A scenario of learning dynamics by reservoir computing

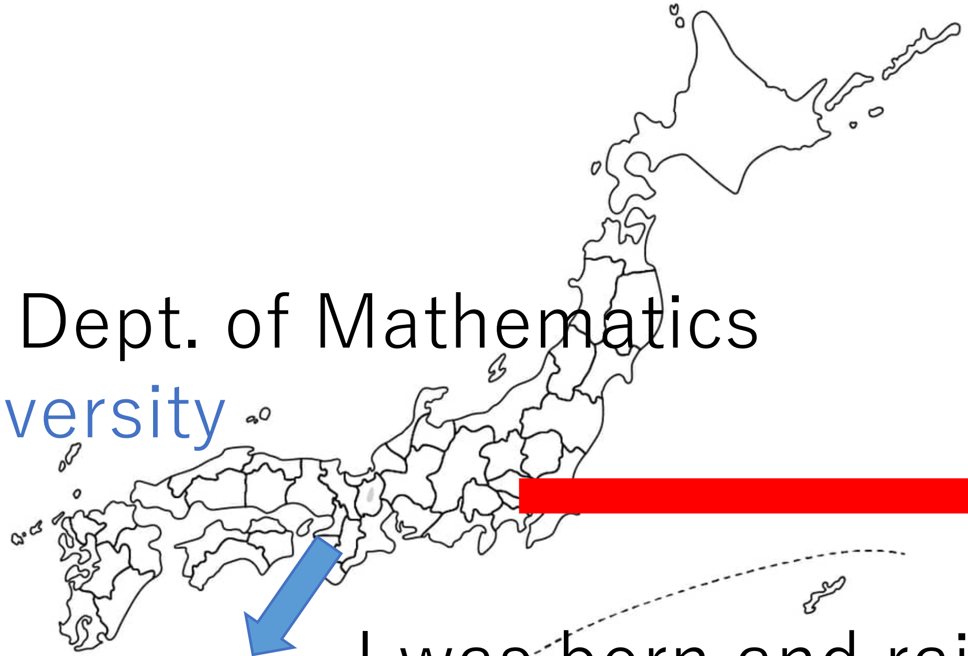
05/28/2024 Boston-Keio-Tsinghua Workshop

Masato Hara (Kyoto U.)

*This is a joint work with Prof. Hiroshi Kokubu (Kyoto U.).

*Related work: A. Hart, *(Thesis) Reservoir Computing With Dynamical Systems*, (2021), *arXiv:2111.14226*

I am from Dept. of Mathematics
Kyoto University



I was born and raised
in Nagoya City

Dept. of Math in Autumn

Nagoya/Boston Museum of
Fine Arts (closing...)

Summary

- "Reservoir computing" can predict time-series.
- This may be because "autonomous reservoir $\mathcal{F}_{W^{\mathrm{out}}}$" become conjugate with the system f behind the time-series.
- In the proof, classical dynamical system theory plays an essential role.

Contents

Contents

What is Reservoir Computing?

- A kind of machine learning (not deep but recurrent neural net)
- Originated with

