

TDA - Fall 2025

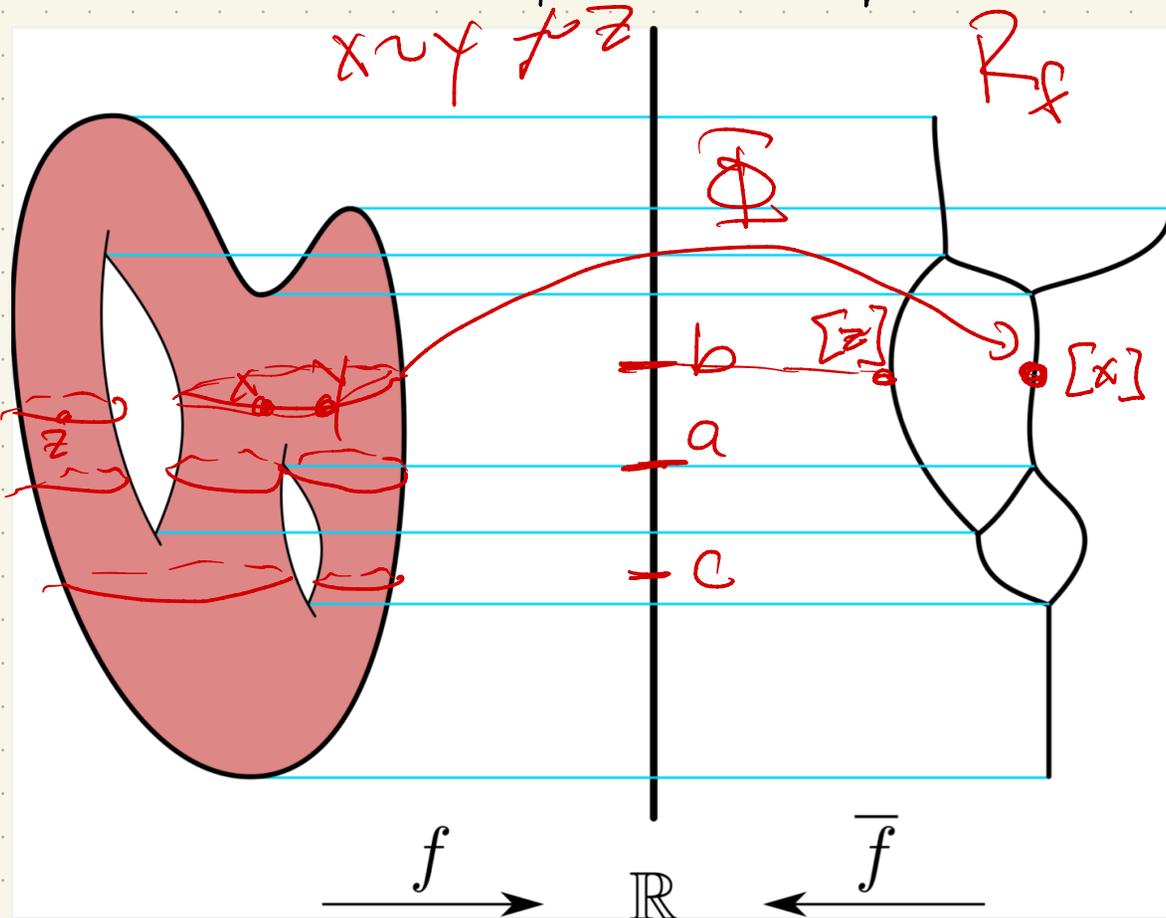
More on Reeb  
Graphs



# Reeb Graph Definition

Given a function  $f: X \rightarrow \mathbb{R}$ ,  
relation  $\sim$  by  $x \sim y$  iff:

- define an equivalence
- $f(x) = f(y) = a$
  - $x + y$  in same connected component of level set  $f^{-1}(a)$ .



Let  $[x]$  be equiv.  
class of  $x$ :  
 $\{y \in X \mid y \sim x\}$

The Reeb graph  $R_f$  is the quotient space  $X/\sim$ .  
Let  $\Phi: X \rightarrow R_f$ ,  $x \mapsto [x]$  be quotient map



