

The coherent matching distance in 2D persistent homology

Andrea Cerri¹, Marc Ethier² and Patrizio Frosini³

¹IMATI – CNR, Genova, Italia

²Faculté des Sciences, Université de Saint-Boniface, Winnipeg, Manitoba, Canada

³Department of Mathematics and ARCES, University of Bologna, Italy

Computational Topology in Image Context,
Marseille, June 15-17, 2016

Outline



Let us recall some concepts in persistent homology

The importance of lines with slope 1 in 2D persistent homology

The need for coherent matchings

The phenomenon of monodromy

The coherent 2-dimensional matching distance



Let us recall some concepts in persistent homology

The importance of lines with slope 1 in 2D persistent homology

The need for coherent matchings

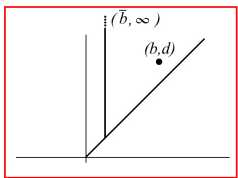
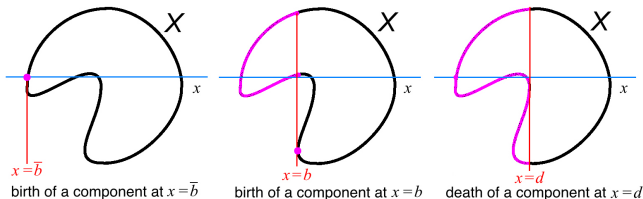
The phenomenon of monodromy

The coherent 2-dimensional matching distance



Let us recall what a persistence diagram is

A persistence diagram is a collection of points associated with a continuous filtering function $\varphi : X \rightarrow \mathbb{R}$. An example for $\varphi = x$:



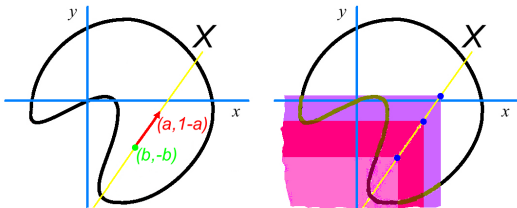
Persistence diagram of the function x

The presence of a point (b, d) in the persistence diagram of $\varphi : X \rightarrow \mathbb{R}$ means that a class of a cycle exists whose birth and death happen at $\varphi = b$ and $\varphi = d$, respectively.

Persistence diagrams can be compared by a matching distance d_{match} .

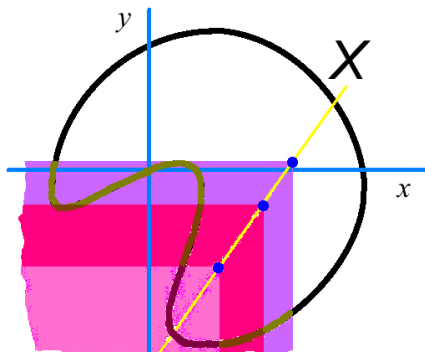
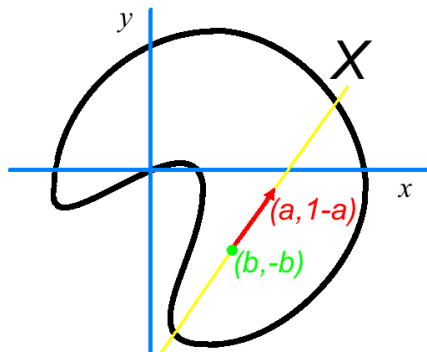
Persistence diagram associated with the pair (P, w)

What can we do in the case $\varphi : X \rightarrow \mathbb{R}^2$?



If we have a bifiltration given by a function $\varphi = (\varphi_1, \varphi_2) : X \rightarrow \mathbb{R}^2$, we can consider a unit vector (w.r.t. $\|\cdot\|_\infty$) $w = (a, 1-a)$ with a positive slope, and a point $P = (b, -b)$. Every choice of P and w defines a filtration $\{X_t\}$ of X , where X_t is the set of points of X that are both under and on the left of the point $P + tw$. As a consequence, each choice of P and w defines a persistence diagram $\mathcal{D}_\varphi(a, b)$.

The 2-dimensional matching distance



By multiplying the coordinates of each point in $\mathcal{D}_\varphi(a, b)$ by $\min\{a, 1 - a\}$ we obtain the **normalized persistence diagram** $\hat{\mathcal{D}}_\varphi(a, b)$.

