Teaspoon: A Comprehensive Python Package for Topological Signal Processing

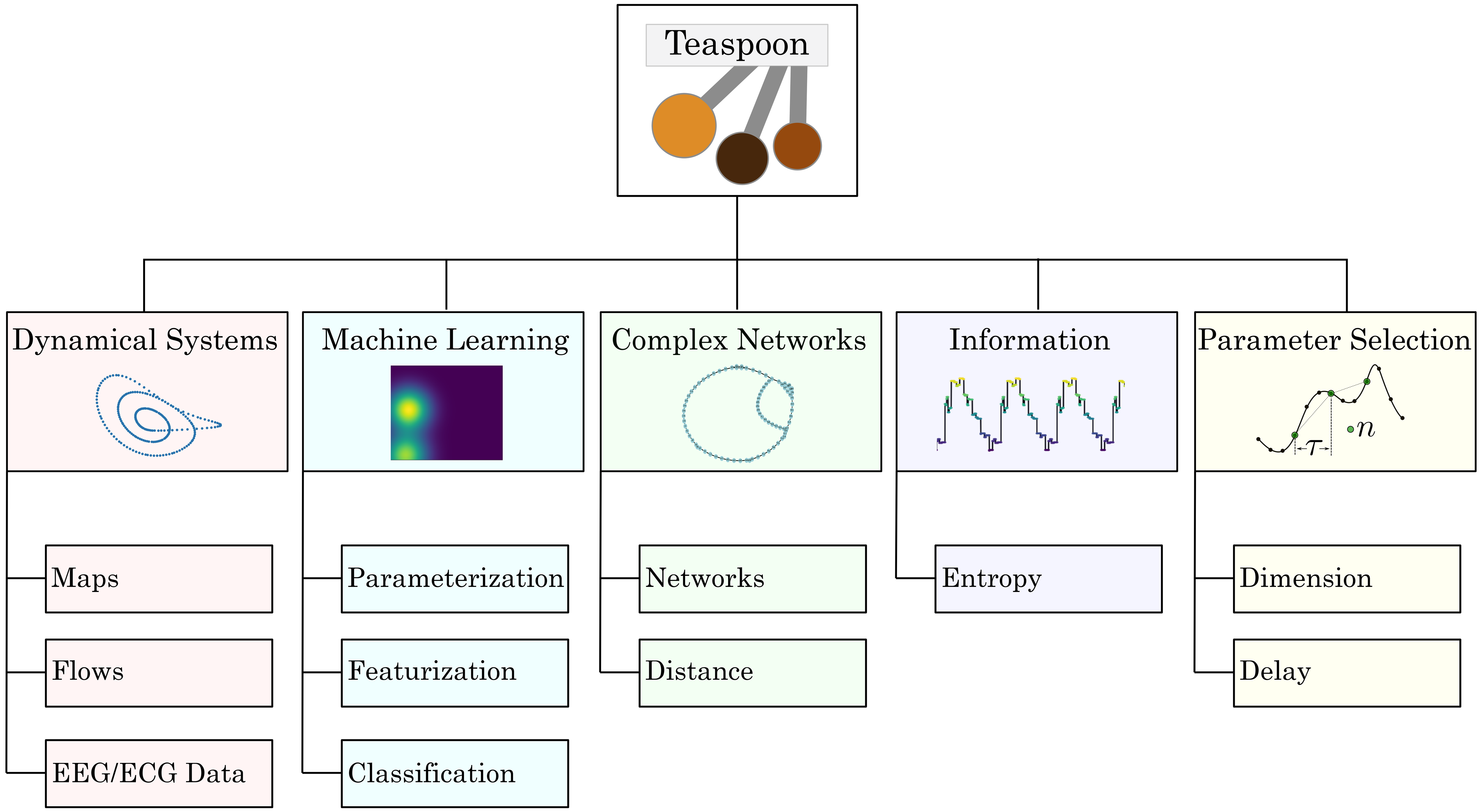
Audun D. Myers[3], Melih C. Yesilli[3], Sarah Tymochko[1], Firas A. Khasawneh[3], Elizabeth Munch[1,2]

Michigan State University, [1]Dept. of Computational Mathematics, Science, and Engineering, [2]Dept. of Mathematics, [3]Dept. of Mechanical Engineering

CMMI-1759823, DMS-1759824, CMMI-1800466, DMS-1800446

In this work, we present an overview of the python package `teaspoon`, which brings together available software for computing persistent homology, the main workhorse of TDA, with modules that expand the functionality of `teaspoon` as a state-of-the-art topological signal processing tool.