# Discrete versus continuous viewpoints

Graph  $G = (V, E)$

$V$ : vertex set

$E$ : edges =  $\{[u, v]\}$

$w \in E \implies f(w) \neq f(v)$

Essentially, a 1-dim  
abstract simplicial  
complex

function  $f$ : defined on  
vertices

"Topological" graph

- Vertices

- Edges  $\simeq$  interval

Essentially, a 1-dim  
stratified space

$\hookrightarrow$  1 manifold  
except at  
singularities



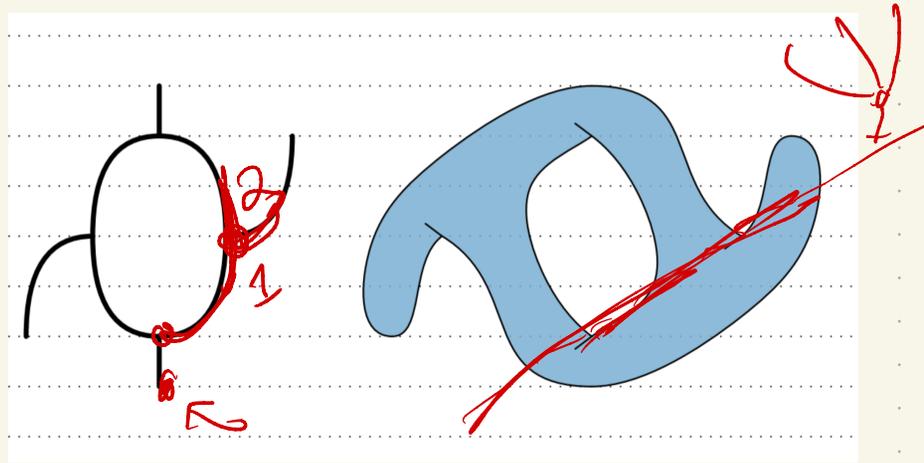
interpolate  
along  
edge  $f(v)$

# Vertices in $R_f$

Given a node  $x \in V$  (where  $V = \underline{V(R_f)}$ ),

• up-degree of  $x$ : # of edges  $xu$   
 where  $f(u) > f(x)$

• down degree of  $x$ :  
 # of edges  $xu$   
 where  $f(u) < f(x)$



A node is

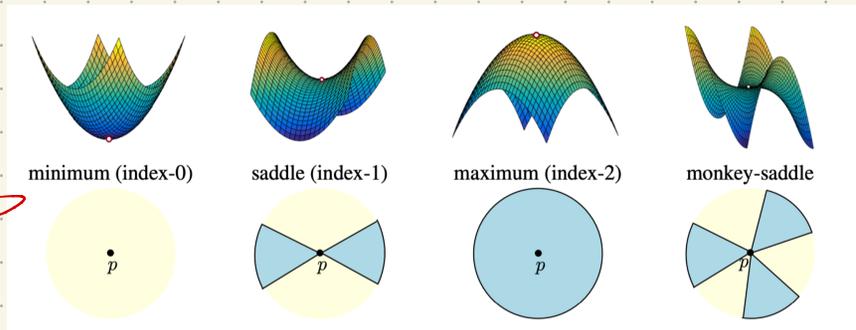
• regular if up & down degree both = 1

• critical otherwise



→ only considers critical vertices

[Aside: Why?]

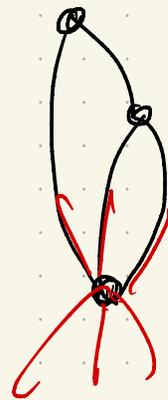
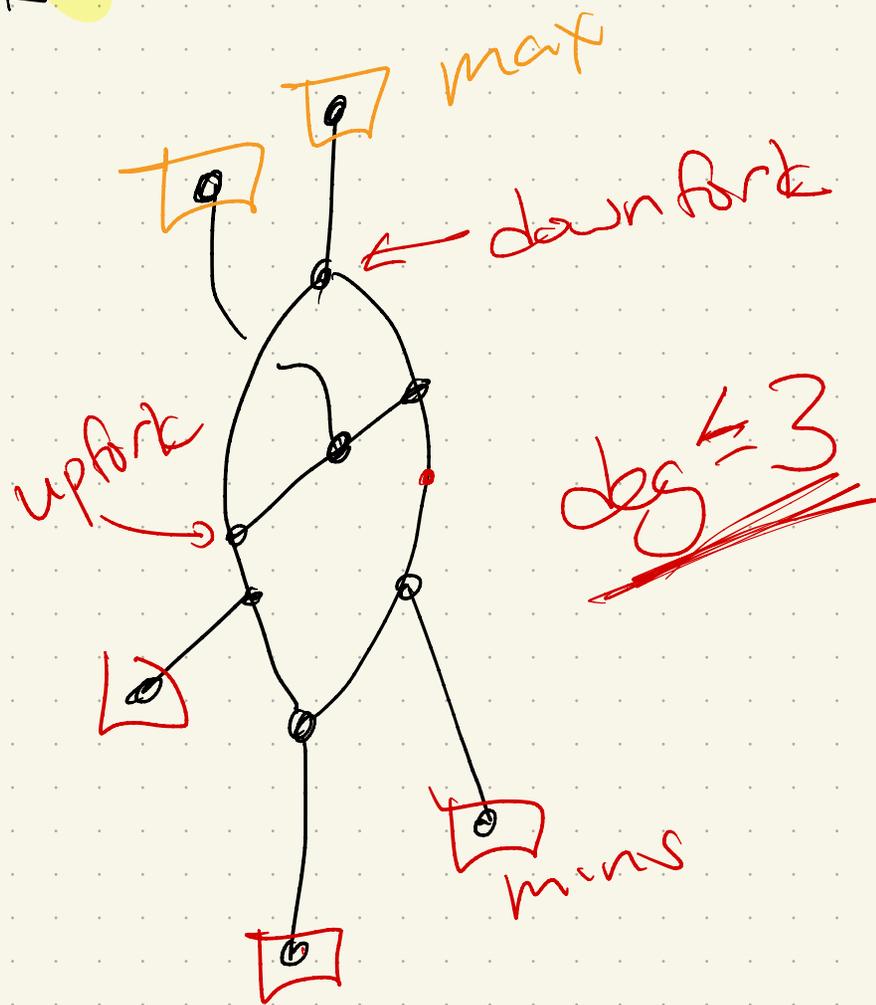


# Types of critical points

A critical point is a

- minimum if it has down-degree = 0
- maximum if it has up-degree = 0
- a downfork if it has down-degree  $> 1$
- an upfork if it has up-degree  $> 1$

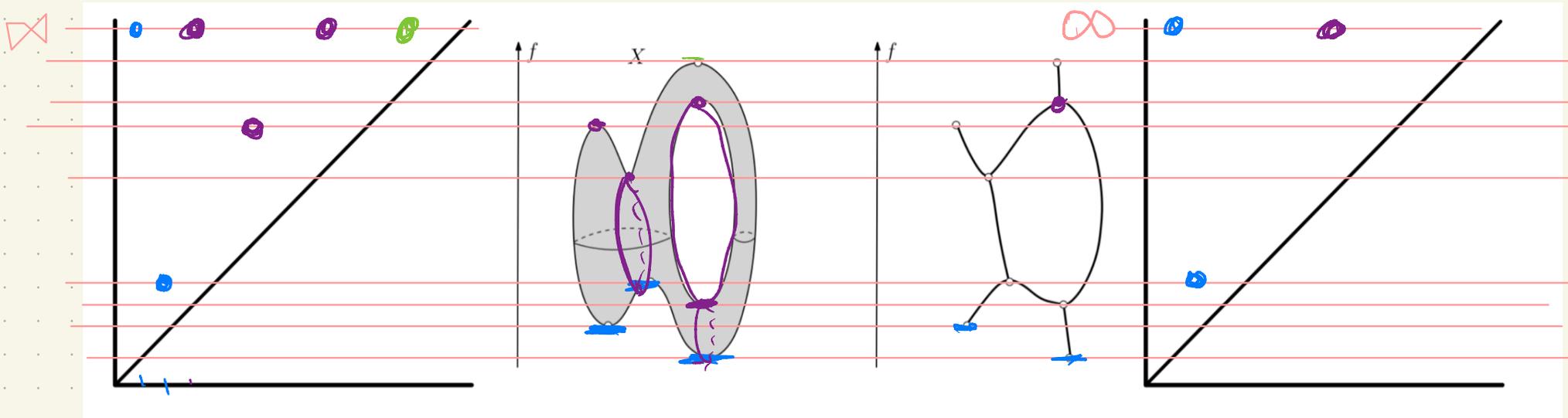
Note: a degenerate critical point may have more than one type of criticality



