JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

No Equations, No Variables, No Parameters No Space and No Time Data and the Modeling of Complex Systems

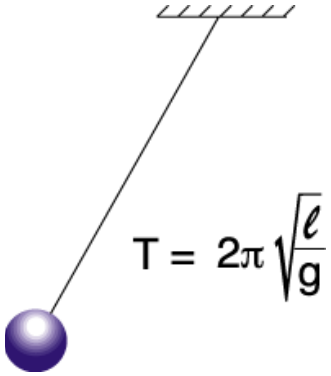
Yannis Kevrekidis and many good people:

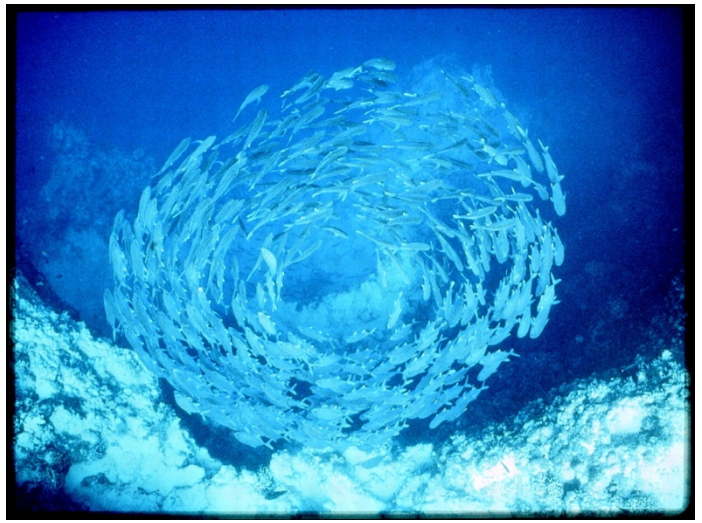
Felix Kemeth, Dave Scroczynski, Felix Dietrich, Tom Bertalan……

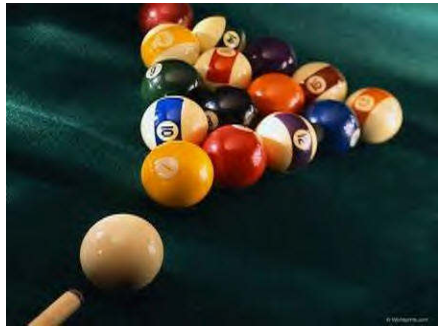
C. William Gear, C. Siettos, Raphy Coifman…

Google, April 28, 202

J. Ottino, NWU







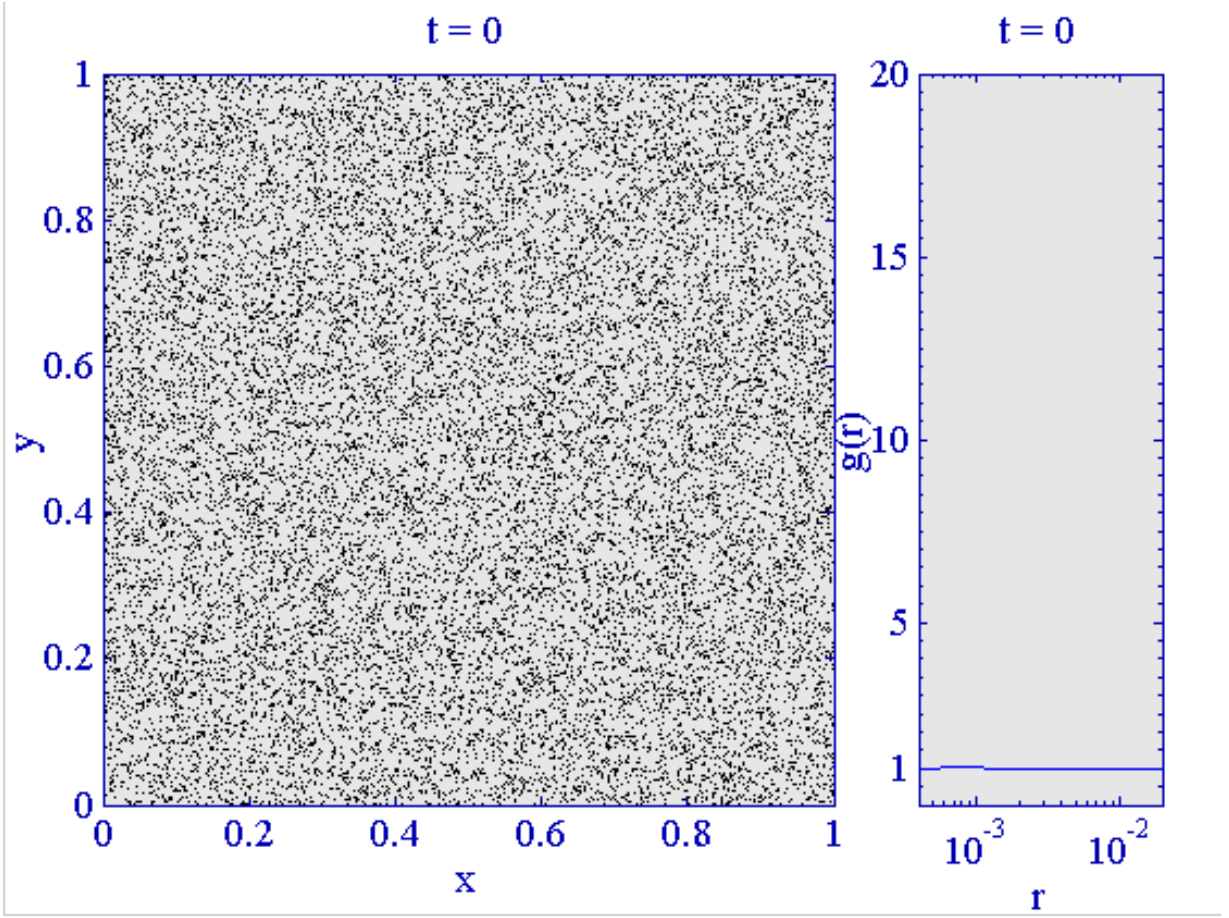
“simple” “complicated” “complex”

Clustering and stirring in a plankton model

An “equation-free” demo

Young, Roberts and Stuhne, *Nature* 2001

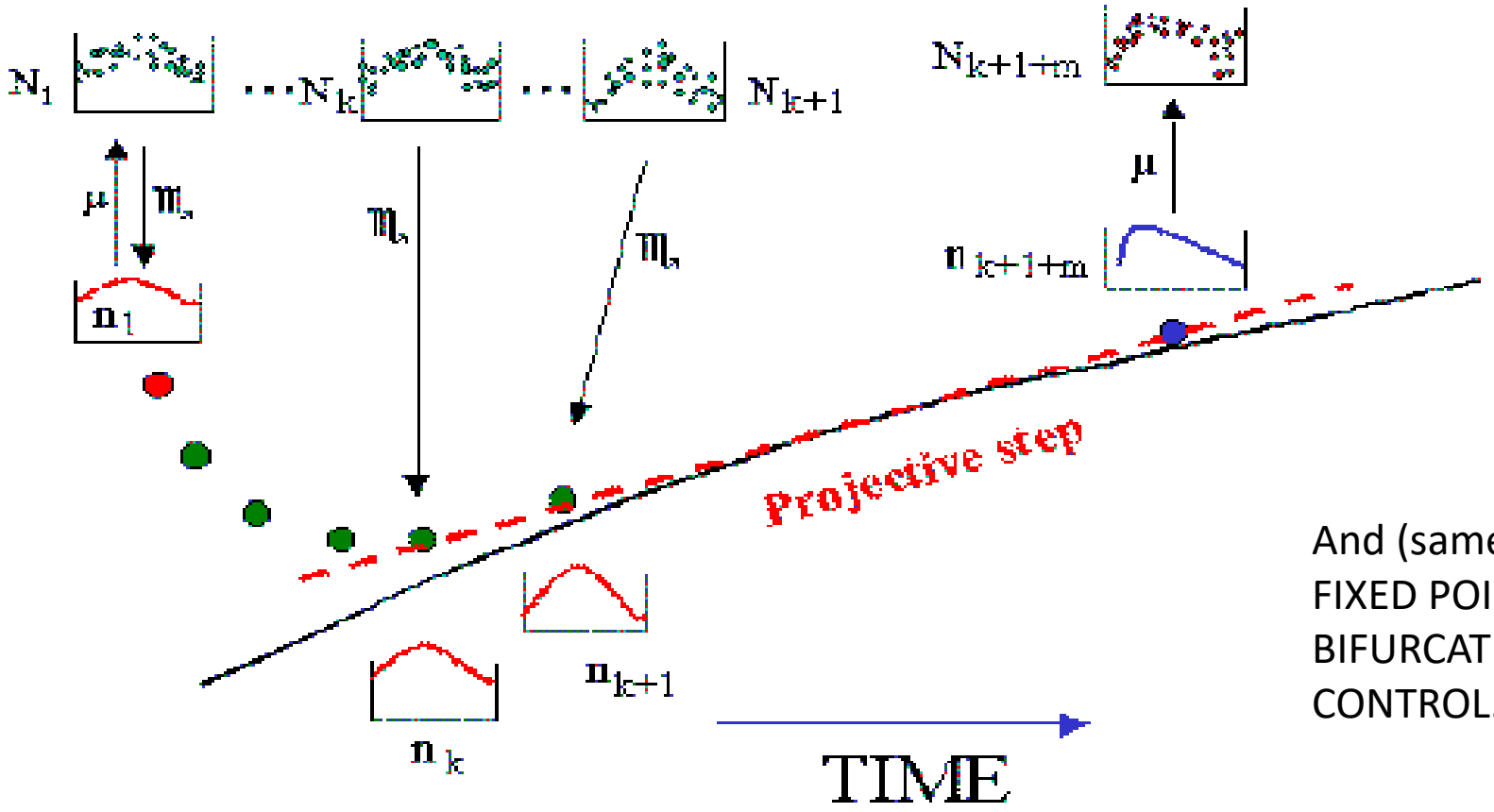
Dynamics of System with convection



Equation-Free (2000) *Coarse Projective Integration*

solve the equation without writing it down

through judicious bursts of particle level simulation



And (same idea as Herb's RPM)
FIXED POINTS,
BIFURCATIONS,
CONTROL….. the world….

Variable-Free (2007)

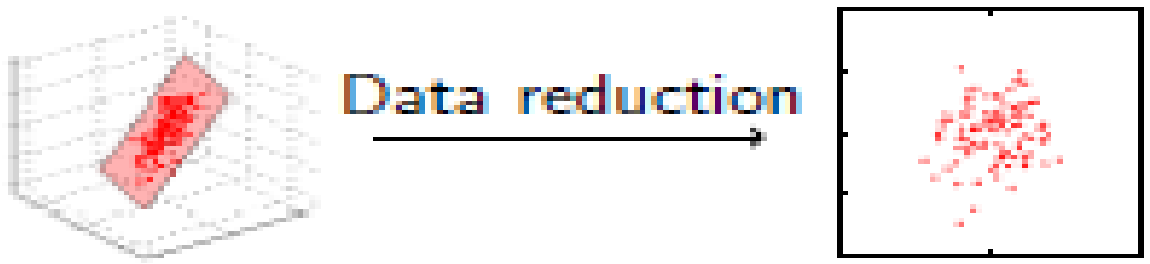
Where do good coarse variables (good descriptors) come from?

- Systematic hierarchy (and mathematics)
  - Fourier modes, moments…..
- Experience/expertise/knowledge/brilliance
  - Human learning (“brain” data mining, observation, phase fields…)
- Machine Learning
  - Data mining, manifold learning (here: mainly diffusion maps)

Our Approach: Data Reduction Techniques

Common *linear* technique: Principal Component Analysis (PCA)

Project onto hyperplane which captures *maximum variance*

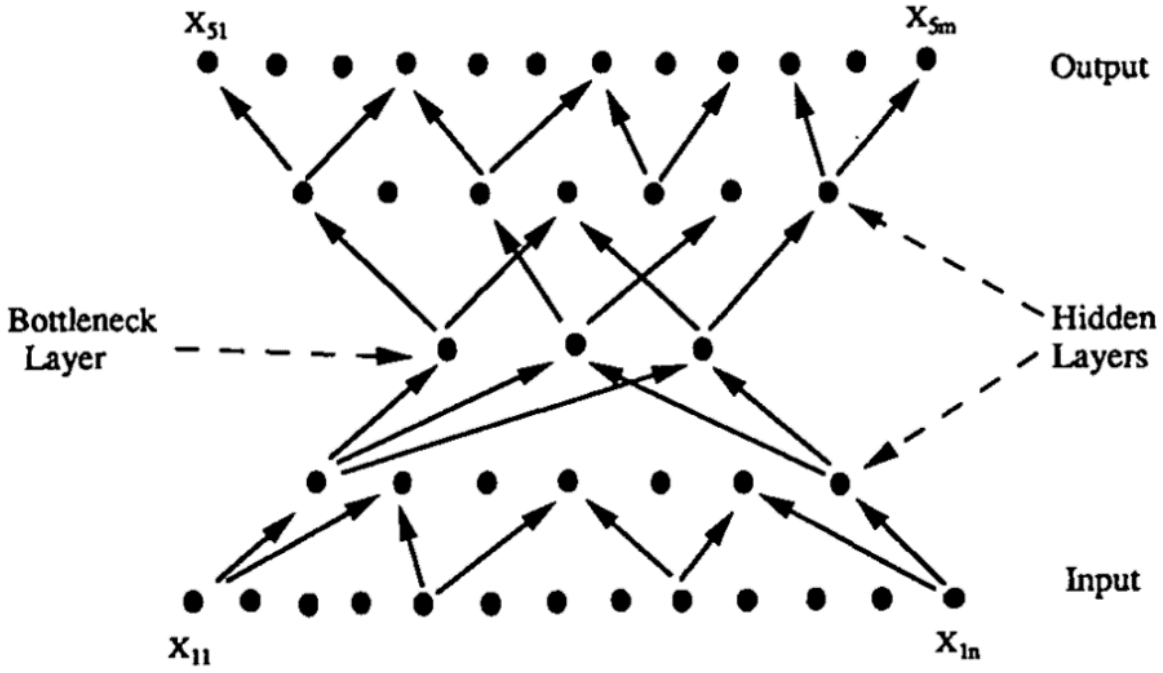


High dimensional data on low-dimensional structure

Reduced dimensionality

“Nonlinear Principal Components” by Autoencoder

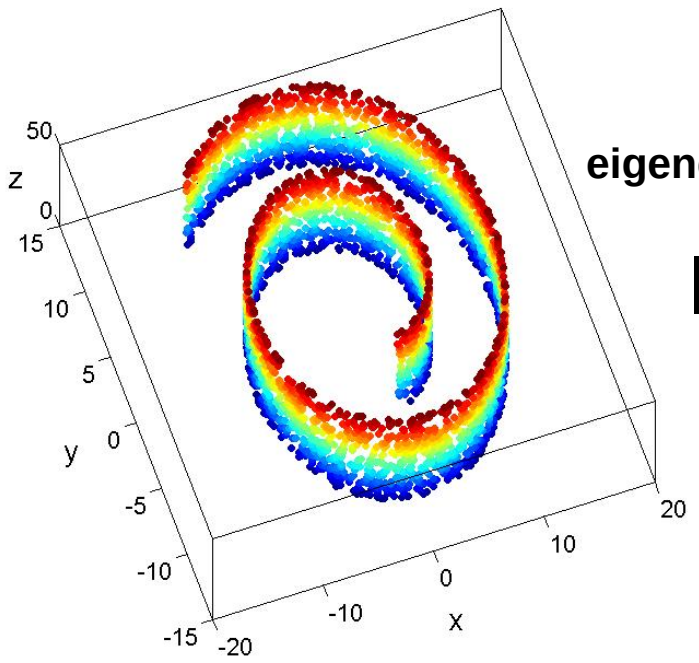
- Train network to reconstruct state vectors $\boldsymbol{x}$ (or delay vectors) (output ≈ input).
- Use a bottleneck layer of an arbitrarily selected small number of neurons.
- Output of bottleneck layer gives NLPC features.
- Input section can be used to restrict data to NLPCs; output section can be used to lift data back to full $\boldsymbol{x}$ space.
- Do ODE network training with time series data viewed NLPC space.



Rico-Martínez, Krischer, & Kevrekidis. (1992). Discrete- vs continuous-time nonlinear signal processing of Cu Electrodissolution Data. Chem. Eng. Comm.

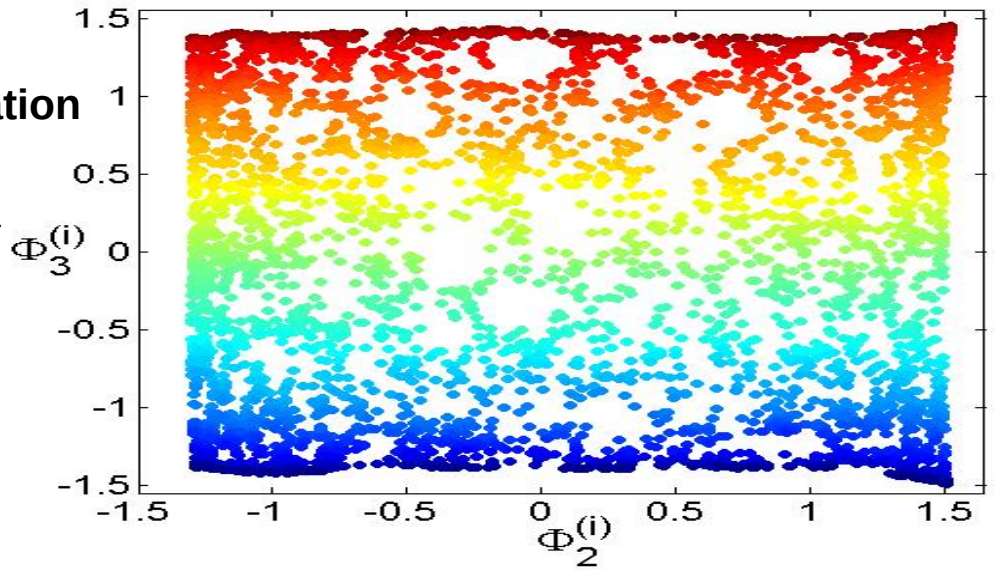
Diffusion Maps

Dataset in x, y, z



eigencomputation

Dataset Diffusion Map



N datapoints

$$\boldsymbol{x}^{(i)} = \left(x_i, y_i, z_i \right), \; i = 1, N$$

N datapoints

$$\boldsymbol{\Phi}^{(i)} = \left(\Phi_2^{(i)}, \Phi_3^{(i)} \right), \; i = 1, N$$

R. Coifman, S. Lafon, A. Lee, M. Maggioni, B. Nadler, F. Warner, and S. Zucker, Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps.
***PNAS* 102 (2005).**

B. Nadler, S. Lafon, R. Coifman, and I. G. Kevrekidis, Diffusion maps, spectral clustering and reaction coordinates of dynamical systems.
***Appl. Comput. Harmon. Anal.* 21 (2006).**

Ancient History (1992)
Runge-Kutta Resnets

Embed network in an explicit time-stepping scheme, such as Runge-Kutta. Only one set of weights and biases is used; shared across the Runge-Kutta stages.

Train to reconstruct next state $\boldsymbol{y}_{n+1}$ given both current state $\boldsymbol{Y}_n$ and parameter vector $\boldsymbol{X}$.

Chem. Eng. Comm. 1992, Vol. 118, pp. 25–48
Reprints available directly from the publisher.
Photocopying permitted by license only.

Printed in the United States of America

DISCRETE- vs. CONTINUOUS-TIME NONLINEAR SIGNAL PROCESSING OF Cu ELECTRODISSOLUTION DATA

R. RICO-MARTÍNEZ, K. KRISCHER and I.G. KEVREKIDIS

Department of Chemical Engineering, Princeton University, Princeton, NJ 08544-5263.

and

M.C. KUBE and J.L. HUDSON

Department of Chemical Engineering, University of Virginia, Charlottesville, VA 22093-2442.

(Submitted August 20, 1991; in revised form January 22, 1992)

Artificial neural networks (ANNs) are often used for short term discrete time predictions of experimental data. In this paper we focus on the capability of such nets to correctly identify long term behavior and, in particular, observed bifurcations. As we show, the usual discrete time mapping approach is (precisely because of its discrete nature) often incapable of reproducing observed bifurcation sequences. If the interest is only in periodic or temporally more complicated behavior, a Poincaré map extracted from the experimental time series can be used to circumvent this problem. A complete dynamic picture including bifurcations of steady states can, however, only be captured by a continuous-time model. We present an ANN configuration which couples a "nonlinear principal component" network for data processing (Kramer, 1991, Usui *et al.*, 1990) with a composite ANN based on a simple integrator scheme. This ANN is able to correctly reconstruct the bifurcation diagram of our experimental data. All time series we process stem from the potentiostatic electrodissolution of Cu in phosphoric acid solution. As the applied potential is varied, the electrodissolution rate changes from steady behavior to periodic oscillations, followed by a sequence of period doublings to apparently chaotic motion, and then returns to simple oscillations via a reverse cascade of period doublings.

KEYWORDS Neural networks Time-series Electrodissolution Bifurcation.

Rico-Martínez, Krischer, & Kevrekidis. (1992). *Discrete- vs continuous-time nonlinear signal processing of Cu Electrodissolution Data*.

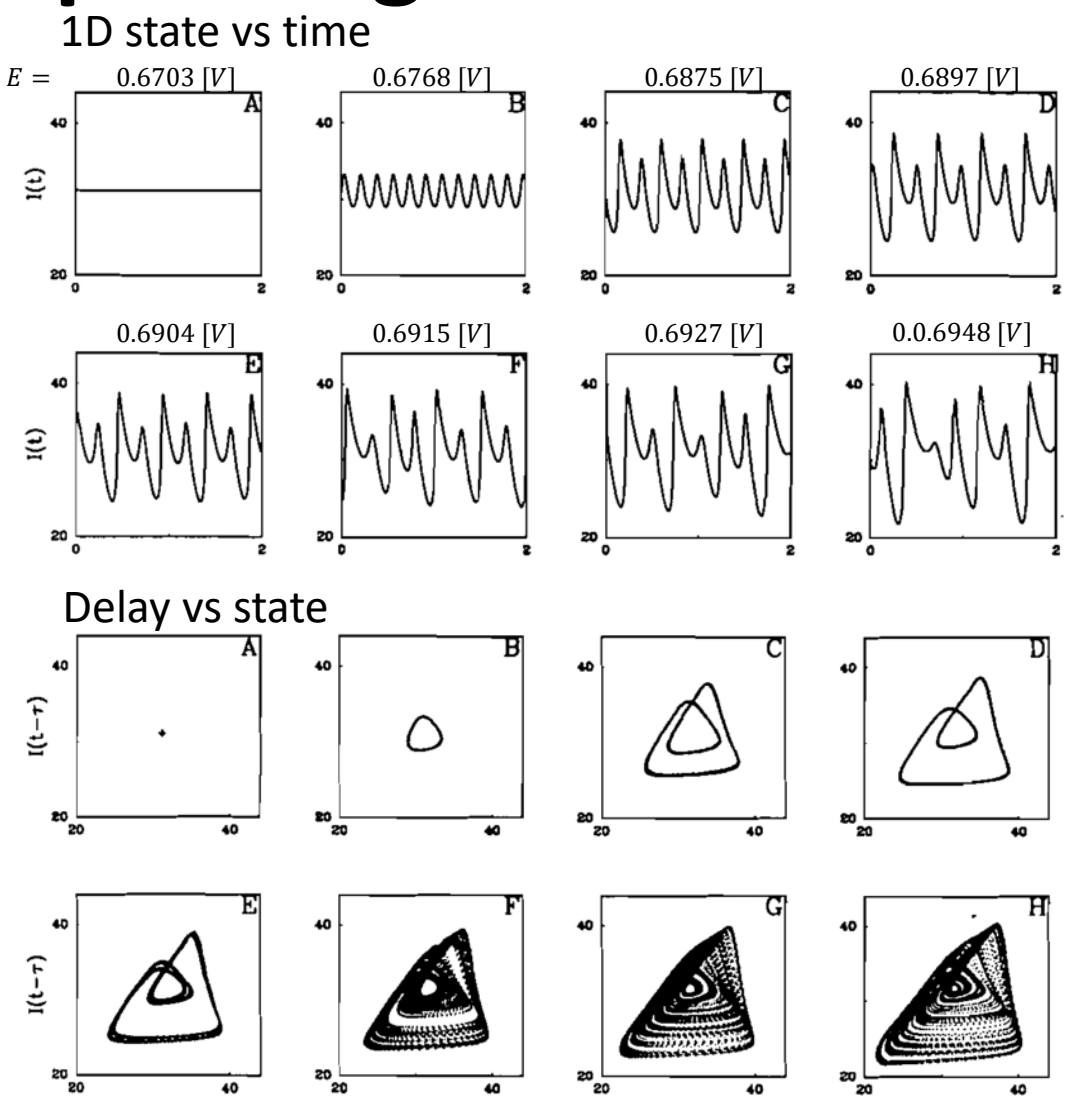
Neural Networks for Identification of Dynamical Systems Depending on Parameters

Discrete vs. Continuous Networks

Training Data

State histories $\boldsymbol{x}_t$, paired with parameter settings $\boldsymbol{p}$.

Trajectories are expected to lie on or near the system attractor, or at least in regions of interest in state space.



Rico-Martínez, Krischer, & Kevrekidis. (1992). Discrete- vs continuous-time nonlinear signal processing of Cu Electrodissolution Data. Chem. Eng. Comm.

Explicit Runge-Kutta / Runge-Kutta ResNet

We can realize such a scheme with a NN with parameter sharing for f.

RK4:

Let $\frac{dy}{dt} = f(y)$.

Step from y_0 to y_τ.

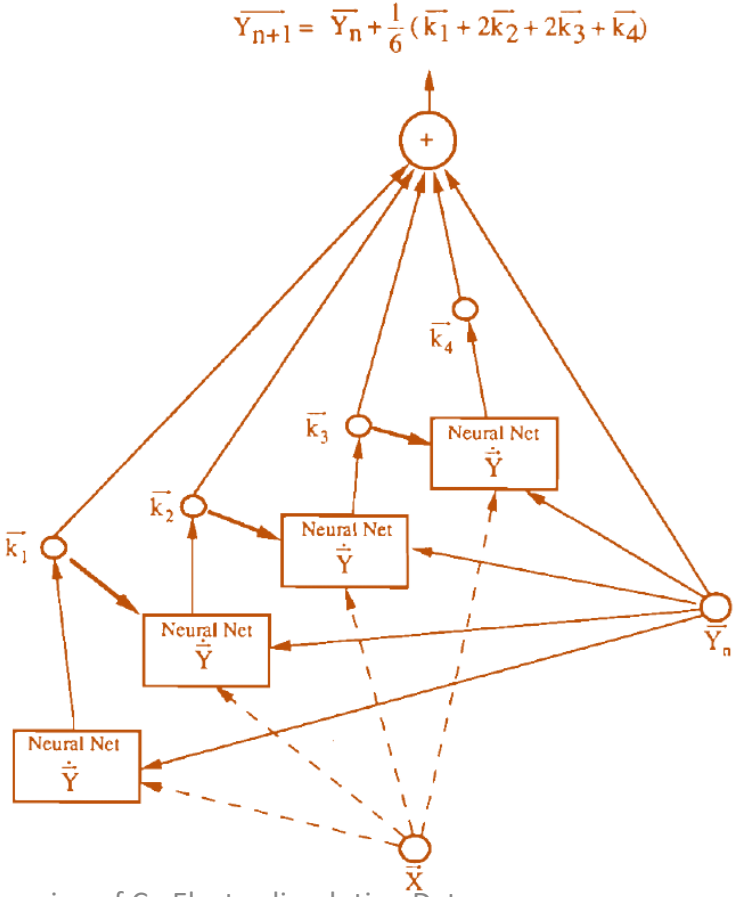
$$k_1 = f(y_t)$$

$$k_2 = f\left(y_0 + \frac{h}{2} k_1\right)$$

$$k_3 = f\left(y_0 + \frac{h}{2} k_2\right)$$

$$k_4 = f(y_0 + h\, k_3)$$

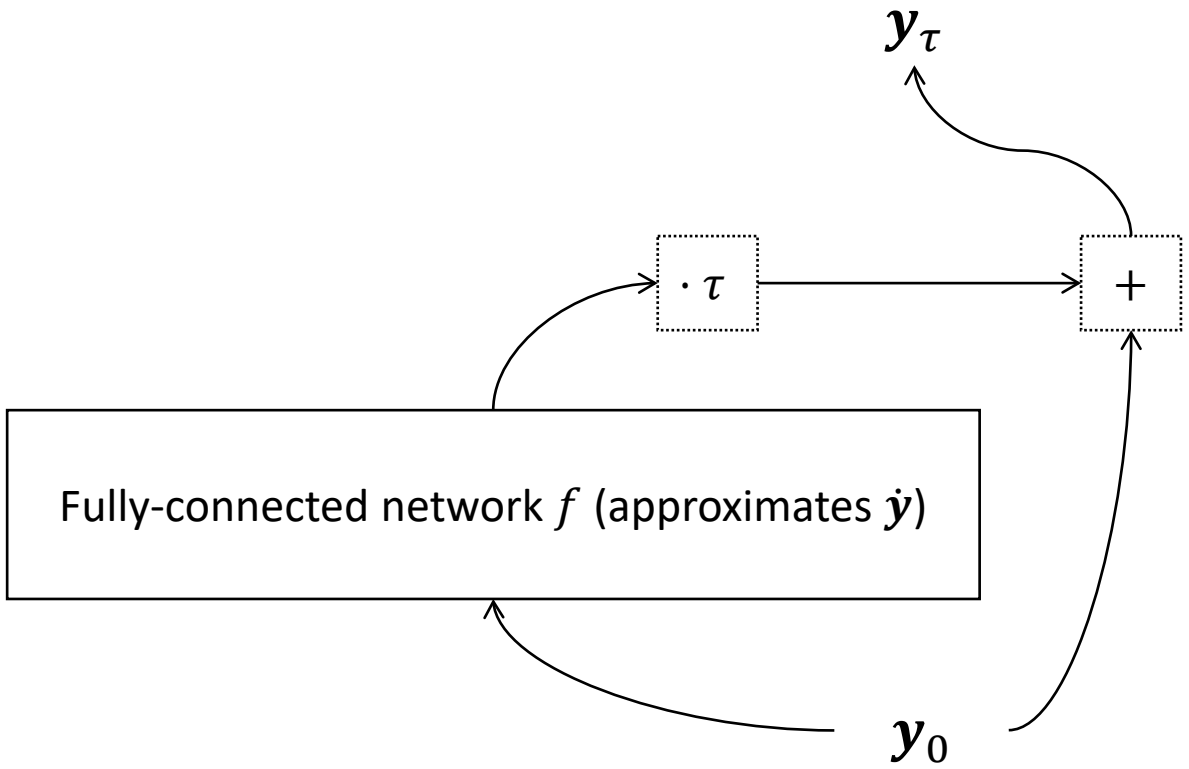
$$y_\tau = \left(\frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6}\right)$$

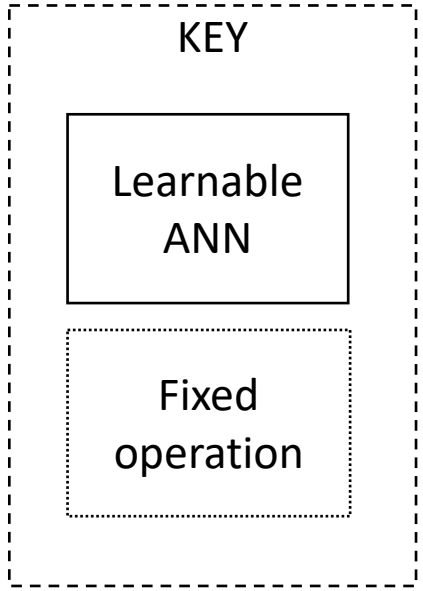


Rico-Martínez, Krischer, Kevrekidis, Kube, & Hudson. (1992). Discrete- vs continuous-time nonlinear signal processing of Cu Electrodissolution Data. Chemical Engineering Communications.

