Topological Data Analysis and Industrial Automation

Stefan Huber

Josef Ressel Center for Intelligent and Secure Industrial Automation
Department for Information Technologies and Digitalisation
Salzburg University of Applied Sciences
Austria

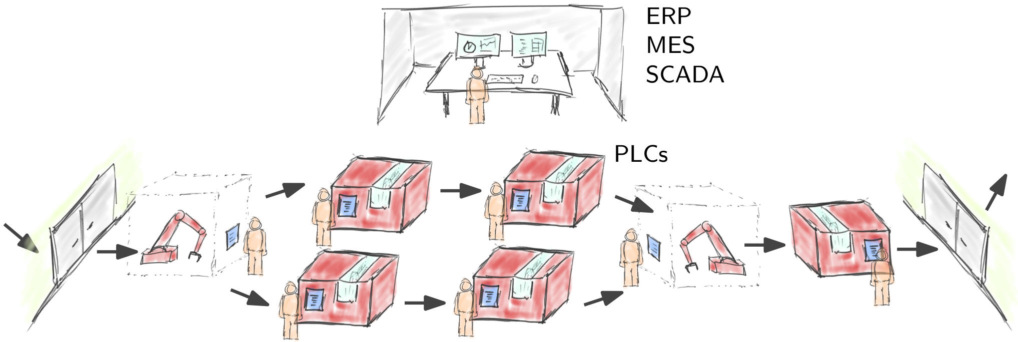
Mar 5, 2025

Salzburg University
of Applied Sciences

The automation industry

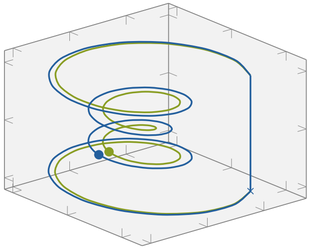
Strict hierarchical structure

- At the shop floor level industrial machines optimized for repetitive tasks.



Discrete automation and state space trajectories

- Discrete automation is about repetitive processes
- Let us consider a "state vector" of an industrial machine
- Intuition: Repetitive processes lead to cyclic state space trajectories
- Record them and form a model of their distribution, or a model of the underlying subspace, for the "benign" trajectories.



Anomaly detection

Detecting trajectories leaving benign subspace, for machine operation and security

Honey Pots

Generative methods for artificial trajectories. (Or for simulation purposes.)

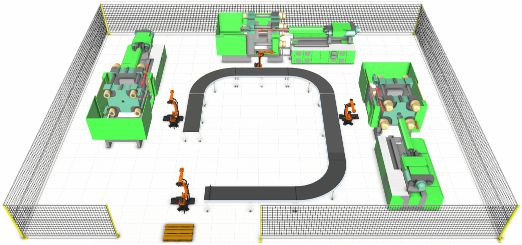
Process control

Learning the control task of keeping trajectories within benign subspace

→ Intuitively, there should be plenty of potential where TDA is useful

Shop floor:

- Three injection molding machines (IMMs)
- Three + one robot
- One conveyer



Testbed implementation:

- Real-world components: PLCs, Automation PCs, HMIs, supervisory control
- Simulation of the physical production processes

TDA applications in manufacturing

TDA promotion and survey paper for manufacturing

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

ELSEVIER

Review

Topological Data Analysis in smart manufacturing: State of the art and future directions

Martin Uray [a,b,*], Barbara Giunti [c], Michael Kerber [d], Stefan Huber [a]

[a] *Josef Ressel Centre for Intelligent and Secure Industrial Automation, Salzburg University of Applied Sciences, Urstein Sued 1, 5412, Puch/Hallein, Austria*
[b] *Department of Artificial Intelligence and Human Interfaces, Paris-Lodron-University of Salzburg, Jakob-Haringer-Str. 2, 5020, Salzburg, Austria*
[c] *Department of Mathematics and Statistics, University at Albany, 1400 Washington Avenue, Albany, 12222, NY, USA*
[d] *Institute of Geometry, Graz University of Technology, Kopernikusgasse 24, 8010, Graz, Austria*

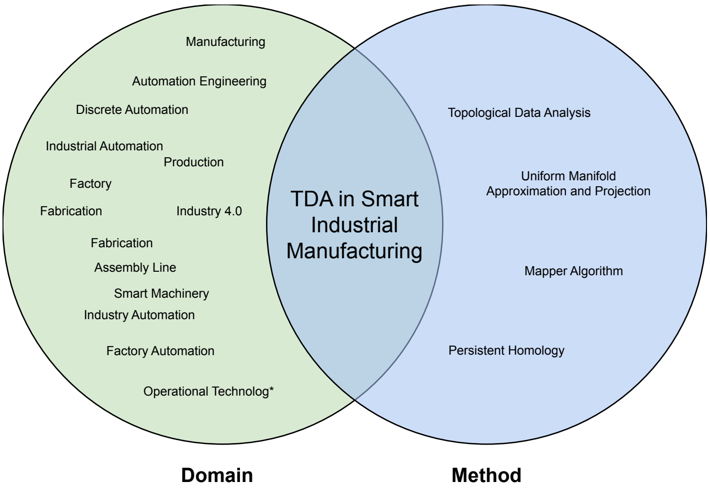
ARTICLE INFO

Keywords:
Industry 4.0
Manufacturing
Topological Data Analysis
Persistent Homology
UMAP
Mapper