The 2-dimensional matching distance $D_{match}(\varphi, \psi)$



Now, we can define the 2D matching distance $D_{match}(\varphi, \psi)$ between φ and ψ by setting

$$D_{match}(\varphi, \psi) := \sup_{(a,b)} d_{match}(\hat{\mathcal{D}}_{\varphi}(a, b), \hat{\mathcal{D}}_{\psi}(a, b)).$$

Theorem

$$D_{match}(\varphi, \psi) \leq \|\varphi - \psi\|_{\infty}.$$

Remark. The previous theorem strongly depends on the **normalization** of persistence diagrams.

The distance $D_{match}(\varphi, \psi)$ has been introduced in the paper [S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, C. Landi, *Multidimensional size functions for shape comparison*, Journal of Mathematical Imaging and Vision, vol. 32 (2008), n. 2, 161–179.]



Computation of the 2D matching distance

An algorithm to compute the 2D matching distance exists:

- S. Biasotti, A. Cerri, P. Frosini, D. Giorgi,
A new algorithm for computing the 2-dimensional matching distance between size functions,
Pattern Recognition Letters, vol. 32 (2011), n. 14, 1735–1746
- A. Cerri, P. Frosini,
A new approximation algorithm for the matching distance in multidimensional persistence,
AMS Acta, 2971 (2011)

The results of our experiments on the computation of the 2D matching distance have revealed a strange phenomenon that we are going to illustrate in the next slides.

Computation of the 2D matching distance



Remarks:

- In principle, 2D persistent homologies can be also compared by using the interleaving distance ([M. Lesnick, *The theory of the interleaving distance on multidimensional persistence modules*, Foundations of Computational Mathematics, vol. 15 (2015), n. 3, 613–650]). Unfortunately, as noted in that paper, the question of if and how the interleaving distance on multidimensional persistence modules can be computed remains open. This fact justifies the interest in the 2D matching distance $D_{match}(\varphi, \psi)$, which is easily computable.
- As for the visualization of 2D persistence modules we point out the interesting paper [M. Lesnick, M. Wright, *Interactive Visualization of 2-D Persistence Modules*, arXiv:1512.00180].



Let us recall some concepts in persistent homology

The importance of lines with slope 1 in 2D persistent homology

The need for coherent matchings

The phenomenon of monodromy

The coherent 2-dimensional matching distance

Our first experiments in 3D shape comparison



Let us have a look at some pictures illustrating the first results that we have obtained in 3D shape comparison by means of the 2D matching distance.

The objects that we compare are displayed on the left of each figure.

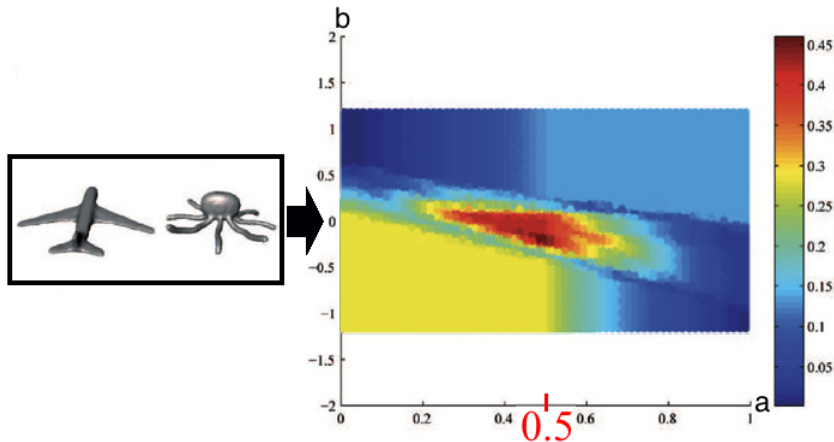
The color at (a, b) represents the value $d_{match}(\hat{\mathcal{D}}_\varphi(a, b), \hat{\mathcal{D}}_\psi(a, b))$.

The largest values are in **red** and **brown**, the lowest ones are in **blue**.

The values at points (a, b) with $|b|$ large are not displayed, since they just represent the 1-dimensional persistence of the two components of $\varphi = (\varphi_1, \varphi_2)$ and are not relevant for our exposition.