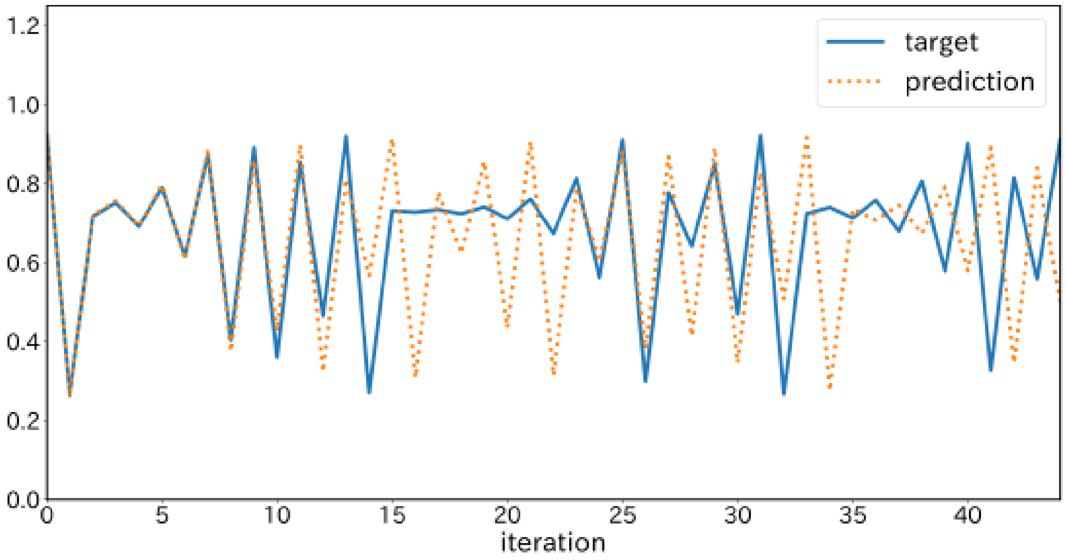
  [Jaeger '01] The "echo state" approach to analysing and training recurrent neural networks

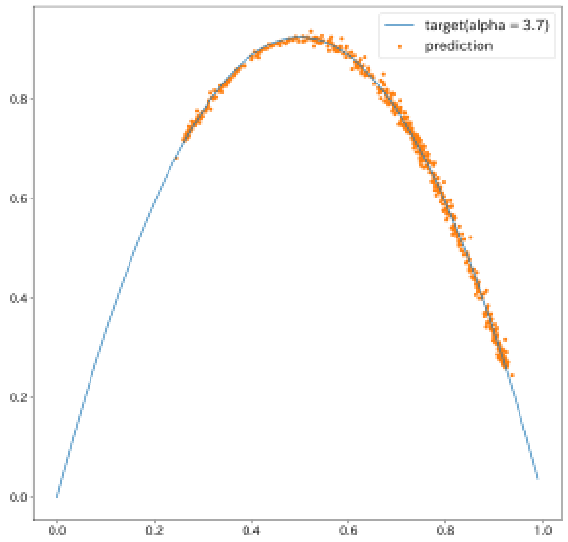
  [Maass et al. '02] Real-time computing without stable states: A new framework for neural computation based on perturbations
- A reservoir itself is a dynamical system.
- RC deal with time series data well.

Time-series prediction

task

Predict time-series generated by a chaotic dynamical system
$x(t+1) = f(x(t))$ (e.g. the logistic map).





Remark

Long-term prediction is impossible because of the chaoticity
→ The dynamics can be reproduced

Contents

1. Introduction
2. Algorithm
3. A scenario of RC

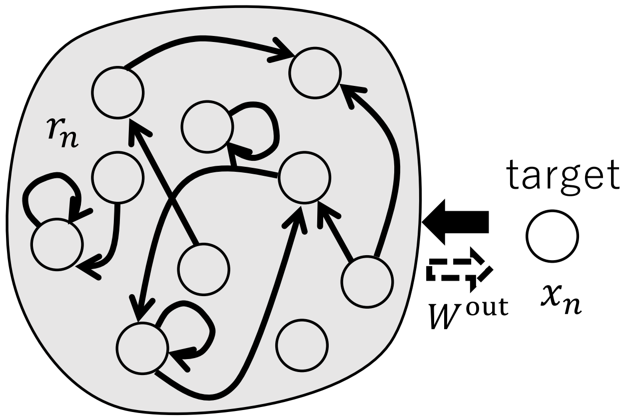
*4. Numerical example（Hyperbolic toral automorphism）

5. Discussion

Algorithm

data $\{x_n\}_{n\geq 0} \subseteq \mathbb{R}^d$
*generated by $f\colon \mathbb{R}^d \circlearrowleft$
reservoir map $F\colon \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^N$