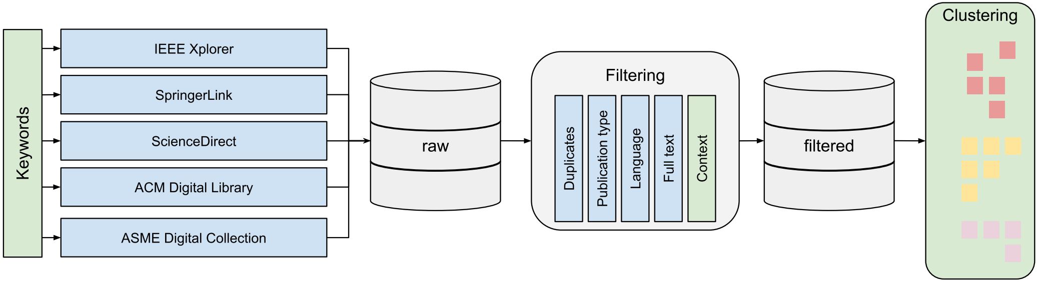
ABSTRACT

Topological Data Analysis (TDA) is a discipline that applies algebraic topology techniques to analyze complex, multi-dimensional data. Although it is a relatively new field, TDA has been widely and successfully applied across various domains, such as medicine, materials science, and biology. This survey provides an overview of the state of the art of TDA within a dynamic and promising application area: industrial manufacturing and production, particularly within the Industry 4.0 context. We have conducted a rigorous and reproducible literature search focusing on TDA applications in industrial production and manufacturing settings. The identified works are categorized based on their application areas within the manufacturing process and the

Exhaustive literature review methodology



Exhaustive literature review methodology



- Passed search phrases of the form "domain AND method" to the search engines.
- 7000 results received on Apr 23, 2024.
- Filtering: Only peer-reviewed conferences and journals, only English, only with fulltext availability.
- Resulted in 34 papers for the survey.

Three clusters

- A: Quality control on product level

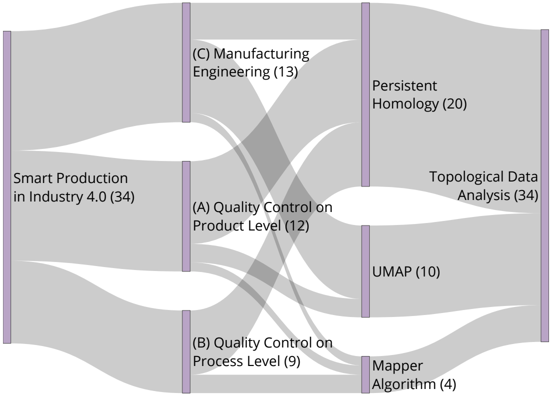
  Finding discrepancies on goods produced. TDA for analysis of structures, surfaces, shapes.

- B: Quality control on process level

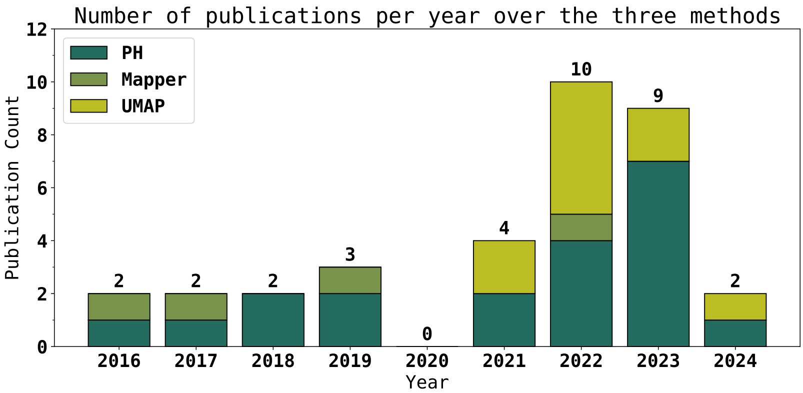
  Observing the production process. TDA for predicting productivity or detecting chatter.

- C: Manufacturing Engineering

  Design, analyze, improve manufacturing systems and processes. TDA for planning problems, predictive/preventing maintenance, process optimization.