Using an Euler Template

Note: This is analogous to the structure of a single ResNet module. See e.g. K. He et al., 2015 (CVPR 2016)

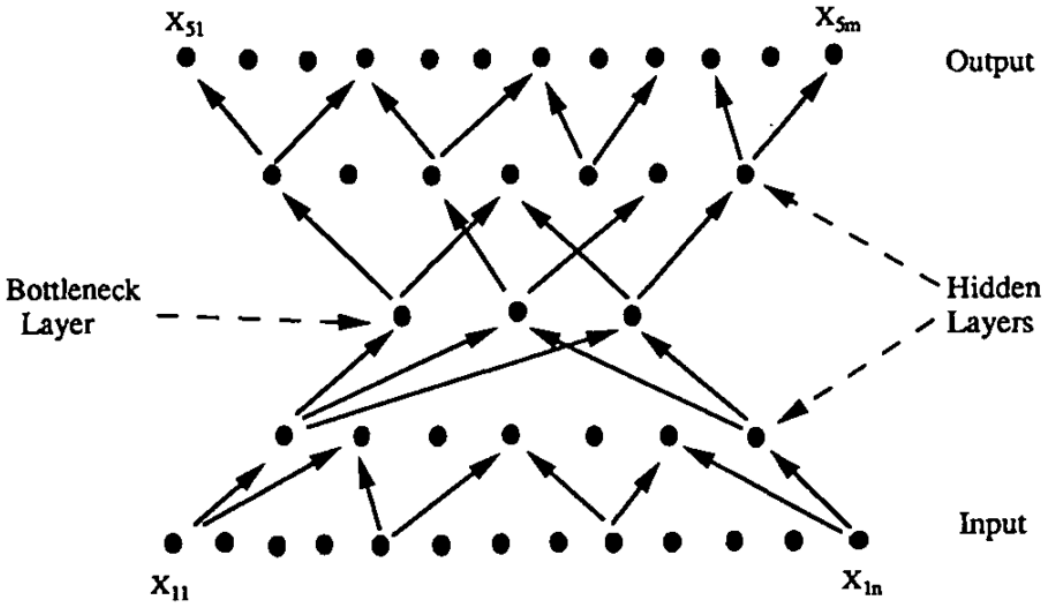




Loss function on supervised output for y_τ allows us to learn ODE $\dot{y}$.

“Nonlinear Principal Components” by autoencoder

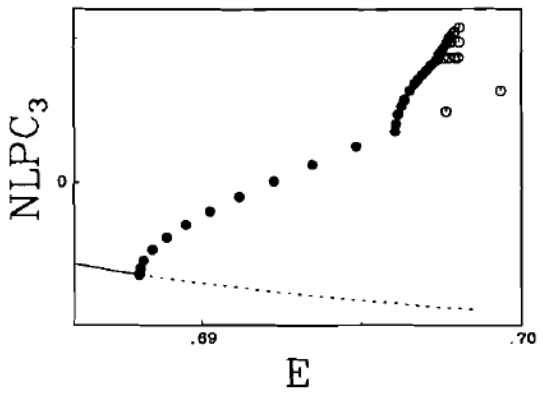
- Train network to reconstruct state vectors $\boldsymbol{x}$ (or delay vectors) (output ≈ input).
- Use a bottleneck layer of an arbitrarily selected small number of neurons.
- Output of bottleneck layer gives NLPC features.
- Input section can be used to restrict data to NLPCs; output section can be used to lift data back to full $\boldsymbol{x}$ space.
- Do ODE network training with time series data viewed NLPC space.

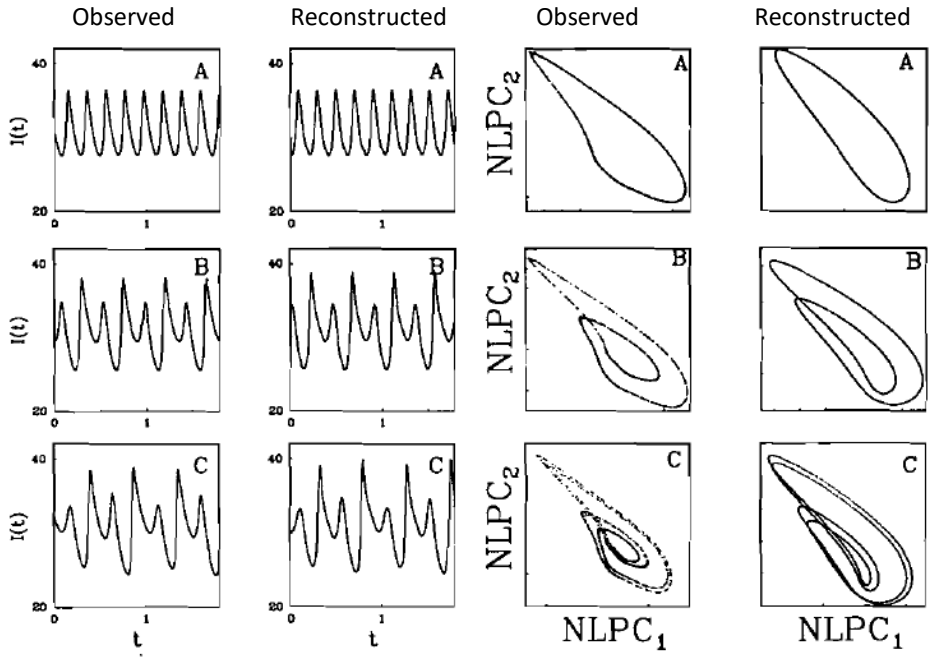


Rico-Martínez, Krischer, & Kevrekidis. (1992). Discrete- vs continuous-time nonlinear signal processing of Cu Electrodissolution Data. Chem. Eng. Comm.

ODE net reconstructs attractors and transitions

- Bifurcation diagram can also be constructed.
- Attractor is observed in **“nonlinear principal components” (NLPC)**—which are constructed via another autoencoder network…

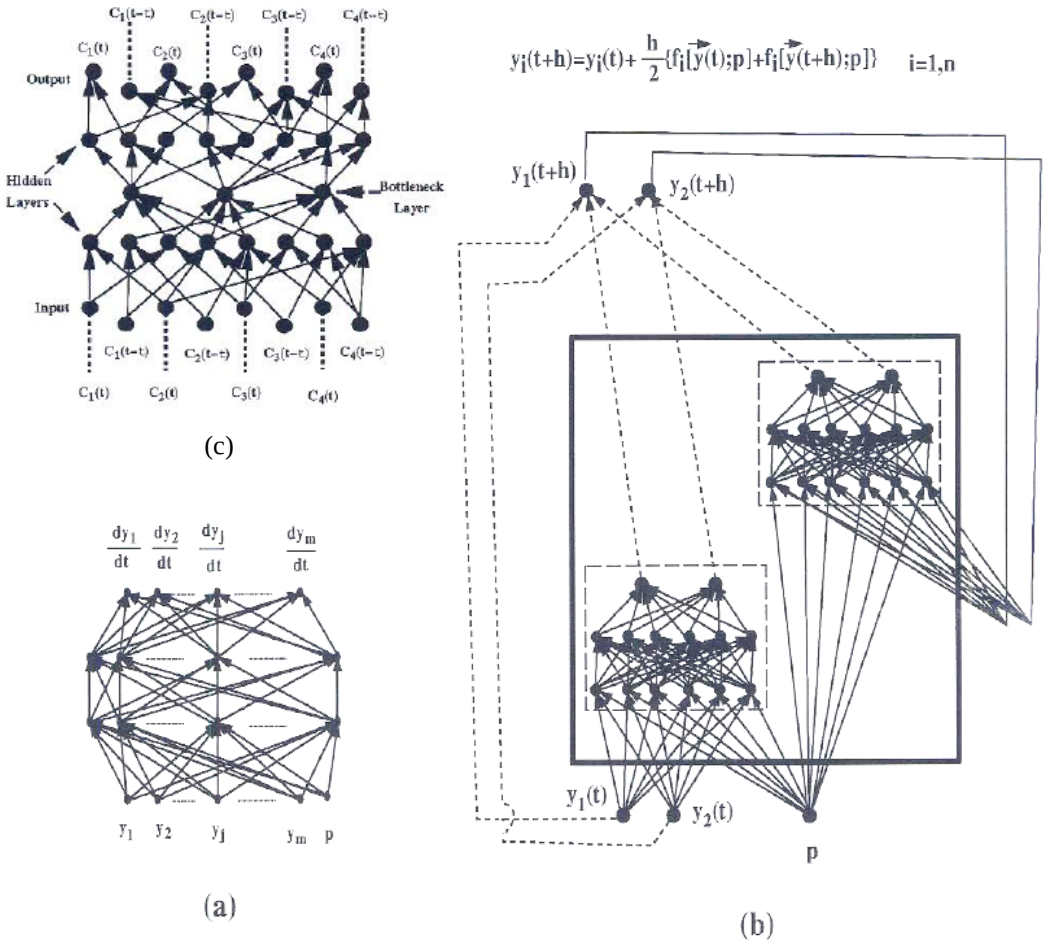




Rico-Martínez, Krischer, & Kevrekidis. (1992). Discrete- vs continuous-time nonlinear signal processing of Cu Electrodissolution Data. Chem. Eng. Comm.

Implicit ODE training

- CO oxidation on Pt, as before, or oscillatory electrode potential during electrochemical oxidation of H_2 on a Pt wire in $HClO_4$ solution.
- For CO oxidation, four KL modes and their four time-τ delayed values are further reduced by a 3-neuron bottleneck layer.
- Train a network to produce $\frac{d\boldsymbol{y}}{dy}$ given $\boldsymbol{y}$ and parameters $\boldsymbol{p}$.
- For training embed the network in an integrator.
- **Implicit integrators require recurrent connections**, whereas explicit integrators are feed-forward, as before.



Network (a) embedded (b) in an implicit trapezoidal-rule integrator.
Reduction of 8 KL+delay coordinates to 3 NLPC coordinates (c).

Rico-Martínez, Kevrekidis, & Krischer (1995). Nonlinear system identification using neural networks: dynamics and instabilities. In A. B. Bulsari (Ed.), Neural Networks for Chemical Eng.

Gray-box identification of a surface-catalyzed reaction ODE

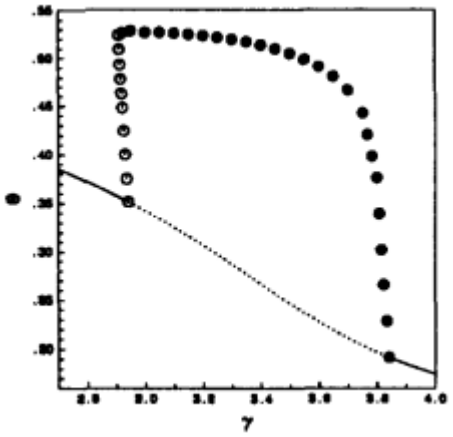
Rather than learning a whole model from scratch, we replace only one term with a function to be learned.

$$\frac{\mathrm{d}\theta}{\mathrm{d}\tau} = K_a\Pi(1-\theta) - f(\theta,\gamma) - K_R\theta e^{-\frac{1}{\tau}}$$

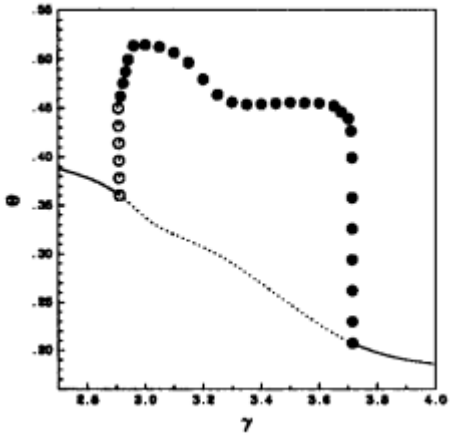
$$\frac{\mathrm{d}\Pi}{\mathrm{d}\tau} = 1 - \Pi + \Pi^*[K_d e^{-\frac{\alpha^*\theta+\beta}{\gamma}} - K_a\Pi(1-\theta)]$$

θ is surface coverage, Π is gas-phase partial pressure, and $f(\theta,\gamma)$ is a learned neural surrogate for the desorption rate term $K_d\theta e^{-(\alpha^*\theta+\beta)/\gamma}$.

Bifurcation Diagram



True

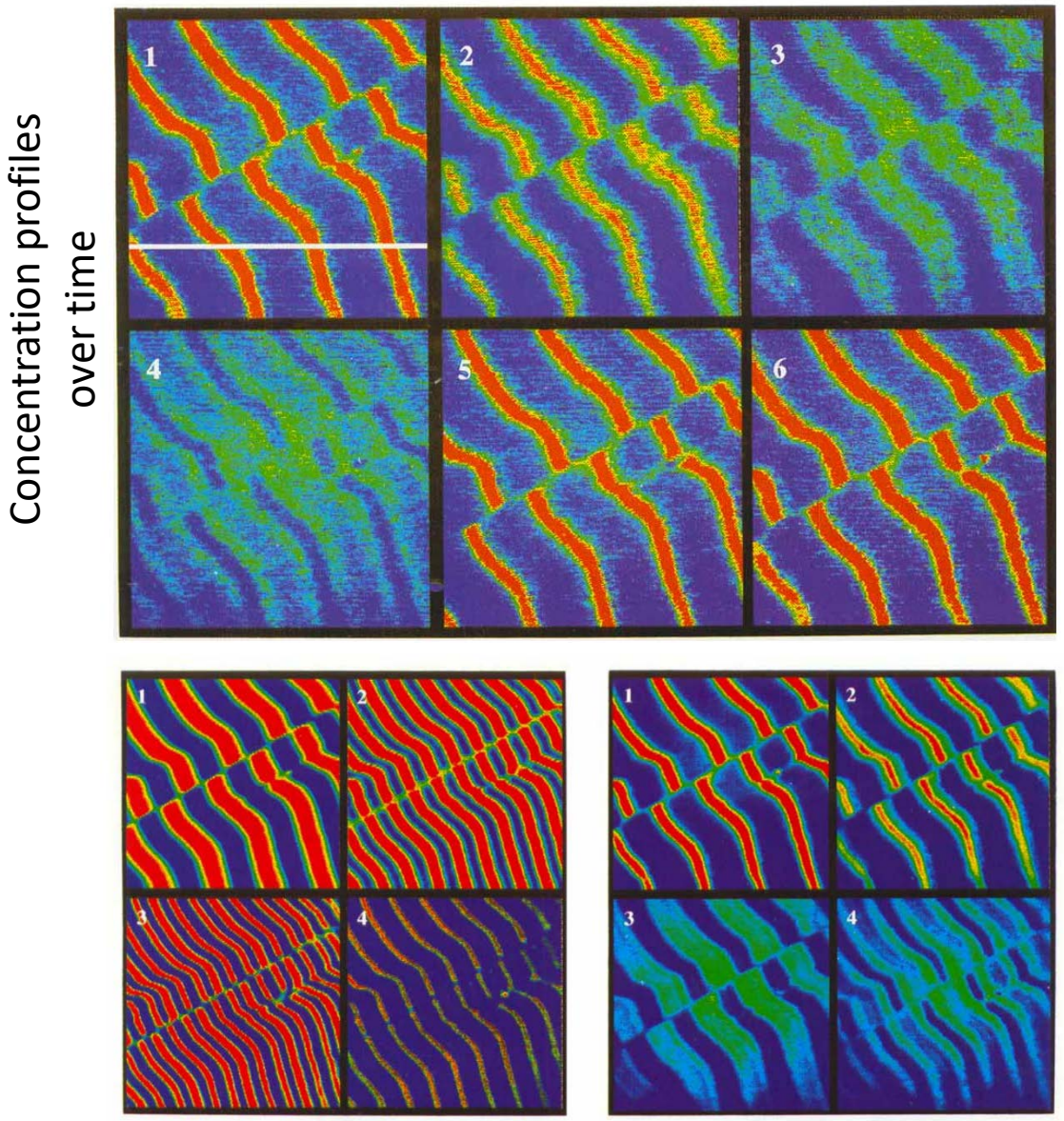


Predicted

Rico-Martínez, Anderson, & Kevrekidis. (1994). *Continuous-time nonlinear signal processing: a neural network based approach for gray box identification.*

Dimension reduction allows ANN prediction of time series

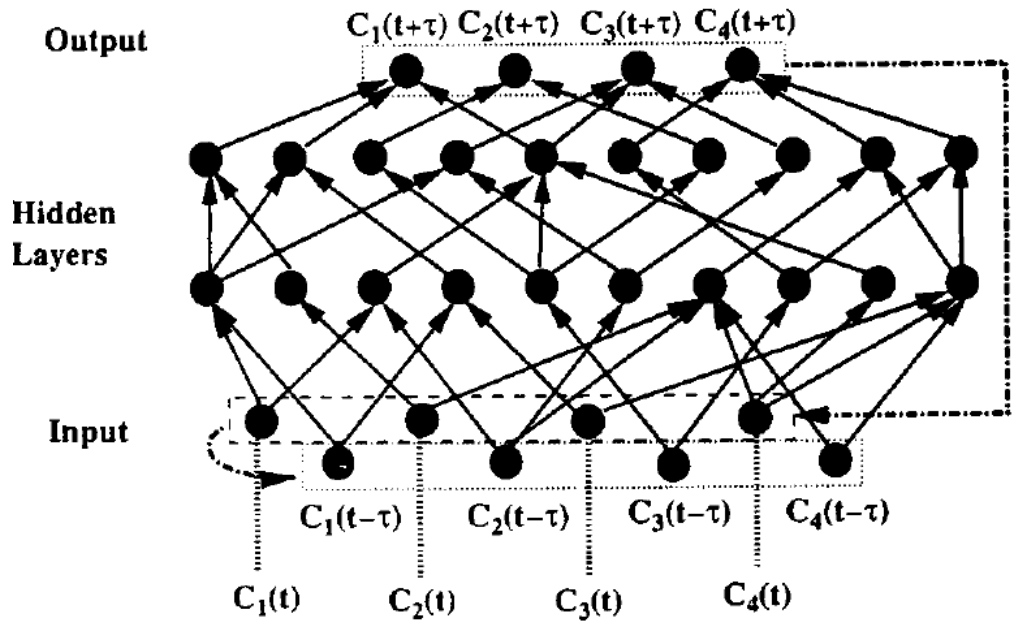
- Source time series are 2D spatial concentration profiles for CO oxidation on a $300 \times 300\ \mu m^2$ platinum catalyst, imaged by photoemission electron microscope.
- Data is first projected onto **Karhunen-Loève (KL) modes**, a.k.a. Proper Orthogonal Decomposition (POD) or Empirical Orthogonal Eigenfunctions (EOF).
- In principle any dimension reduction could be used—such as a bottleneck network.



Krischer, Rico-Martínez, Kevrekidis, Rotermund, & Hudson (1993). Model identification of a spatiotemporally varying catalytic reaction. AIChE J.

Dimension reduction allows ANN prediction of time series

ANN is trained to predict first four KL modes from their current and time-τ delayed values.

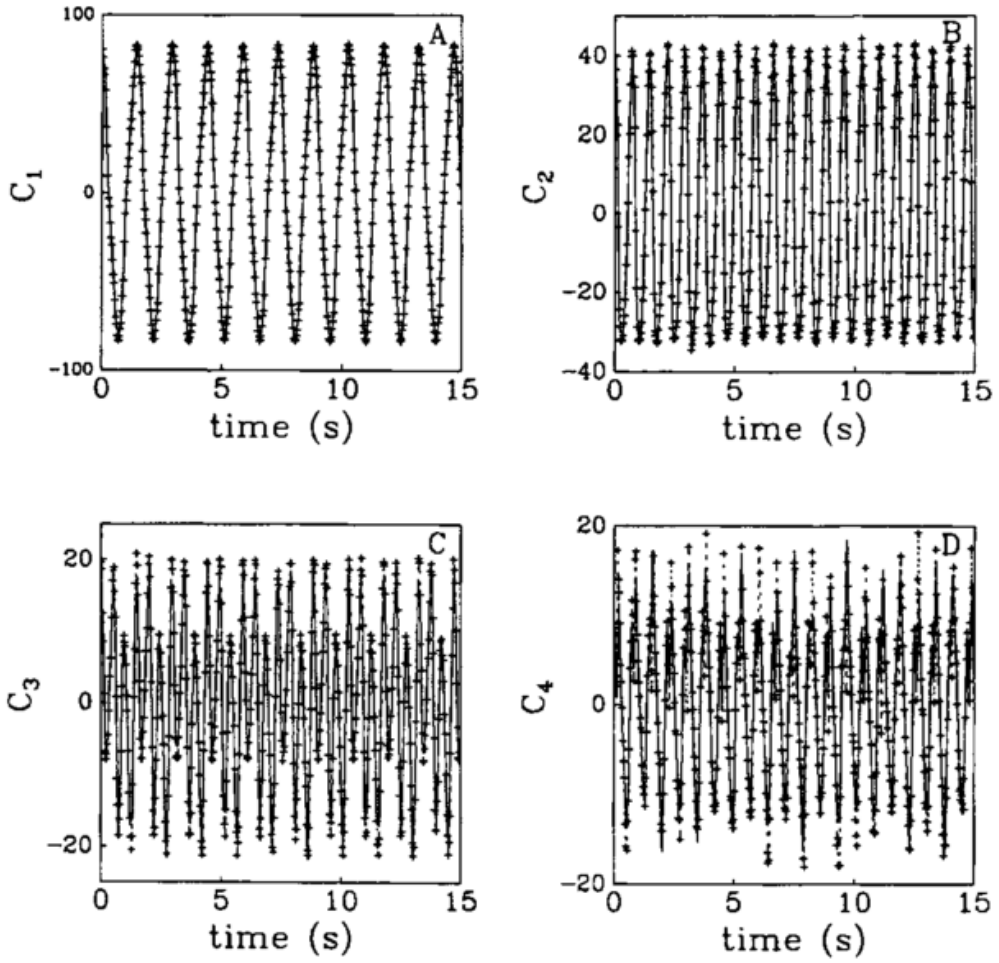


Krischer, Rico-Martínez, Kevrekidis, Rotermund, & Hudson (1993). Model identification of a spatiotemporally varying catalytic reaction. AIChE J.

Dimension reduction allows ANN prediction of time series

- Dimension reduced with autoencoder.
- Flow-map fit to dynamics.

Experiment (points) vs reconstruction (curves)

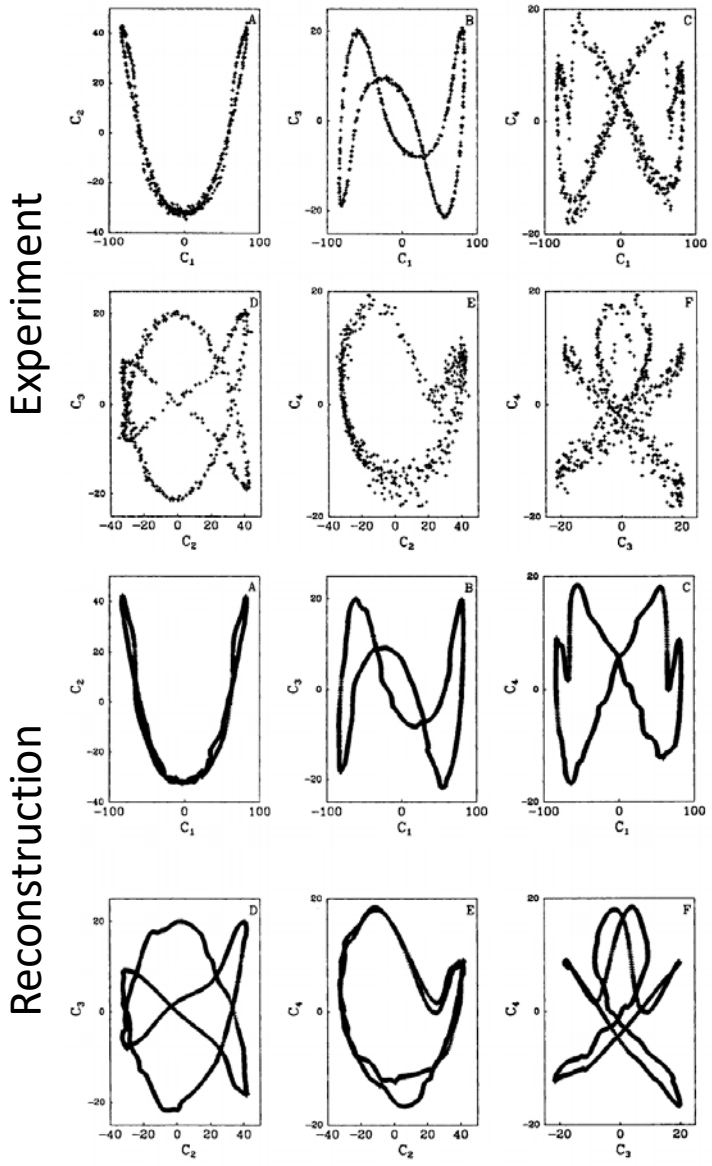


Krischer, Rico-Martínez, Kevrekidis, Rotermund, & Hudson (1993). Model identification of a spatiotemporally varying catalytic reaction. AIChE J.

Dimension reduction allows ANN prediction of time series

Learned flow map reproduces attractor, but with wrinkles characteristic of learning a finite-time map

Attractor is plotted in various KL-pair projections.

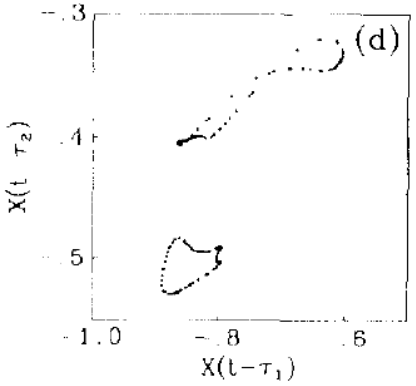


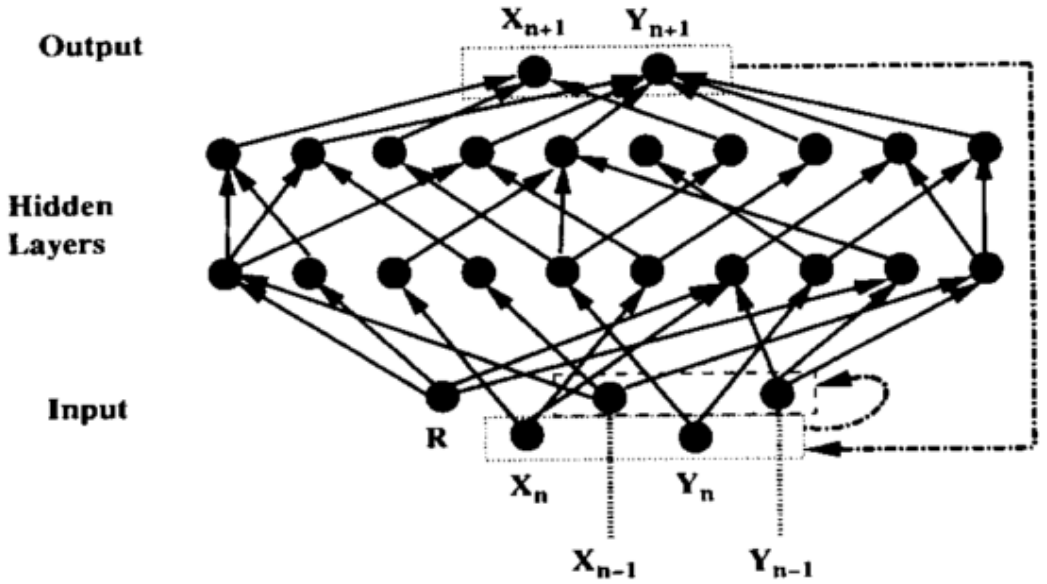
Krischer, Rico-Martínez, Kevrekidis, Rotermund, & Hudson (1993). Model identification of a spatiotemporally varying catalytic reaction. AIChE J.