Step1: Listening phase $(n = 0, \ldots, T)$



Generate $\{r_n\}_{n=0}^{T} \subseteq \mathbb{R}^N$ by

$$r_{n+1} = F(x_n, r_n)\ .$$

Step2: Training

Compute W^{out} to minimize the MSE between $\{W^{\mathrm{out}} r_n\}_{n=T_0}^{T}$ and

$\{x_n\}_{n=T_0}^{T}$, i.e., $W^{\mathrm{out}} = \underset{W\in\mathbb{R}^{d\times N}}{\operatorname{argmax}} \sum_{n=T_0}^{T} \|W r_n - x_n\|_2^2$

Algorithm

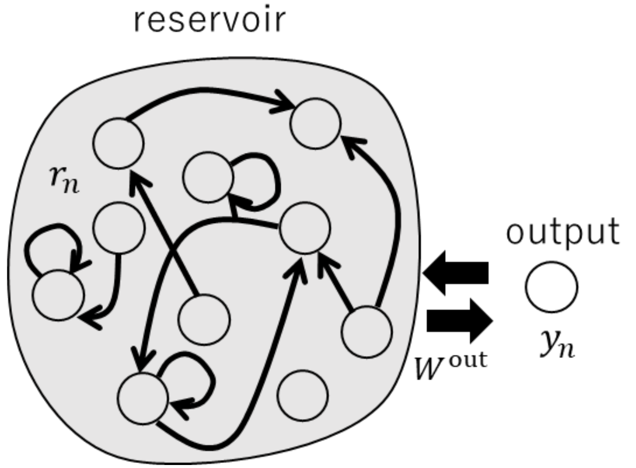
data $\{x_n\}_{n\geq 0} \subseteq \mathbb{R}^d$
*generated by $f: \mathbb{R}^d \to \mathbb{R}^d$
reservoir map $F: \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^N$

Step3: Prediction phase ($n \geq T+1$)

Generate $\{r_n\}_{n\geq T} \subseteq \mathbb{R}^N$ by

$$r_{n+1} = F(W^{\text{out}} r_n, r_n)$$
$$=: \mathcal{F}_{W^{\text{out}}}(r_n)$$

and get $\{y_n := W^{\text{out}} r_n\}_{n\geq T+1}$ as prediction.



$\mathcal{F}_{W^{\text{out}}}: \mathbb{R}^N \circlearrowleft$
: autonomous system
("autonomous reservoir")

Contents

1. Introduction
2. Algorithm
3. A scenario of RC

*4. Numerical example（Hyperbolic toral automorphism）

5. Discussion

What we want to prove?

"Thm." $d \ll N$, $M \subseteq \mathbb{R}^d$: C^1 cpt. mfd.,

$f: M \circlearrowleft$: target (diffeo. & C^1 str. stb.),

$F: M \times \mathbb{R}^N \to \mathbb{R}^N$: reservoir map ($F$: strongly fiber contracting).

We also suppose the [Hypothesis] (explained later).

Let $W^{\mathrm{out}} \in \mathbb{R}^{d \times N}$ be the "output" obtained by successful learning.

Then the autonomous reservoir $\mathcal{F}_{W^{\mathrm{out}}}(r) \coloneqq F(W^{\mathrm{out}} r, r): \mathbb{R}^N \circlearrowleft$

and f are topologically conjugate.

How can we prove?

Notations

- $f \sim_{C^0} g$: f and g are top. conj.
- $f \sim_{C^1} g$: f and g are C^1 conj.

Outline of prf.

- Step1(listening phase)

target → f, ideal reservoir → $\mathcal{F}_\gamma$

$$\exists U \subseteq \mathbb{R}^N,\ \exists \gamma : U \to M \text{ (ideal output), s.t., } f \sim_{C^1} \mathcal{F}_\gamma \quad \cdot\cdot\cdot \ (*)$$

Where $\mathcal{F}_\gamma(r) \coloneqq F(\gamma(r), r)$.