# Persistence & Betti Numbers

Theorem: For a tame function  $f: X \rightarrow \mathbb{R}$   
 $\beta_0(X) = \beta_0(K_f)$  and  $\beta_1(X) \geq \beta_1(K_f)$

Why? Look at "standard" persistence:



0-dim: 2 points  
1-dim: 3 points  
2-dim: 1 point

2

# Special settings: 2-manifolds

Theorem: Let  $f: X \rightarrow \mathbb{R}$  be a Morse function on a connected compact 2-manifold

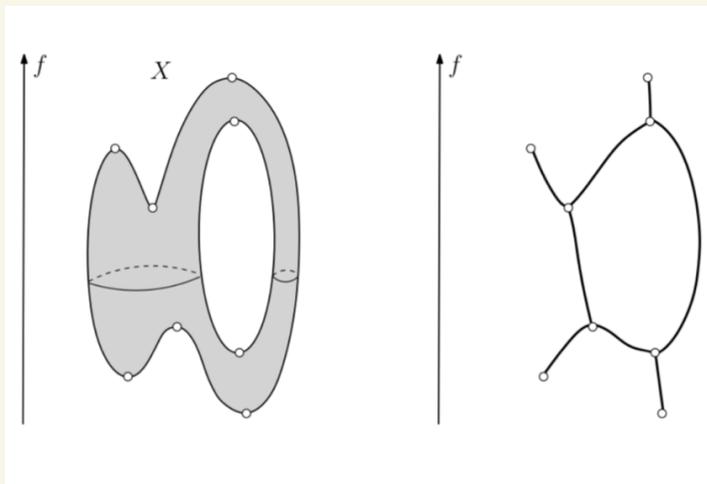
Then if  $X$  is orientable,

$$\beta_1(\mathbb{R}_f) = \beta_1(X)/2$$

If  $X$  is non orientable,

$$\beta_1(\mathbb{R}_f) \leq \beta_1(X)/2$$

Recall:



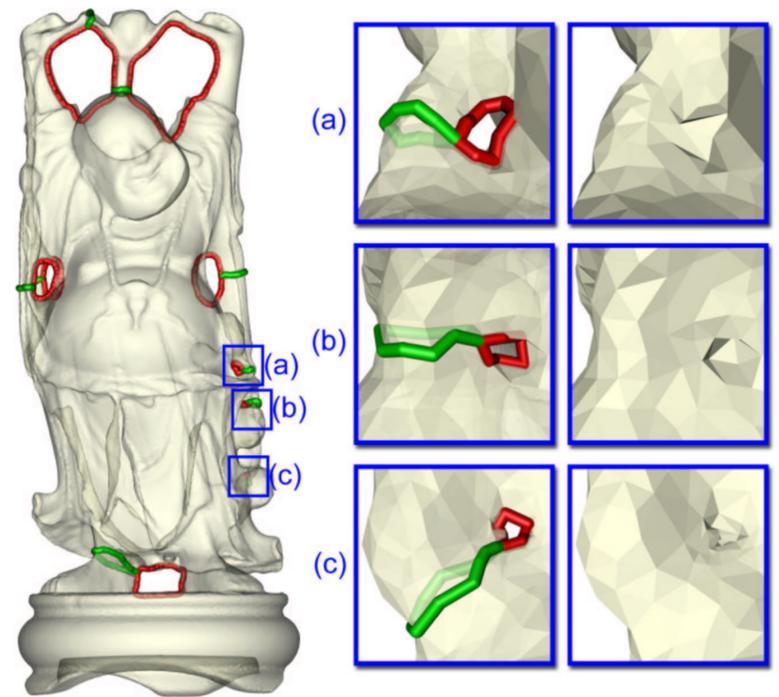
Leads to an interesting classification  
of cycles:  
vertical versus horizontal  
(with respect to  $f$ )

- gives improved algorithms Dey-Wang 2012
- & applications to surface reconstruction



**Figure 7:** From left to right: handle and tunnel loops computed on Casting, tunnel features are identified, small tunnel features are filled.

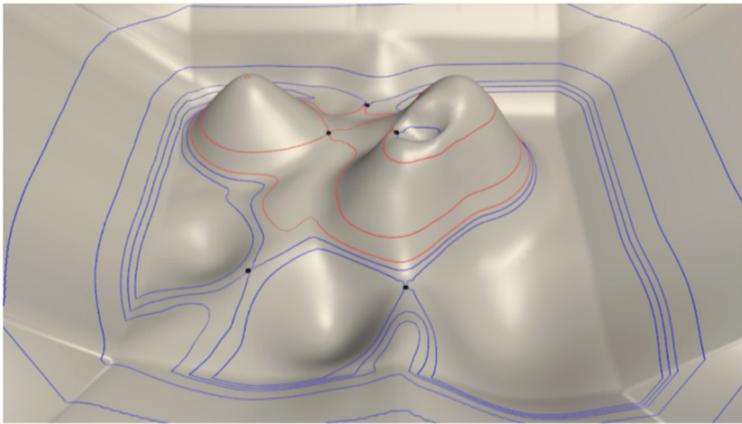
Dey et al 2008



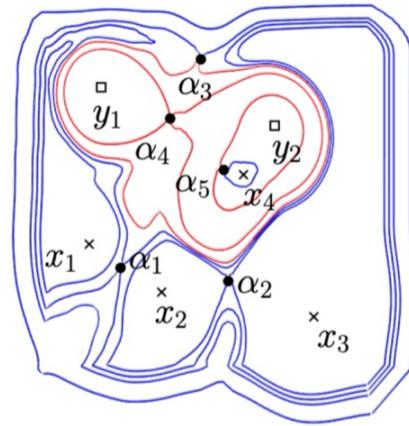
**Figure 8:** Three small tunnels in Buddha are filled.