Training on Poincaré maps

- Rayleigh-Bénard convection experiments: 1.46% solution of ^{3}He in superfluid ^{4}He, with heated top plate and constant-temperature bottom plate.
- Bifurcation parameter is R, the Rayleigh number (proportional to temperature difference across the fluid layer).
- Single temperature probe near center of top plate.
- ANN trained to reproduce Poincaré sections: Given point (x_n, y_n), its time-τ delayed preimage, and R, predict (x_{n+1}, y_{n+1})

Sample Poincaré section

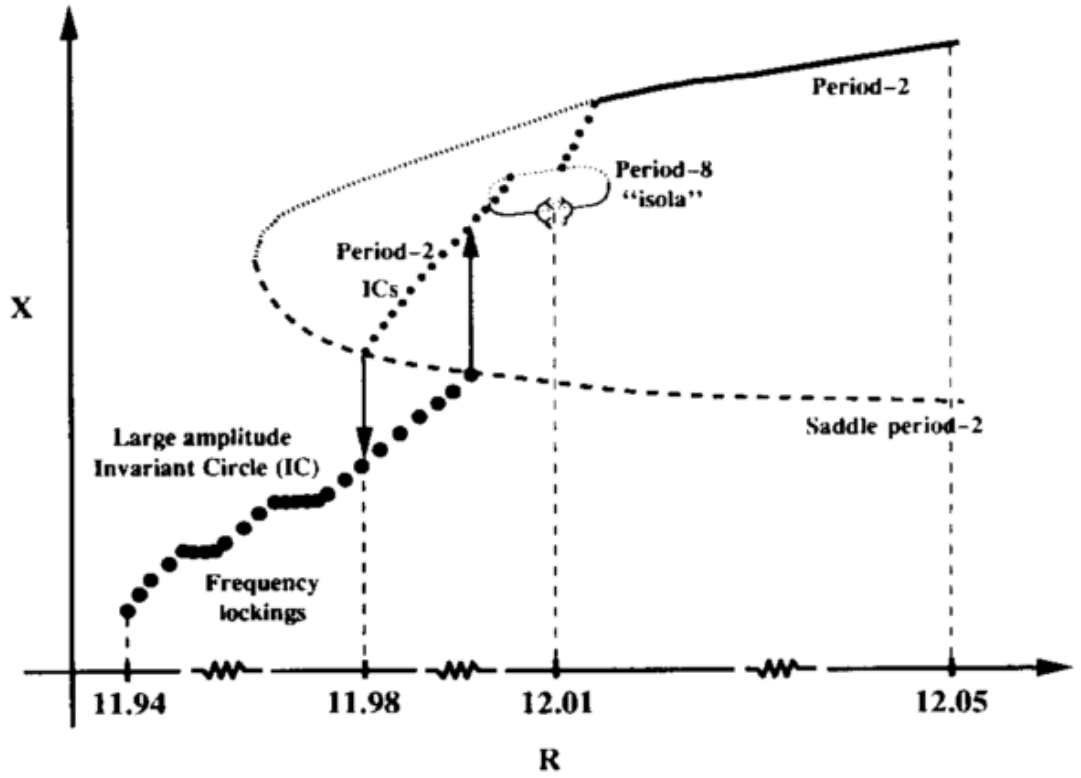




Kevrekidis, Rico-Martinez, Ecke, Farber, and Lapedes (1994) .Global bifurcations in Rayleigh-Bénard convection. Experiments, empirical maps and numerical bifurcation analysis. Physica D.

Training on Poincaré maps

Rayleigh-Bénard ANN can be used to make a detailed bifurcation diagram.



Kevrekidis, Rico-Martinez, Ecke, Farber, and Lapedes (1994) .Global bifurcations in Rayleigh-Bénard convection. Experiments, empirical maps and numerical bifurcation analysis. Physica D.

Identification of PDEs from data

Identification of distributed parameter systems: A neural net based approach

R. González-García[a], R. Rico-Martínez[a] *, and I. G. Kevrekidis[b]
[a] Depto. de Ingeniería Química, Instituto Tecnológico de Celaya, Celaya, Gto. 38010 México
[b] Department of Chemical Engineering, Princeton University, Princeton, N. J. 08544 U.S.A.

Abstract
Advances in scientific computation and developments in spatially resolved sensor technology have, in recent years, critically enhanced our ability to develop modeling strategies and experimental techniques for the study of the spatiotemporal response of distributed nonlinear systems. The usual alternatives for the modeling of these systems, simplifying techniques that seek to capture the distributed system dynamics through lumped parameter models, can be drastically underresolved, and miss important features of the true system response. Robust implementations of distributed system identification algorithms based on detailed spatiotemporal experimental data have, therefore, an important role to play. In this contribution we present a methodology for the identification of distributed parameter systems, based on artificial neural network architectures, motivated by standard numerical discretization techniques used for the solution of partial differential equations.

González-García, Rico-Martínez, & Kevrekidis. (1998). *Identification of distributed parameter systems: A neural net based approach*.

Idea

Assumption: “particle-based” trajectory data available with low dimensional dynamics

- fine-scale simulations
- sensors

but no effective/macroscopic description of the dynamics

Here: Data-driven discovery of effective descriptions for agent-based systems

$$\frac{\partial v}{\partial t} = f\left(v, \frac{\partial v}{\partial x}, \frac{\partial^2 v}{\partial x^2}, \ldots, \frac{\partial^n v}{\partial x^n}, p\right)$$

Hassan Arbabi et al, ACC21 arXiv:2011.08138

Seungjoon Lee et al., Chaos 30, 013141 (2020)

R. González-García et al., Computers & Chemical Engineering, 22 (1998)

S. Brunton et al, SINDy, PNAS (2016)

Zichao Long et al, PDE-Net, arXiv:1710.09668 (2017)

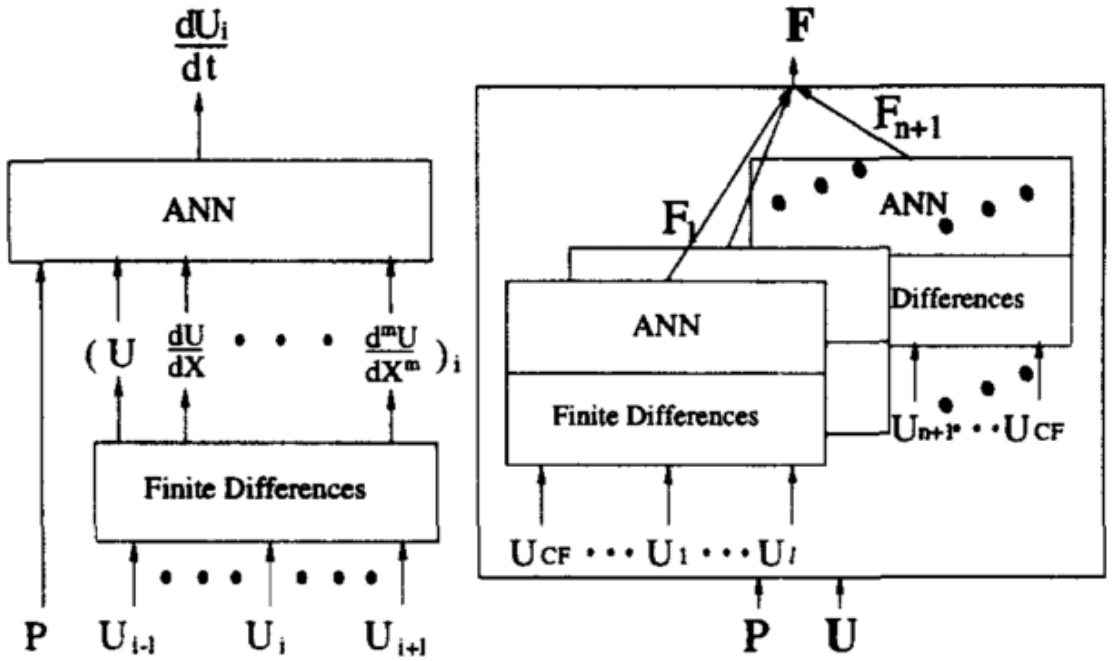
Identification of PDEs from data

Train a network that produces time-derivatives

from space-derivatives.

Use parameter sharing across space.

Note: Cross-correlation with a stencil for FD is identical to (and implemented as) a convolutional neural network (CNN).



Evaluation of ODE discretizations of PDEs

González-García, Rico-Martínez, & Kevrekidis. (1998). *Identification of distributed parameter systems: A neural net based approach*.

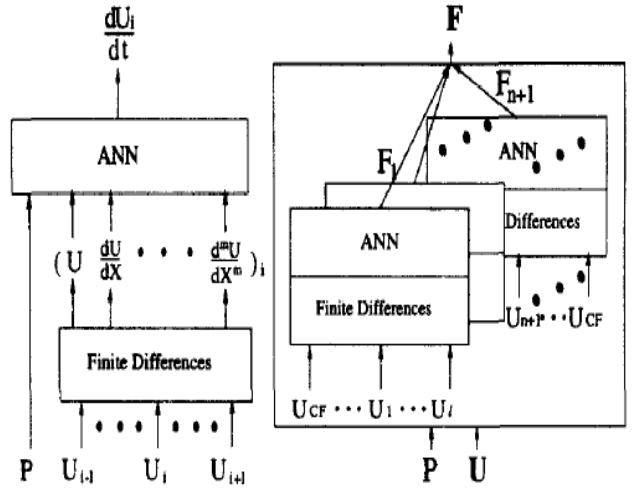


Figure 2: Evaluation of the right-hand-side (RHS) of the set of ODEs: On the left, at a point i in the grid, and on the right, for all points of the grid to form the vector of RHS $\mathbf{F}$.

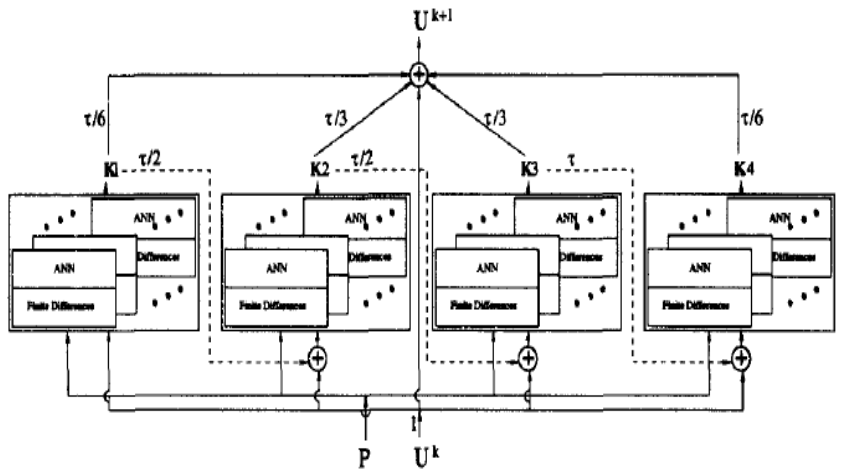
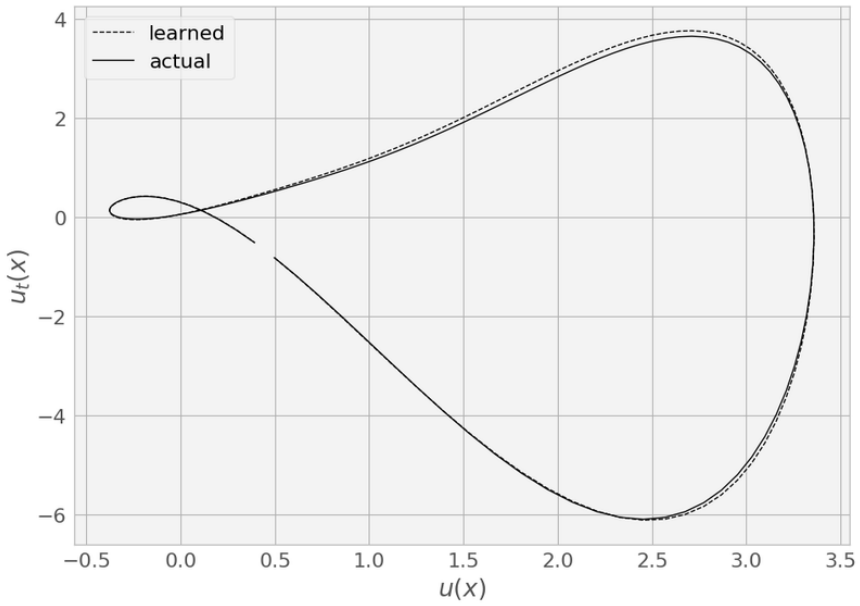
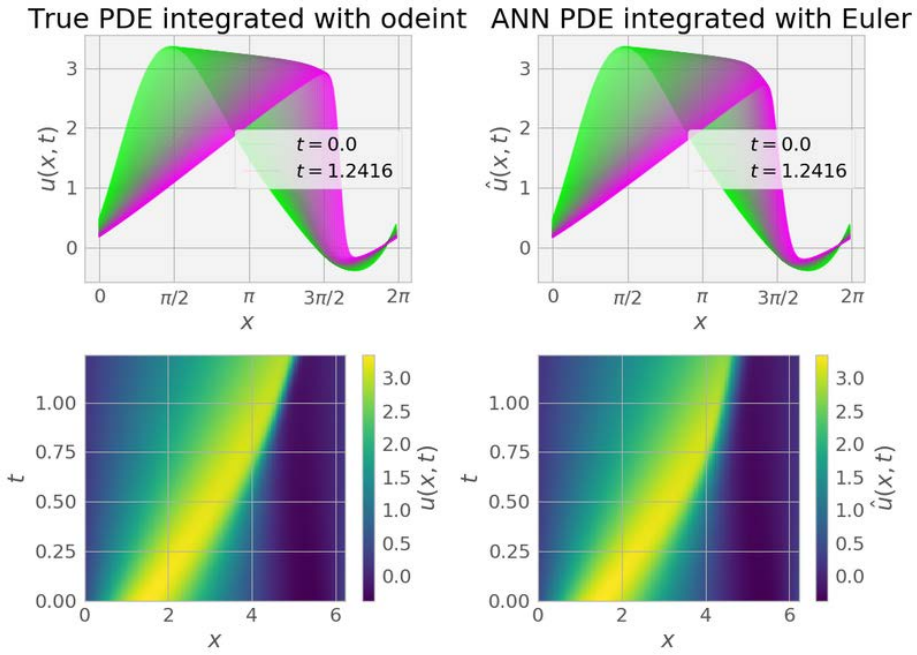


Figure 3: Network template for the identification of 1-D PDEs using the Runge-Kutta integrator.

Identification of PDEs from data

The result is a reasonable approximation of the Burgers' equation RHS





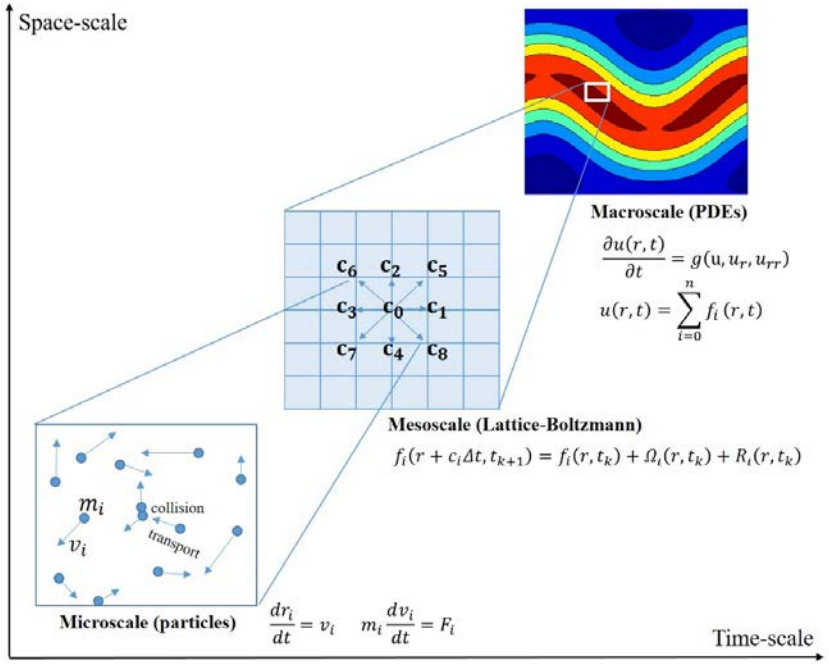
Narrow down to a fine-scale reaction-diffusion system

We collect fine-scale data from a Lattice Boltzmann model (LBM) for a reaction-diffusion system.

1) On a given lattice, LBM provides particle distribution functions (f_i).
2) Using the zeroth moment of particle distribution functions on every grid, we find the concentration of particles.

$$u(x_j, t_k) = \sum_{i=0}^{n} f_i(x_j, t_k).$$

3) Usually, the three particle distribution functions (n=3) for 1D and the nine particle distribution functions (n=9) for 2D.
4) We can calculate higher moments but we expect these moments become quickly slaved to few lower moments – we can understand how many coarse-scale observables we need.
5) Through manifold learning, we also check this tendency effectively.



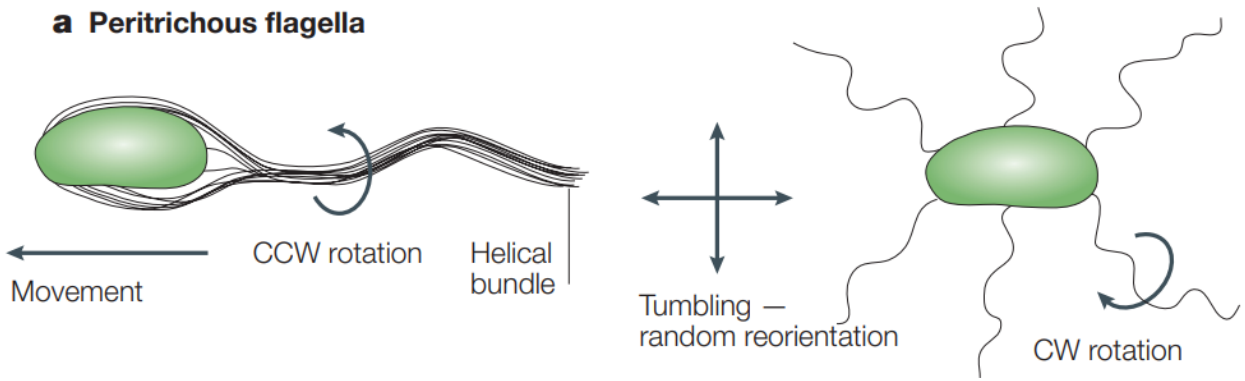
Successful extraction of coarse-scale observables!

Theodoropoulos, C., Qian, Y.H. and Kevrekidis, I.G., 2000. "Coarse" stability and bifurcation analysis using time-steppers: A reaction-diffusion example. Proceedings of the National Academy of Sciences, 97(18), pp.9840-9843.

Chemotactic Migration of Bacteria

Swimming in a straight line + Tumbling in space = Random Walk

- Counter-clockwise rotation aligns the flagella into a single rotating bundle.
- Clockwise rotation breaks the flagella bundle apart such that each flagellum points in a different direction.
- No memory for the direction due to tumbling.
- **Random walk**: relatively straight swims interrupted by random tumbles that reorient the bacterium → We can construct a mathematical model as a Diffusion equation.



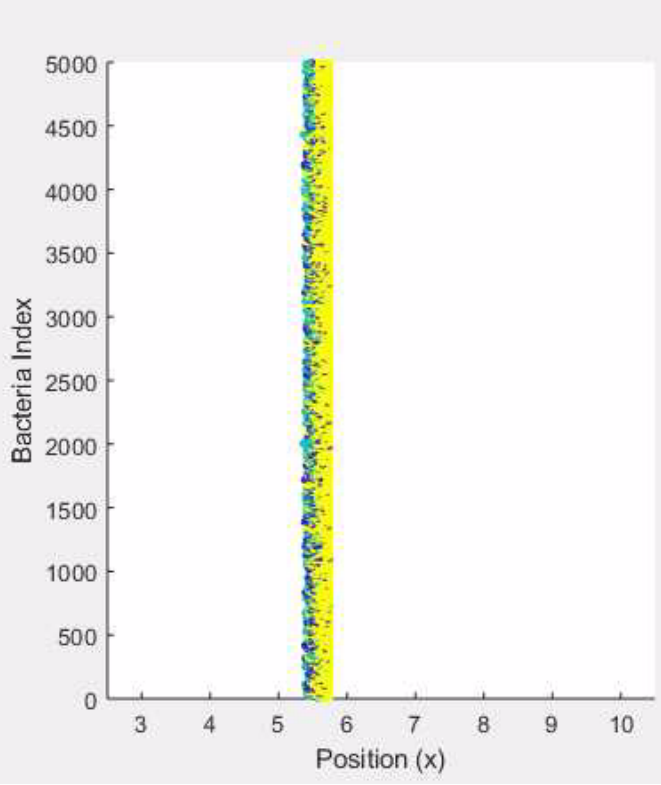
Butler, S.M. and Camilli, A., 2005.
Going against the grain: chemotaxis and infection in Vibrio cholerae.
Nature reviews Microbiology, *3*(8), pp.611-620.

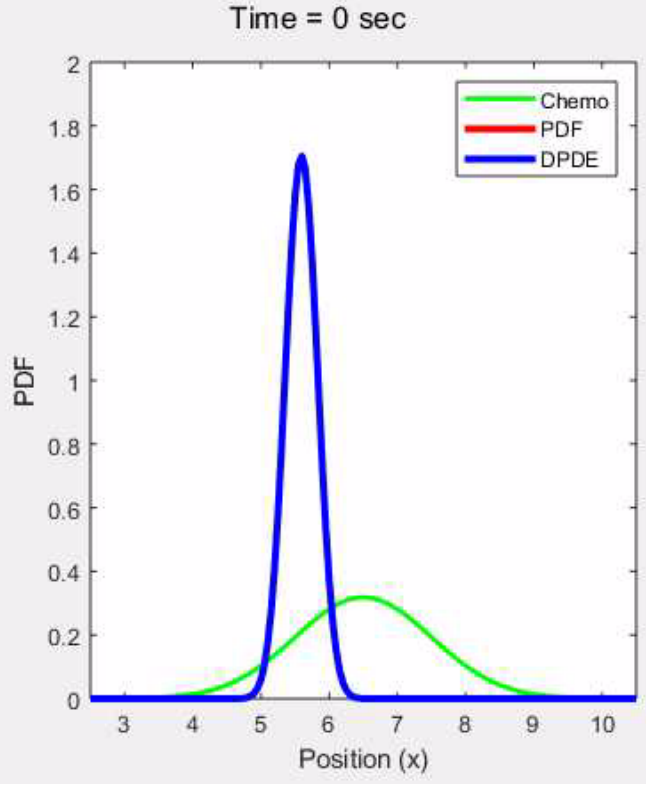
Results

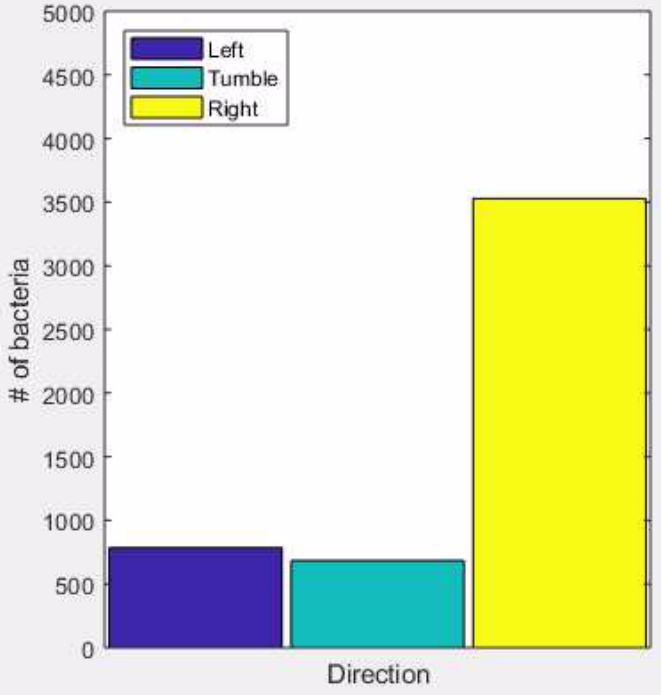
Green: fixed profile of the chemo-nutrient. $c(x) = \frac{1}{\pi} e^{\left(-\frac{(x-6.5)^2}{2}\right)}$

Blue: Data-driven PDEs (the entire black box right hand side)

Red: Probability density function from the agent-based model





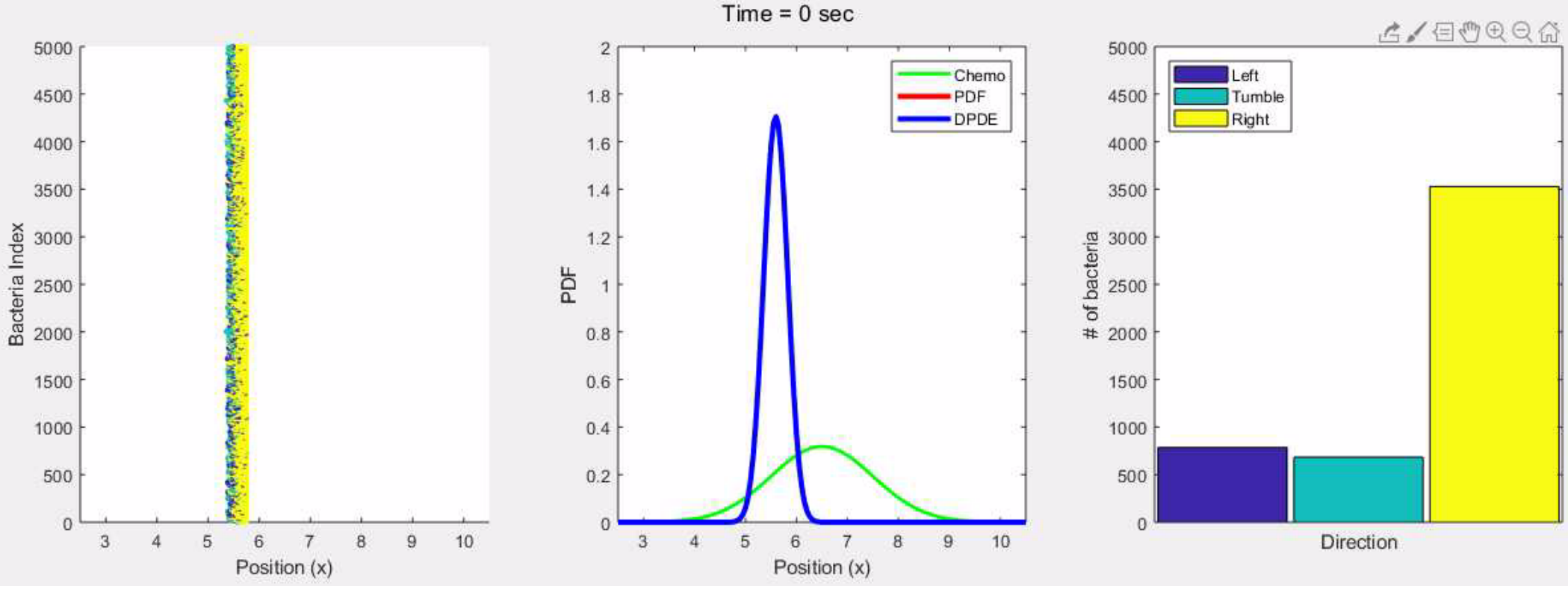


GRAY BOXES – just learn what you don't know

Green: fixed profile of the chemo-nutrient. $c(x) = \frac{1}{\pi} e^{\left(-\frac{(x-6.5)^2}{2}\right)}$

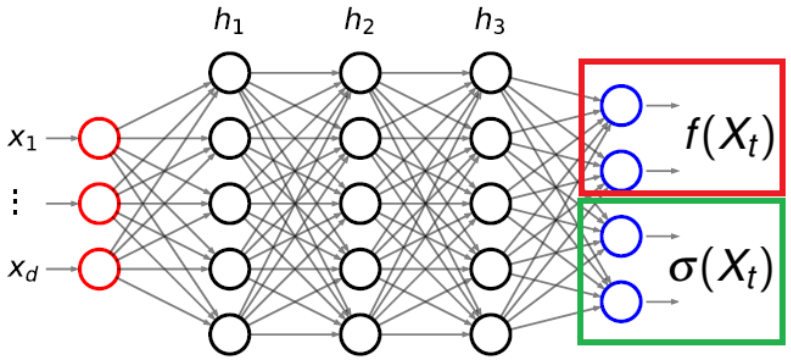
Blue: Data-driven PDEs with approximated closure of chemotactic term.

Red: Probability density function from the agent-based model



New twist: estimating SDEs using SDE solvers as templates (Dietrich, Reich)

- **Data**: snapshots $(x_t, x_{t+\Delta t})$ all over the state space, possibly with different Δt for each snapshot.
- **Goal**: approximate drift f and diffusivity σ.
- **Improvement**: compared to other current approaches, we **do not need** (a) long time series or (b) constant time steps.
- **Current assumptions**:
  - (a) diffusivity is an SPD matrix everywhere;
  - (b) drift and diffusivity are continuous w.r.t. the input;
  - (c) our dataset samples the state space well.



Network architecture to learn drift and (diagonal) diffusivity.

Our approach to estimate SDE

Main idea: We assume the data is generated by the SDE:

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t$$

This SODE can be approximated with the **Euler-Maruyama** scheme:

$$X(t+\Delta t) - X(t) \approx f\big(X(t)\big)\Delta\mathrm{t} + \sigma\big(X(t)\big)\xi,\ \xi \sim \mathcal{N}(0, \Delta t\ I_n).$$

Given this scheme, we can assume that

$$X(t+\Delta t) \sim \mathcal{N}\left(X(t) + f\big(X(t)\big)\Delta\mathrm{t}, \sigma\big(X(t)\big)^T \sigma\big(X(t)\big)\Delta t\ \right).$$

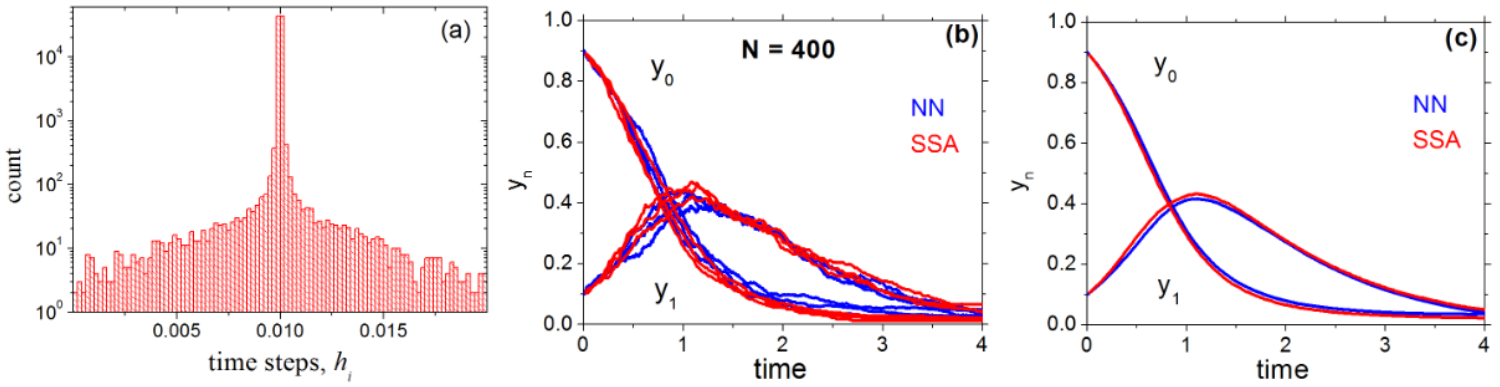
Setting p_θ as the probability density of this normal distribution, and f_θ, σ_θ our neural networks, then

$$\theta \coloneqq \arg\max_\eta \mathbb{E}\big[\log p_\eta\ (X(t+\Delta t)|X(t))\big] \approx \arg\max_\eta \frac{1}{N}\sum_{i=1}^{N}\big[\log p_\eta\ (X_1^i|X_0^i)\big]$$

If we use the **Euler-Maruyama** template, the logarithm of p_η can be computed easily.