Publications over time



- Data for 2024 only up to Apr 23 (and pre 2024-04-23 publications might not be indexed yet)
- UMAP introduced 2018, software released 2017-11-20

Publications over data type

Cluster	Input Data Format								
	fin. $\subset \mathbb{R}^n$	$\mathbb{N} \to \mathbb{R}^n$	Scalar Field ($\mathbb{R}^2 \to \mathbb{R}$)					Graph	
	Point Cloud	Time Series	Wafer Map	Sign. Dist. Func.	Microscope Imag.	Surface	Logs	Funct. Block Netw.	Task Graph
A[1]	PH PH	PH PH PH	PH Mapper UMAP		PH UMAP	PH PH			
B[1]	PH	PH PH PH PH PH PH Mapper Mapper							
C[1]	UMAP PH PH	Mapper UMAP UMAP UMAP UMAP UMAP		PH			UMAP	UMAP	PH
Count	6	17	3	1	2	2	1	1	1
			8					2	

[1] A: Quality control on product level, B: Quality control on process level, C: Manufacturing Engineering

Summary

- Low numbers of applications compared to biology and medicine.
- Most prominent data type: 17 out of 34 on time series data
- Mapper is underrepresented
- No UMAP for quality control on process level
- PH and UMAP on a diversity of data types
  - Mapper on time series (and once on a scalar field)

Industry is about signals and systems

Key scientific disciplines in industry:

- Physics, mostly mechanics, electrodynamics, thermodynamics
- Electrical engineering
- Control theory
- Signal processing

Summary

Industry is a about studying, controlling and designing dynamical systems and processing signals.

How can we leverage TDA for

- time series (aka. signals, trajectories, . . .)
- and dynamical systems?

Learning dynamics with Persistent Homology

Neural Persistence Dynamics

Sebastian Zeng[†/‡], Florian Graf[†], Martin Uray[†/‡], Stefan Huber[‡], Roland Kwitt[†]
[†]University of Salzburg, Austria
[‡]Josef Ressel Centre for Intelligent and Secure Industrial Automation,
University of Applied Sciences, Salzburg, Austria
{sebastian.zeng, florian.graf, roland.kwitt}@plus.ac.at,
{martin.uray, stefan.huber}@fh-salzburg.ac.at

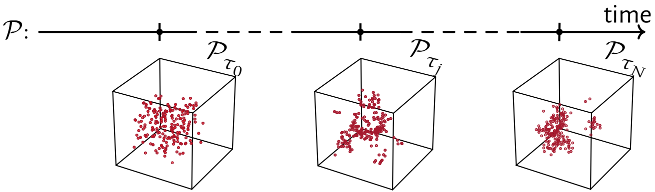
Abstract

We consider the problem of learning the dynamics in the topology of time-evolving point clouds, the prevalent spatiotemporal model for systems exhibiting collective behavior, such as swarms of insects and birds or particles in physics. In such systems, patterns emerge from (local) interactions among self-propelled entities. While several well-understood governing equations for motion and interaction exist, they are notoriously difficult to fit to data, as most prior work requires knowledge

Motivation

Problem formulation

Consider a dynamic point cloud governed by a parameterized dynamical system. Given a finite number of observations, what were the parameters?



- Think of bird flocking, cell motions in biology, particle systems in physics, ...
- Interaction between the points leads to emerging patterns $\rightarrow$ TDA for the win?

Problem setting

Dynamic point set $\mathcal{P}_t \overset{\text{fin.}}{\subset} \mathbb{R}^3$, with Euclidean metric on $\mathbb{R}^3$ and time t

- $\mathcal{P}_t = \{x_1(t), \dots, x_M(t)\}$, with $x_i \colon \mathbb{R} \to \mathbb{R}^3$ denoting the trajectory of i-th point.
- Dynamical system: The $\dot{x}_i$ are controlled by a small number of parameters $\beta_1, \dots, \beta_P$, i.e.,

$$\dot{x}_i = f_{\beta_1,\dots,\beta_P}(x_1, \dots, x_P, \dot{x}_1, \dots, \dot{x}_P),$$

  typically modeling local interactions between points, possibly with stochasticity.

Observations $\mathcal{P}_{\tau_0}, \dots, \mathcal{P}_{\tau_N}$ at (not necessarily equidistant) time points τ_i:

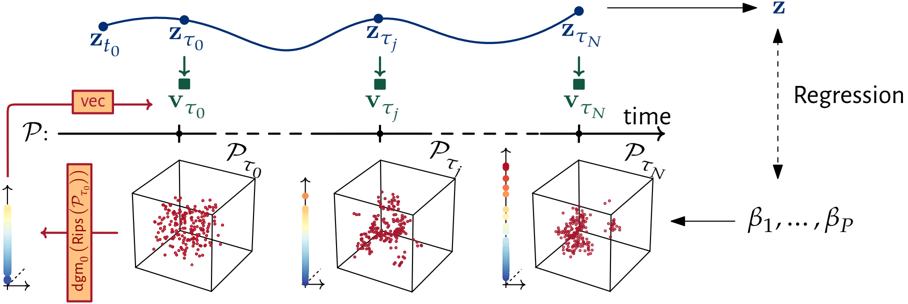
- No point tracks: We have no correspondences of points between the τ_i
- Changing cardinalities: think of occlusion or points leaving and reentering scenery.