# Simplifications

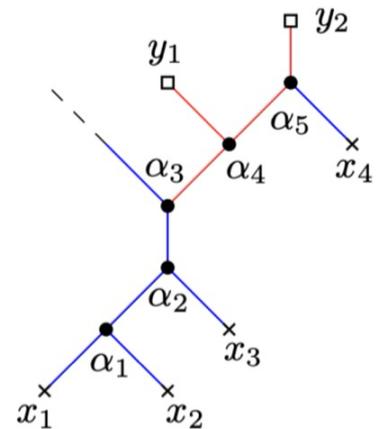
If  $X$  has no  $H_1$ , then resulting graph is called a contour tree!



(a)



(b)

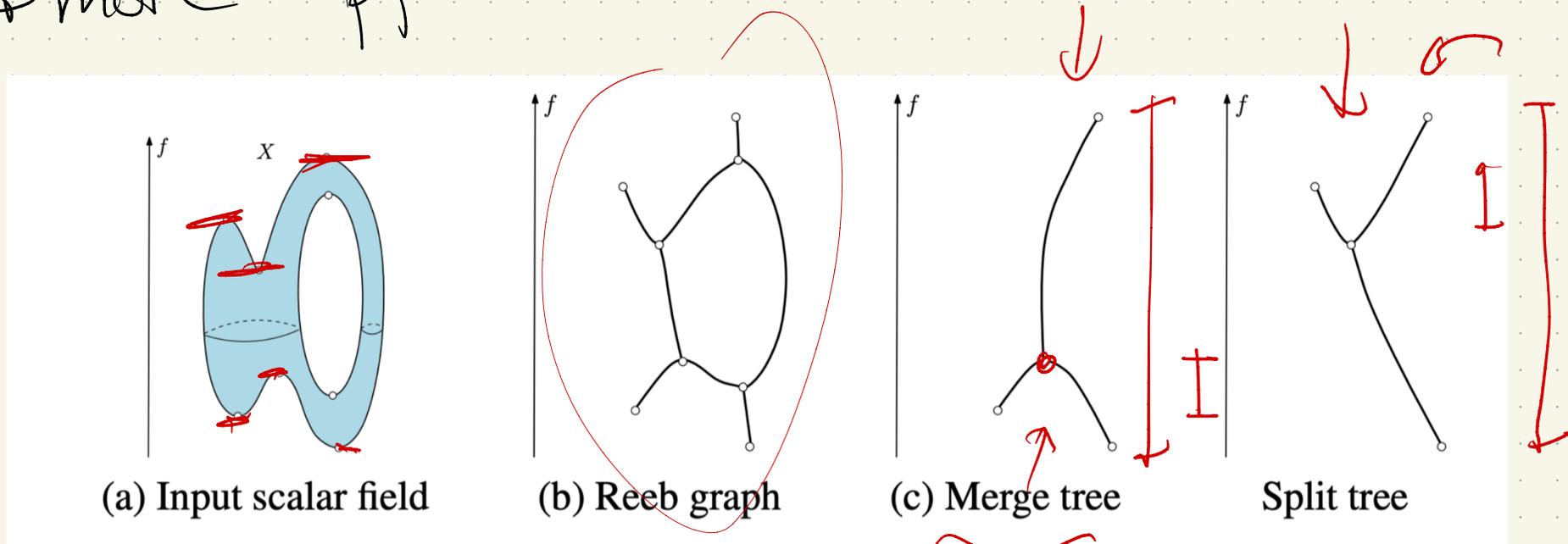


(c)

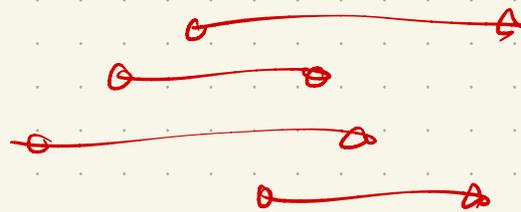
Agarwal et al 2014

Why we care: usually faster

In fact, simpler versions of Reeb graphs are very popular in visualization & more applied work!



