Time series analysis using persistent homology methods

Alexander Khe

Lavrentyev Institute of Hydrodynamics, Novosibirsk, Russia

In collaboration with

Ya. Bazaikin (Sobolev Institute of Mathematics, Novosibirsk, Russia)
A. Chupakhin (Lavrentyev Institute of Hydrodynamics, Novosibirsk, Russia)
A. Cherevko (Lavrentyev Institute of Hydrodynamics, Novosibirsk, Russia)

British–Russian Workshop

Mathematical and Computational Modelling in Cardiovascular Problems

Moscow, April 15–17, 2014

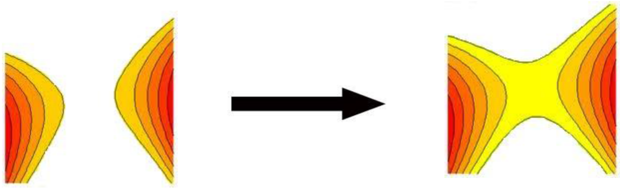
Processing and analysis of digital information

- How to analyse and compare digital data?

Definitions

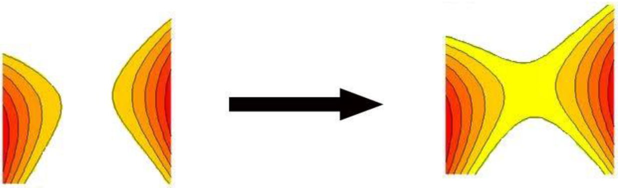
Introduce the following notation:

- X — topological space,
- $f : X \mapsto \mathbb{R}$ — continuous mapping
- $X_a = \{p \in X \mid f(p) \leqslant a\}\ \forall a \in \mathbb{R}$
- $C_a = C(X_a)\ \forall a \in \mathbb{R}$ — set of components of connectivity $X_a \subset X$
- $f_a^b : C_a \to C_b$ — mapping of the sets, induced by embedding $X_a \subset X_b$ $\forall a < b$.



Critical values

Value $a \in \mathbb{R}$ is called *critical* for function f, if for any sufficiently small $\varepsilon > 0$ the mapping $f_{a-\varepsilon}^{a+\varepsilon} : C_{a-\varepsilon} \to C_{a+\varepsilon}$ is not bijection.
Let $a_1, \ldots, a_n$ be all critical values of function f.
Denote by $C_i = C_{a_i}$ a set of connectivity components, $i = 1, \ldots, n$

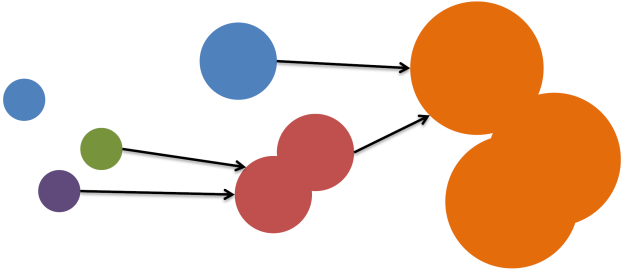


Merge tree

Construct graph Γ in the following way.

1. Set of vertices is $\cup_{i=1}^{n} C_i$.
2. Connect every vertex $c \in C_i$ with an edge to vertex $d = f_i(c) = f_{a_i}^{a_{i+1}}(c)$.

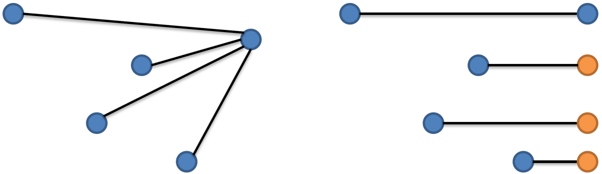
Graph Γ is a tree, which we call a *merge tree*.



Barcode. Persistent diagram

For every $Y \in C_i$ we define $f(Y) = a_i$ and $w(Y) = inf_{p \in Y} f(p)$.
Construct a graph Γ' from graph Γ in the following way.

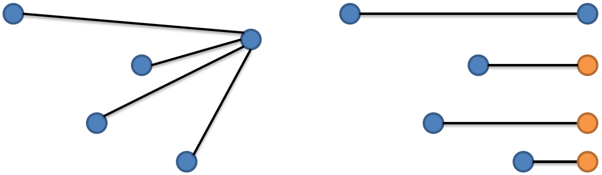
1. Let $d \in C_{i+1}$ be a vertex of Γ and $f_i^{-1}(d) = \{c_1, c_2, \ldots, c_k\} \subset C_i$.
2. Renumber vertices $c_j, j = 1, \ldots, k$ such that $w(c_1) \geq w(c_2) \geq \ldots \geq w(c_k)$.
3. Disjoin every edge $c_j d$, $j = 2, \ldots, k$ from graph Γ, and add a new vertex d_j, which will be the end of that edge.