Question

Given $\mathcal{P}_{\tau_0}, \dots, \mathcal{P}_{\tau_N}$, what are $\beta_1, \dots, \beta_P$?

Neural persistence dynamics: Core strategy

1. Consider vectorizations v_{τ_j} of the point clouds $\mathcal{P}_{\tau_j}$.
   - Vecotrizations of $\mathrm{dgm}_0(\mathrm{Rips}(\mathcal{P}_{\tau_j}))$, or PointNet++ of $\mathcal{P}_{\tau_j}$, or a joint version
2. Instead of thinking of dynamics of v_{τ_j}, impute some dynamics in a latent space, leading to a path z on which z_{τ_j} are latent encodings of v_{τ_j}.



3. Build a generative model for the underlying latent dynamics, which can entirely reconstruct z.
   - mTAN for the latent encoding, NeuralODE as generative model
4. Perform auxiliary regression task for (a summary of) the latent path z.

Experimental evaluation

- D'Orsogna with 4 parameters
- Vicsek with 4 parameters
- `10k` is the number of sequences in the dataset
- VE: Variance explained, SMAPE: symmetric mean absolute percentage error

		⌀ VE ↑	⌀ SMAPE ↓
`dorsogna-10k`	Ours (joint, v3)	**0.689**$\pm$0.021	**0.088**$\pm$0.004
	Ours (PH-only, v1)	**0.680**$\pm$0.025	**0.090**$\pm$0.005
	PSK	0.647$\pm$0.005	0.100$\pm$0.003
	Crocker Stacks	0.343$\pm$0.016	0.145$\pm$0.001
`vicsek-10k`	Ours (joint, v3)	**0.576**$\pm$0.030	**0.144**$\pm$0.006
	Ours (PH-only, v1)	**0.579**$\pm$0.034	**0.146**$\pm$0.006
	PSK	0.466$\pm$0.009	0.173$\pm$0.003
	Crocker Stacks	0.345$\pm$0.005	0.190$\pm$0.001

Summary

- Outperforms SOTA
- Scales to large number of observation sequences[2]
- Ablation study shows: PH gives complementary information

[2] Overall training time on `dorsogna-1k`: Ours 190 s, PSK: 646 s, Crocker Stacks: 24 600 s.

Maybe future application: anomaly detection

Assumption: Industrial machine, or production line, or a factory displays behavior that can be (sufficiently) modeled by a dynamical system.

- The behavior depends on parameters $\beta_1, \ldots, \beta_P$.
- The observation is a time series $x\colon [0, T] \to \mathbb{R}^n$.

parameters $\beta_1, \ldots, \beta_P$ $\longrightarrow$ $\boxed{\text{system}}$ $\longrightarrow$ observation $x(t)$

Assumption: Anomaly is caused by changing parameters $\beta_1, \ldots, \beta_P$, e.g., friction parameter increased.

Approach

If you can estimate $\beta_1, \ldots, \beta_P$ from observations x then you can detect anomalies, and name its root cause.

Multi-variate time series with PH

Decomposing multi-variate time series

Multi-variate time series $x \colon I \to \mathbb{R}^m$ from machines of discrete automation display cyclic behavior, with $I = [0, T] \subset \mathbb{R}$.

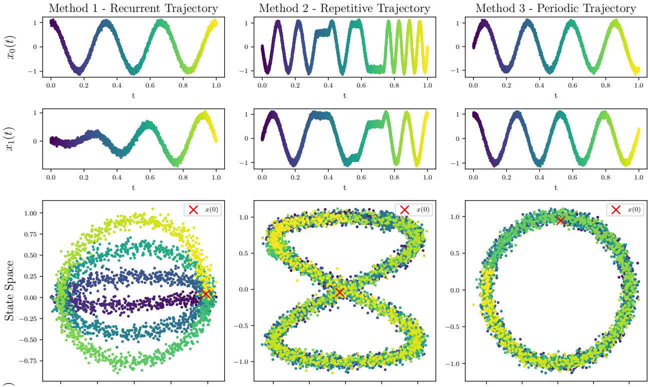
Question

How can we decompose it into its "cycles"?[3]

Find $0 = T_0 < \cdots < T_k = T$ such that $I_i = [T_{i-1}, T_i]$ is such a "cycle" and call $\tau_i = |I_i|$ the cycle length.