Documentation: `http://elizabethmunch.com/code/teaspoon/index.html`
Code: `https://github.com/lizliz/teaspoon`
Installation: `pip install teaspoon`



Dynamical Systems Library

The dynamical systems library provides 60 dynamical systems to simulate. These systems are separated into three categories: (1) flows, (2) maps, and (3) ECG/EEG data. Please see the documentation page for more details.

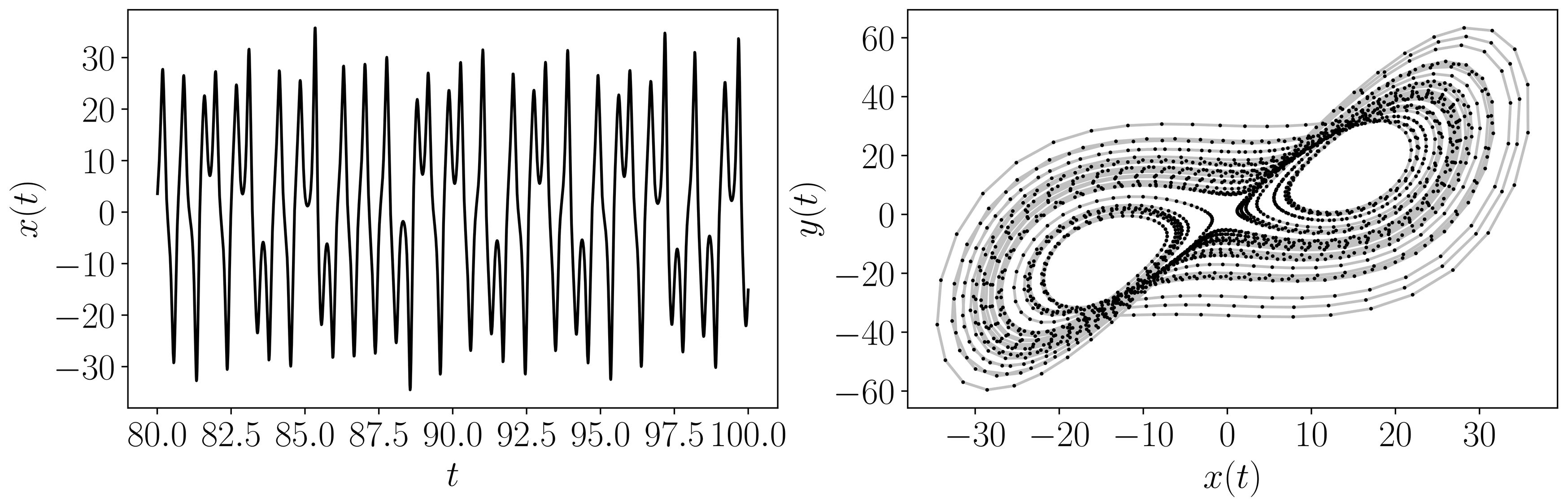


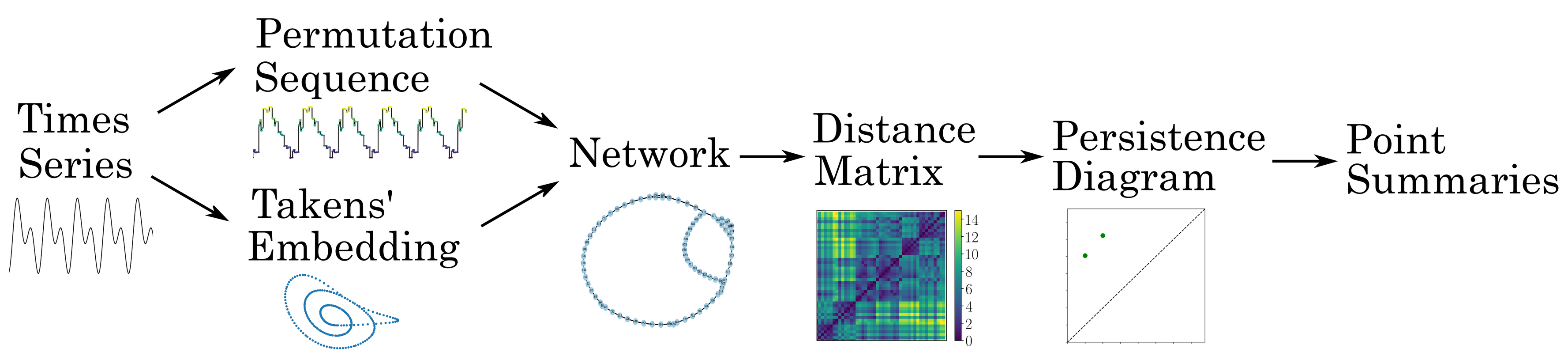
Figure: Default simulation of the chaotic Lorenz system using the dynamical systems library module.