Definition

A set B of all such intervals $[c, d)$ is called *barcode* of function f on X.

Persistent diagram of function f is a set $D(f)$ of points $(c, d) \in \mathbb{R}^2, [c, d) \in B$, united with the set of diagonal points $\Delta = \{(x, x) | x \in \mathbb{R}\}$.



Stability

Define a distance between two sets D_1 and D_2:

$$d_B(D_1, D_2) = \inf_{\gamma} \sup_{p \in D_1} \|p - \gamma(p)\|_\infty, \quad \gamma : D_1 \to D_2 \text{ — bijection.}$$

Theorem

Let X be a topological space, $f, g : X \to \mathbb{R}$. Then

$$d_B(D(f), D(g)) \leqslant \|f - g\|_\infty.$$

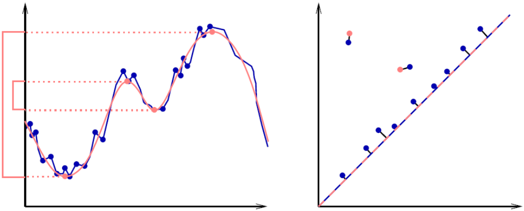
Thus, persistent diagram $D(f)$ is stable with respect to pertubations of the function f.

[D. Cohen-Steiner, H. Edelsbrunner, J. Harer, 2007]

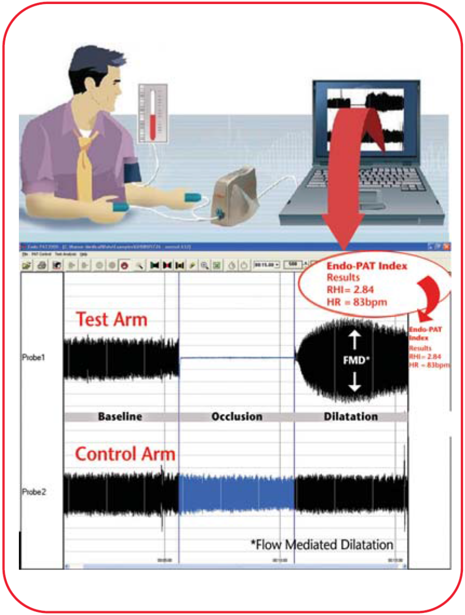
Stability

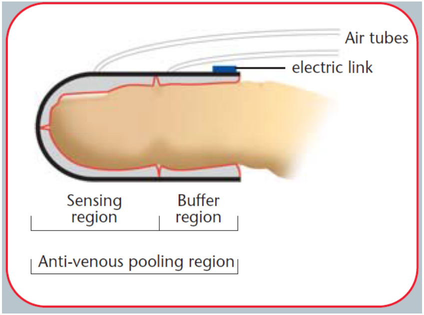
When the function f with diagram $D(f)$ is pertubed:

1. some points move over a short distance;
2. a number of points close to diagonal, moves onto the diagonal;
3. a part of points comes out from the diagonal.