- Merge trees: just Ho of sublevel sets
- Split trees: same but superlevel sets



Some examples: Yan et al 2021 ← survey

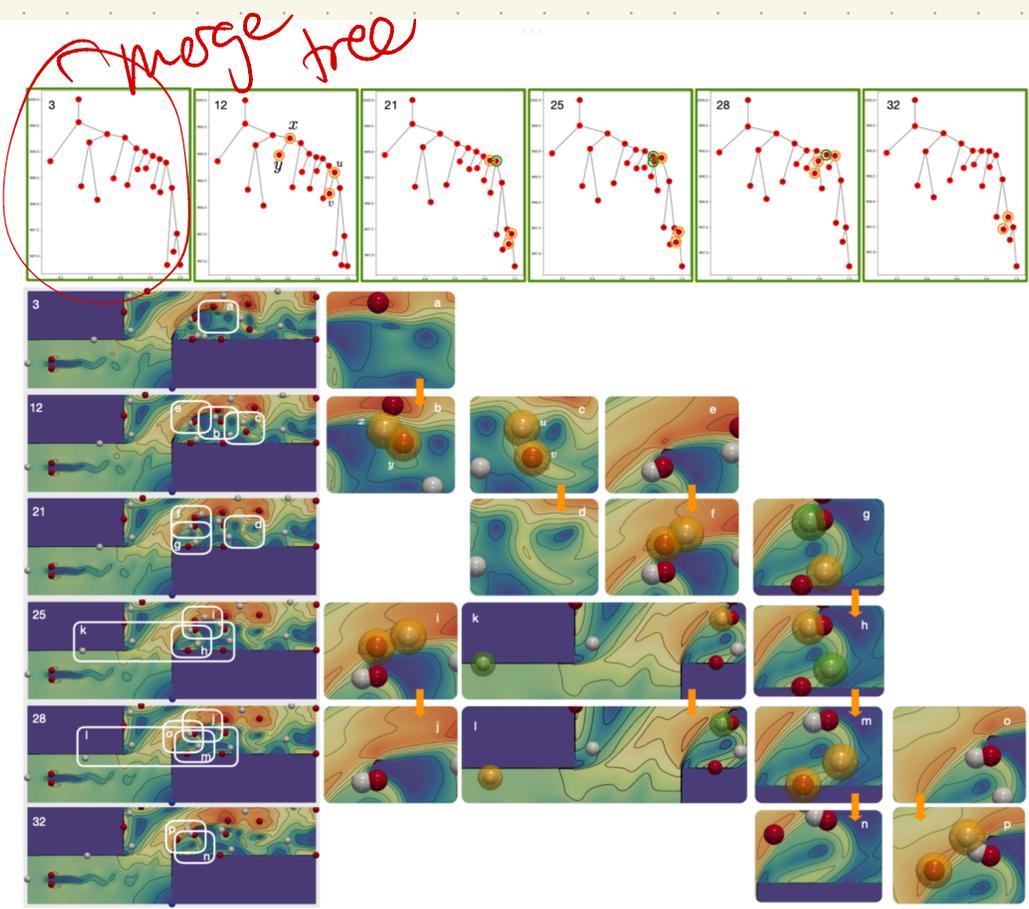


Fig. 11: Sketching the *Corner-Flows* dataset with 15 basis trees with IFS: (Top) first 6 basis trees where orange circles highlight topological changes w.r.t. near basis trees, (Bottom Left) scalar fields that give rise to these basis trees, areas with critical points appearances/disappearances are shown with zoomed views in (Bottom Right).

Fasy et al 2024

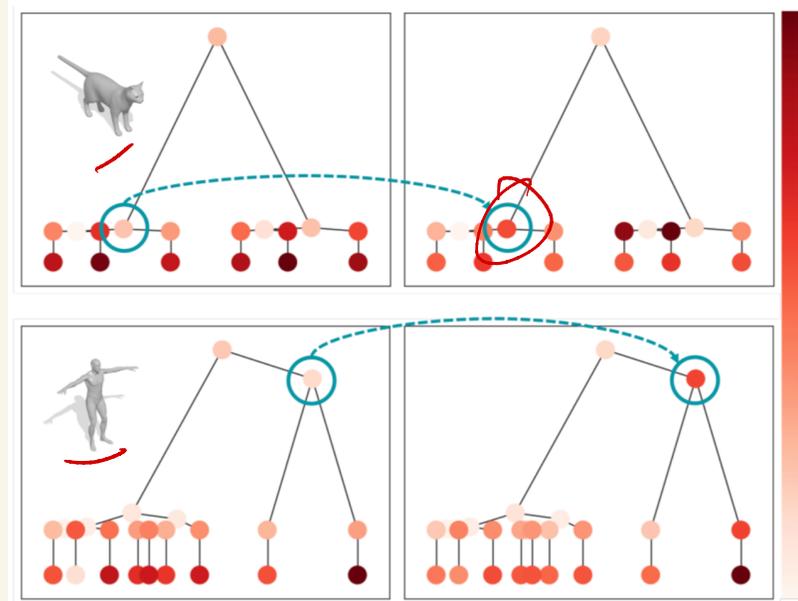
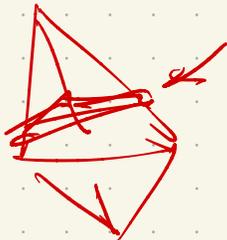


Fig. 13: Topological attention on two merge trees from the TOSCA dataset. Each row displays the merge tree visualized by the relative importance, with example inserts within the merge trees. Left: node importance without topological attention, using the GIN model. The node colors indicate their relative importance when comparing the two examples (the darker with red, the higher the weight). Right: node importance with topological attention using the MTNN model. The GIN model already emphasizes structural differences between the two merge trees. But, the introduction of topological attention further emphasizes these differences, re-weighting the nodes in the high persistence feature highlighted in cyan.