[3] Joint work with Simon Schindler, Elias Reich, Simon Hoher, Saverio Messineo.

- Machines may run at varying speed.
- May display different modes or abnormalities, so we have "forking" or otherwise spatially diverging behavior in the state space.



Three levels of cyclic behavior

From strictest to weakest:

- x is periodic if $x(t+\tau) = x(t)$ for all t, i.e., $\tau_i = \tau$.
  - Natural habitat for convolution-based methods (Fourier, auto correlation), at least in univariate case
- x is repetitive if there are non-decreasing maps $\gamma_i \colon I_i \to I_1$ with $\gamma_i(T_{i-1}) = T_0$, and for $i < k$ also surjective, such that $x(\gamma_i(t)) = x(t)$, i.e., $(x \circ \gamma)\big|_{I_i} = x\big|_{I_1}$
  - Allows for reparameterizations of x, i.e., changes of speed at a fixed image of x.
- x is recurring if $\{T_0, \ldots, T_k\} = \ker(x - x(0)) \cup \{T\}$ is finite.
  - Allows for variations in the image of x, i.e., production modes or other variations in the process.

Three levels of cyclic behavior

From strictest to weakest:

- x is periodic if $x(t+\tau) = x(t)$ for all t, i.e., $\tau_i = \tau$.
  - Natural habitat for convolution-based methods (Fourier, auto correlation), at least in univariate case
- x is repetitive if there are non-decreasing maps $\gamma_i \colon I_i \to I_1$ with $\gamma_i(T_{i-1}) = T_0$, and for $i < k$ also surjective, such that $x(\gamma_i(t)) = x(t)$, i.e., $(x \circ \gamma)\big|_{I_i} = x\big|_{I_1}$
  - Allows for reparameterizations of x, i.e., changes of speed at a fixed image of x.
- x is recurring if $\{T_0, \dots, T_k\} = \ker(x - x(0)) \cup \{T\}$ is finite.
  - Allows for variations in the image of x, i.e., production modes or other variations in the process.

Measurements: We need to cope with noise or some sort of "approximate" notions.

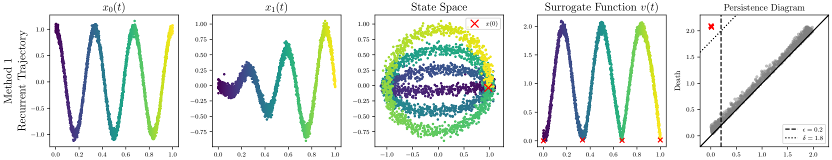
- x is ε-approximately periodic (repetitive) if there is a periodic (repetitive) $\hat{x}$ with $\|x - \hat{x}\|_\infty \leq \varepsilon$.
- x is ε-δ-approximately recurring if for all $i \in \{1, \dots, k-1\}$
  - $x(T_i)$ is a ε-close to $x(0)$ and
  - farthest point $x(t)$ for $t \in I_j$ is farther from $x(0)$ than $\delta + \|x(T_i) - x(0)\|$ for $j \in \{i, i+1\}$.

Method 1: Recurring time series

Let x be an ε-δ-approximately recurring time series $I \to \mathbb{R}^m$.

Method

- Define $v_x\colon I \to \mathbb{R}\colon t \mapsto \|x(t) - x(0)\|$. (Or $v_x(t) = d(t, 0)$ for some metric d in $\mathbb{R}^m$.)
- Compute H_0 of the sublevel-set filtration of v_x
- Take as T_i the local minima of v_x with persistence pairs (b, d) such that $b < \varepsilon, d - b > \delta$.



Good: Stability as $\|v_x - v_{x'}\|_\infty \leq 2\|x - x'\|_\infty$.

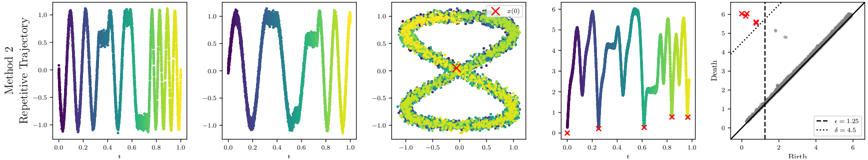
Bad: Breaks when x is (ε-approximately) repetitive but (ε-approximately) self-intersecting (at $x(0)$).

Method 2: Repetitive time series

Let x be an ε-approximately repetitive time series $I \to \mathbb{R}^m$.