*c.f. $\mathcal{F}_{W^{\mathrm{out}}}(r) \coloneqq F(W^{\mathrm{out}} r, r)$

- Step2(training)

[Hypothesis] $d_{C^1}(\gamma, W^{\mathrm{out}}) \ll 1$

- Step3(prediction phase)

ideal reservoir → $\mathcal{F}_\gamma$, actual reservoir → $\mathcal{F}_{W^{\mathrm{out}}}$

$$\mathcal{F}_\gamma \sim_{C^0} \mathcal{F}_{W^{\mathrm{out}}} \quad \cdot\cdot\cdot \ (**)$$

⟹ By $f \sim_{C^1} \mathcal{F}_\gamma$ (*) & $\mathcal{F}_\gamma \sim_{C^0} \mathcal{F}_{W^{\mathrm{out}}}$ (**), we obtain $\boxed{f \sim_{C^0} \mathcal{F}_{W^{\mathrm{out}}}}$!

Key ideas

■ Step1(listening phase)

$\exists U \subseteq \mathbb{R}^N$, $\exists \gamma: U \to M$ (ideal output), s.t., $f \sim_{C^1} \mathcal{F}_\gamma$ ・・・(*)

✓ $(f, F): M \times \mathbb{R}^N \circlearrowleft$ has ∃! C^1 invariant section $s \coloneqq s_{(f,F)}: M \to Y$.

✓ Moreover, s is an embedding (especially, inj.) .

*To be precise, this is a generic property.

*We can prove it by using Whitney/Takens's technique.

✓ Because s is inj., $\gamma: s(M) \to M$ can be defined.

✓ To prove $f \sim_{C^1} \mathcal{F}_\gamma$ is easy.

Key ideas

- Step2(training)

$$[\text{Hypothesis}]\ d_{C^1}(\gamma, W^{\text{out}}) \ll 1$$

- Step3(predicting phase)

$$\mathcal{F}_\gamma \sim_{C^0} \mathcal{F}_{W^{\text{out}}} \quad \cdot\ \cdot\ \cdot\ (**)$$

✓ By [Hypothesis], $d_{C^1}(\mathcal{F}_\gamma, \mathcal{F}_{W^{\text{out}}}) \ll 1$ is easily obtained.

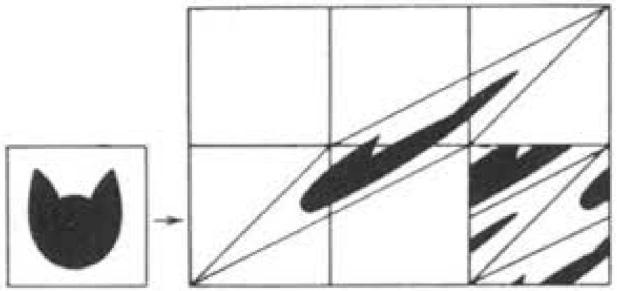
✓ After some arguments, we obtain $\mathcal{F}_\gamma$: C^1str. stb.

*This stability is inherited from f.

✓ $\mathcal{F}_\gamma \sim_{C^0} \mathcal{F}_{W^{\text{out}}}$ follows from the definition of str. stb.