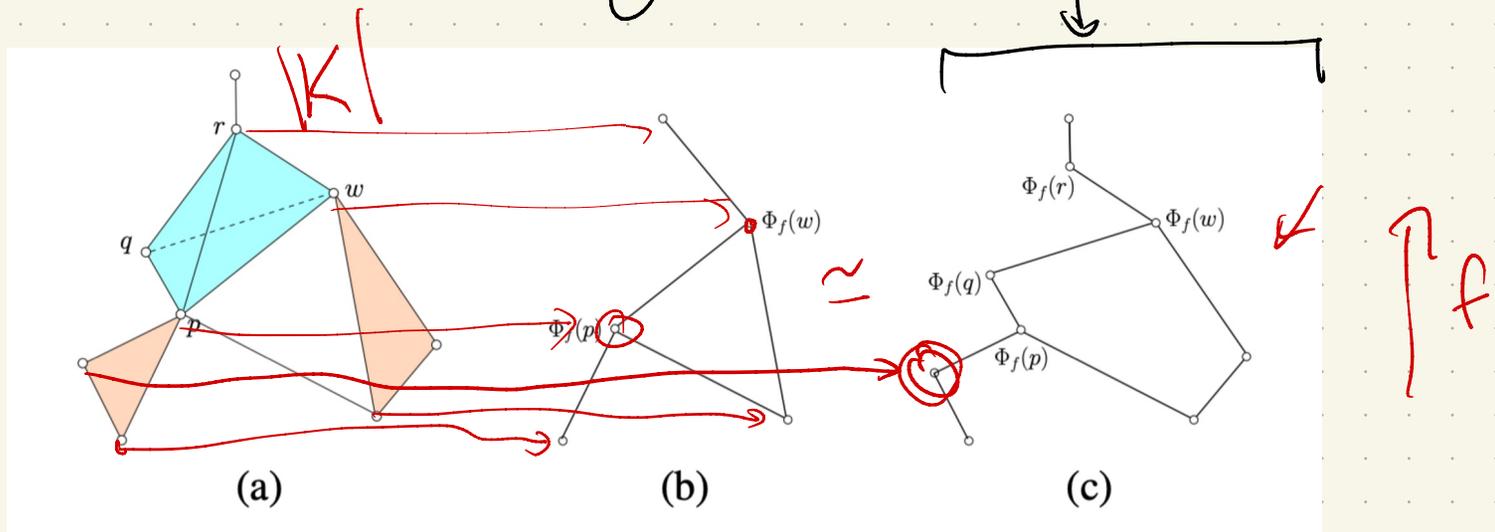
Li et al 2025

# Computation

Given a simplicial complex  $K$  & a PL function  $f: |K| \rightarrow \mathbb{R}$ ,  $R_f$  depends only on 2-skeleton.  
Why?  $\Rightarrow$  won't affect levelset



Augmented Reeb graph:



Sweep-line algorithm:

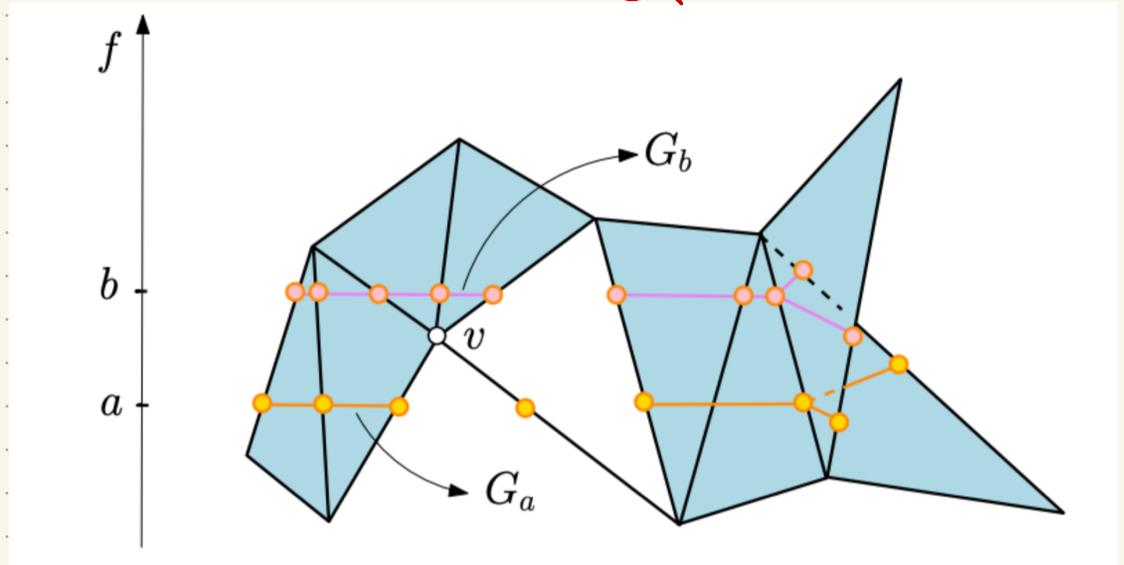
- Sort vertices by  $f$

- Track connected components

- Graph changes only at vertices;  $\neq$  then only at simplices adjacent to vertex

$$O(m \log m)$$

$\uparrow$  # of simplices of dim  $\leq 2$



Data structure:

Track "forest", allowing edge insertion & removal

$\hookrightarrow$  dynamic trees, so

$O(\log m)$  each update

Question: Given 2 Reeb graphs, how to compare them?

Goal: At least an extended pseudo metric:

$$1) d(X, X) = 0$$

$$2) d(X, Y) = d(Y, X)$$

$$3) d(X, Z) \leq d(X, Y) + d(Y, Z)$$

Weaker than metric:

- extended: could be  $\infty$

- pseudo:  $d(x, y) = 0 \not\Rightarrow x = y$

## More Considerations:

- ① Stability: "Perturbed" spaces should have small distance:  $d(R_f, R_g) \leq \|f - g\|$
- ② Discriminativity: How well can it capture differences compared to alternatives, or some baseline?
- ③ Isomorphism invariance:  
Can we have  $R_g = R_f \Leftrightarrow d(R_g, R_f) = 0$ ?
- ④ Path component tolerance:  
For some measures,  $d(R_g, R_f) = \infty$   
 $\Leftrightarrow$  different # of path components.
- ⑤ Universality: (more later...)

Options Several have been studied.

### Edit distances

Based on more combinatorial algorithms, modified to use weights.

### Interleaving distances

Based on more categorical approach.