Machine Learning Module

Machine learning module provides automated feature matrix generation and classification, and it is suitable for the applications where persistence diagrams can be computed. This module involves template function, persistence landscapes, persistence images, Carlsson coordinates, persistence paths and signature and multi-scale kernel method as featurization methods and it is capable of performing supervised classification.

Complex Networks

The complex networks module provides methods for analyzing signals using complex network representations (ordinal partition or k-nearest neighbor networks) and then providing a distance matrix representation of the network to be used with persistent homology (see pipeline below).



Information Theory Module

The following functions are currently available in the information theory module.

- **Permutation Entropy:** a time series complexity measure using the information (Shannon) entropy of the ordinal patterns within the signal.

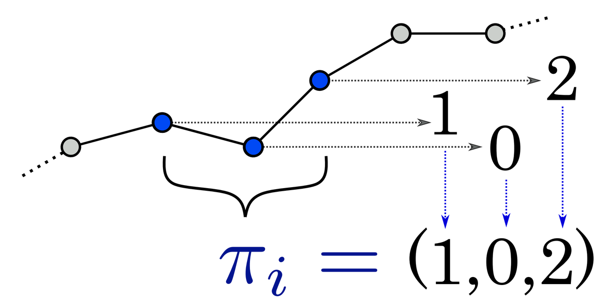


Figure: Example ordinal pattern of dimension $n = 3$.

- **Multiple-scale Permutation Entropy:** An implementation of permutation entropy over multiple time scales.
- **Persistent Entropy:** A method for calculating the complexity of a persistence diagram using a unique implementation of Shannon entropy and probabilities based on the persistence pairs.

Parameter Selection Module

The parameter selection module provides a wide variety of algorithms for automatically calculating the dimension n and delay τ parameters for both permutation entropy and Takens' embedding (state space reconstruction). Specifically, given a time series, $[x_1, \dots, x_N]$, a choice of dimension n and lag τ, the delay embedding is the point cloud, $\{\mathbf{x}_i := (x_i, x_{i+\tau}, \dots, x_{i+(n-1)\tau})\} \subset \mathbb{R}^n$. with an example shown below.

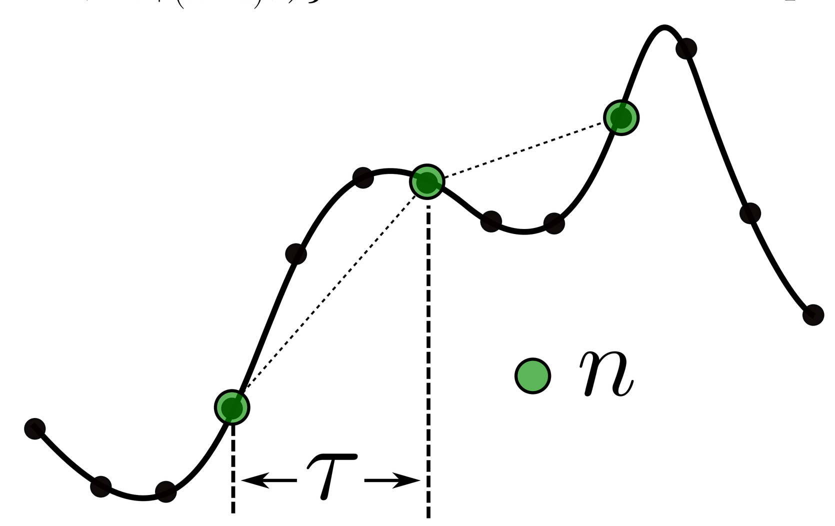


Figure: Example demonstrating the embedding dimension n and delay τ are used to create point cloud vectors, which can be used for permutations or state space reconstruction.