Contents

Numerical example: Hyperbolic Toral Automorphism

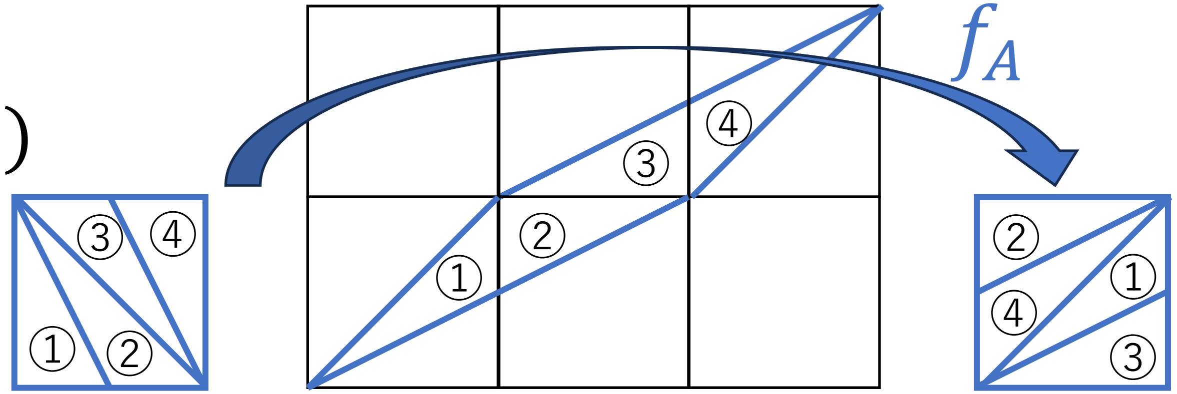


Integer matrix A : $\det(A) = 1$, |eigen value| $\neq 1$

HTA = an auto. on torus induced from $x \mapsto Ax$

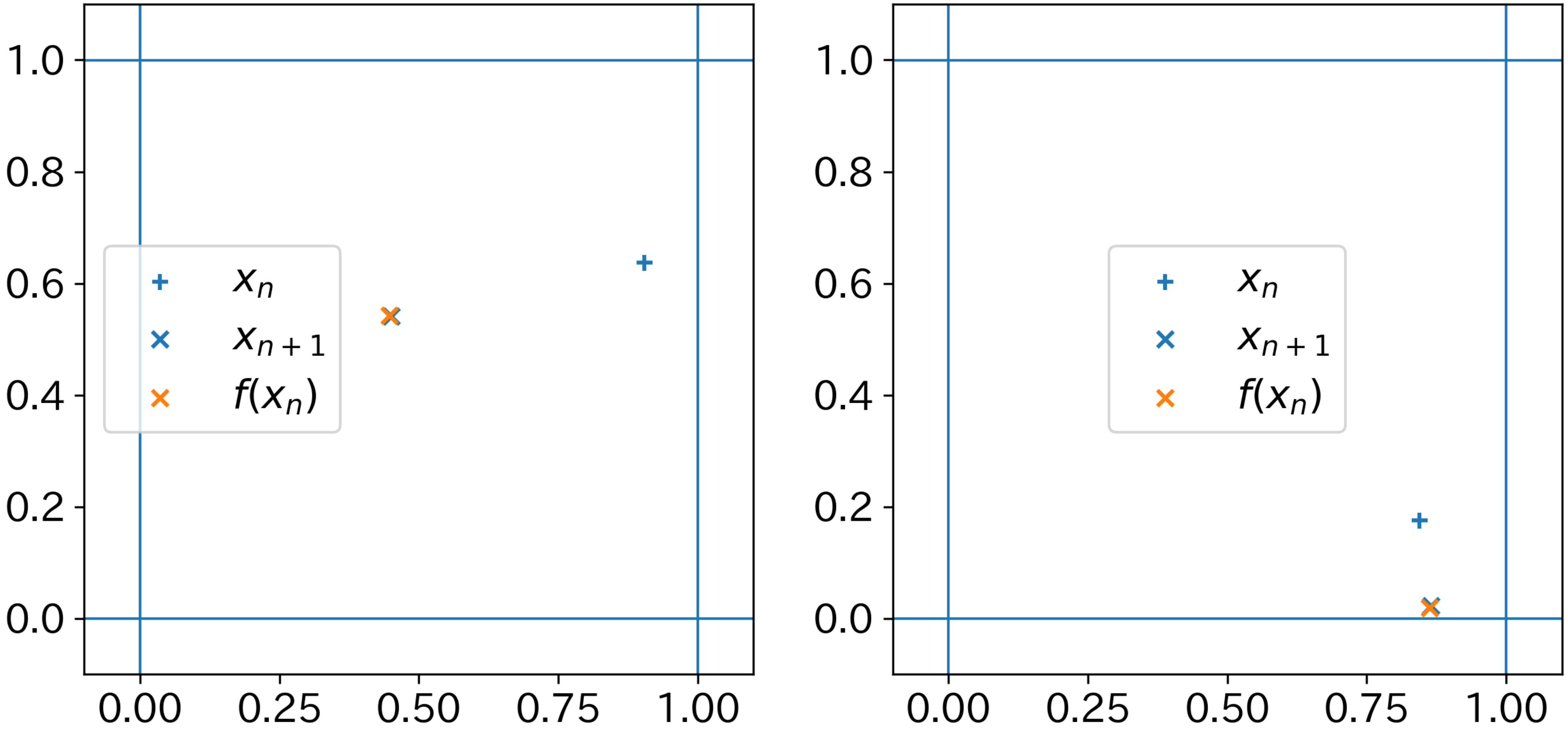
ex) $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $f_A(x) := Ax \pmod 1$