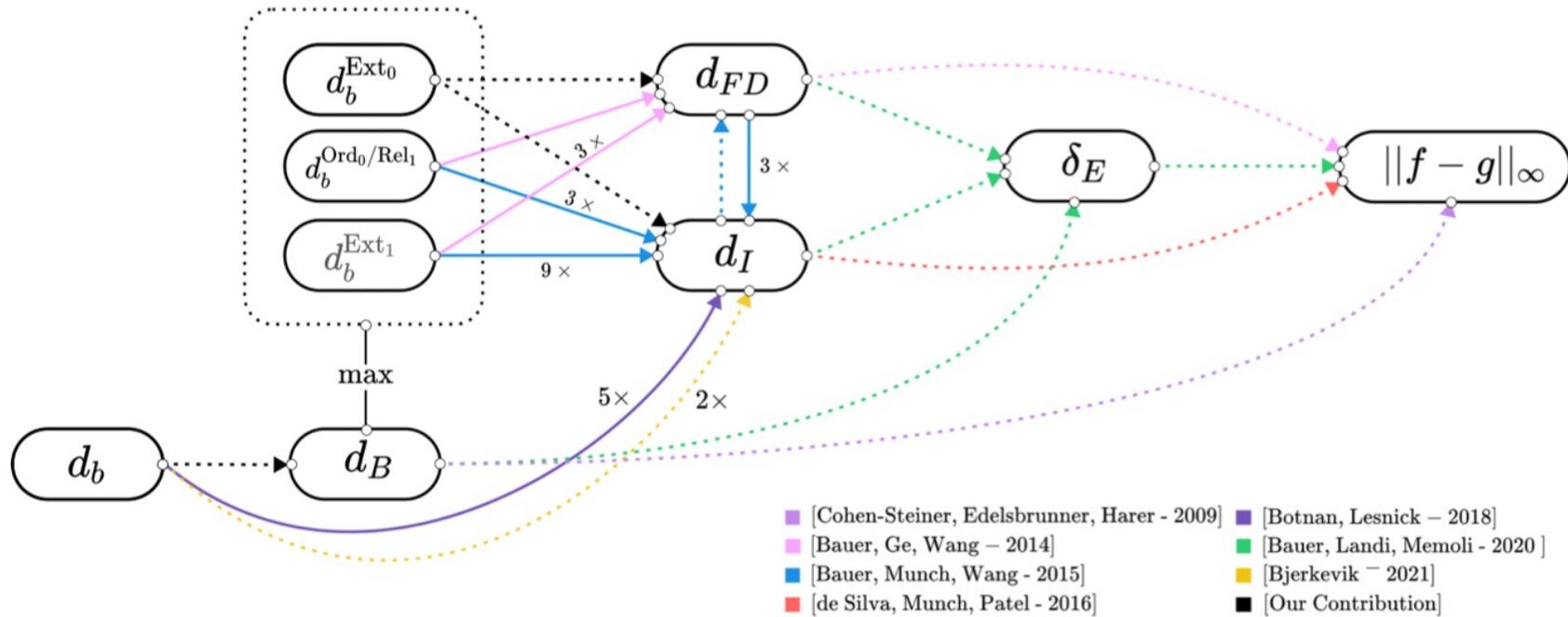
### Functional distortion

Based on ideas from Gromov-Hausdorff distances.

### Bottleneck distances

Based on persistent homology of the Reeb graph.

Surprisingly, these are often related!

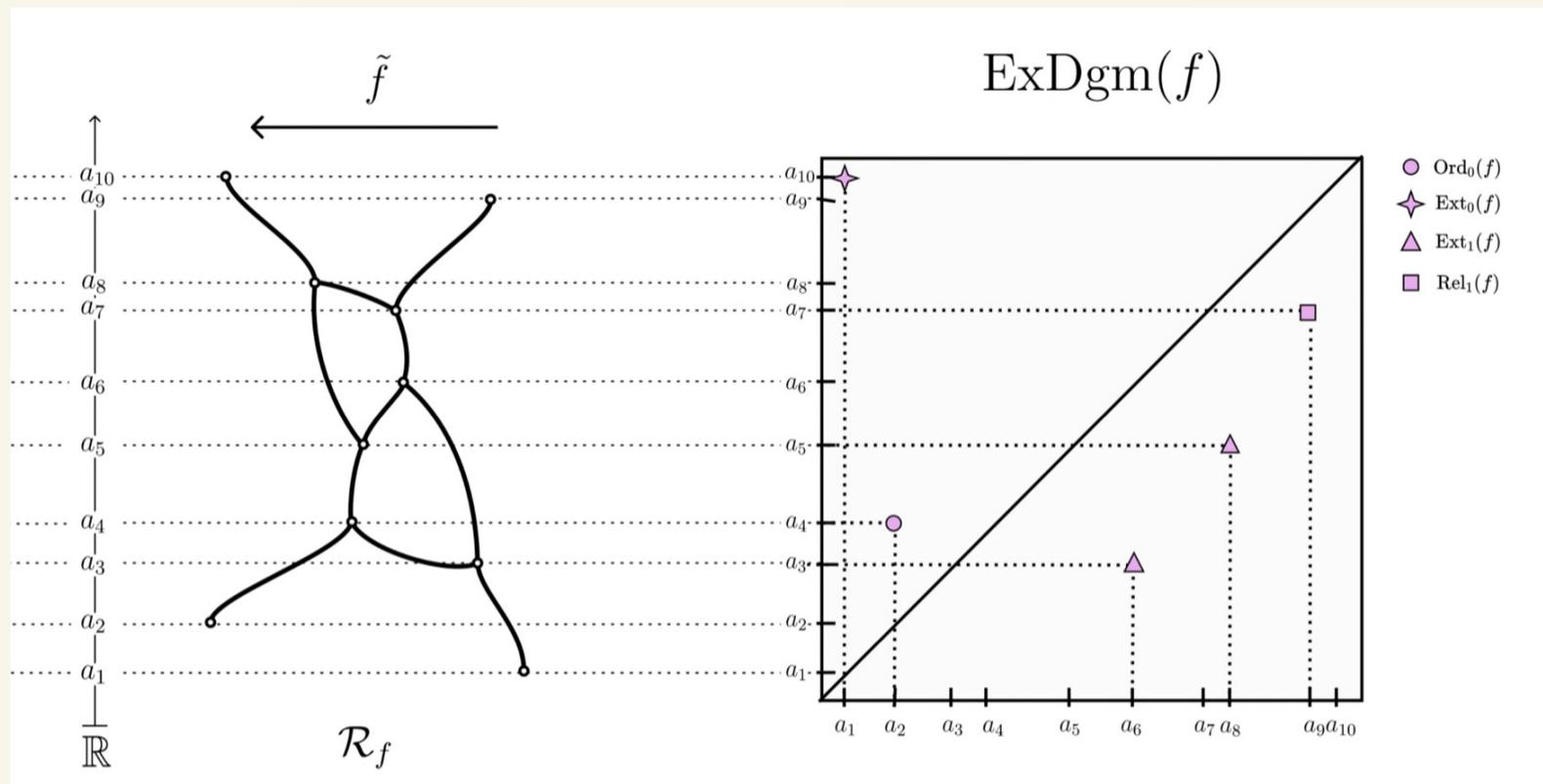


Lets look more carefully...