What about other numerical integration schemes? (Milstein, Heun, stochastic Runge-Kutta, …)

Coarse-graining results



Histogram of the time steps used in the training data set (a). Three sample paths of y0 and y1 (b); Average of 200 paths (c). Blue (red) lines correspond to NN (SSA) results. Parameters: N = 400; Ntr = 10000, tmax = 0.1, h = 0.01.

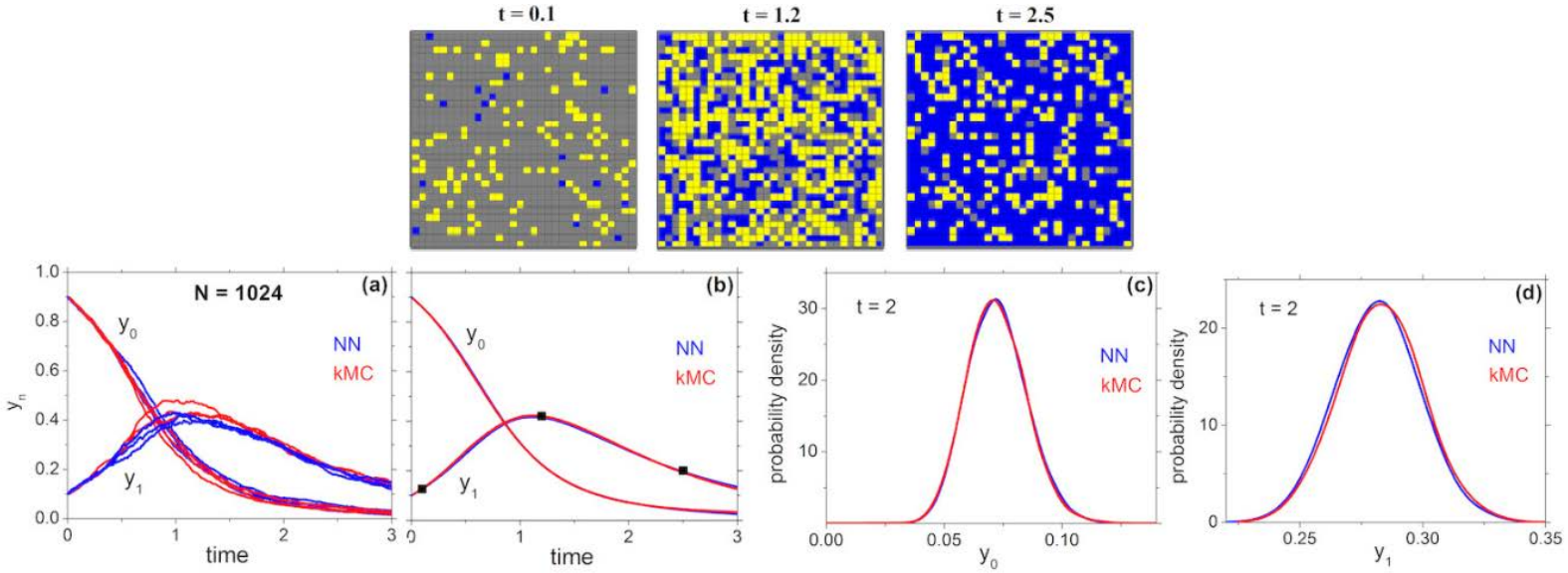
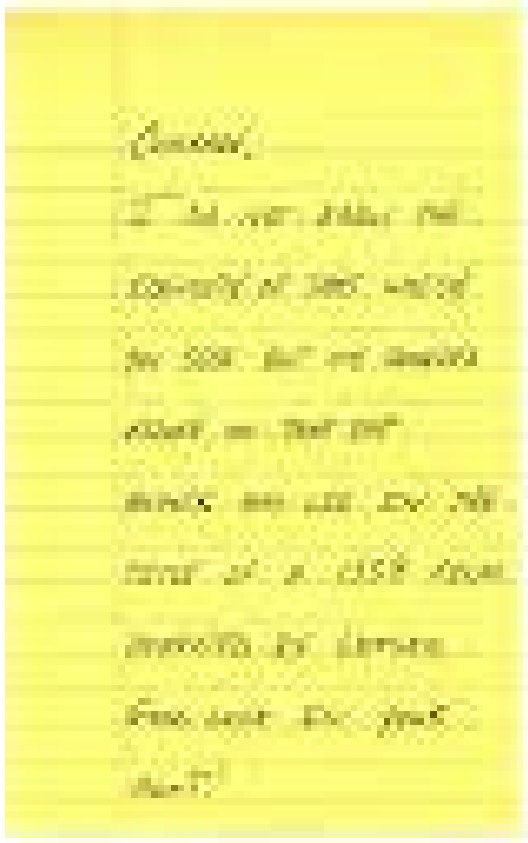


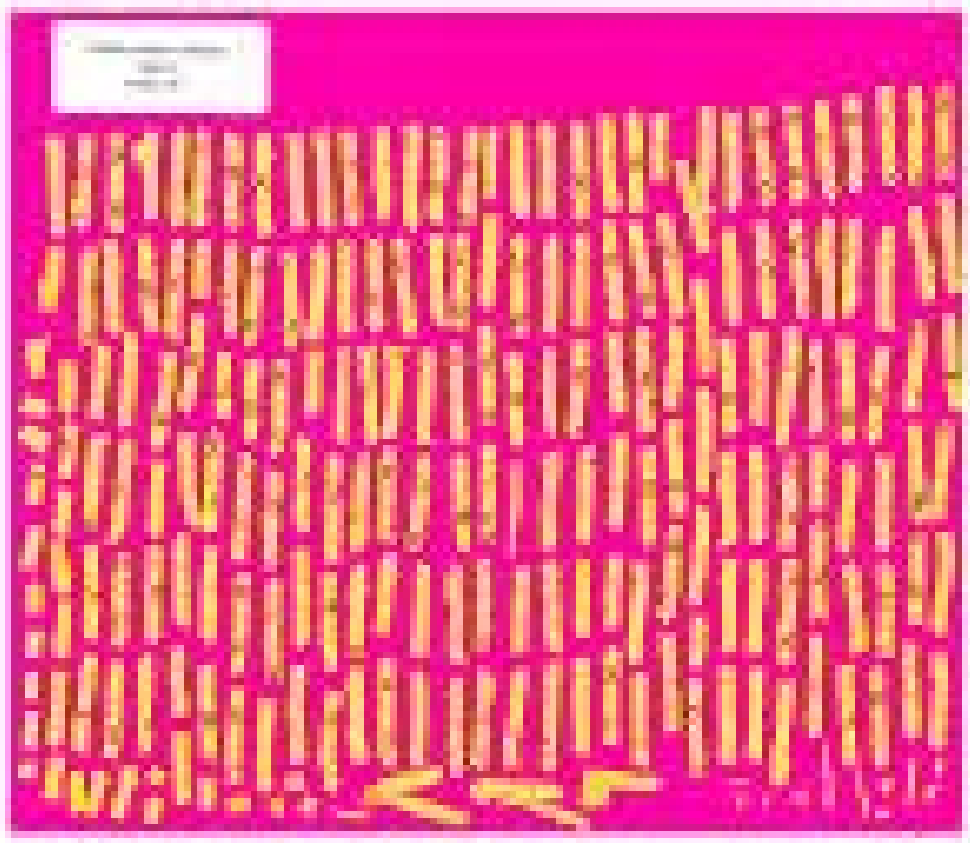
Illustration of the kMC lattice (top row), physically relevant values y0,y1 measured over time resp. simulated with the identified SDE from the network (a), averaged paths over 200 simulations (b), and propagated densities from the initial condition until t = 2 (c,d). Lattice snapshots show S, I and R-type species as grey, yellow and blue squares, respectively. Parameters: 32 ×32 lattice, d = 50; Ntr = 4000, tmax = 0.05, h = 0.01.

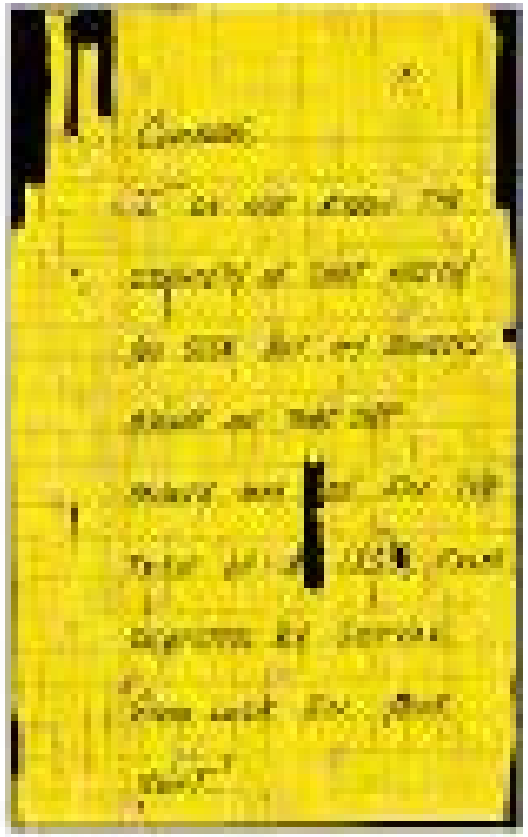
Emergent Space, Emergent Time (no space, no time)

DARPA Shredder Challenge problem 1

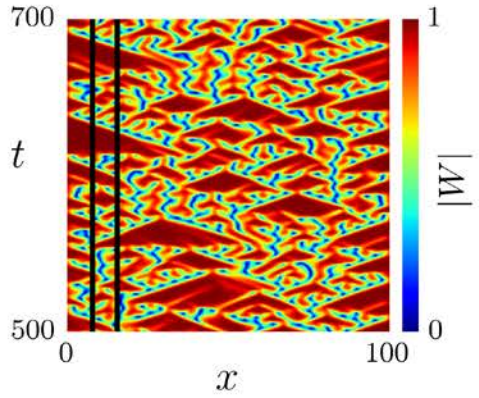


Original Document

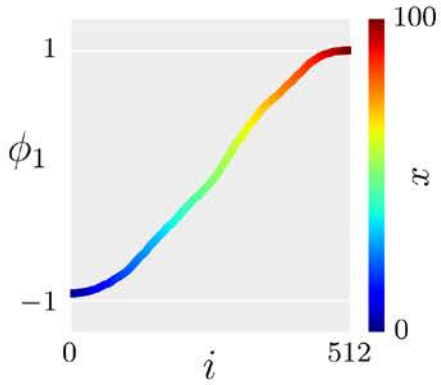




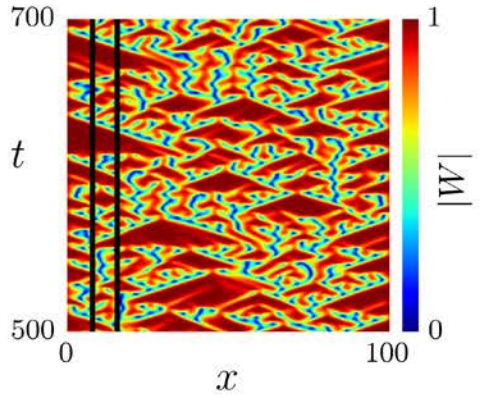
Finding the right space from data



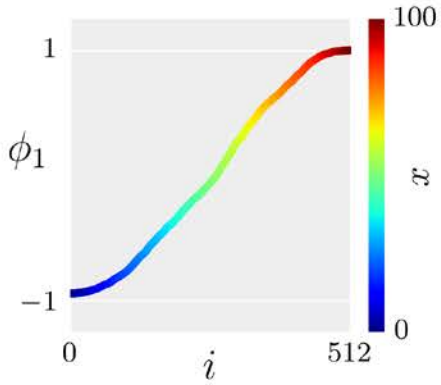
Simulate the 1D model (GL) PDE.



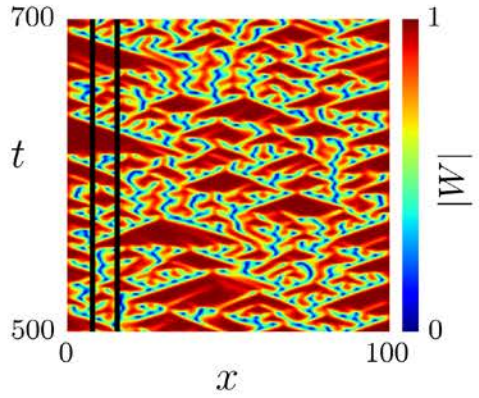
Finding the right space from data

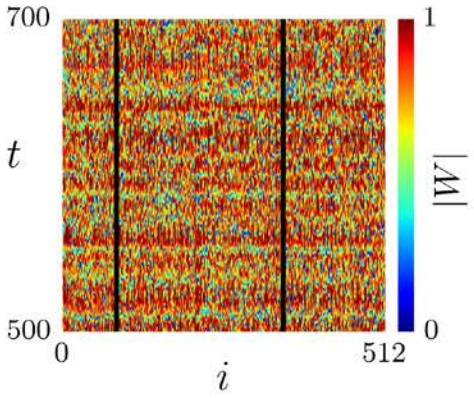


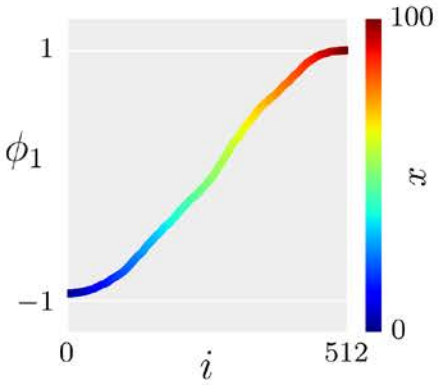
Shuffle time series.

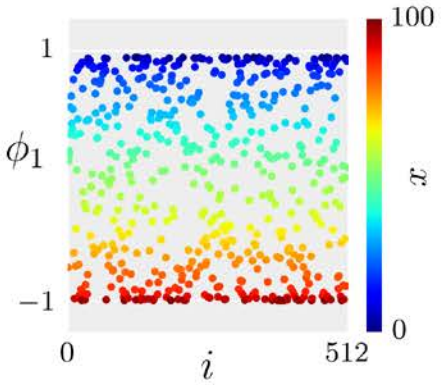


Finding the right space from data



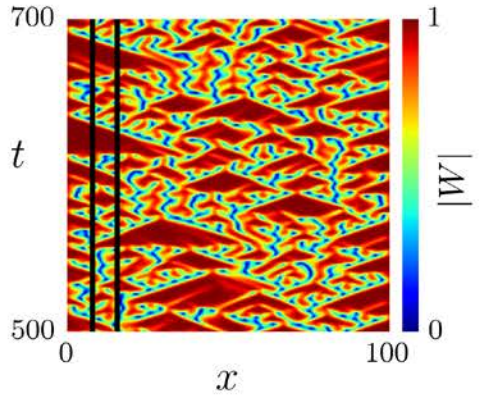


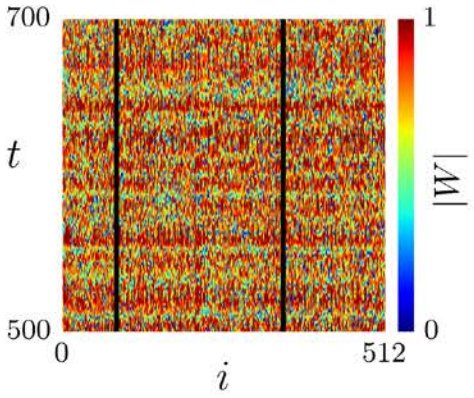


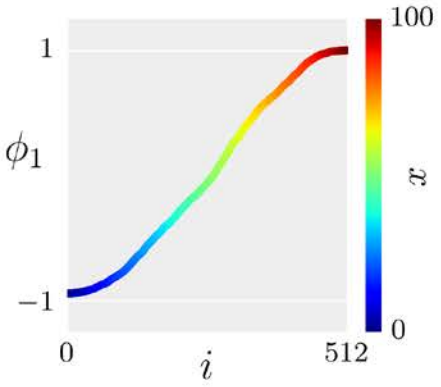


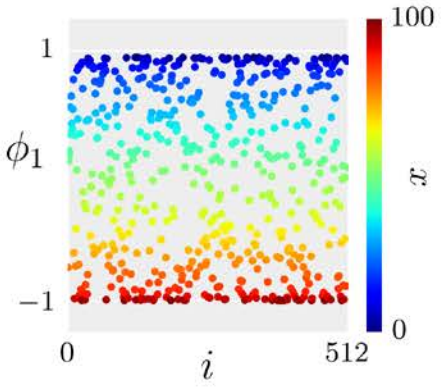
Use metric between time series to create diffusion embedding.

Finding the right space from data



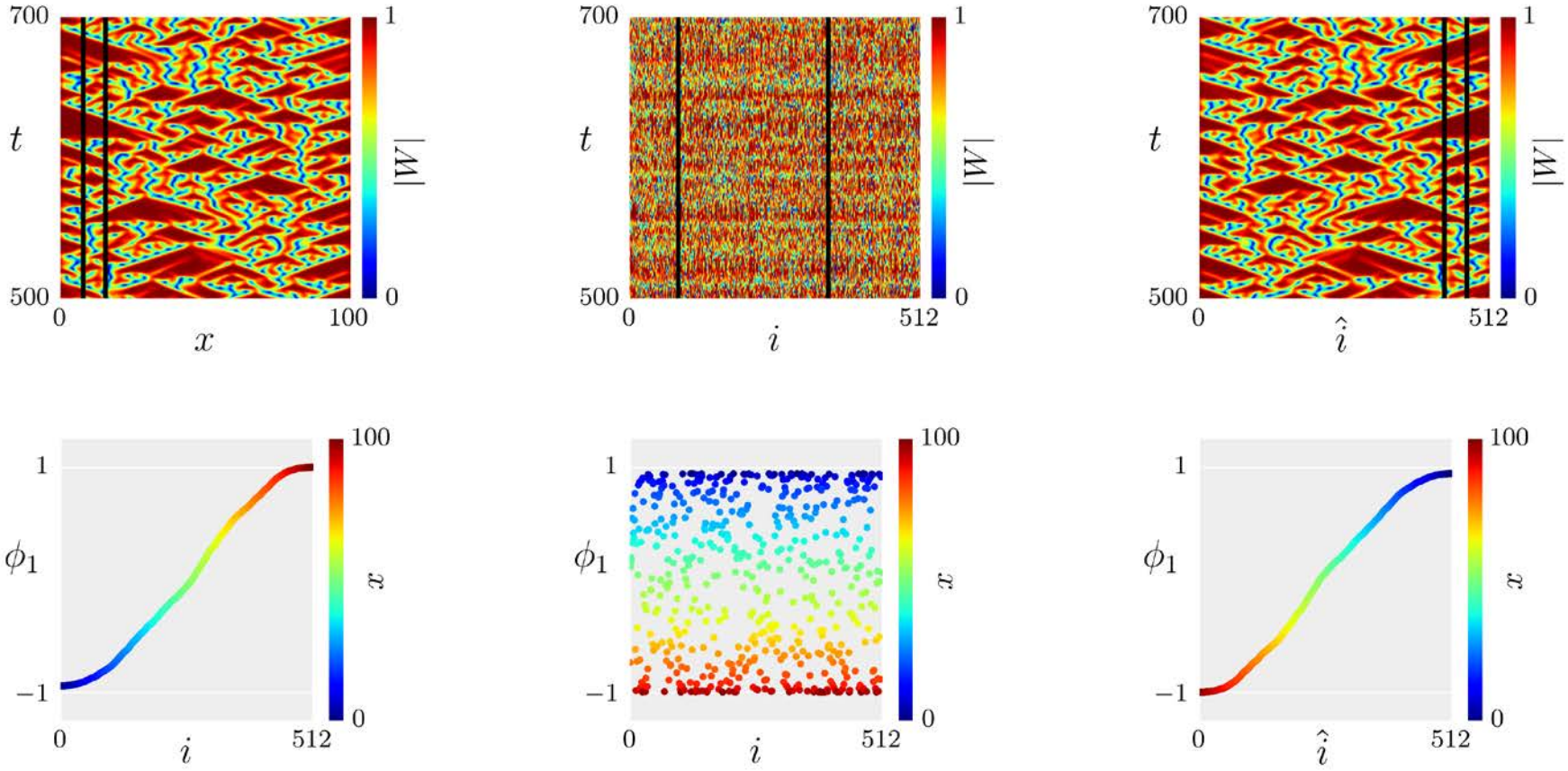




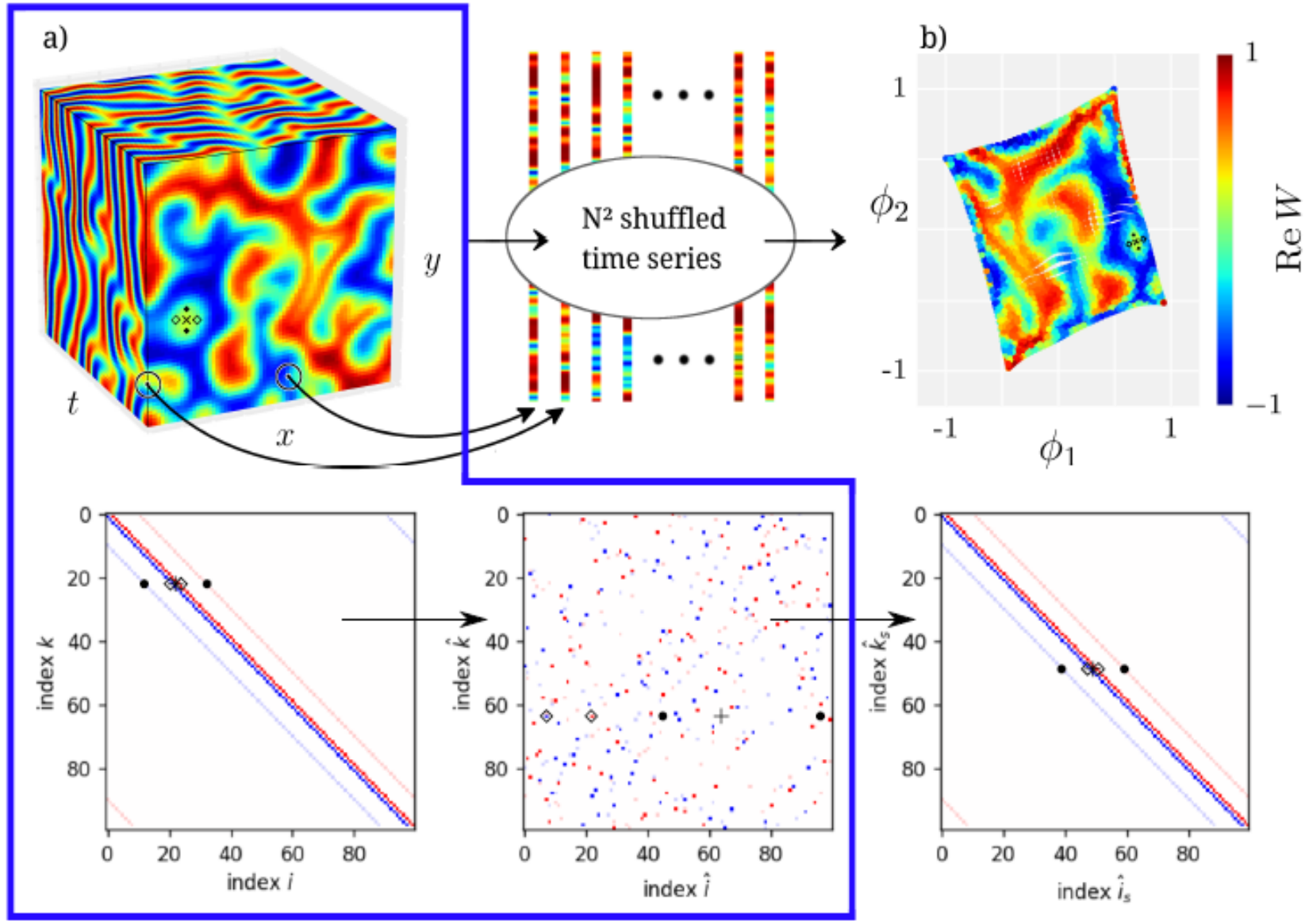


Sort by diffusion mode.

Finding the right space from data

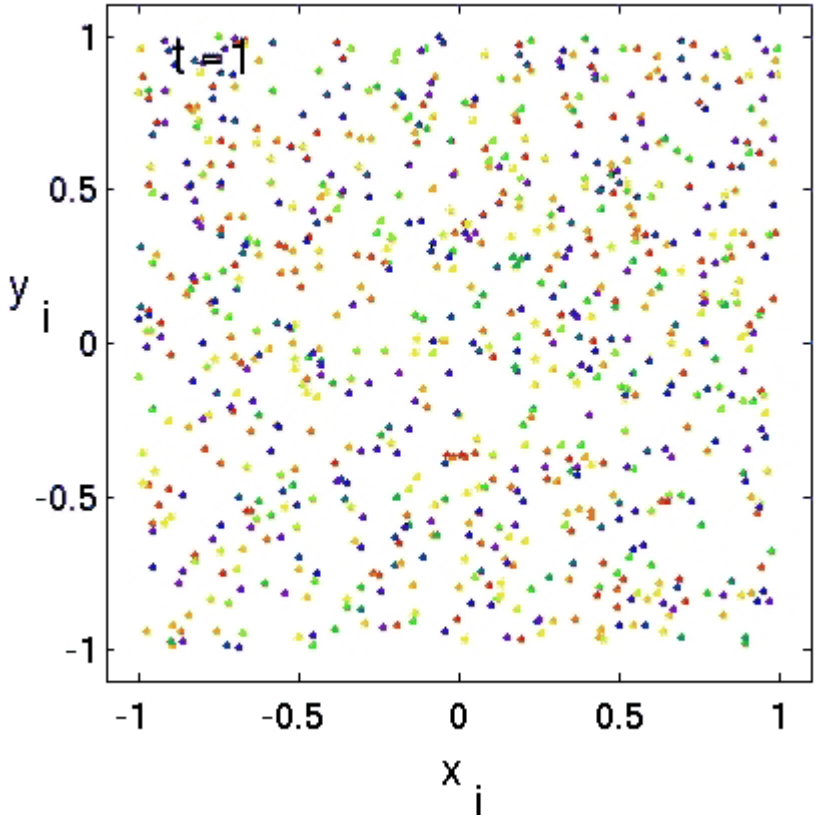


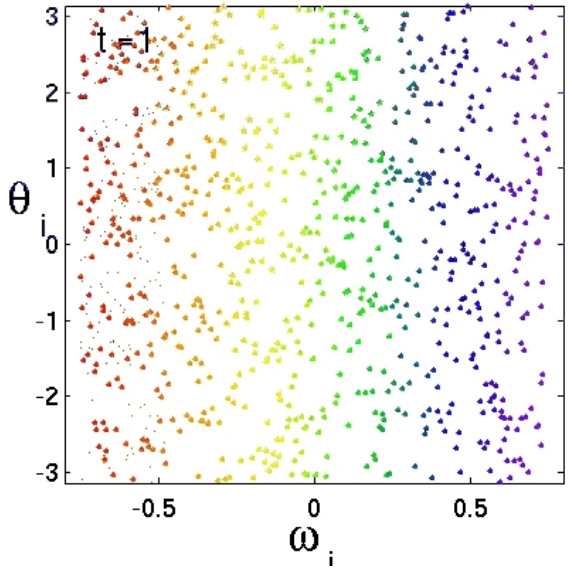
Recover original space, modulo reflection.