(Arnold's cat map)

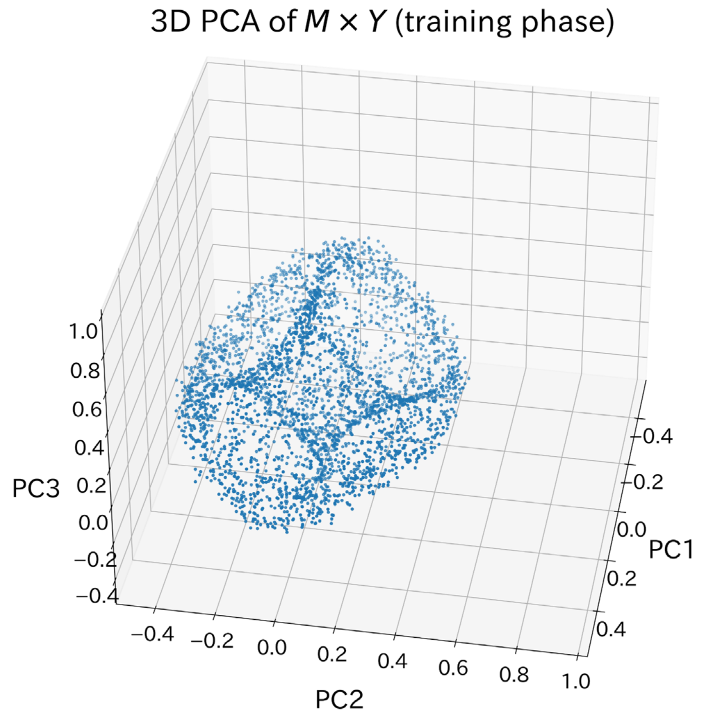


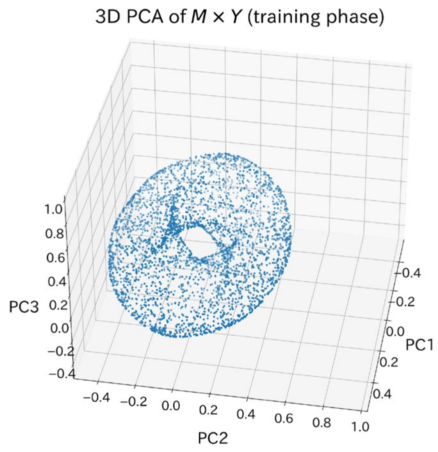
- A very GOOD system! (diffeomorphic, structurally stable, ergodic w.r.t. Lebesgue measure, ...)
- A system on $\mathbb{T}^2$ → some numerical technics are required ($\mathbb{T}^2 \cong \mathbb{R}^2/2\pi\mathbb{Z}^2 \cong S^1 \times S^1 \hookrightarrow \mathbb{R}^4$)

Time series prediction (HTA)