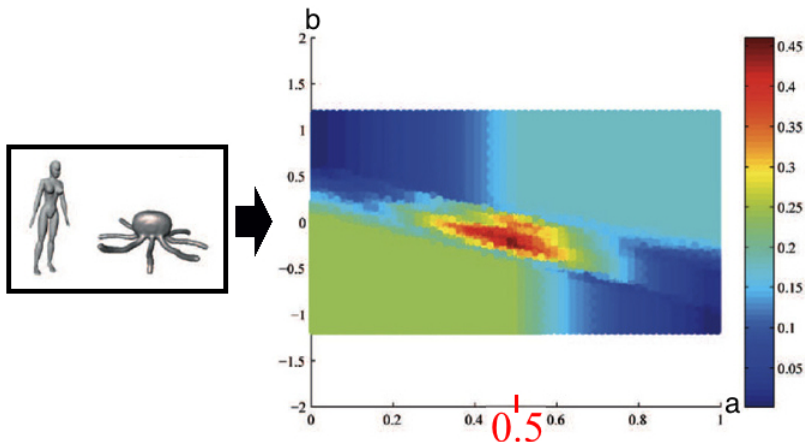
We recall that we are interested in $D_{match}(\varphi, \psi)$, i.e. the supremum of $d_{match}(\hat{\mathcal{D}}_\varphi(a, b), \hat{\mathcal{D}}_\psi(a, b))$, for $0 < a < 1$ and $b \in \mathbb{R}$.

Result 1



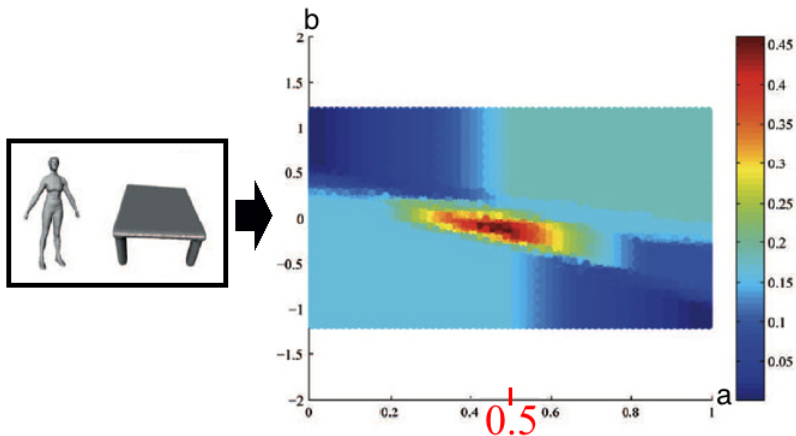
Please note that the largest value (i.e. the 2-dimensional matching distance) is taken at a point (a, b) with a approximately equal to $1/2$.

Result 2



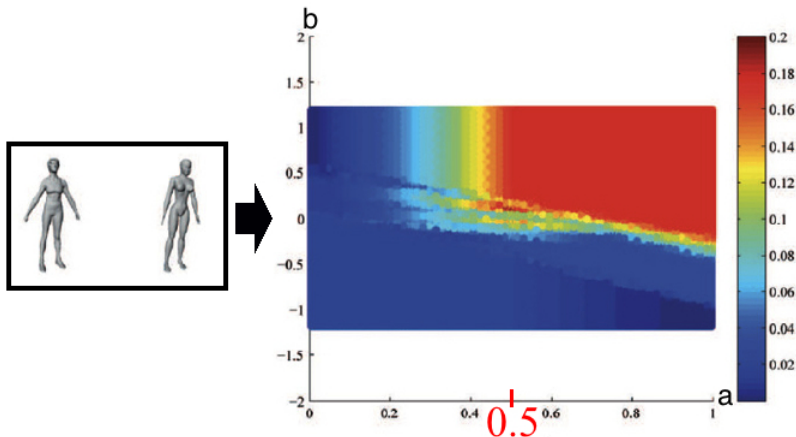
Please note that the largest value (i.e. the 2-dimensional matching distance) is taken at a point (a, b) with a approximately equal to $1/2$.

Result 3



Please note that the largest value (i.e. the 2-dimensional matching distance) is taken at a point (a, b) with a approximately equal to $1/2$.

Result 4



Please note that the largest value (i.e. the 2-dimensional matching distance) is taken at a point (a, b) with a approximately equal to $1/2$.



A conjecture

In our experiments, it seems that all the relevant information is given by lines with $a = 1/2$, i.e. lines with slope 1.