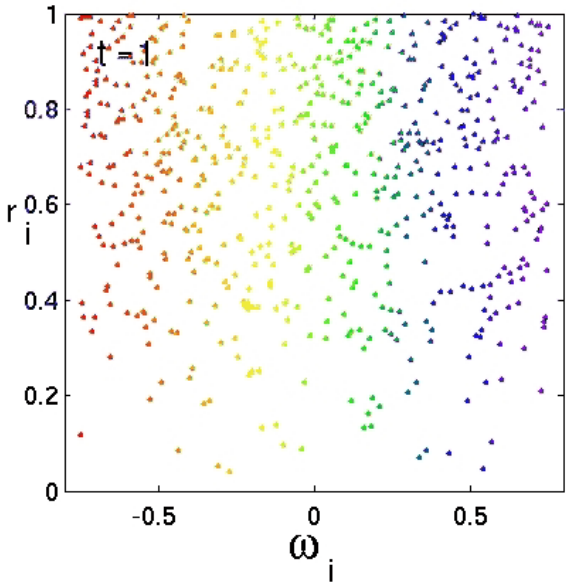

a)
b)
y
t
x
N² shuffled
time series
ϕ_2
ϕ_1
1
-1
Re W
index k
index i
index $\hat{k}$
index $\hat{i}$
index $\hat{k}_s$
index $\hat{i}_s$
0
20
40
60
80
hidden
practically impossible
to interpret
reconstructed through
data-mining

Emergent Spaces

Large, collective oscillations



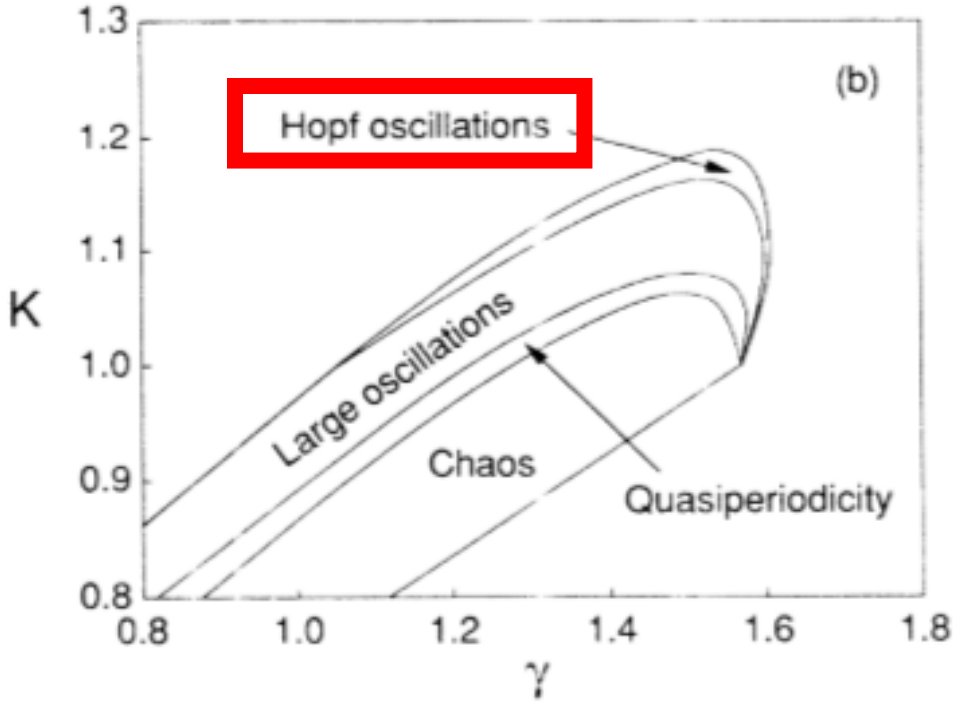




ω: Uniform distribution in [-0.75,0.75]

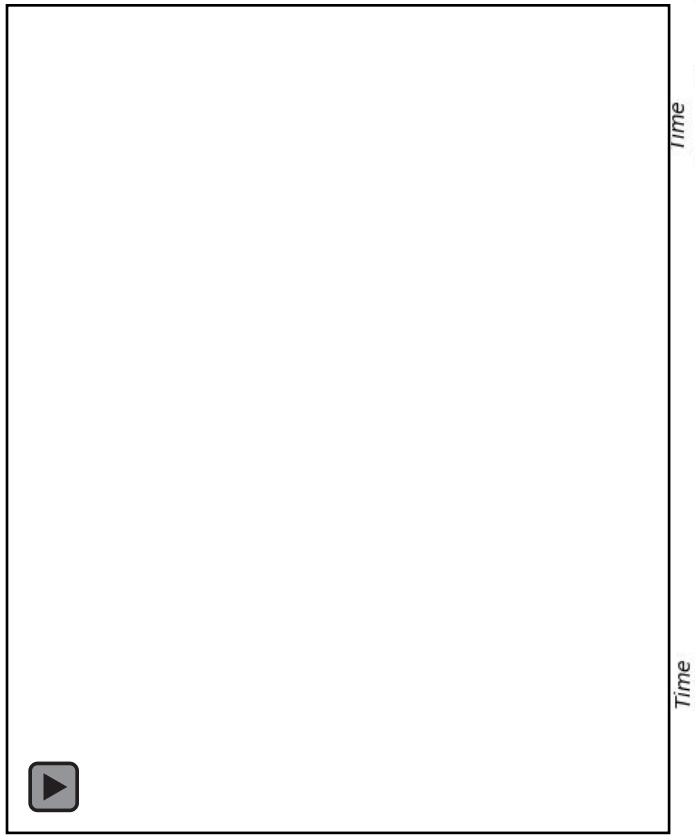
(K =0.8; N = 800)

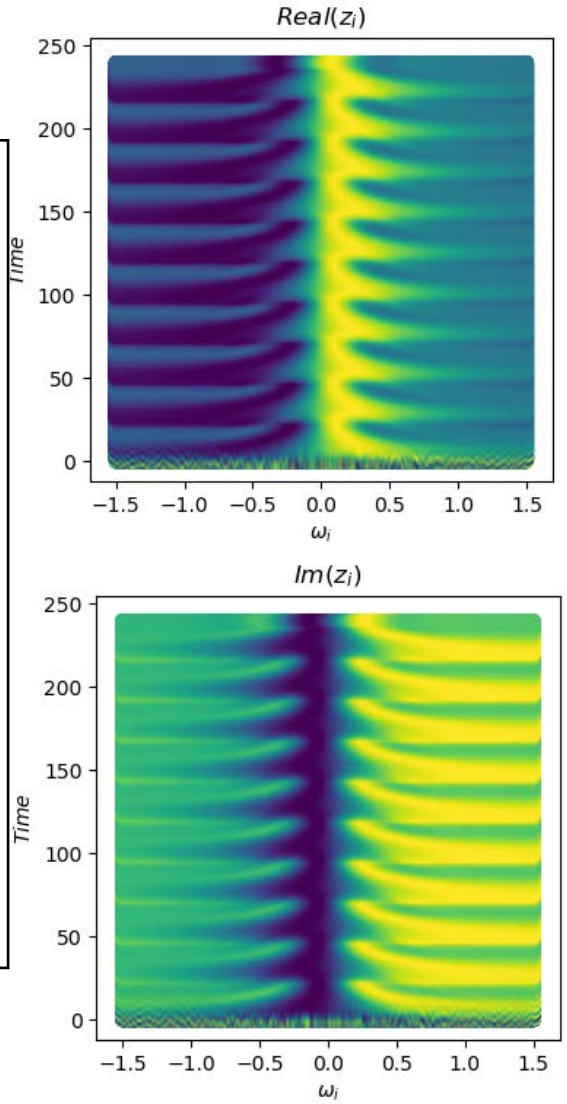
Other phase-amplitude oscillators exhibit multiple regimes



$$\dot{z}_j = z_j\left(1 - |z_j|^2 + i\,\omega_j\right) + \frac{K}{N}\sum_{i=1}^{N}(z_i - z_j)$$

$$\omega_j = -\gamma, \ldots, \gamma$$



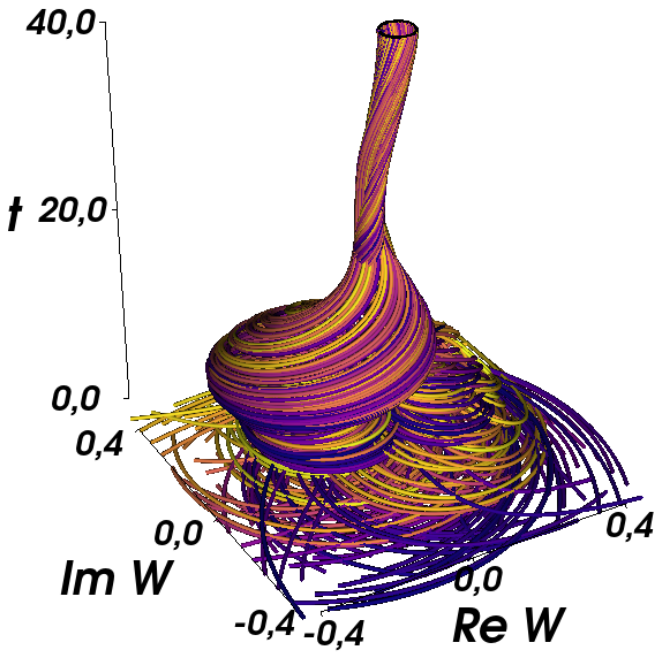


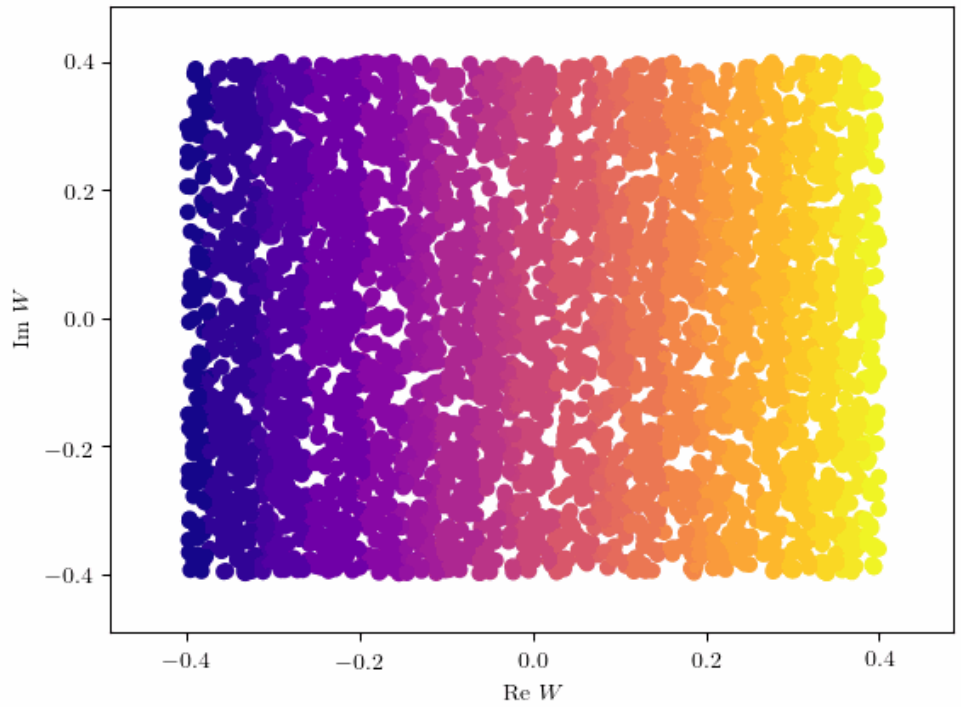
Matthews and Strogatz. (1990). *Phase diagram for the collective behavior of limit-cycle oscillators.*

Example

$$\frac{\partial W_k}{\partial t} = (1 + i\omega_k) W_k - |W_k|^2 W_k + \frac{K}{N} \sum_j (W_j - W_k)$$

Heterogeneous ensemble of mean-coupled oscillators
Frequencies ω_k drawn uniformly from the interval $[\omega_0-\gamma, \omega_0+\gamma]$[1]



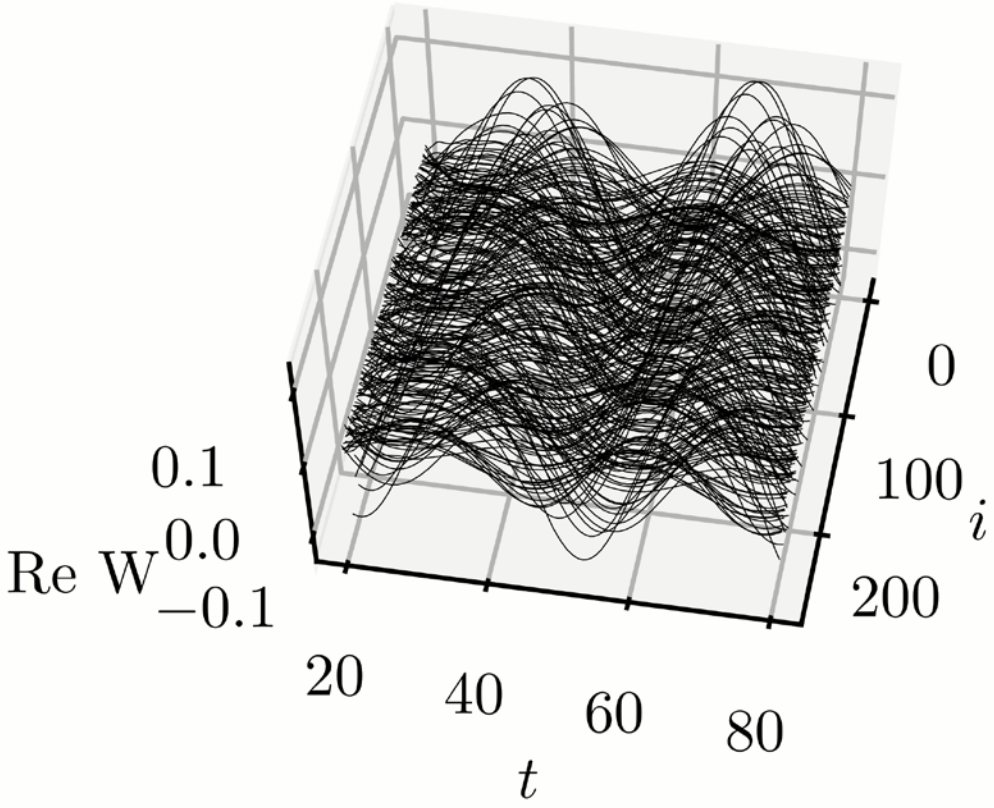


[1]P.C. Matthews and S.H. Strogatz; Phase Diagram for the Collective Behavior of Limit-Cycle Oscillators; PRL 65 (14), 1990

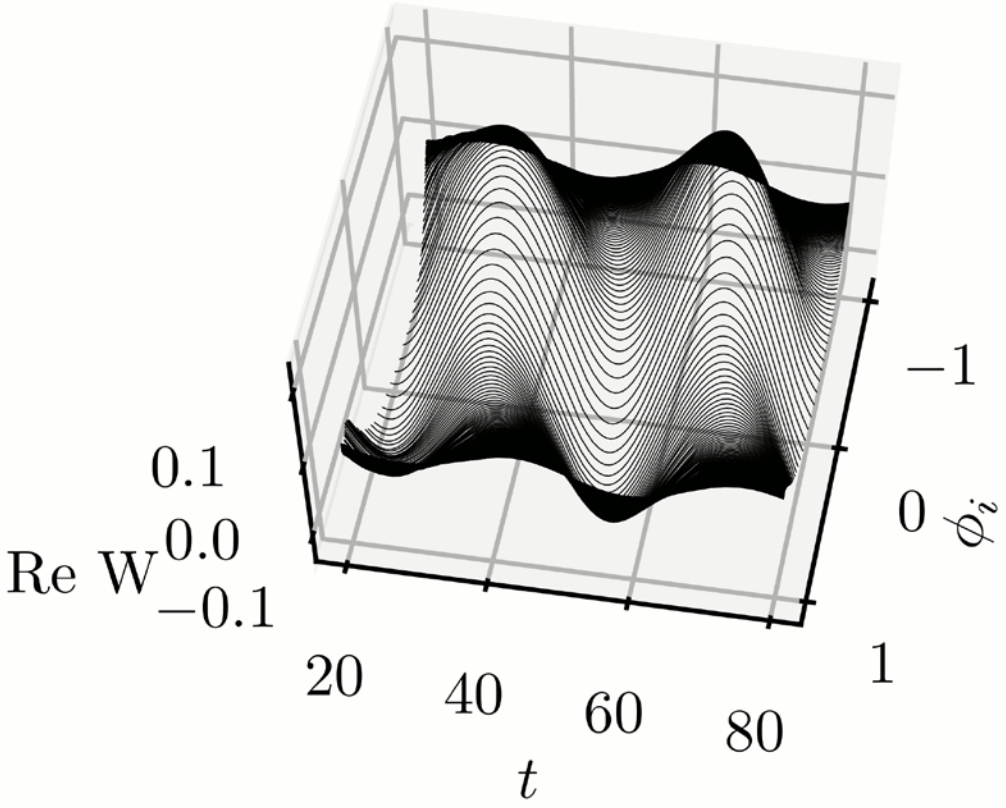
JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Globally coupled heterogeneous Stuart-Landau oscillators

Time series in the original parametrization i

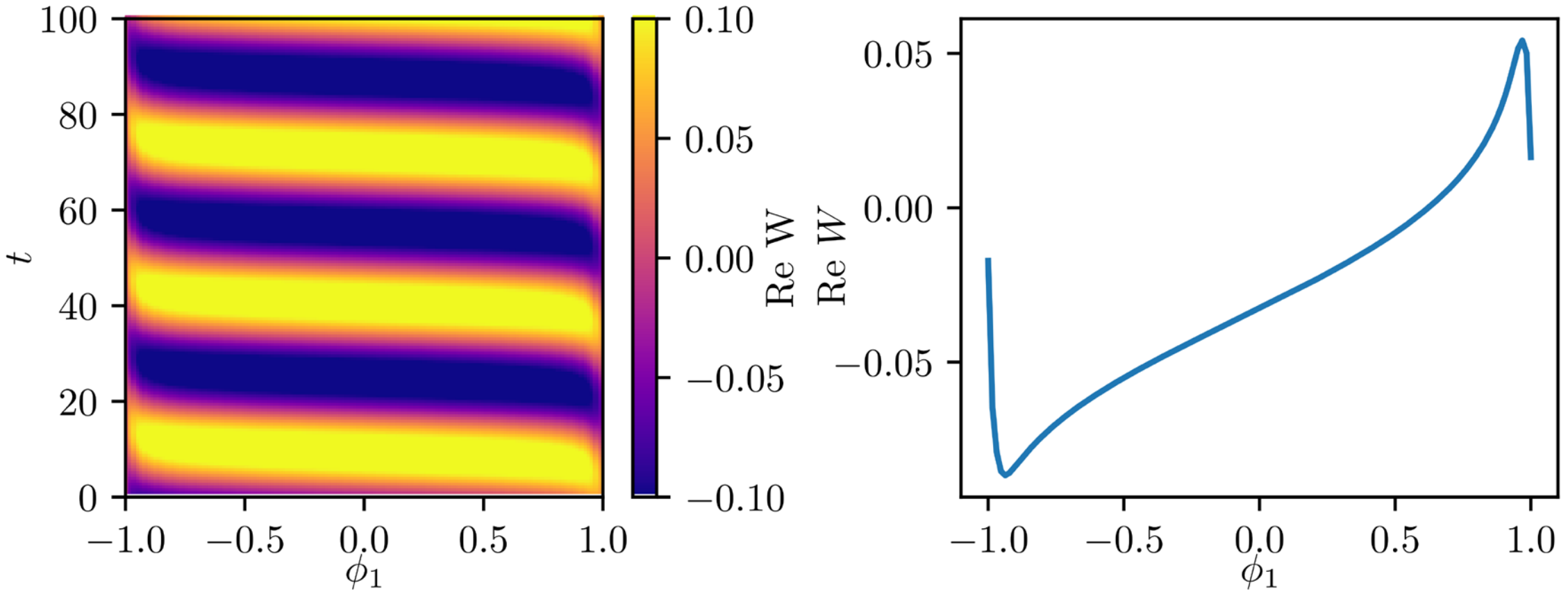


Time series in the new (emergent) coordinate



JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

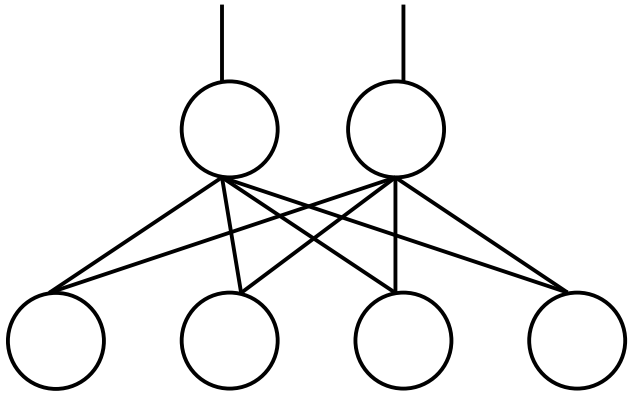
Dynamics and snapshot in the emergent coordinate



NN PDE Description

Find a partial differential equation in the emergent space

$$\frac{\partial W}{\partial t} = f\left(W, \frac{\partial W}{\partial \phi_1}, \frac{\partial^2 W}{\partial \phi_1^2}, \ldots, \frac{\partial^n W}{\partial \phi_1^n}\right)$$



by learning f based on data.

Represent f through a neural network, parametrized by weights Θ[1]

Optimize f by minimizing the mean-squared error between the time derivative and the output of f

$$L = \frac{1}{N_d} \sum \left| \frac{\partial W}{\partial t} - f(\cdot) \right|^2$$

Derivative approximations are obtained using finite differences

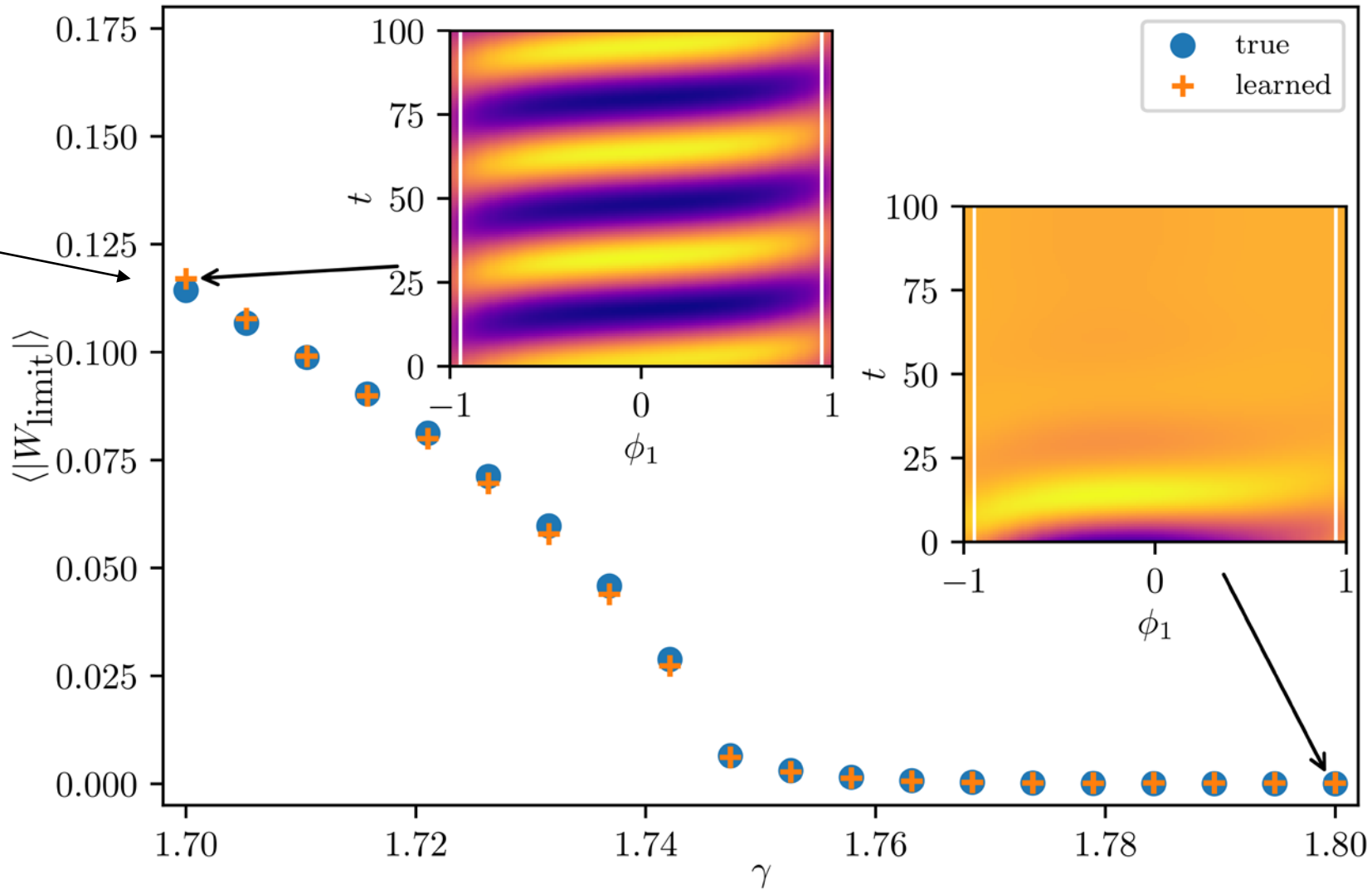
$$\frac{\partial W}{\partial t} \approx \frac{W(t + \Delta t) - W(t)}{\Delta t}$$

[1]K. Krischer et al., Model identificationof a spatiotemporally varying catalytic reaction. AIChE Journal, 39(1):89–98, 1993.
R. Rico-Martínez et al. Discrete- vs. continuous-timenonlinear signal processing of cu electrodissolution data.Chemical Engineering Communications, 118(1):25–48,1992.
R. González-García et al., Identification of distributed parameter systems: a neural net based approach.Computers & Chemical Engineering, 22(nil):S965–S968, 1998.

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Parameter dependent PDE description

Mean amplitude after integrating with the true model and the learned PDE for different γ values



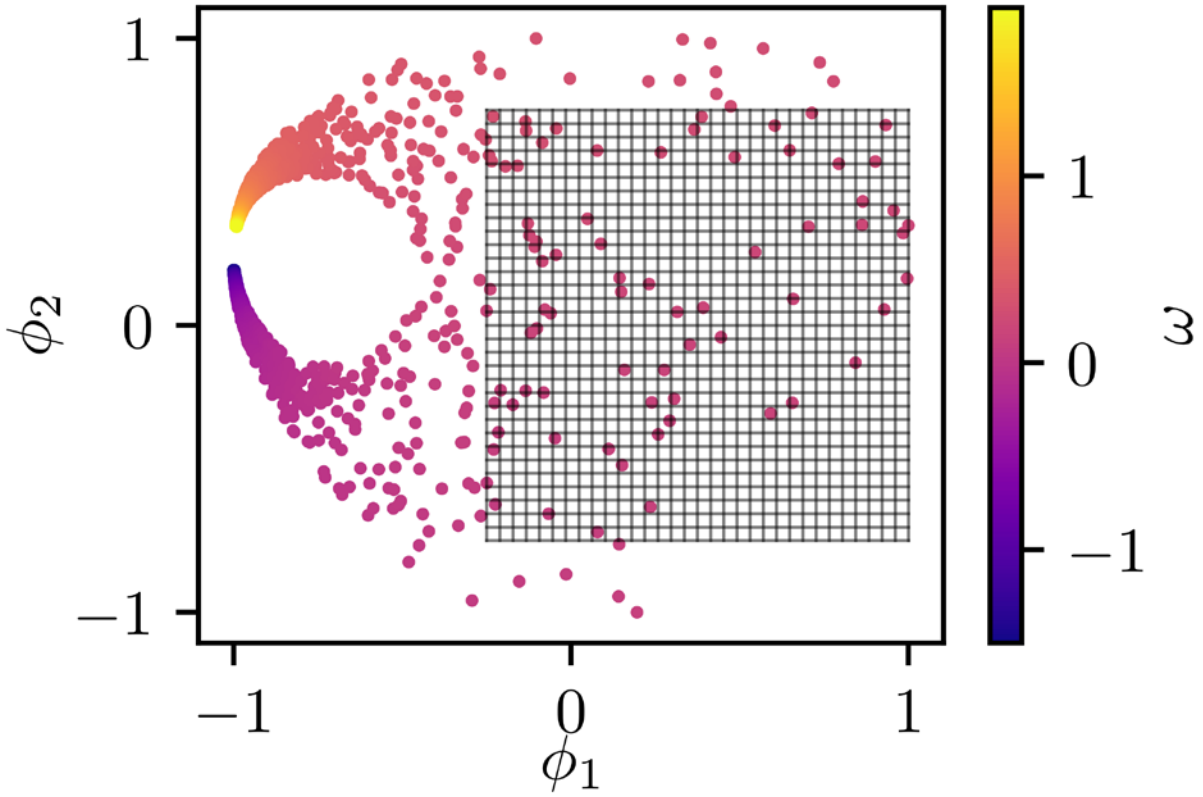
JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

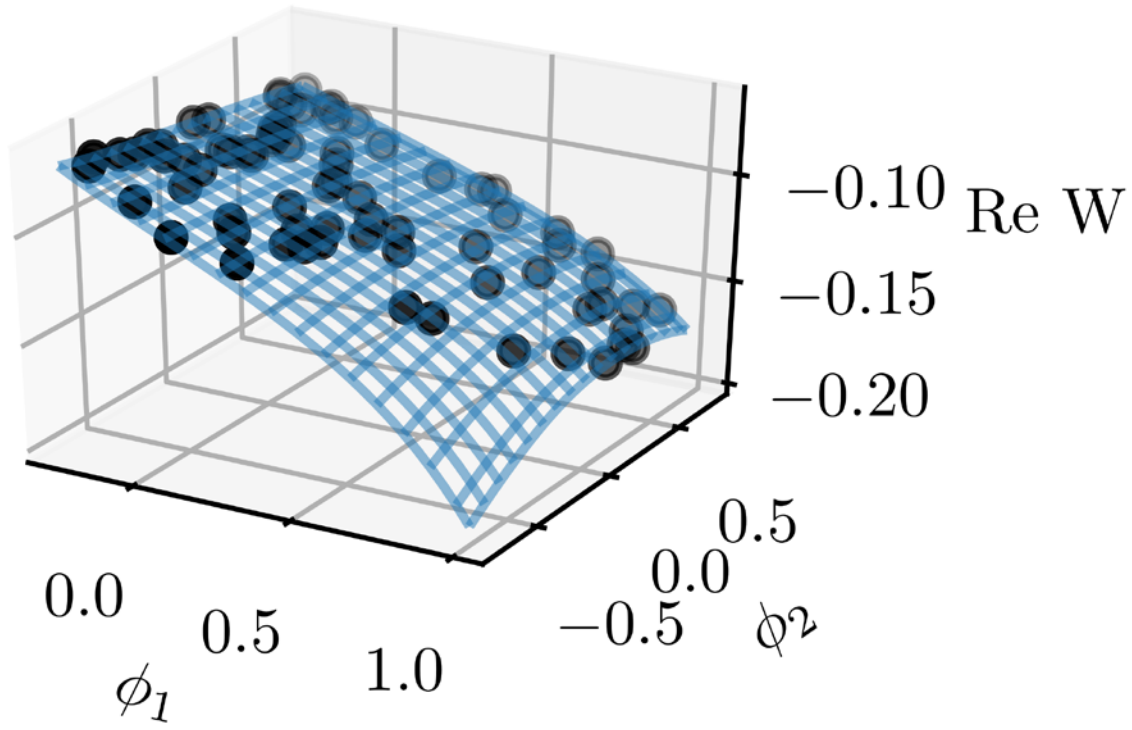
Globally coupled heterogeneous Stuart-Landau oscillators with two heterogeneous parameters