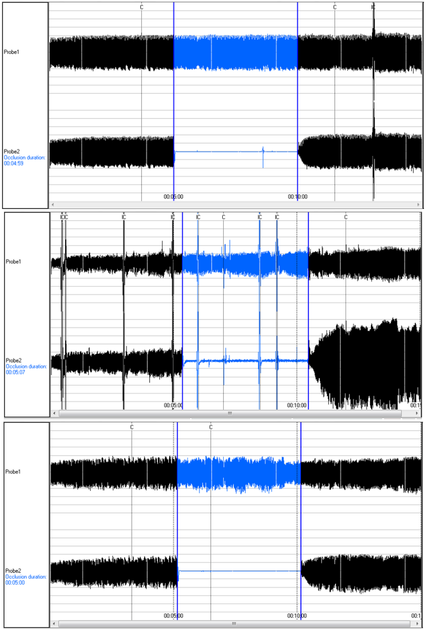


Endo-PAT2000

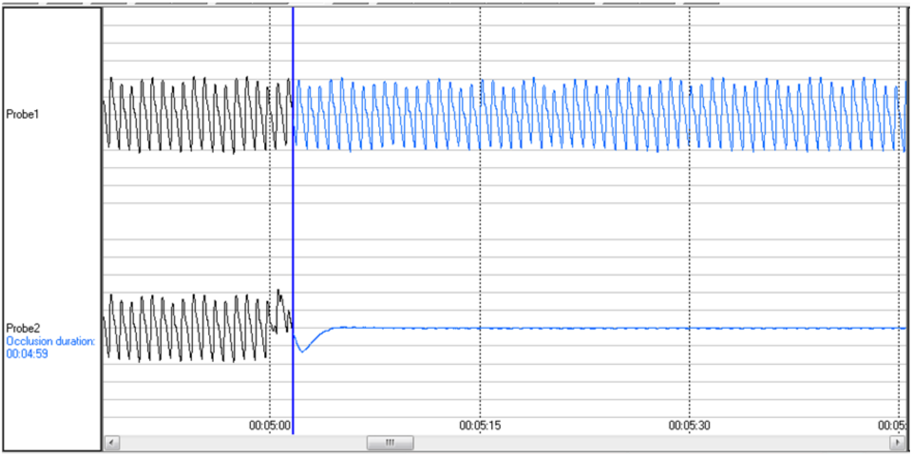




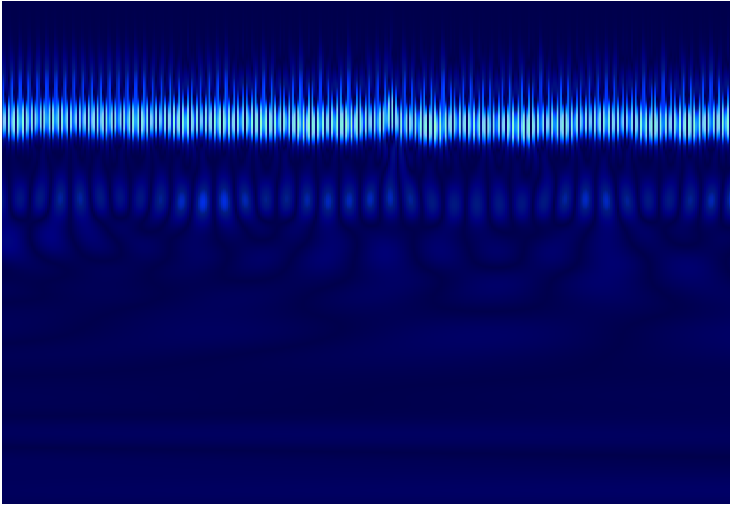
Endo-PAT2000 data



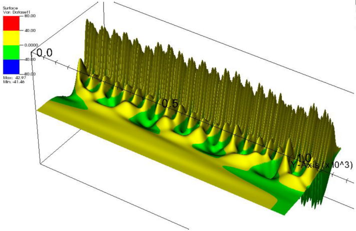
Endo-PAT2000 data



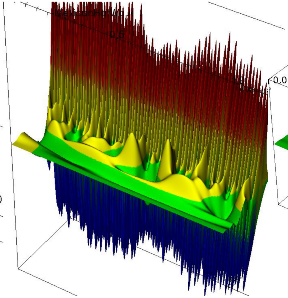
Wavelet scalograms



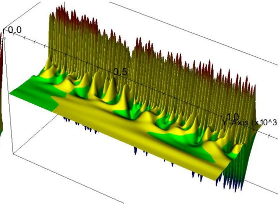
Wavelet scalograms



A

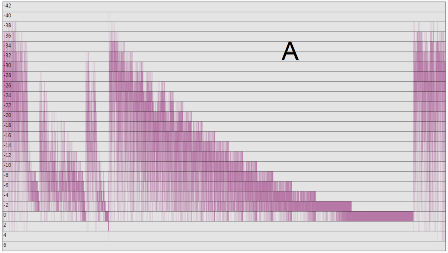