Method

- Consider the time-delay embedding $U_x \colon I \to (\mathbb{R}^m)^d \colon t \mapsto (x(t), x(t+\Delta), \dots, x(t+(d-1)\Delta))$ to disentangle ambiguities at self-intersections. (Also works with some metric on $\mathbb{R}^m$.)
- Define $v_x \colon I \to \mathbb{R} \colon t \mapsto \|U_x(t) - U_x(0)\|_p = \sqrt[p]{\sum_{i=0}^{d-1} \|x(t+i\Delta) - x(i\Delta)\|^p}$.



Good: Stability as $\|v_x - v_{x'}\|_\infty \leq 2\sqrt[p]{d}\|x - x'\|_\infty$.

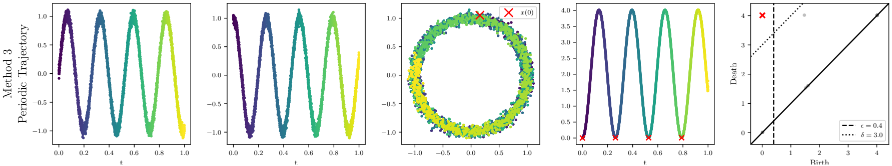
Bad: Needs a parameterization for the time-delay embedding.

Let x be an ε-approximately periodic time series $I \to \mathbb{R}^m$.

Idea: Generalize from one "pivot" $x(0)$ to all possible "pivots". Not unlike considering a "holistic" time delay embedding covering entire x.

Method

- Consider the recurrence function $w_x \colon I \times I \to \mathbb{R} \colon (s,t) \mapsto \|x(s) - x(t)\|$.
- Average along diagonals (a fixed time delay Δ): $v_x(\Delta) = \operatorname{avg}_t w_x(t, t+\Delta)$.



Good: Stability as $\|v_x - v_{x'}\|_\infty \leq \|w_x - w_{x'}\|_\infty \leq 2\|x - x'\|_\infty$.

Bad: Deviations from periodicity are bad for the averaging.

A novel dataset

Dataset

We generated a multi-variate dataset from an injection molding machine (IMM):

- 4.15 M samples, 23 variables, each cycle between 25 k to 35 k samples, sampled at 1 kHz

Periodic (I), repetitive (II) and recurring (III) behavior with 120 cycles grouped in 18 modes:

- 2×20 periodic cycles
- 10×5 repetitive cycles: Varying injection or plastification speed, delayed clamping or ejection
- 6×5 recurring cycles: Increased friction on plastification and injector servo drive axis

Summary:

- Simple idea enhanced with PH, but very useful.
- Time-delay embeddings and recurrence functions improve usability.

Detecting peaks in dual spaces

A thought on applications of PH

Observation

PH is useful for a variety of applications, and complementary to existing techniques.

But: Dissemination in industry is lacking behind

- Group of people in problem space and solution space are largely distinct.

Maybe one source of interesting applications could be this:

- Different application fields have natural dual spaces they operate in. How about applying TDA (or PH) in there?

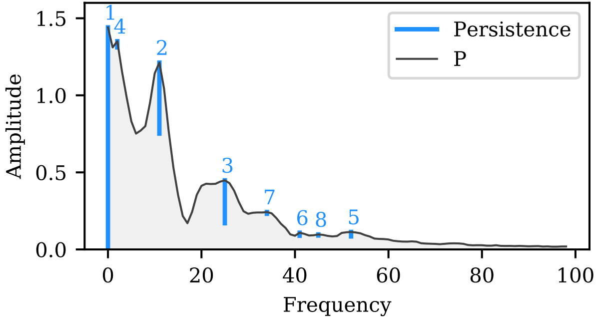
Two examples:

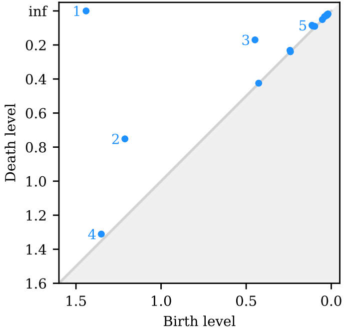
- Peak detection in Fourier spectrum of signals
- Peak detection in the Hough space for line detection

Practical low-hanging fruit: System identification for controller tuning

Closed-loop controllers (e.g., in motion control) can be tuned when the natural frequencies of the system are known.

- Excitation signal of serve drive $\rightarrow$ system's response $\rightarrow$ Fourier transform
- Task: Detect dominant peak at non-zero frequency in the frequency domain

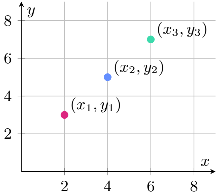


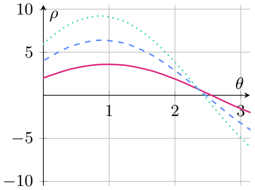


See [Hub20] for details.

Hough transform detects lines in an image (or point cloud) by dualizing the problem:

- Parameterize a line by Hesse normal form (r, ϕ). This gives the Hough space of lines.
- Then all lines through a point form a sinusoid in the Hough space.
- Detecting lines $\to$ detecting peaks in the Hough space.