In addition to the intrinsic frequencies ω_k take a second parameter λ_k as a second heterogeneous parameter:

$$\partial_t W_k = (\lambda_k + i\omega_k) W_k - |W_k|^2 W_k + \frac{K}{N} \sum_j (W_j - W_k)$$

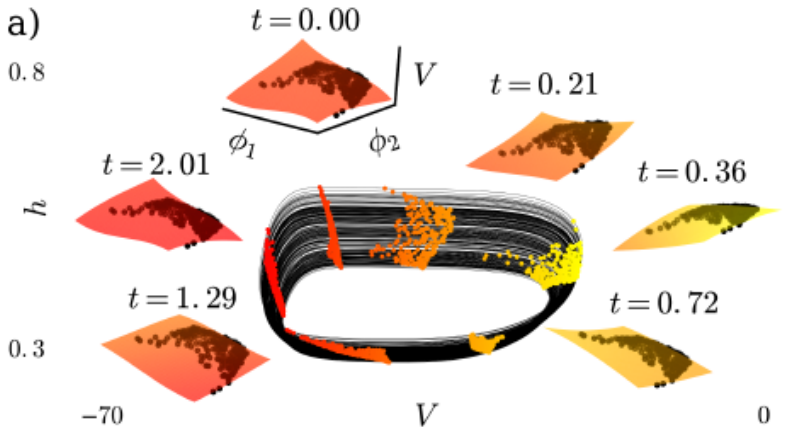
→ Using manifold learning (diffusion maps), one finds **two independent components** parametrizing the oscillators

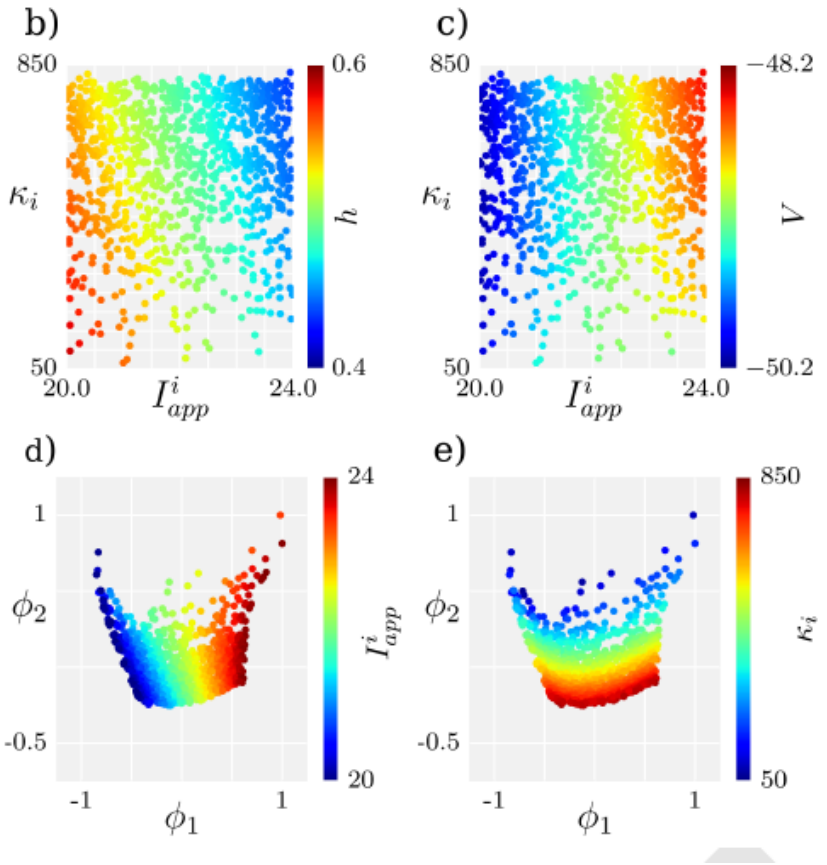




Finding a space for a PDE

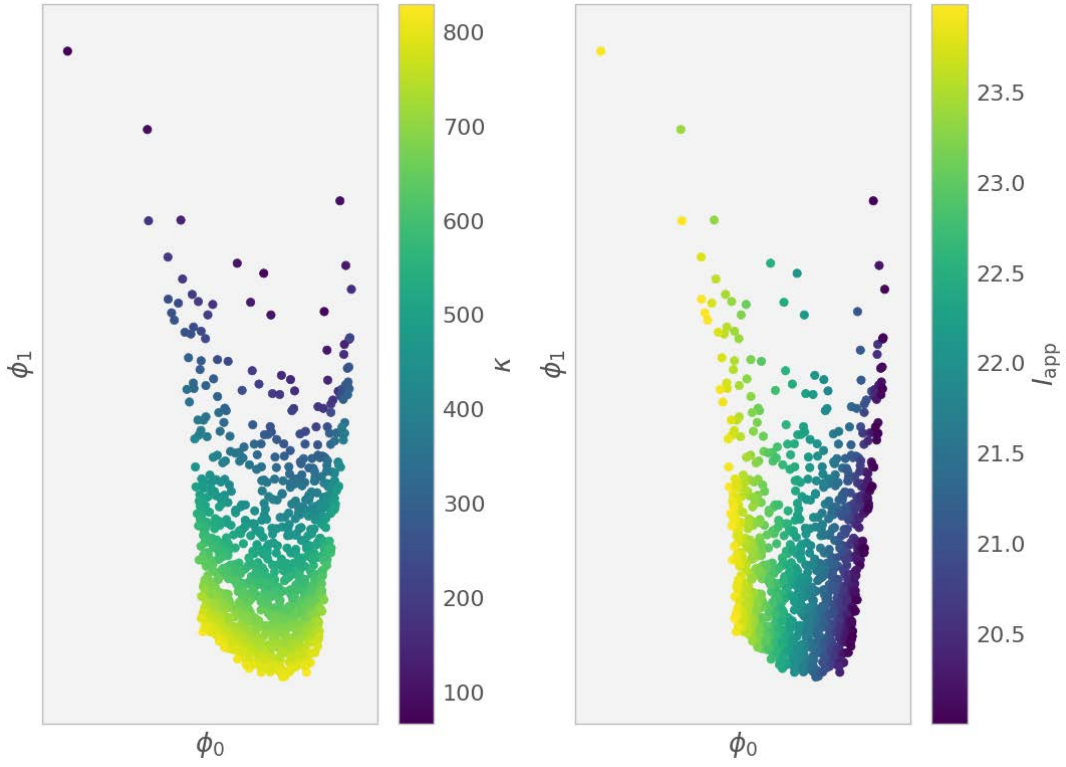
We simulate a Chung-Lu network of 1024 synchronized Hodgkin Huxley neurons.
This looks like a two-parameter family of dynamic behaviors, i.e., like a PDE in two dimensions
But we pretend these “space” dimensions need to be discovered from the data.

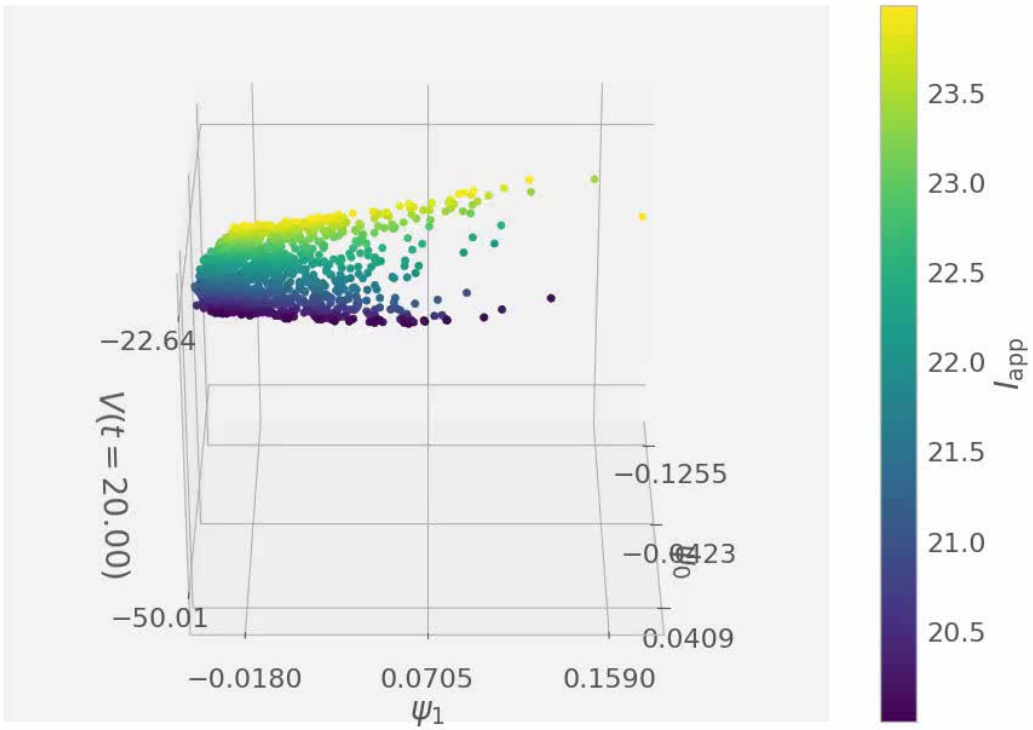




Coupled oscillators can also be treated as a PDE

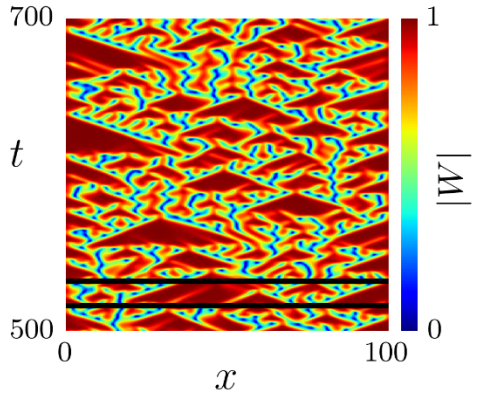
- Single-ion-channel HH neurons with network degrees κ and applied currents I_{app}.
- Use metric between timeseries to create embedding.
- “Spatial” parameterization by heterogeneities is recovered.

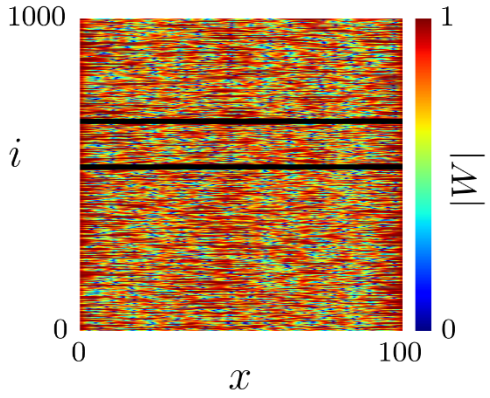


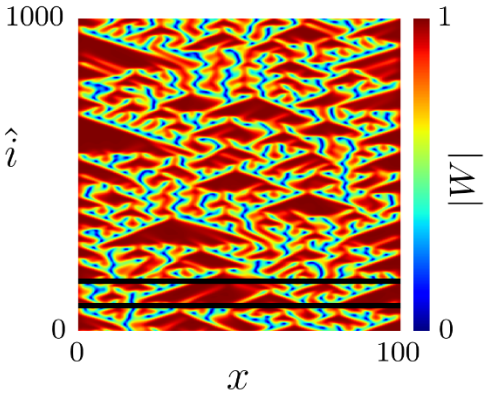


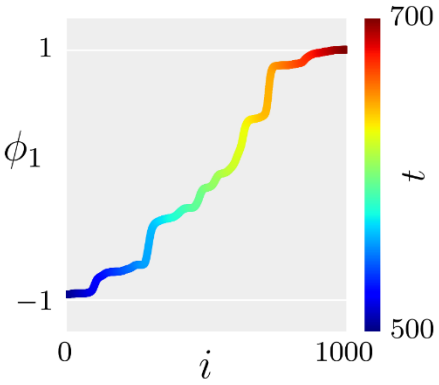
...

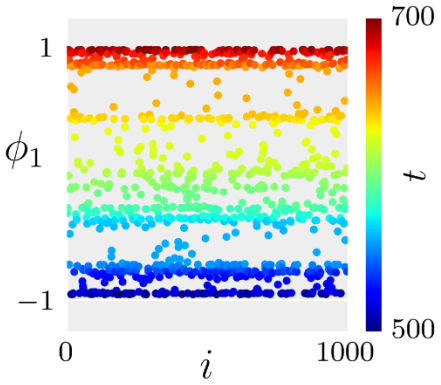
Finding the right Time

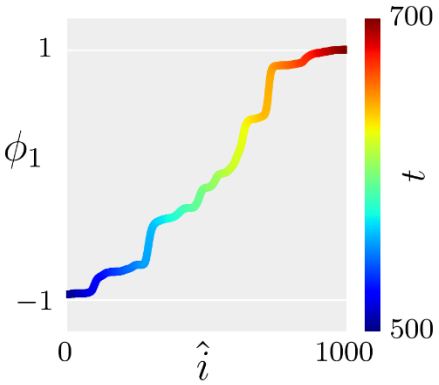




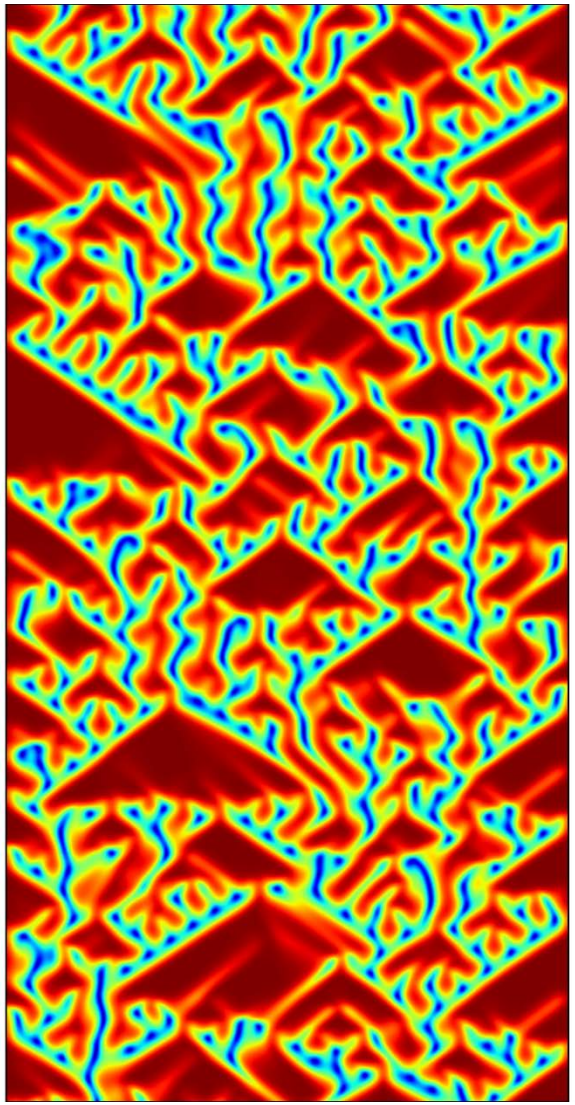




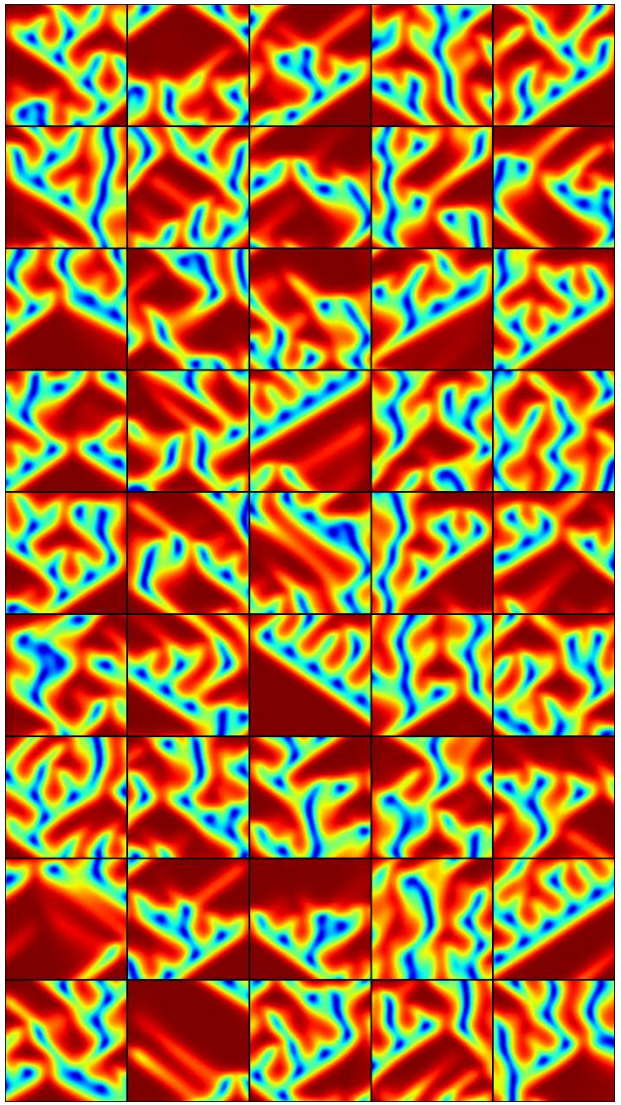




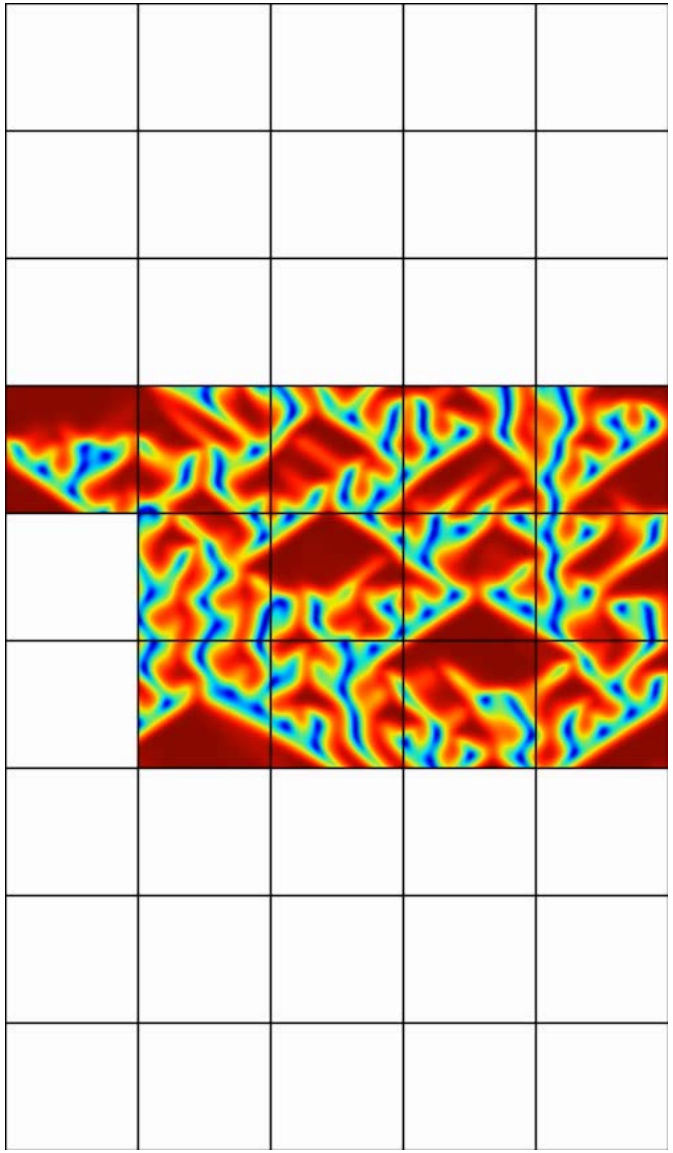
Finding the right Space and Time



Finding the right Space and Time



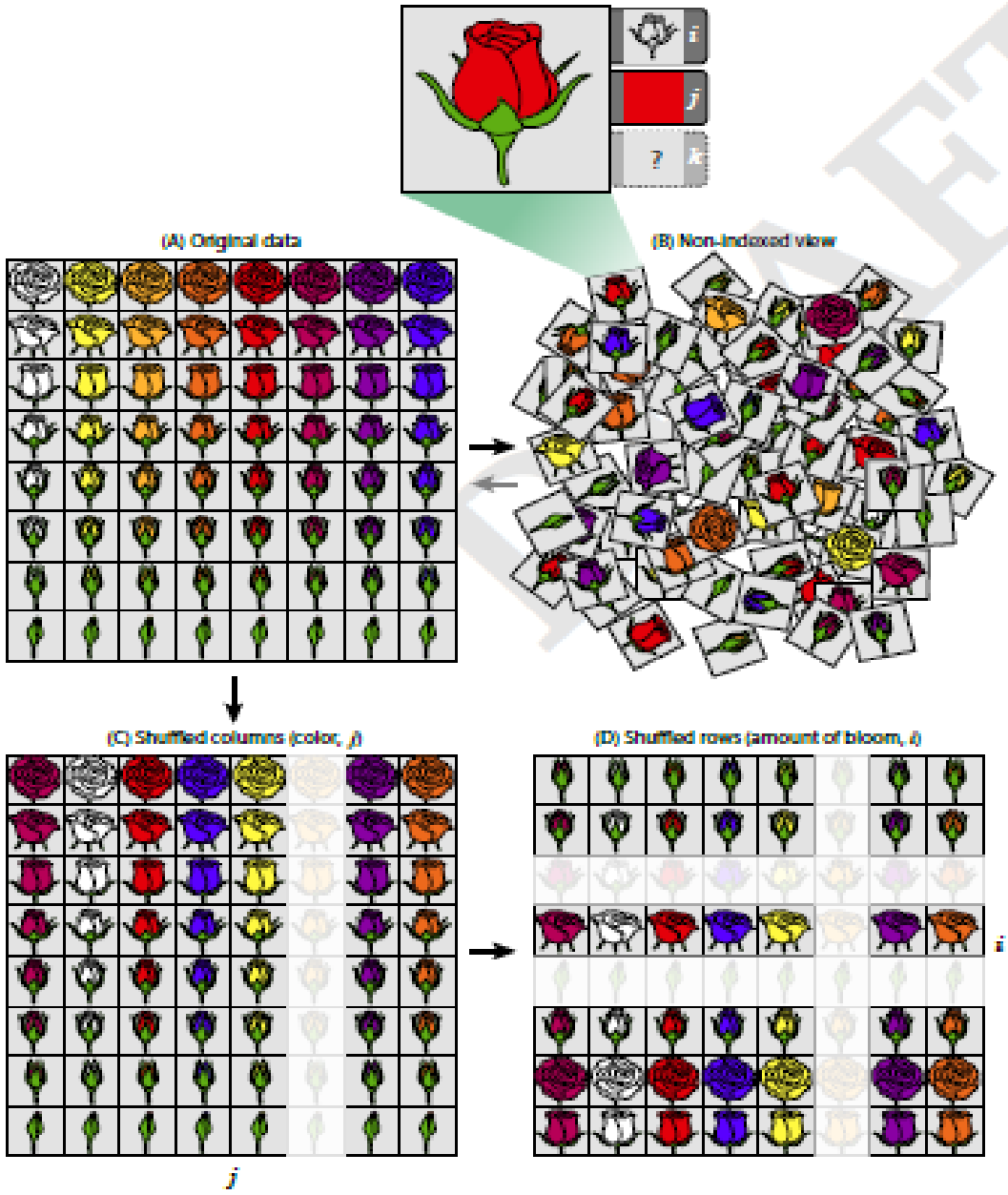
Finding the right Space and Time



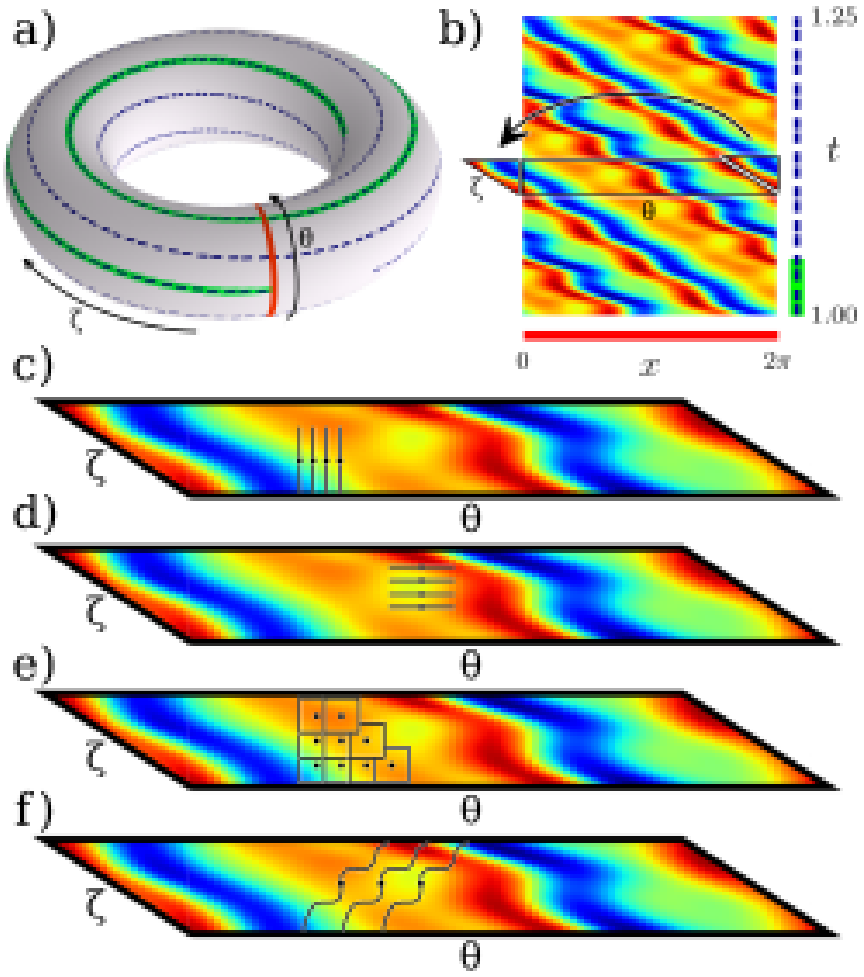
PRINCETON UNIVERSITY

“Shredded” data

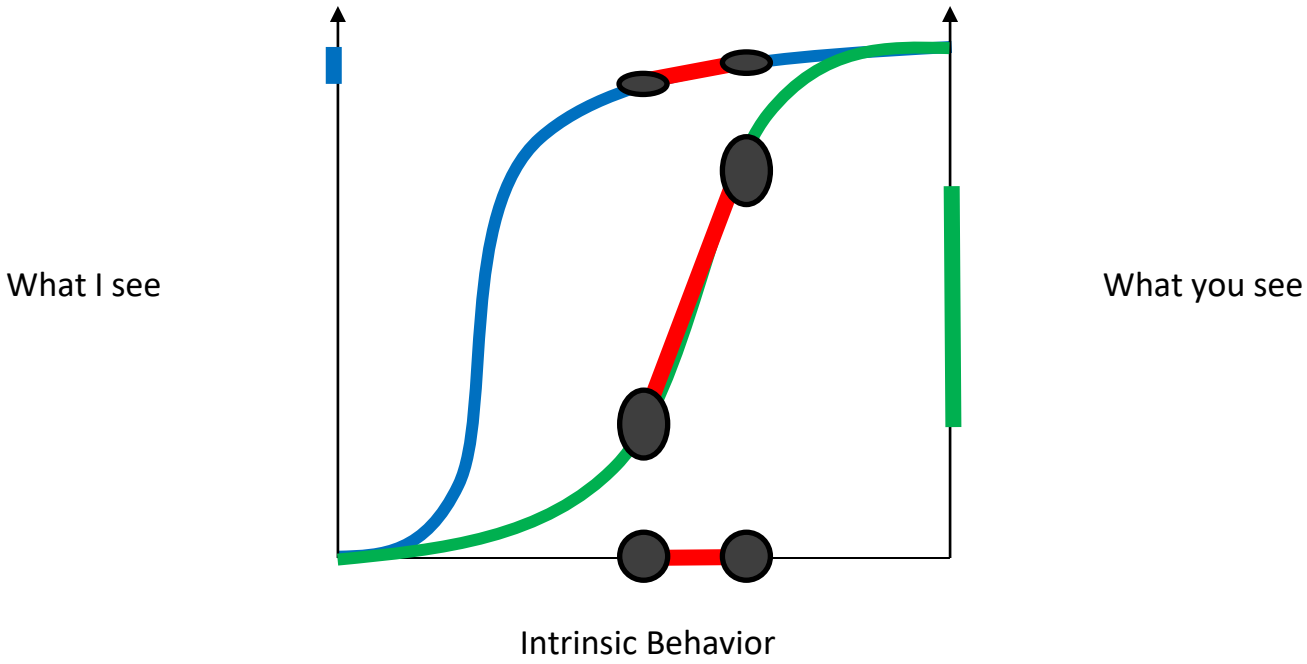
- Each element in the grid represents a measurement
- We don’t get measurements everywhere (faded)
- We don’t know the ordering of rows or columns, but we know who stays together



Different observables. Same data mining

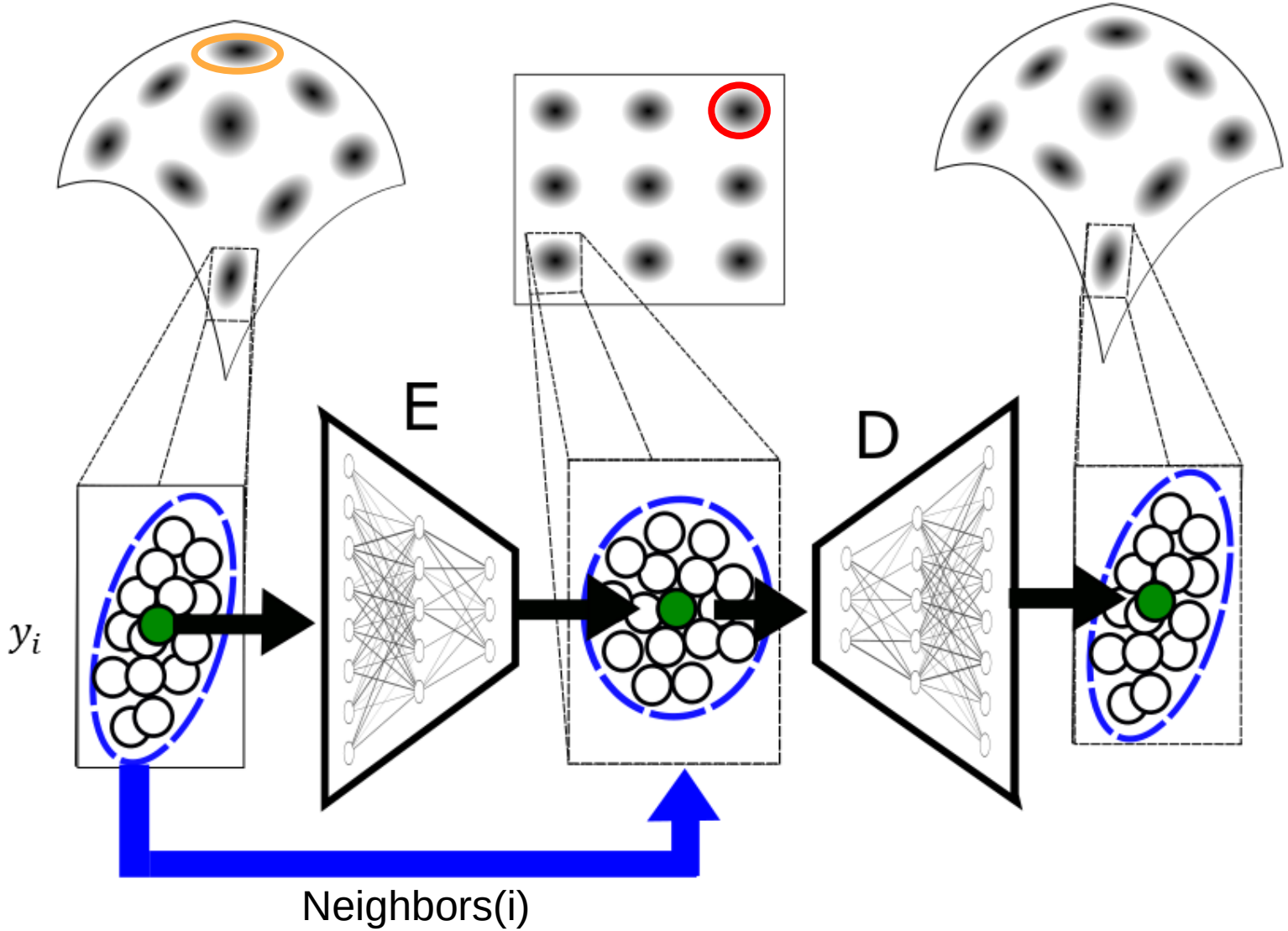


Combining data from different observations

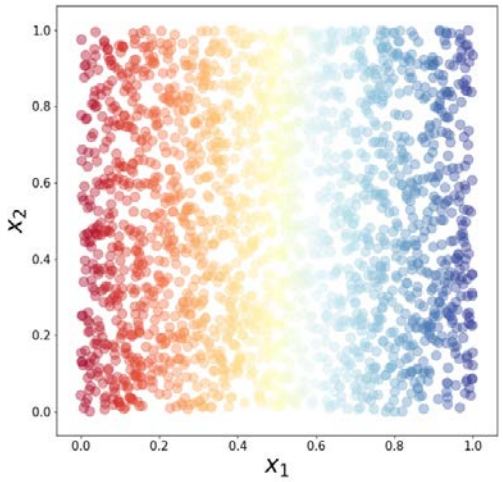


Distorted Data bursts (little ellipses) are observed and encoded as disks (little circles) to be decoded as the original points.

