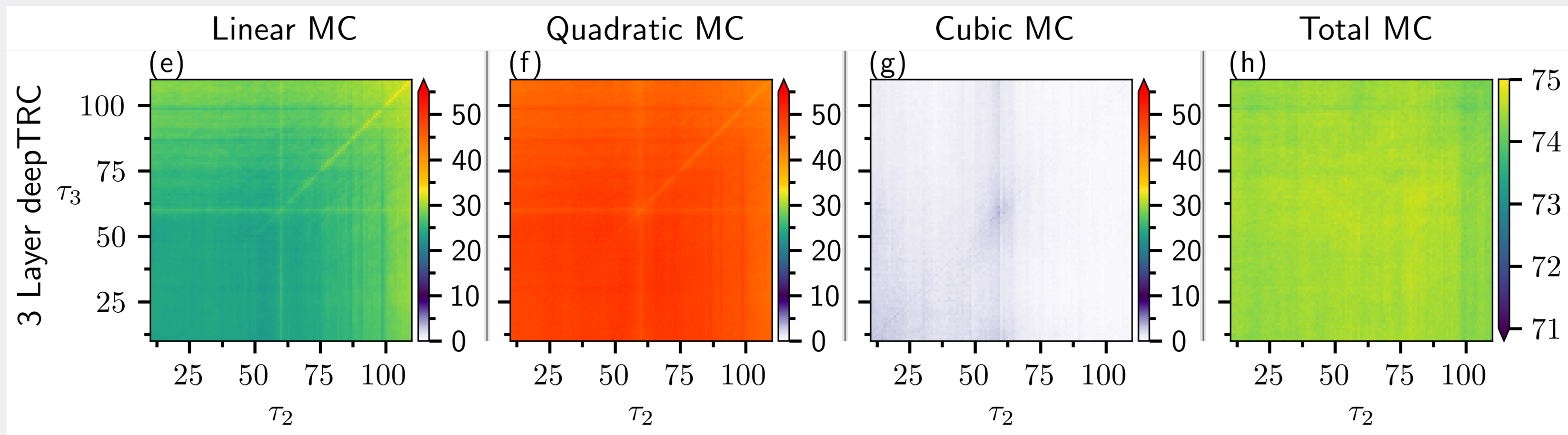




4. Delay Resonances



- $m\tau_l = nT$, $n, m \in \mathbb{Z}$ resonances in all MC degrees
- degradation of total MC
- linear & quadratic MC invert $m\tau_l = nT$ resonance at main diagonal
- resonance at main diagonal $\tau_1 = \tau_2$
  - decreased linear MC
  - increased nonlinear MC
- weak off-diagonal resonances



- $\tau_1 = 60$ kept fixed
- resonances with τ_1
- resonances at main diagonal $\tau_2 = \tau_3 > \tau_1$
- cubic MC high for $\tau_1 = \tau_2 = \tau_3$

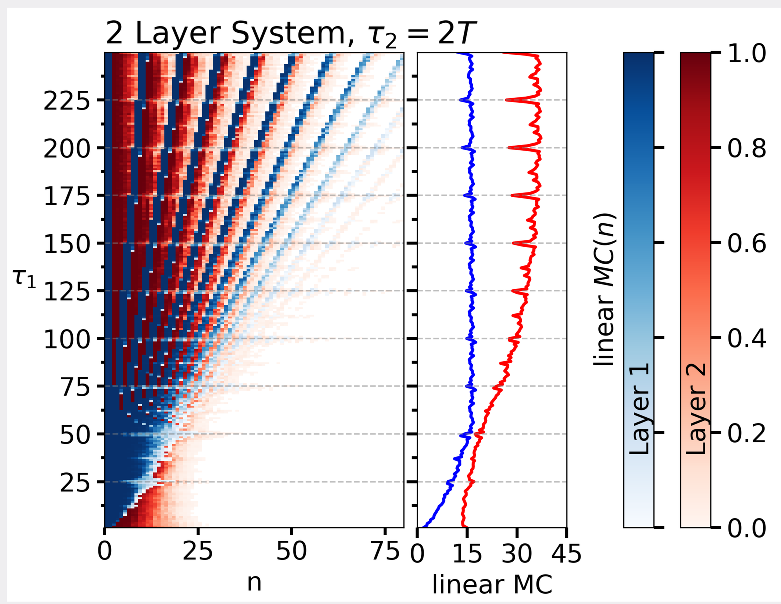
5. Augmentation

Single Layer [5]

- $\tau_1 \gtrsim 2T \rightarrow$ linear MC forks into rays
- frequency of rays $\approx \tau_1/T$
- at $\tau_1 = nT, n \in \mathbb{Z}$ slower decay of the recallability

2 Layer

- $\tau_1 \approx \tau_2$ weak extension of linear MC
- $\tau_2 < 2T < \tau_1$ augments ray structure
- envelope of linear MC in layer 2 decays faster



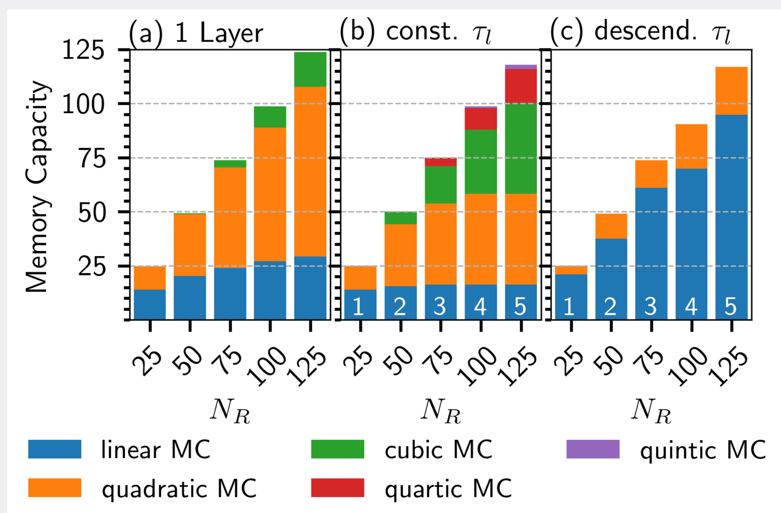
6. Deep Configurations

Single Layer

- increasing N_V increases $MC_d = 2$
- slow enhancing higher nonlinear MC

Multiple Layer

- constant delays boost nonlinear MC
- loss of total MC
- descending delays near clock cycle resonances $\tau_i = 0.5\tau_{i+1}$ increase linear MC



7. Conclusion

- adding layer increase computational capabilities
  - more virtual nodes by constant clock cycle
  - more variability in Memory Capacity
- reaching total MC up to L ≤ 3
- negative close zero CLE indicates high linear MC
  - strong relation between MC and dynamics
- variation of delays provides balancing of linear and nonlinear MC

Literatur

[1] L. Appeltant et al., Nature Communications 2.1 (2011):1-6.
[2] J. Dambre et al., Scientific reports 2.1 (2012): 1-7.
[3] B. Penkovsky et al., Physical review letters 123.5 (2019): 054101.
[4] F. Stelzer et al., Neural Networks 124 (2019): 158-169 .
[5] F. Köster et al., Cognitive Computation (2020): 1–8.
[6] M. Goldmann et al., Chaos 30(9), (2020).