At the beginning we thought that this phenomenon was just a coincidence. So we looked for other examples, showing global maxima at points (a, b) with $a \neq 1/2$, but our search was unsuccessful.

After that, we started to think that some interesting principle might be hidden in the results of our experiments.

We formulated this conjecture:

Conjecture

$$\sup_{0 < a < 1, b \in \mathbb{R}} d_{\text{match}}(\hat{\mathcal{D}}_{\varphi}(a, b), \hat{\mathcal{D}}_{\psi}(a, b)) = \sup_{b \in \mathbb{R}} d_{\text{match}}(\hat{\mathcal{D}}_{\varphi}(\frac{1}{2}, b), \hat{\mathcal{D}}_{\psi}(\frac{1}{2}, b)).$$

In other words: Can we confine ourselves to assume $a = 1/2$?



Let us recall some concepts in persistent homology

The importance of lines with slope 1 in 2D persistent homology

The need for coherent matchings

The phenomenon of monodromy

The coherent 2-dimensional matching distance



Looking for a proof of our conjecture

We started to look for a proof of our conjecture but we realized that **we needed to change our definition of 2D matching distance** in order to proceed.

Indeed the classical definition of 2D matching distance does not require that the matchings between the persistence diagrams $\hat{\mathcal{D}}_\phi(a, b)$, $\hat{\mathcal{D}}_\psi(a, b)$ are COHERENT to each other as a and b change. This means that the classical definition does not require that the matchings change “continuously” when a and b change continuously. To proceed in our proof we needed this property.

A naïve (and wrong) solution to the coherence problem



Our first (and naïve) idea to get coherent matchings was the following one.

- We should fix a pair (\bar{a}, \bar{b}) and a matching $\sigma_{\bar{a}, \bar{b}}$ between $\hat{\mathcal{D}}_{\varphi}(\bar{a}, \bar{b})$ and $\hat{\mathcal{D}}_{\psi}(\bar{a}, \bar{b})$.
- Then we should construct a coherent family of matchings $\sigma_{a,b}$ between $\hat{\mathcal{D}}_{\varphi}(a,b)$ and $\hat{\mathcal{D}}_{\psi}(a,b)$, by “transporting $\sigma_{\bar{a}, \bar{b}}$ to any other point (a,b) in the parameter space by continuity”.
- This transportation would use the stability of the normalized persistence diagrams $\hat{\mathcal{D}}_{\varphi}(a,b)$ and $\hat{\mathcal{D}}_{\psi}(a,b)$, i.e. the fact that they depend continuously on a and b .



A naïve solution to the coherence problem

The previously proposed approach has a problem. If along the transportation we meet a point (a', b') at which one of the two normalized persistent diagrams contains a multiple point, then our transportation is not well-defined. Indeed, during “collisions” the identity of points is not preserved and we are not able to follow them.

The consequent natural idea is the one of following the movements of points of $\hat{\mathcal{D}}_\varphi(a, b)$ and $\hat{\mathcal{D}}_\psi(a, b)$ avoiding the pairs (a, b) for which points with multiplicity greater than one exist. We call these pairs **singular**. The other pairs are called **regular**.



Let us recall some concepts in persistent homology

The importance of lines with slope 1 in 2D persistent homology

The need for coherent matchings

The phenomenon of monodromy

The coherent 2-dimensional matching distance



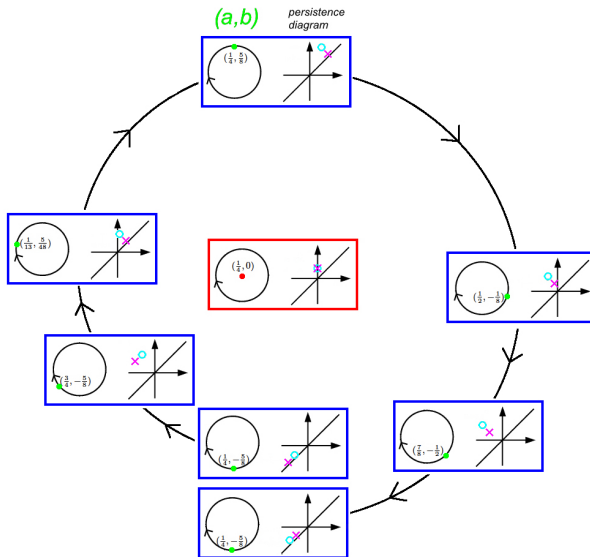
A key difficulty in our naive solution

Unfortunately, the previous approach still has a problem. Indeed, the definition of transport depends on the choice of the path c that we follow in the parameter space. Precisely, it depends on the homotopy class of c relative to the startpoint (\bar{a}, \bar{b}) and the endpoint (a, b) .