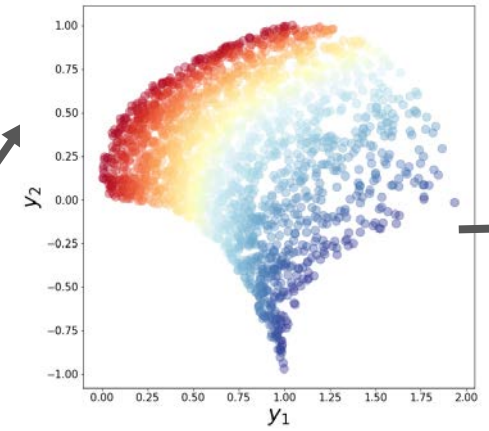
Had we observed the surface through some distorted lense we would have obtained the same central natural coordinates.

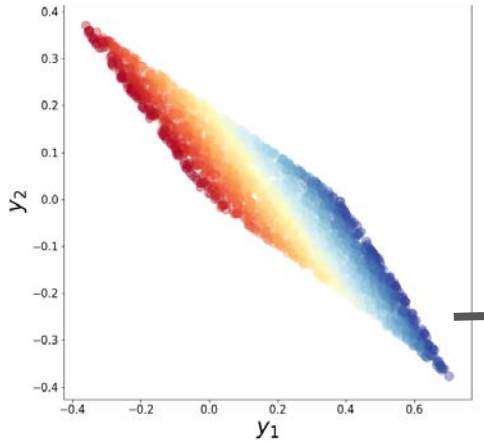


x

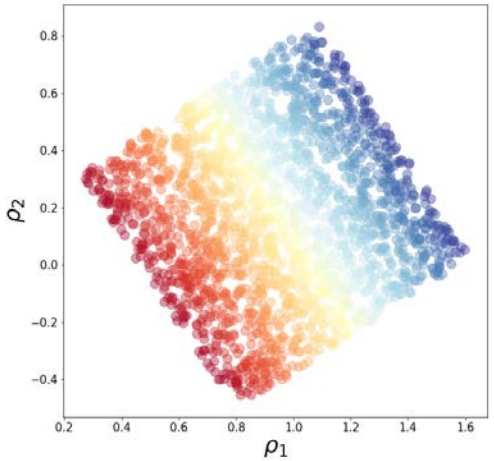


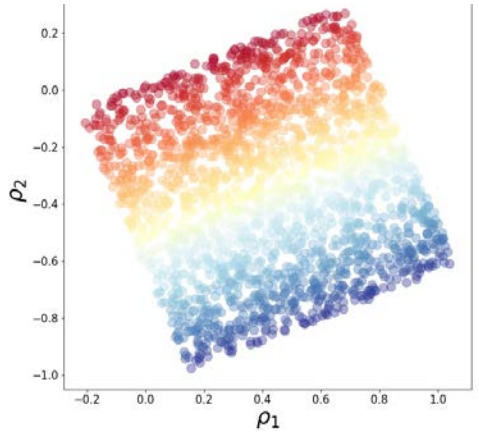
y



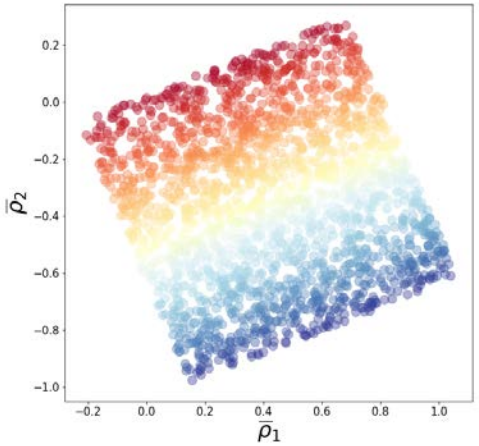


Loca:





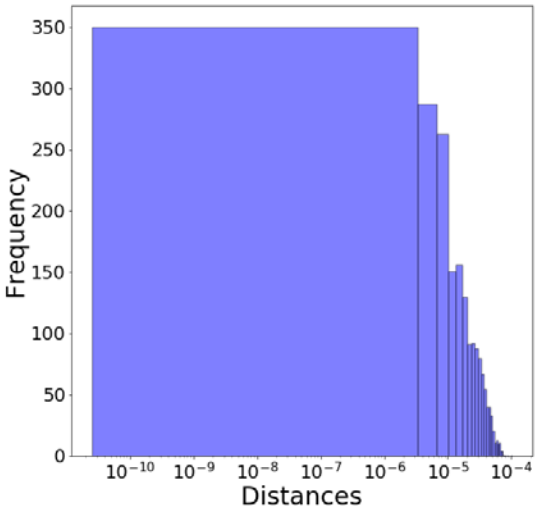
Shift + orthogonal tranformation:



Diamond transformation:

$$y_1 = \sin(2x_1 - x_2)$$
$$y_2 = \sin(x_2 - x_1)$$

L2 distance between the two embeddings (after alignment)

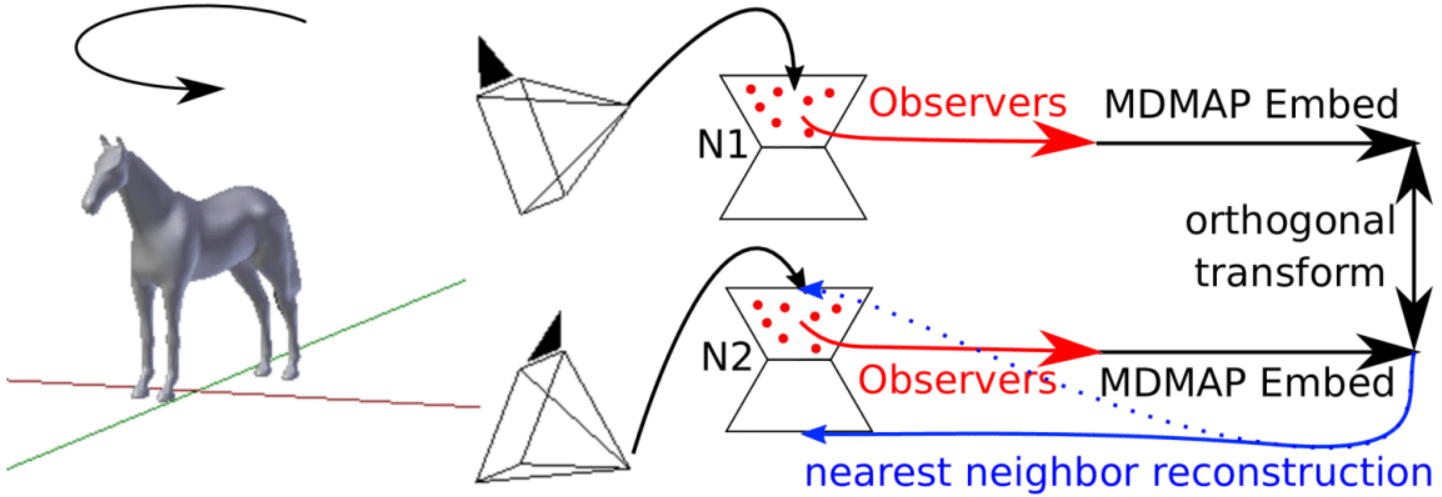


Matching Networks by Mahalanobis Embedding

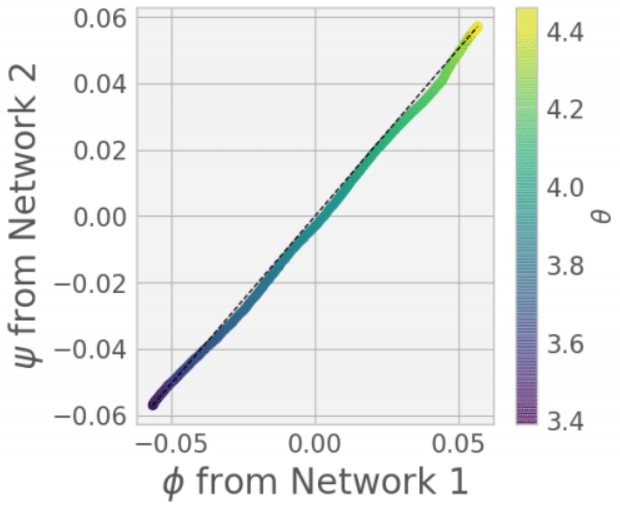
a) Two cameras observe a rotating horse. Tasks are to encode the images in six dimension
using two convolutional autoencoders N_1 and N_2.

b) MDMAP applied to selected activations yields embeddings invariant up to an orthogonal transformation (here, equality).

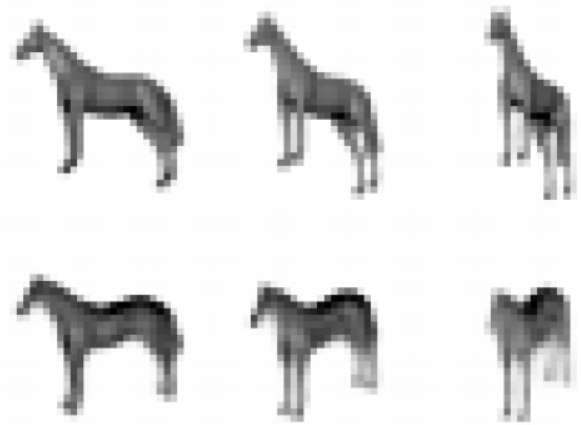
c) Inputs to N_1 and corresponding cross-reconstructions of decodings (predictions) from N_2.



(a)



(b)



(c)

TRS Yannis Kevrekidis, February 8, 2022

The Right Space

Since we cannot change reality, let us change the eyes with which we see reality

N. Kazantzakis

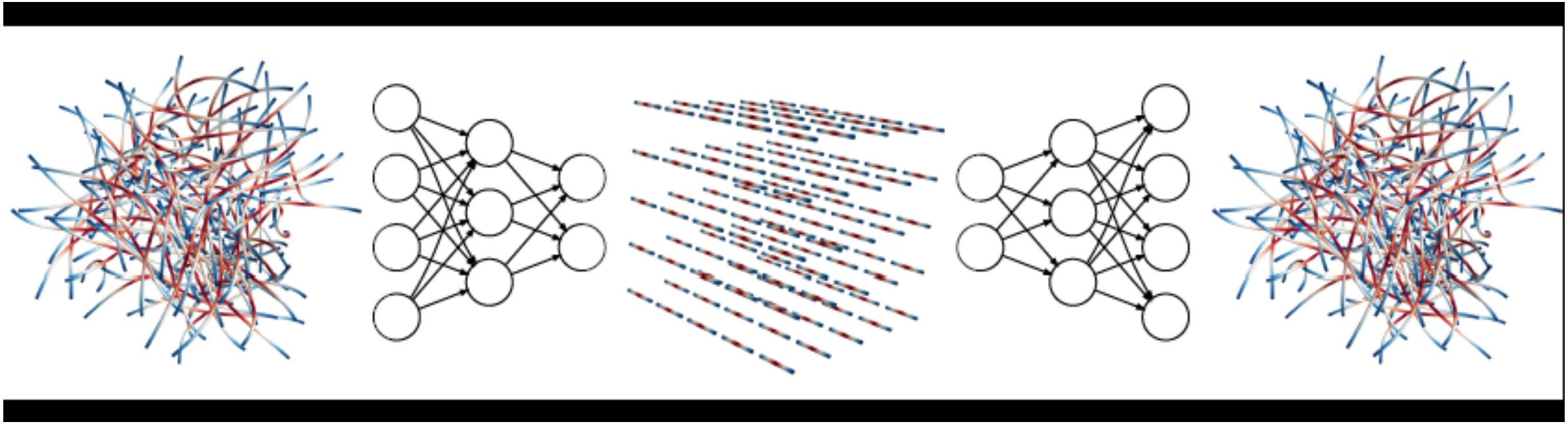
It's the wrong time and the wrong place; though your space is charming, it's the wrong space…

(with apologies to Cole Porter)

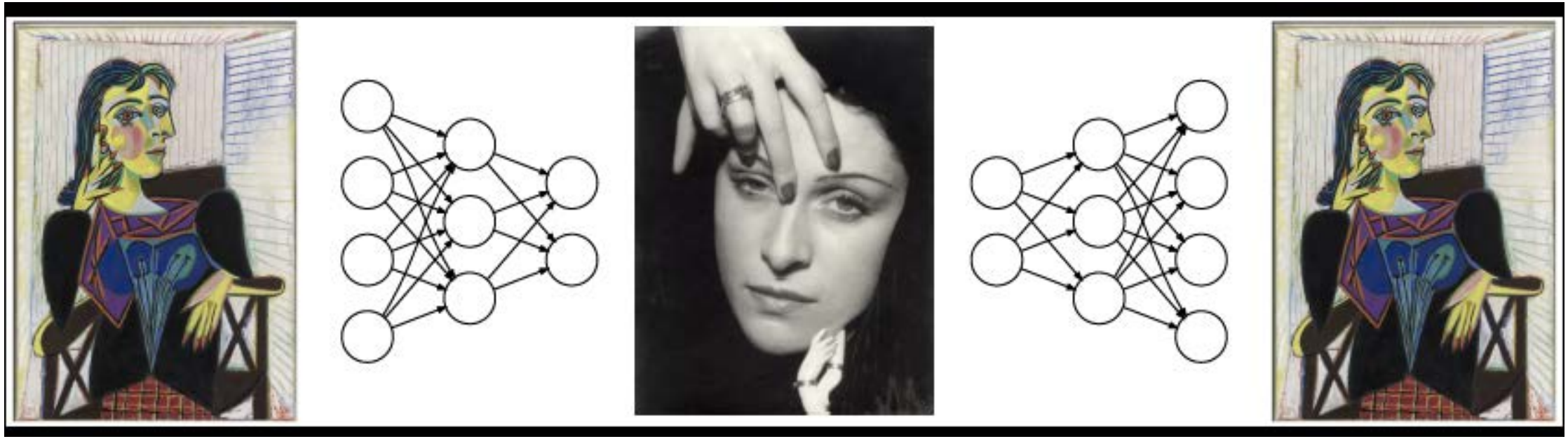
On ne voit bien qu’avec le cœur. L’essentiel est invisible pour les yeux

A. de Saint-Exupery

Rectifying ABC flow streamlines



Dora Maar (Picasso, Man Ray)



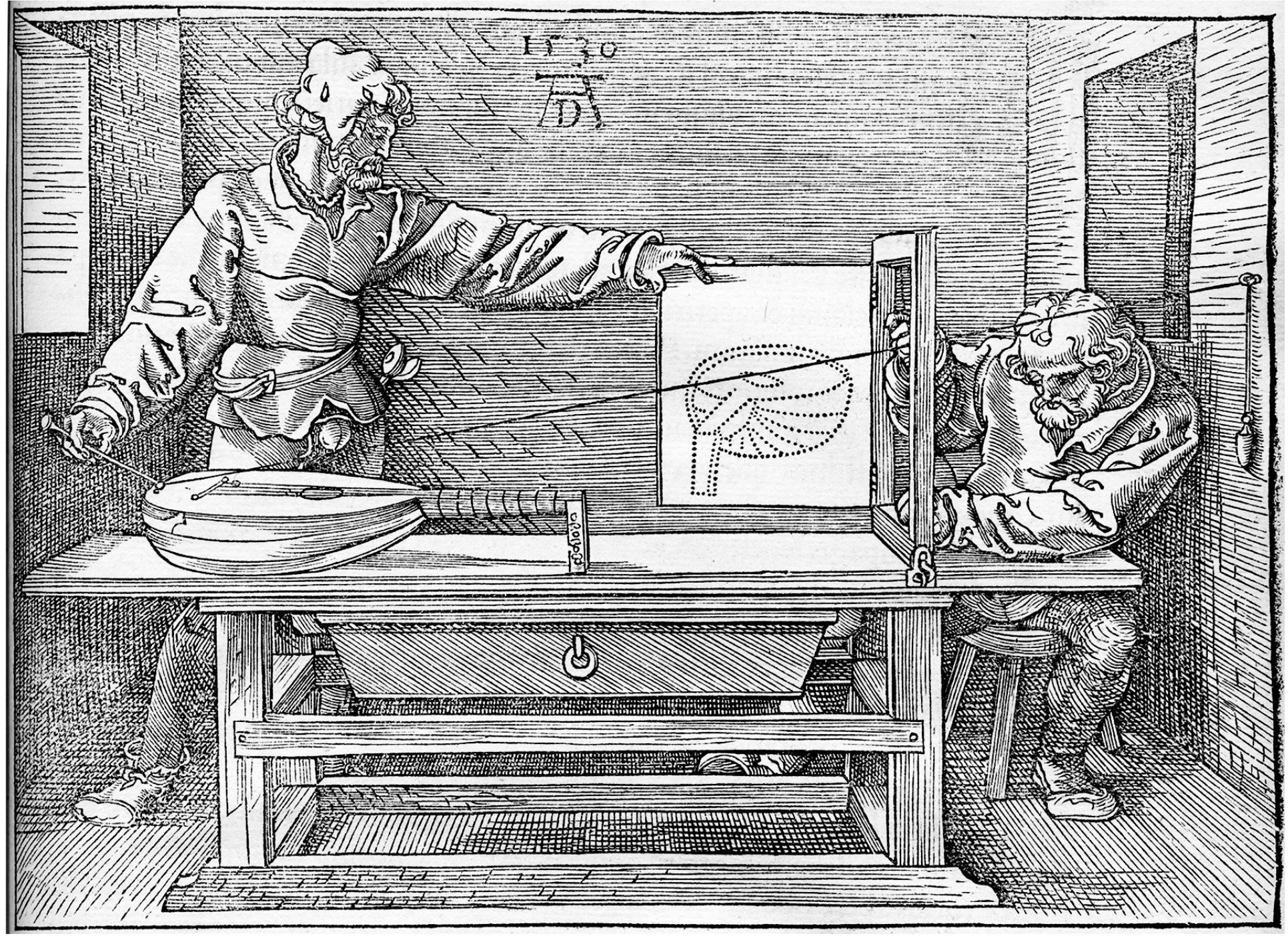
Isaac Newton Institute, Cambridge, June 2016 ROTHSCHILD LECTURE

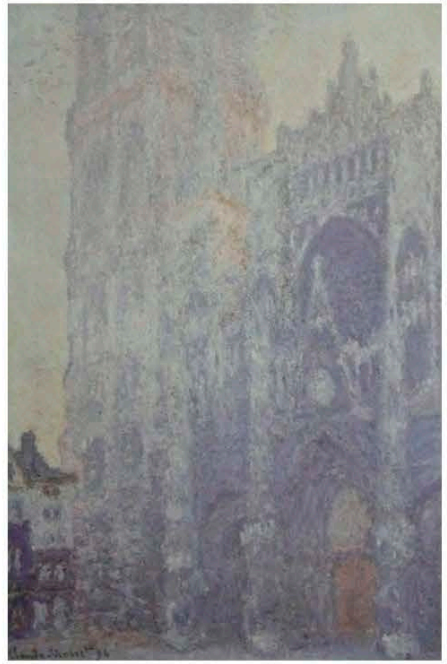
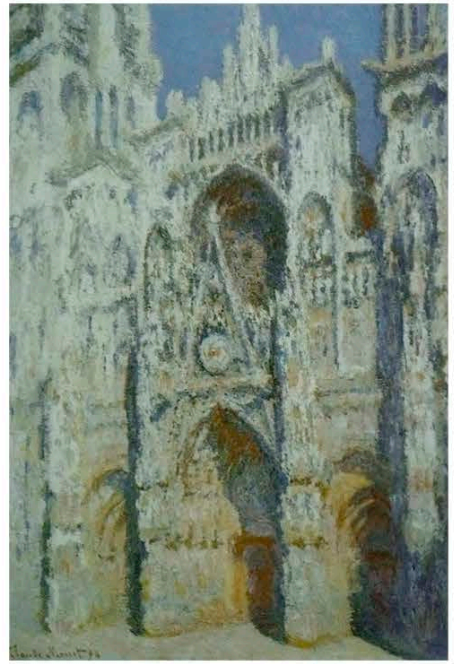
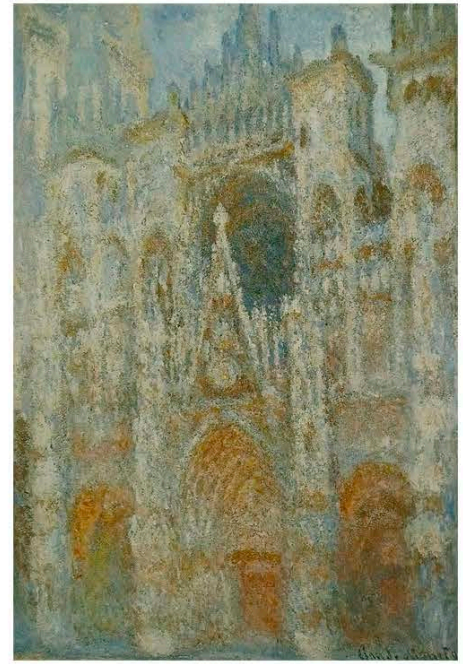
Mathematics for Data Driven Modeling
The Science of Crystal Balls

[Yannis Kevrekidis]

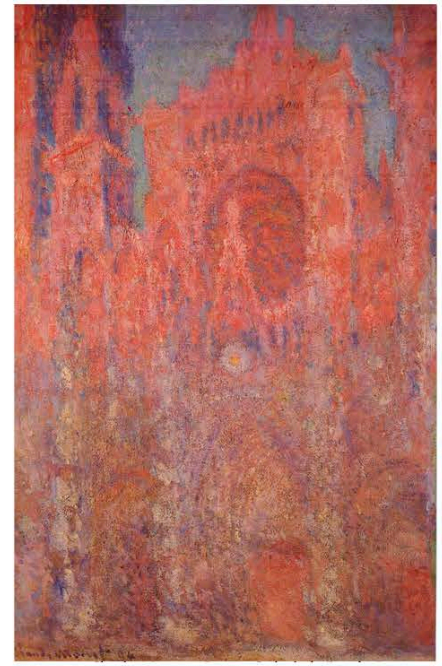
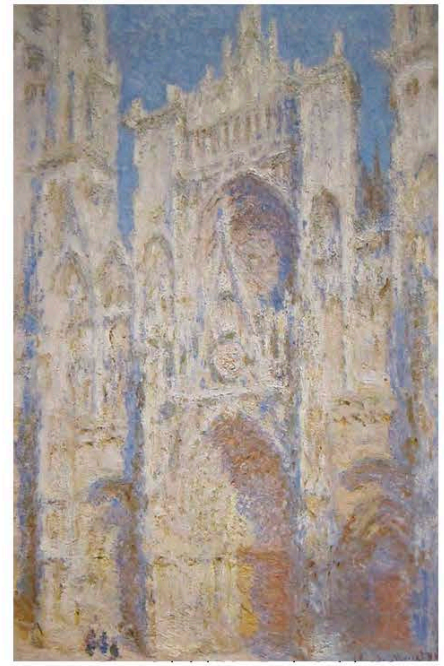
Departments of Chemical and Biomolecular Engineering &
Applied Mathematics and Statistics & Urology

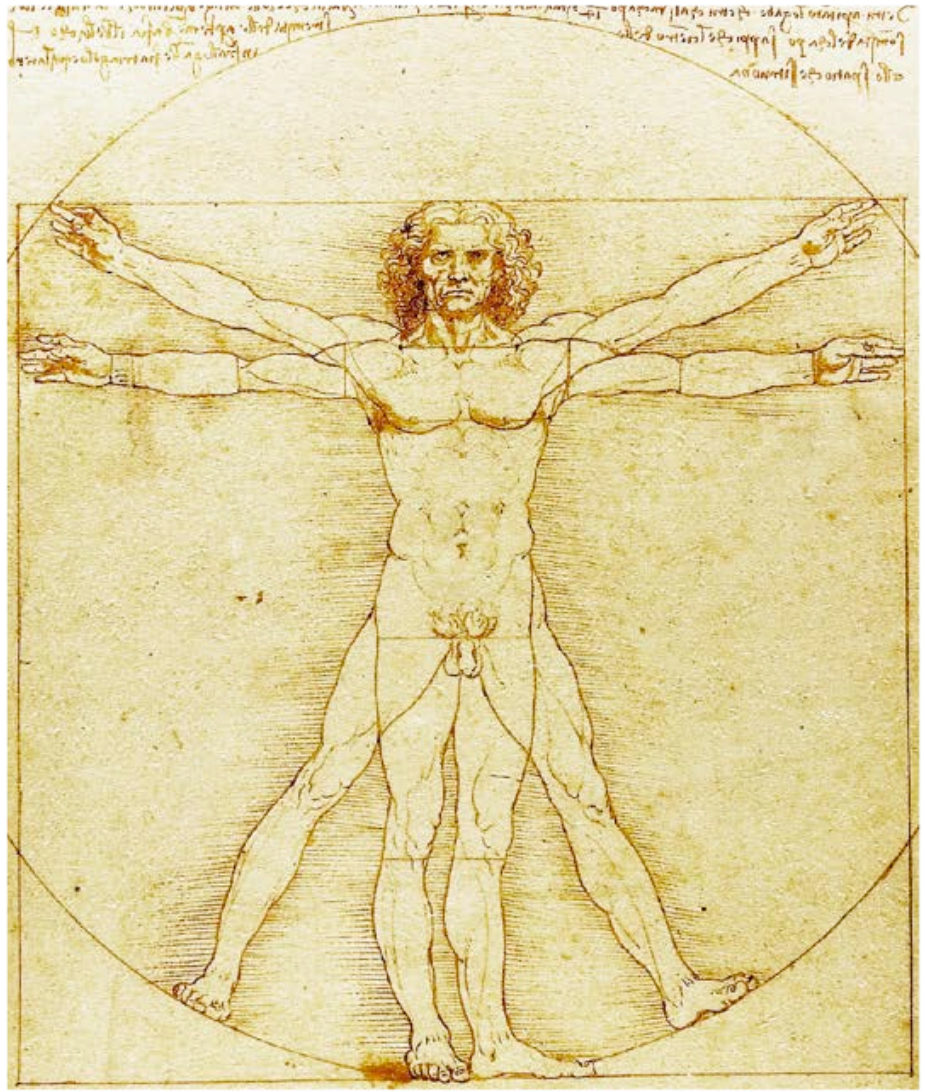
This year: EinsteinVisitor, ZIB/FU & Fischer Fellow, IAS-TUMuenchen