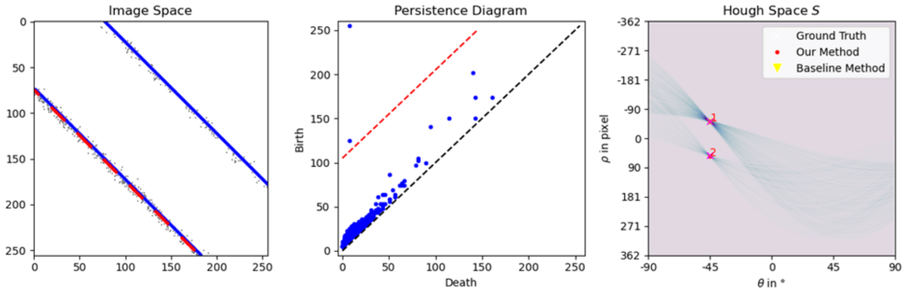




- Classic approach (opencv): Thresholding. Bad at presence of noise and uneven sampling of lines
- Our approach: Persistence based peak detection[4]

4 Joint work with Martin Uray, Johannes Ferner, Angel Pop, Saverio Missineo.

Persistence-based Hough transform



- Noise adds points in persistence diagram close to diagonal
- Uneven sampling of lines leads to less higher peaks $\rightarrow$ bad for birth thresholding
- Different noise levels leads to a diffused peak in Hough space $\rightarrow$ bad for birth thresholding

Final slide

Industrial automation is an important field for the wealth of our societies.

1. TDA has potential in engineering and industry, but dissemination is lacking behind.
2. For TDA to be useful for foundations of industry, focus on dynamical systems and time series appears natural.
   - Forecasting, classification, pattern recognition, system identification, optimization, ...
   - TDA for Reinforcement Learning, as the ML-counterpart to control theory
3. Looking at established dual spaces to apply TDA at could lead to valuable applications.

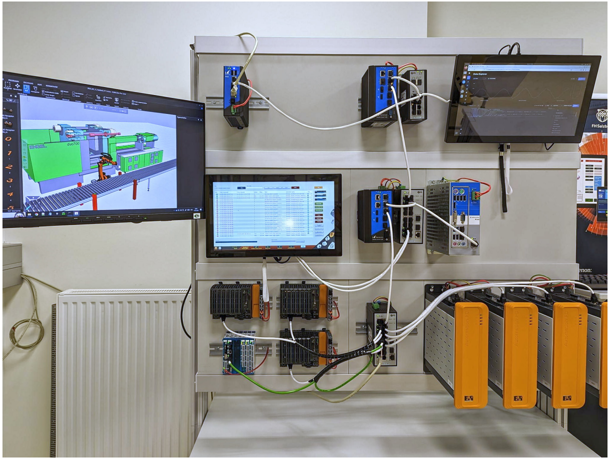
What's next?

- We have a realistic testbed, we can generate interesting data, we have interdisciplinary team.
- Interested to cooperate?
- One of our primary research interests are: Anomaly detection and generative models for multi-variate time series (stemming from discrete automation)

Thank you

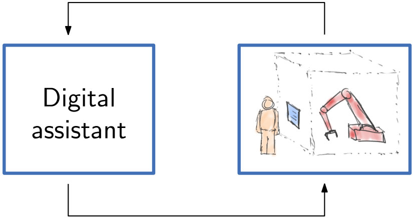
Testbed of a production scenario

Implementation as mobile testbed



Digital assistant

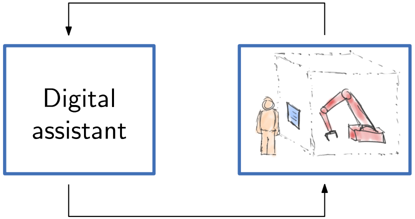
- Supports human operator in supervision & operation
- An intelligent agent in the sense of Russel & Norvig
- Security important due to permeability of IT and autonomy
- Shall be largely machine agnostic



Josef Ressel Center for Intelligent and Secure Industrial Automation

Digital assistant

- Supports human operator in supervision & operation
- An intelligent agent in the sense of Russel & Norvig
- Security important due to permeability of IT and autonomy
- Shall be largely machine agnostic



Constitutional question

How can an intelligent, secure, machine-agnostic digital assistant for a variety of industrial machines of discrete automation been formed?