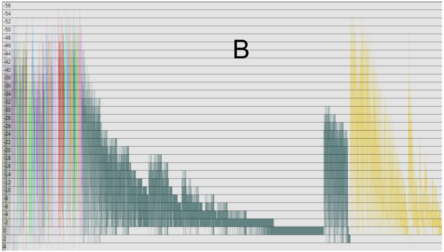
B

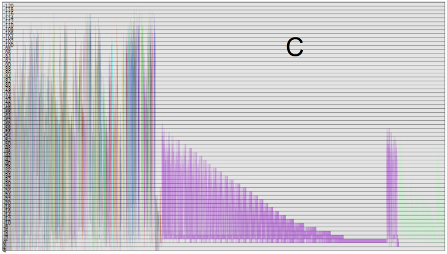


C

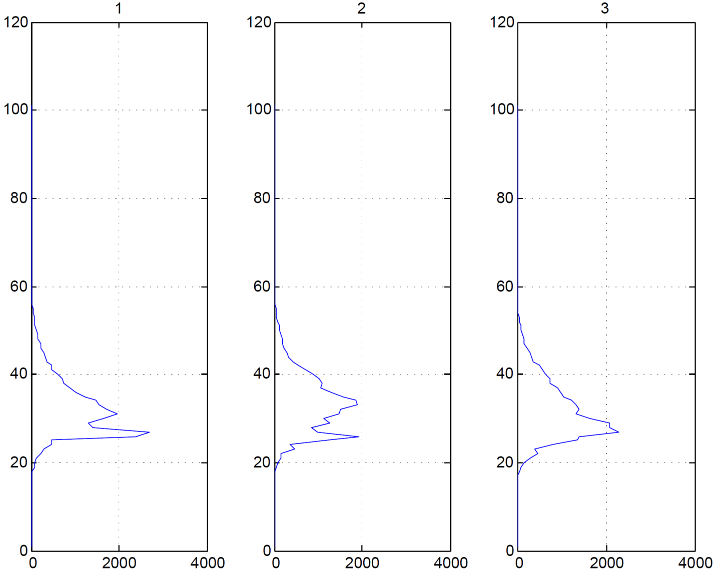
Barcodes


A
-42
-40
-38
-36
-34
-32
-30
-28
-26
-24
-22
-20
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
2
4
6

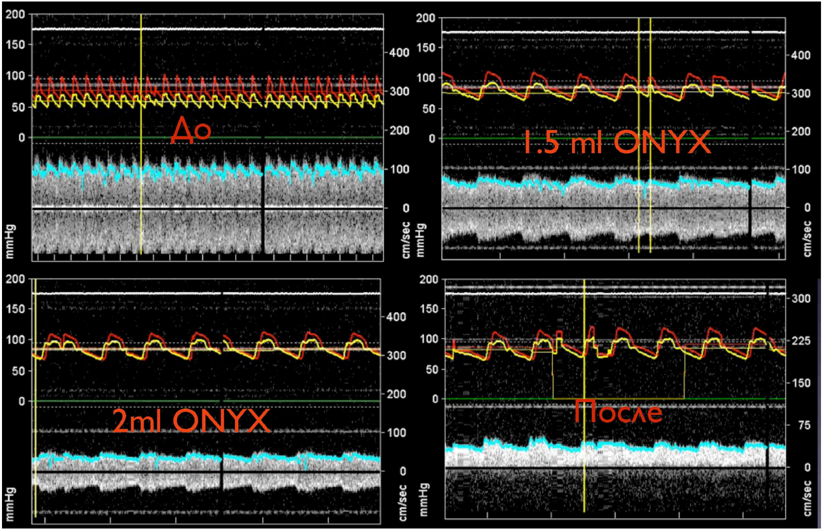

B
-56
-54
-52
-50
-48
-46
-44
-42
-40
-38
-36
-34
-32
-30
-28
-26
-24
-22
-20
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
2
4