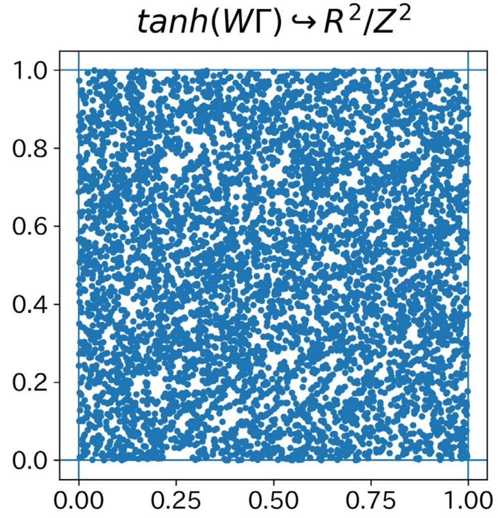
Listening phase





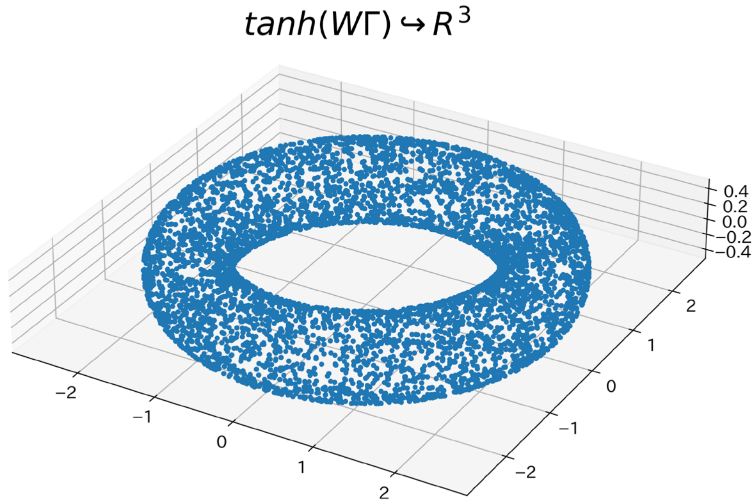
1st principle：24.214 %
2nd principle：22.307 %
3rd principle：19.827 %
4th principle：17.756 %
5th principle：3.306 %

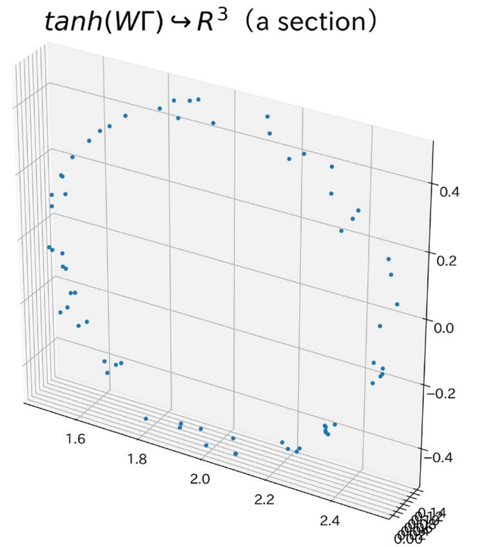
Predicting phase



	Lyapunov spectrum of reservoir
1	0.962338
2	-0.96136
3	-3.10875
4	-3.51422
5	-4.33653
6	-4.78709

$\approx \pm 0.962423 \ldots$





Contents

Summary

- RC can predict time-series.
- This may be because autonomous reservoir $\mathcal{F}_{W^{\mathbf{out}}}$ become conjugate with the system f behind the time-series.
- In the proof, classical dynamical system theory plays an essential role.

Discussion

- [Hypothesis] $d_{C^1}(\gamma, W^{\mathrm{out}}) \ll 1$ may not holds in general.
  - ✓ I believe $d_{C^0}(\gamma, W^{\mathrm{out}}) \ll 1$ is more reasonable and similar theorem holds even in this setting.

    *In this case, the conclusion may become not "top. conj." but "semi-conj.".

    This may be proved by using pseudo orbit tracing arguments.
- The condition "f: diffeo., C^1str. stb." is too strong to explain typical numerical examples (e.g., logistic map).
  - ✓ I want a theory including the logistic map case. (feature work)