We call this fact the **monodromy phenomenon** in 2-dimensional persistent homology.

For more details about the monodromy phenomenon in 2-dimensional persistent homology we refer the interested reader to the paper [A. Cerri, M. Ethier, P. Frosini, *A study of monodromy in the computation of multidimensional persistence*, Proceedings of the 17th IAPR International Conference on Discrete Geometry for Computer Imagery, LNCS 7749, 2013, 192–202].

An example of monodromy in 2D persistent homology





Let us recall some concepts in persistent homology

The importance of lines with slope 1 in 2D persistent homology

The need for coherent matchings

The phenomenon of monodromy

The coherent 2-dimensional matching distance

The monodromy phenomenon as an obstruction



The existence of monodromy implies that each loop in the space of regular pairs induces a permutation on $\hat{\mathcal{D}}_\varphi(\bar{a}, \bar{b})$. In other words, we cannot establish which point in $\hat{\mathcal{D}}_\varphi(\bar{a}, \bar{b})$ corresponds to which point in $\hat{\mathcal{D}}_\varphi(a, b)$, since the answer depends on the path that we follow from (\bar{a}, \bar{b}) to (a, b) in the space of regular pairs. As a consequence, different paths going from (\bar{a}, \bar{b}) to (a, b) can “transport matchings in different ways”.

Does monodromy prevent us from proceeding in our research and from defining a coherent 2-dimensional matching distance?

Fortunately, the answer is no, as we will show in the next slides.

(We will skip the mathematical details. The interested reader can find them in the paper.)



Coherent 2D matching distance

The Coherent 2D matching distance $CD_{match}(\varphi, \psi)$ between φ and ψ is defined by this procedure:

- Let us fix a pair (\bar{a}, \bar{b}) that is regular both for φ and ψ ;
- For each matching $\sigma_{(\bar{a}, \bar{b})}$ between the normalized persistence diagrams $\hat{\mathcal{D}}_{\varphi}(\bar{a}, \bar{b})$ and $\hat{\mathcal{D}}_{\psi}(\bar{a}, \bar{b})$ and every regular pair (a, b) , we consider the set $T_{(a,b)}(\sigma_{(\bar{a}, \bar{b})})$ of every possible transportation of $\sigma_{(\bar{a}, \bar{b})}$ to (a, b) . In order to do that, we recall that we have to consider only one path for each homotopy class relative to the startpoint (\bar{a}, \bar{b}) and the endpoint (a, b) , in the set of paths not touching singular pairs. We set $\text{cost } T_{(a,b)}(\sigma_{(\bar{a}, \bar{b})}) := \text{maximum cost of a matching in } T_{(a,b)}(\sigma_{(\bar{a}, \bar{b})})$.
- We define $CD_{match}(\varphi, \psi) = \inf_{\sigma_{(\bar{a}, \bar{b})}} \sup_{(a,b)} \text{cost } T_{(a,b)}(\sigma)$.



Properties of CD_{match}

Proposition

The definition of $CD_{match}(\varphi, \psi)$ does not depend on the choice of the regular pair (\bar{a}, \bar{b}) .

Proposition

$CD_{match}(\varphi, \psi)$ is a pseudo-distance.

Theorem

$$D_{match}(\varphi, \psi) \leq CD_{match}(\varphi, \psi) \leq \|\varphi - \psi\|_{\infty}.$$



Recap

- Examples suggest that in 2-dimensional persistent homology just filtrations defined by **lines with slope 1** might be relevant.
- The attempt of proving the previous statement leads to modify the definition of 2-dimensional matching distance by requiring that the matchings associated with the lines are **coherent** to each other.
- The search for coherent families of matchings leads to discover the phenomenon of **monodromy** in 2-dimensional persistent homology, as an obstruction to construct such families.
- The transport of matchings along paths allows to overcome that obstruction and to introduce the **coherent 2D matching distance**.
- We are working on the proof that just filtrations defined by **lines with slope 1** are relevant for the computation of the coherent 2D matching distance.



THANKS
FOR YOUR
ATTENTION!