C

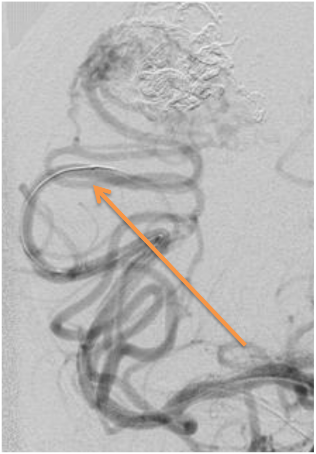
Barcodes. Distribution

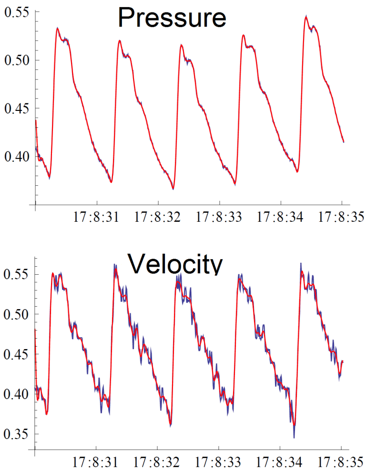


Endovascular measurements

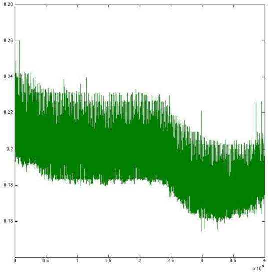


Measurement data

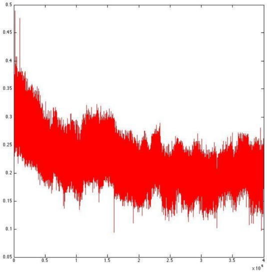




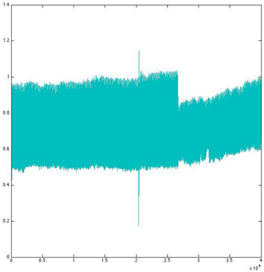
Data analysis



Pressure

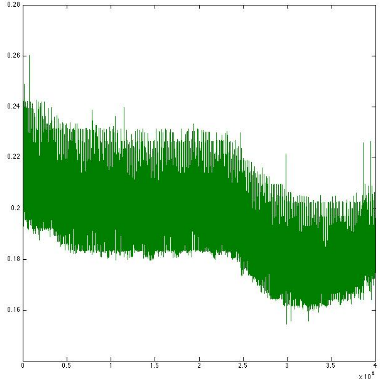


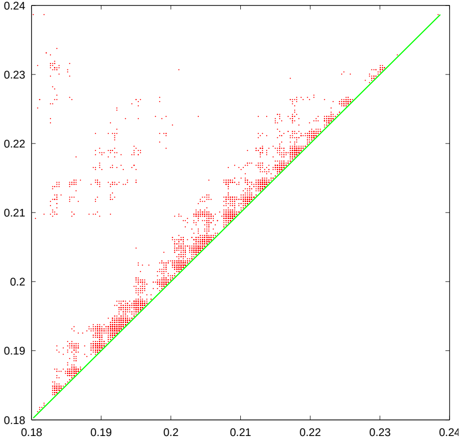
Velocity



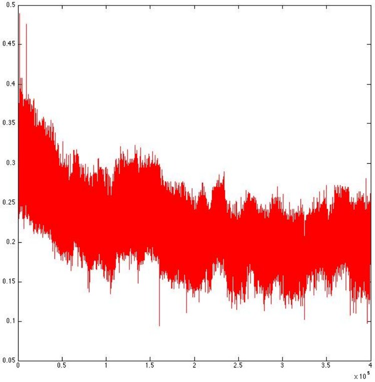
Systemic pressure

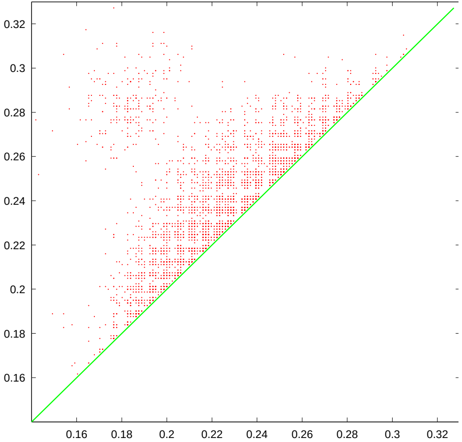
Persistent diagram: Pressure



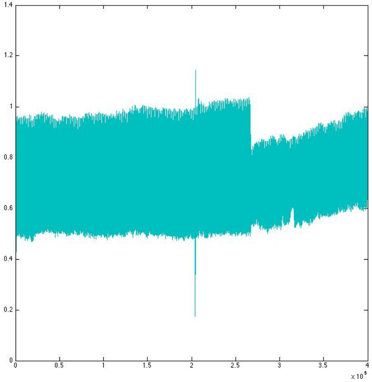


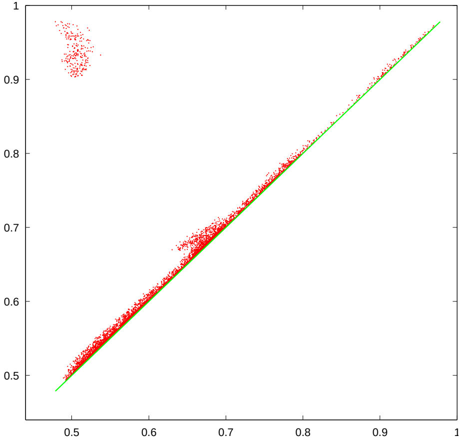
Persistent diagram: Velocity





Persistent diagram: Systemic pressure





Summary

- Methods of persistent diagrams were applied to medical data, obtained during examinations.
- Comparison of different study cases shows that barcodes and persistent diagrams can be used for data analysis.