TDA survey: Examples for time series

Cluster A:

- $3\times$ PH: three-phase motor current signals: H_0, H_1 of sampled time series for eccentricity detection

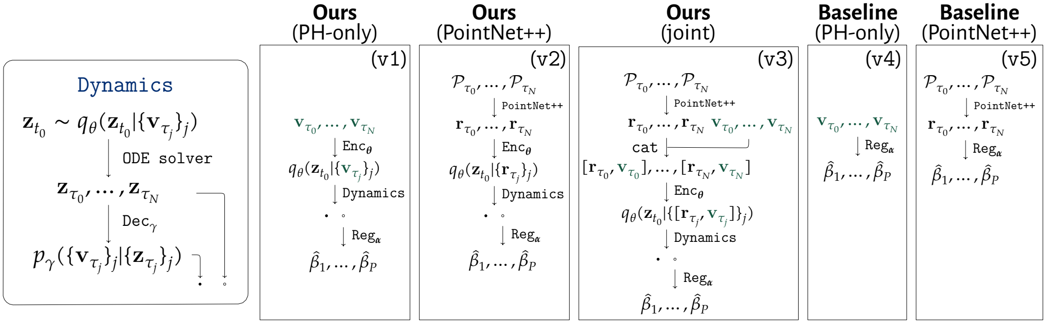
Cluster B:

- $5\times$ PH: Chatter detection and classification using sampling of Taken embedding of time series
- $1\times$ PH: Porosity detection
- $2\times$ Mapper: Predicition of productivity

Cluster C:

- $1\times$ Mapper: Demand forecasting
- $5\times$ UMAP: Variety of applications for dimensionality reduction

- v_{τ_j} are vectorizations of point clouds at observation time τ_j, e.g., from PointNet++ or PH.



- Dynamics via ELBO maximization:
  - $q_\theta(z_{t_0} \mid \{v_{\tau_j}\})$ variational posterior via an mTAN
  - latent z_t for any t via Euler integration from z_{t_0} of the neural ODE
  - decoder network via 2-layer MLP
- Regression via MSE: Linear maps from latent states z_t at equidistant $_i \in [0, T]$ to $\hat\beta_1, \dots, \hat\beta_P$.

Repetition detection: Related work

Classical approaches:

- Frequency domain and auto correlation methods require evenly spaced samplings of periodic signals

Perea papers [PH15]; [Per16]; [Per19]:

- Examine characteristics of delay embeddings of (quasi-)periodic functions
- Starting idea: Detecting homological cycles in time-delay embedding space

Bonis et al. [Bon+24]:

- Determine γ^{-1} of a reparameterized series $f \circ \gamma$ of a uni-variate periodic f to find number of periods

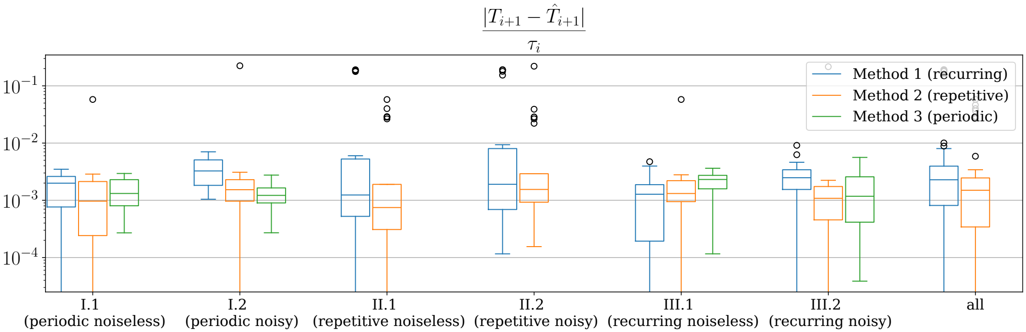
Bauer et al. [Bau+24]:

- Embedding of a multi-variate time series along with its first derivative into a higher dimensional space. H_1 of Rips filtration from subsequences are used to find the number of cycles.

Ichinomiya et al. [Ich23]:

- Super- and sublevel set filtration of recurrence function of multi-variate time series to study dynamical systems.

Repetition detection: Experimental evaluation



Two different natures of sequential data?

In Machine Learning we deal with different forms of sequential data:

- Audio and video streams
- Token sequences forming text in a language, like log data, prosa text, source code, genomes
- Physical or economical quantities over time, like from weather, finance, healthcare, supply chains
- Et cetera

Two different natures of sequential data?

In Machine Learning we deal with different forms of sequential data:

- Audio and video streams
- Token sequences forming text in a language, like log data, prosa text, source code, genomes
- Physical or economical quantities over time, like from weather, finance, healthcare, supply chains
- Et cetera

Token sequences

A sequence of tokens forming an abstract "sentence" of a "language", formalized by some sort of "grammar".

- Sequences a la Noam Chomsky
- Language theory
- Attention-based methods more natural?

Signals, time series

Signals as temporal evolutions of quantities following some "laws" of some "process" or "system".

- Sequences a la Norbert Wiener
- Signals and systems theory
- Filter-like methods more natural?