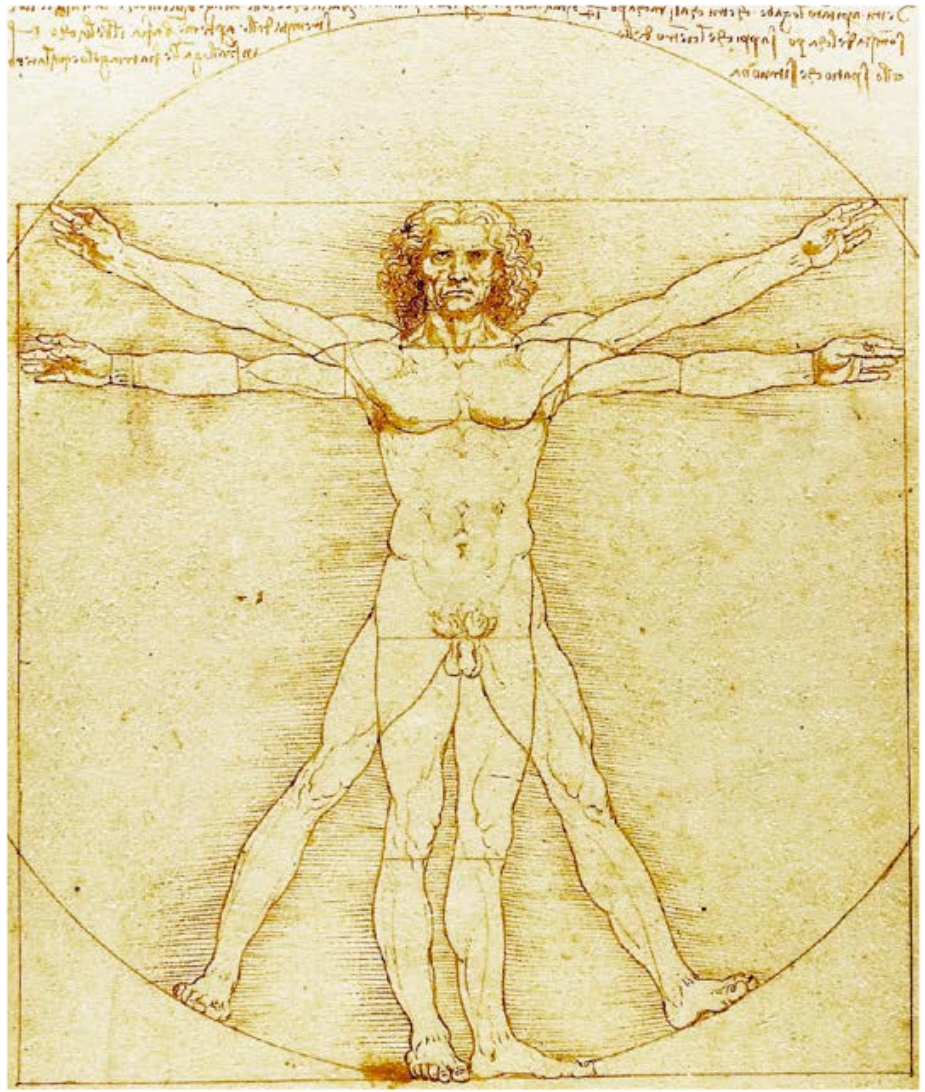

1530
AD

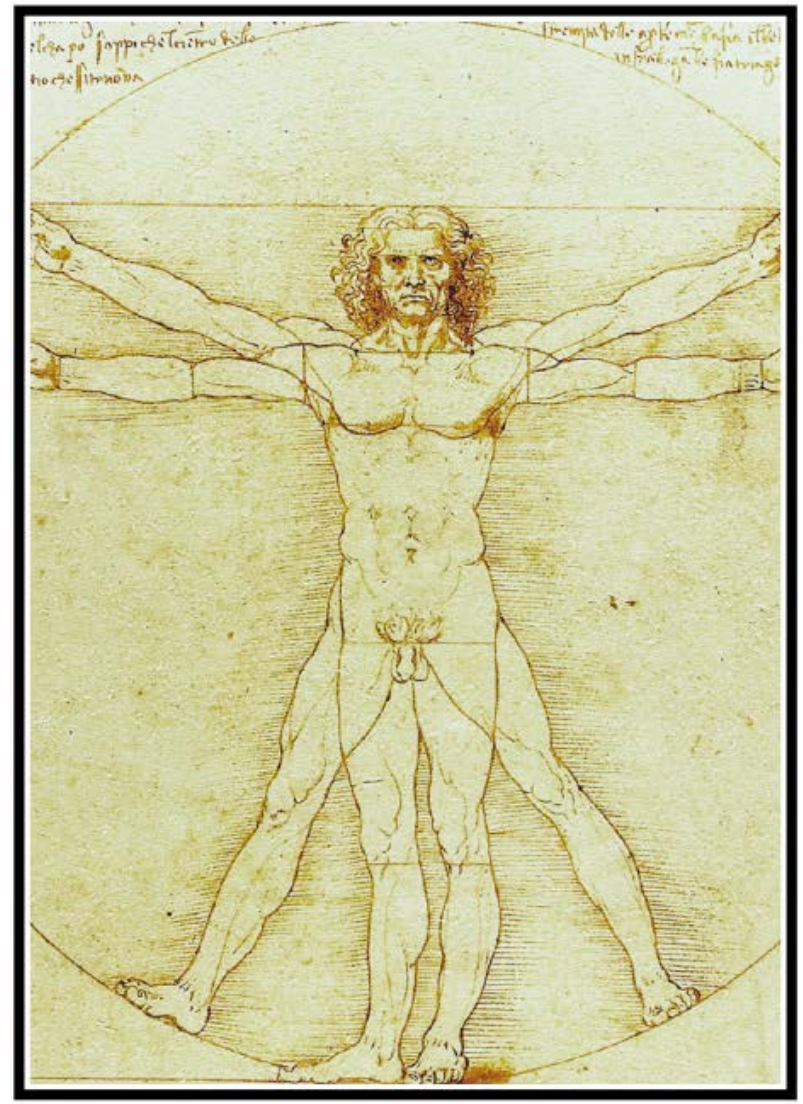




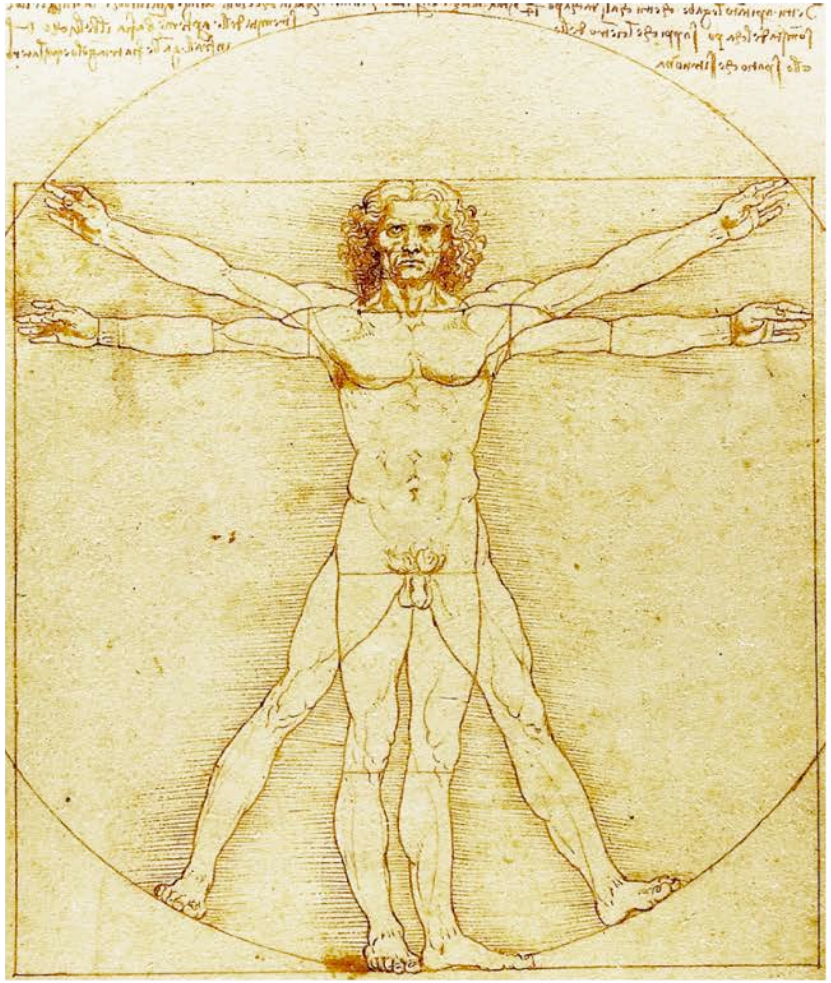


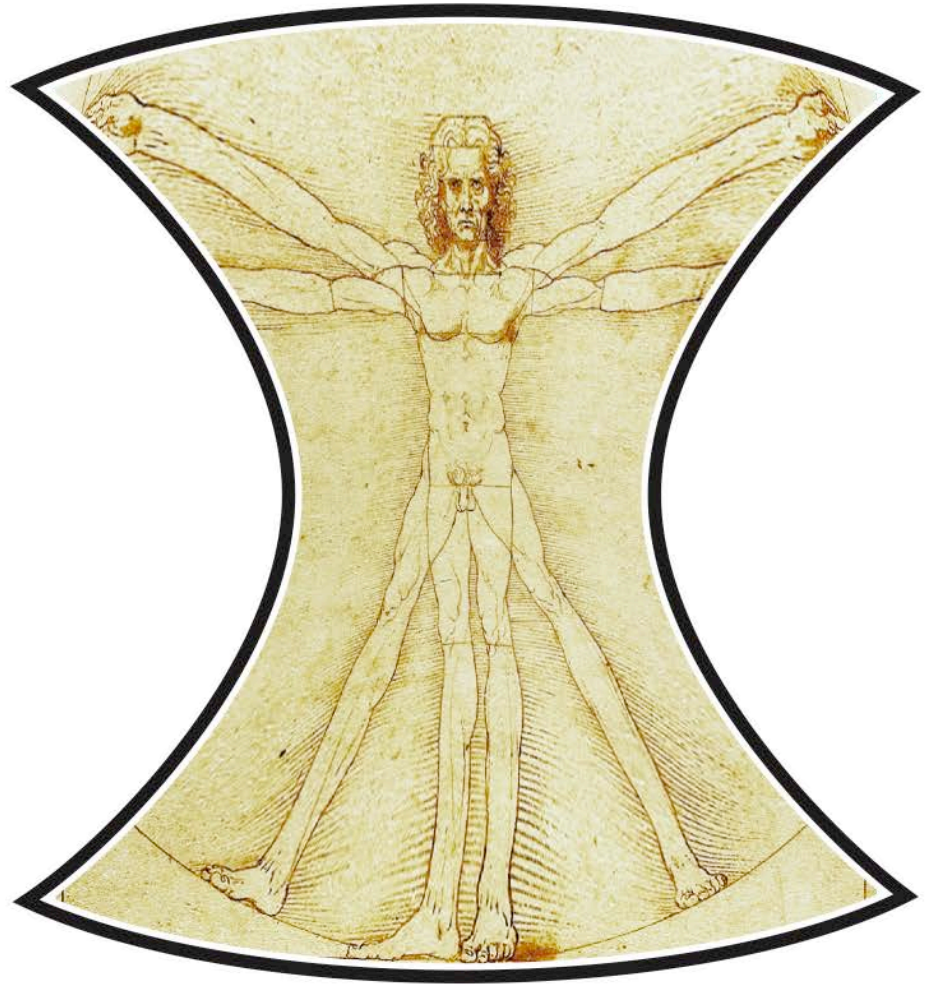



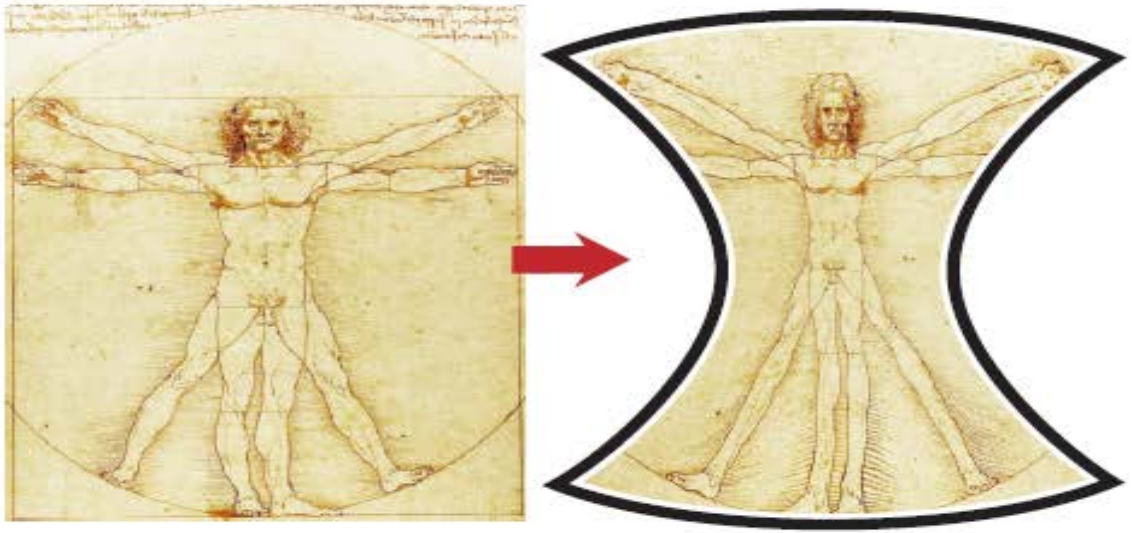


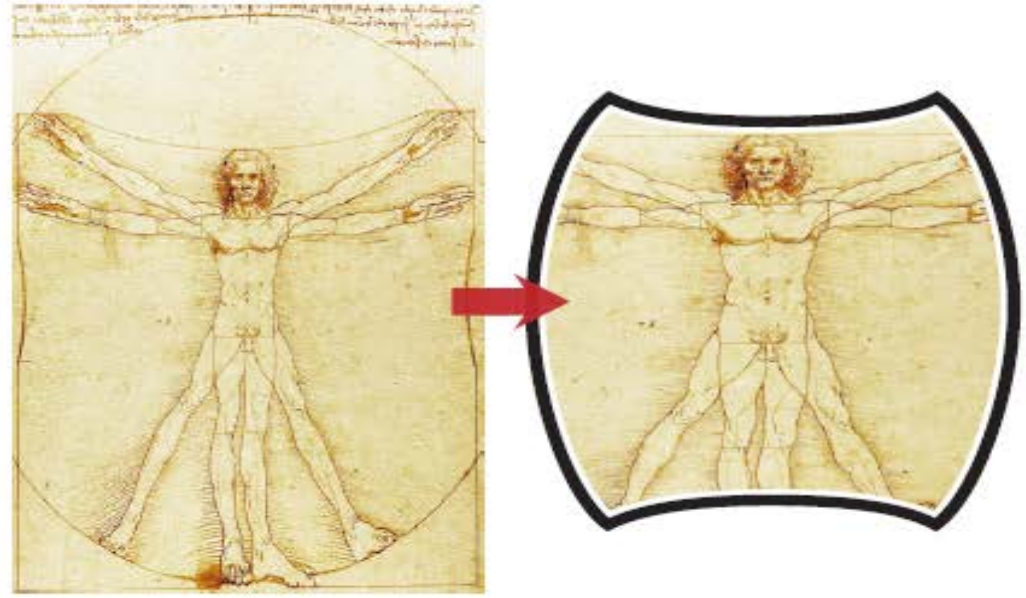


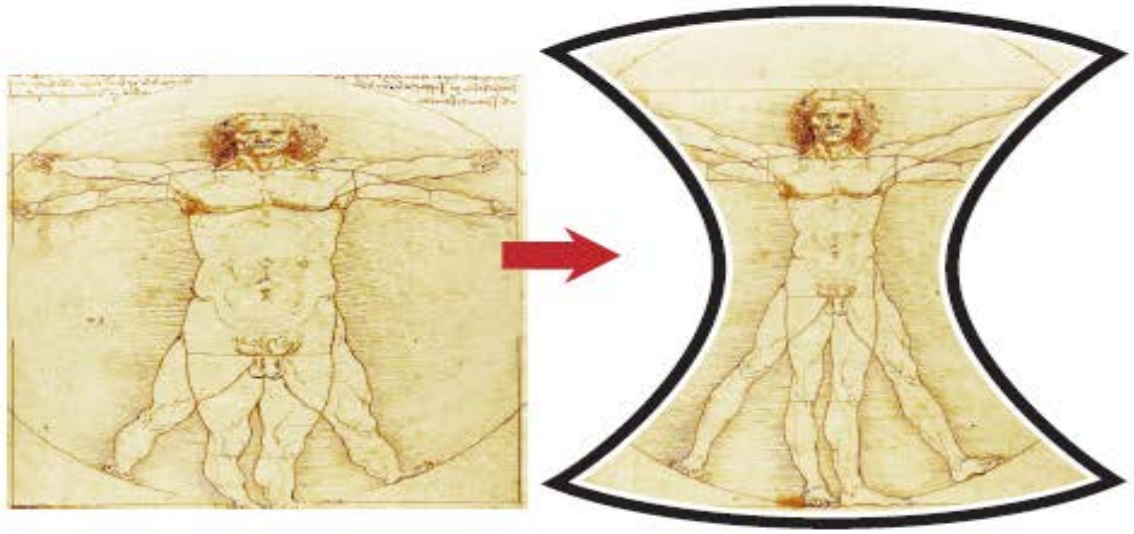


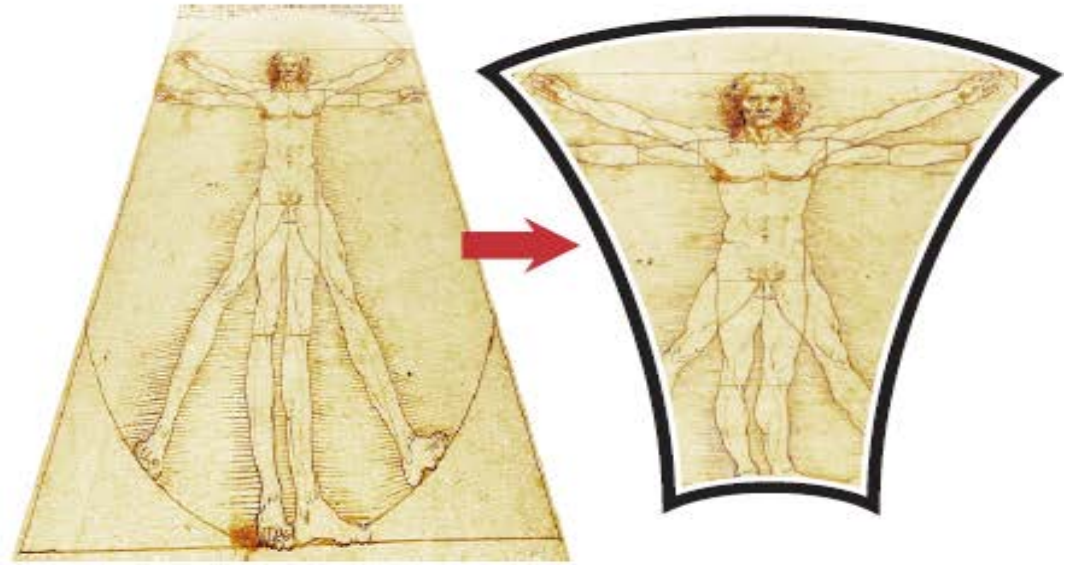


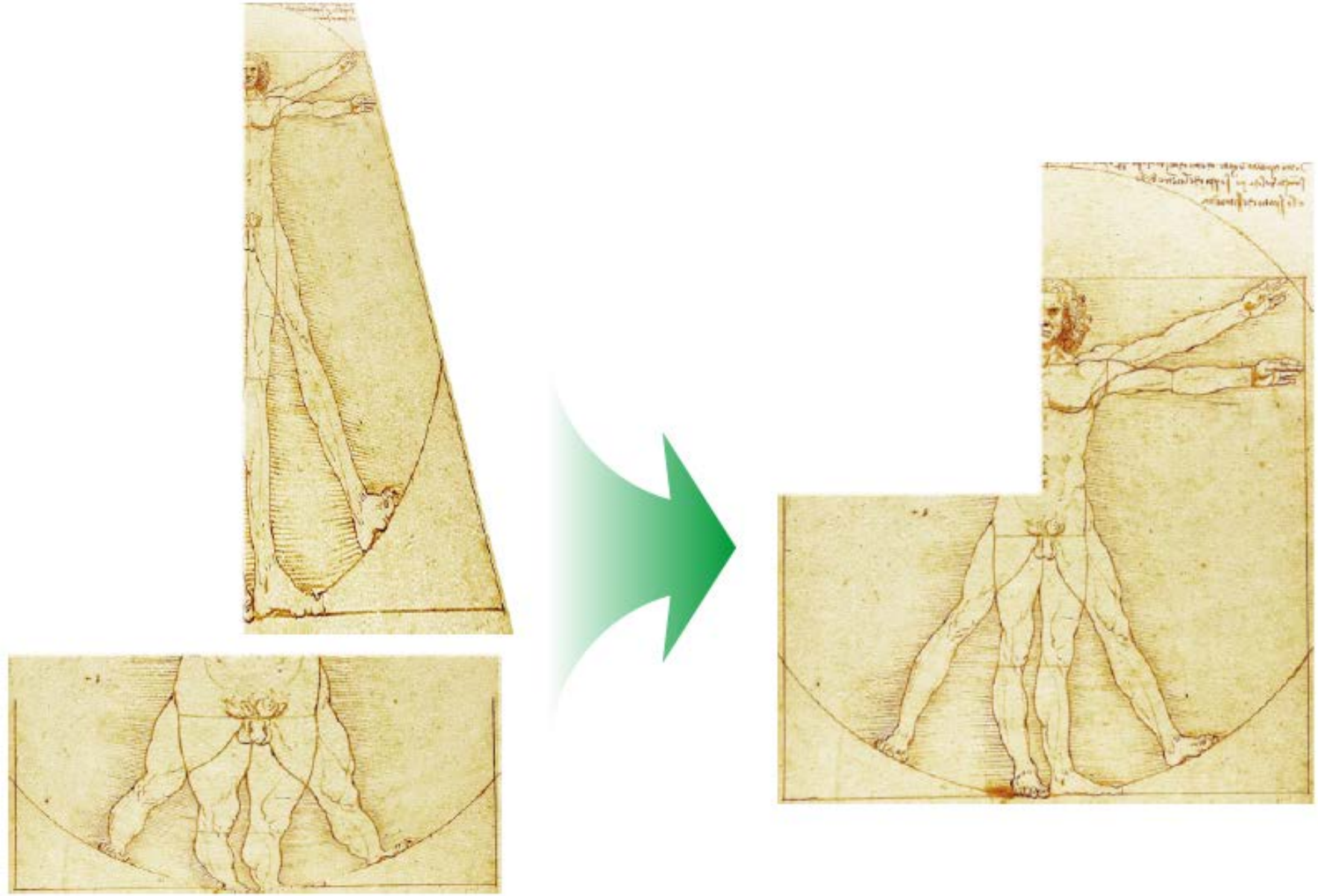






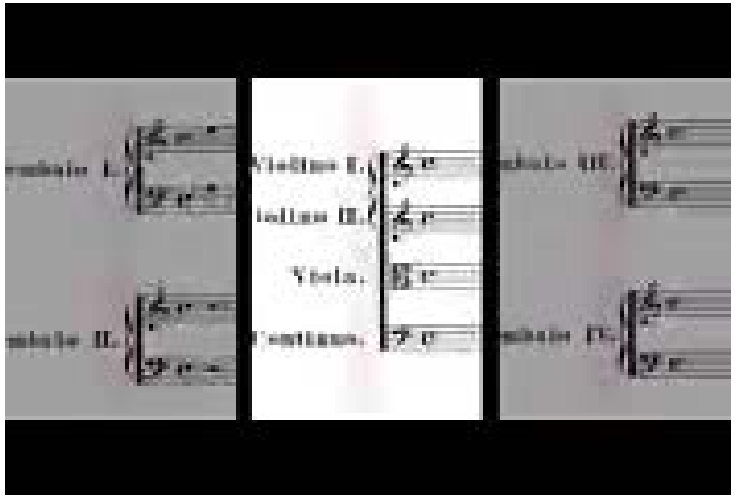






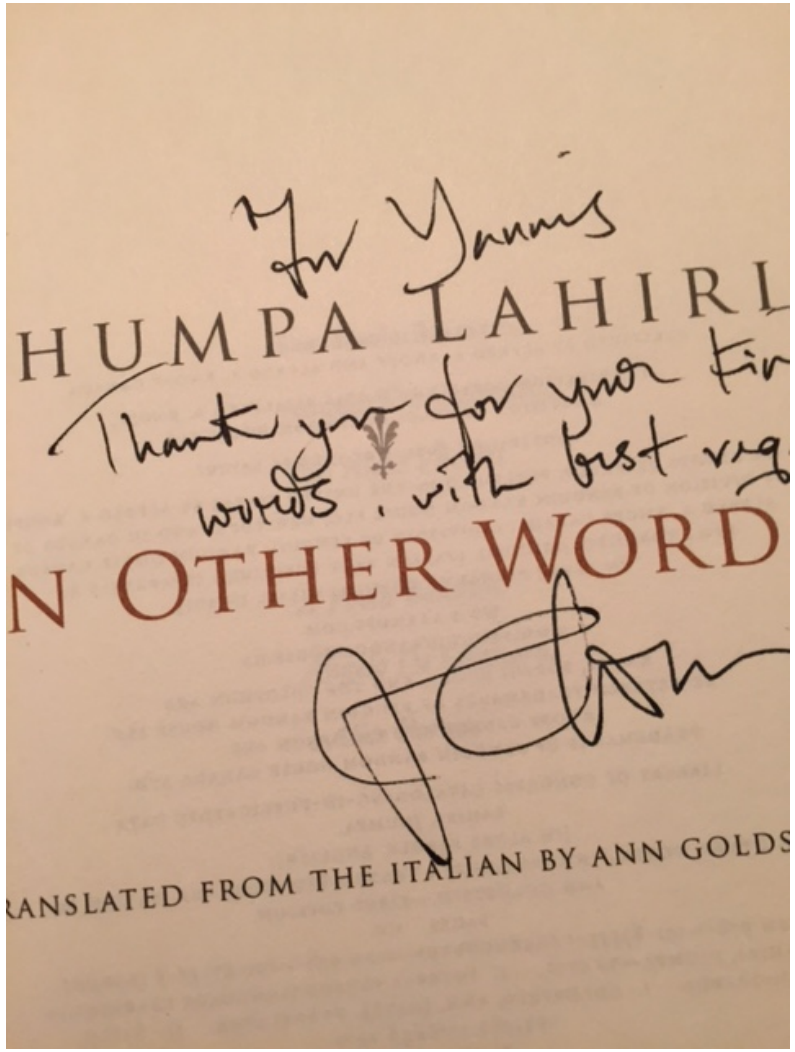
The Music of Gauge Invariance

BWV 1065 Concerto for Four Harpsichords and Strings



Vivaldi – Concerto for Four Violins in B Minor RV 580 – II Giardino Armonico

Gauge Invariance in Literature (in Princeton)



AFTERWORD

In 1939, fifteen years before he died, Henri Matisse began to move away from traditional painting and develop a new artistic technique. It involved cutting up pieces of paper that had been painted in gouache, in various colors. Matisse then combined and arranged the different pieces to create an image. He fixed the elements first with pins, then with paste, often directly on the wall. He stopped using the easel, the canvas. His main tool became a pair of scissors rather than the brush.

The method, a sort of synthesis of collage and mosaic, arose out of certain limitations. The eyesight of the seventy-year-old painter, which had greatly deteriorated, was one factor. Further, after a serious illness in 1941 he used a wheelchair, and was often forced to stay in bed. One day he was inspired to make a "garden" in the house, an exuberant jumble of leaves and fruit attached to the walls of his studio. It was a collective process: Matisse had his assistants paint the paper. He was no longer able to execute his works by himself.

The result was a distinctive form, a hybrid style, notably more abstract than his painting. He continued to play

with the same elements that he had always portrayed: nature, the human figure. But suddenly another energy emerged, a different language.

The images on paper were more simplified, crude compared to the ones on canvas, but they required painstaking, complex workmanship. One recognizes the hand and the eye of the painter, but they have changed. We follow the thread between the new method and the earlier paintings, and are aware of a turning point, a radical move.

For Matisse, cutting was not only a new technique but a system for thinking about and expanding the possibilities of shape, color, and composition. A rethinking of his artistic strategy. The painter said: "The conditions of this journey are a hundred percent different." He compared his method—which he called "painting with scissors"—to the experience of flying.

Matisse's new approach was at first received with distrust, with skepticism. One critic found it, at best, "a pleasant distraction." The artist, too, was unsure. Cutting, for Matisse, began as an exercise, an experiment. Without knowing what it meant, he followed an unknown path, exploring on an increasingly vast scale. In spite of the difficulties, this was a period of intense, fertile work. Gradually he embraced this method completely; it remained, until his death, a definitive step.

Last year, as I was finishing *In Other Words*, I saw a show, in London, devoted to Matisse's final creative stage. I encountered a series of lyrical, bold, wide-ranging images. I observed a surprising dialogue between negative and positive space. I understood how white space, like silence, can have a meaning.

I was struck by the essential effect of the images on paper. There is nothing superfluous. They show the seams, the cracks. Being literally cut into pieces, the images communicate a sort of deconstruction, an almost violent act of demolition. And yet they are harmonious, balanced. They express a new beginning. Every image, first cut out, then reconstructed, suggests something temporary, suspended, vulnerable. It evokes other permutations, other possibilities.

As I went through the show, I recognized an artist who at a certain point felt the need to change course, to express himself differently. Who had the mad impulse to abandon one type of vision, even a particular creative identity, for another. I thought of my writing in Italian: a similarly intricate process, a similarly rudimentary result compared with my work in English.

Writing in another language represents an act of demolition, a new beginning.

◆

In Other Words is the first book I've written directly in Italian. It originated in the fall of 2012, in a private, fragmented, spontaneous way. I had just moved to Rome, after spending almost my whole life in America. I spoke Italian, but my knowledge was elementary. I wanted to master