# Bottleneck distance:

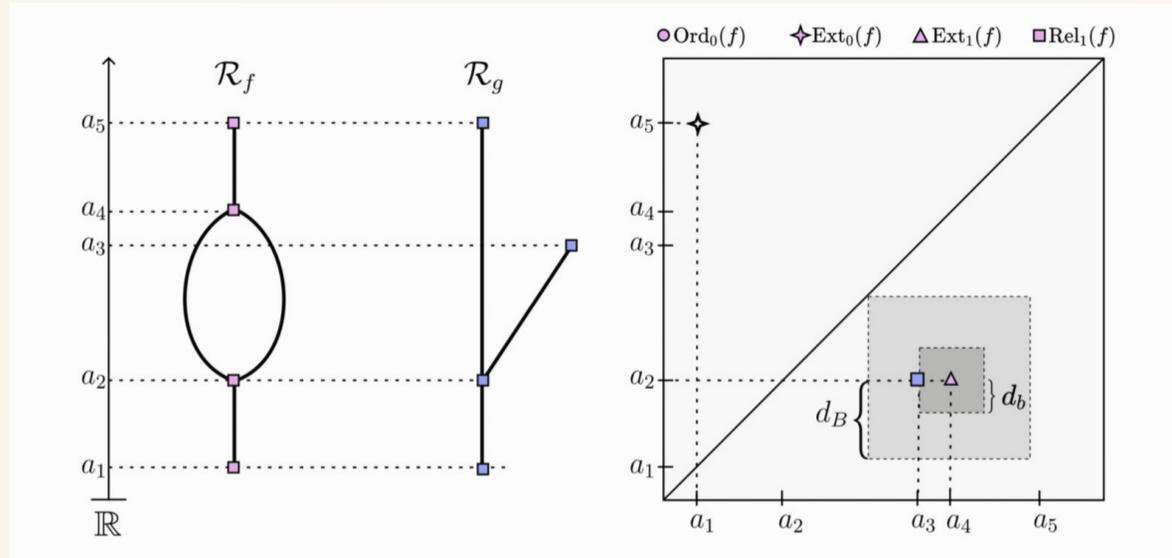
Based on persistence or extended persistence of the Reeb graph (NOT input space):

Recall:



# Back to bottleneck distance:

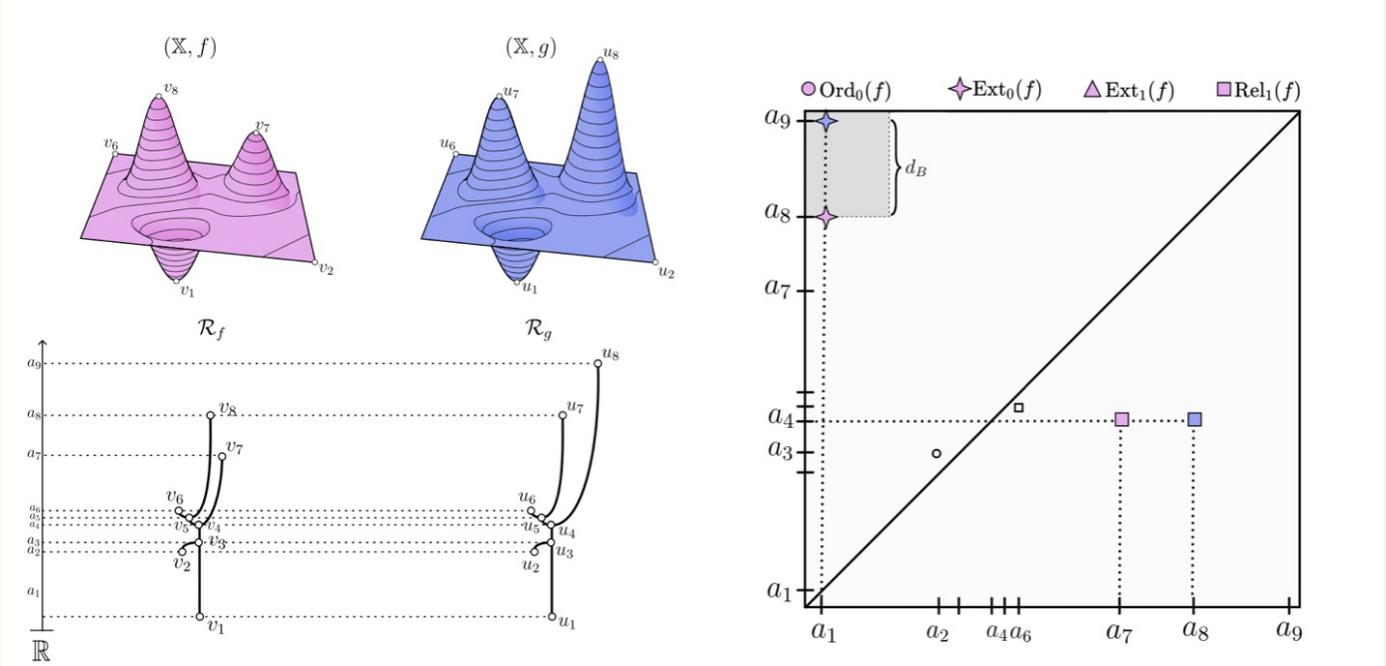
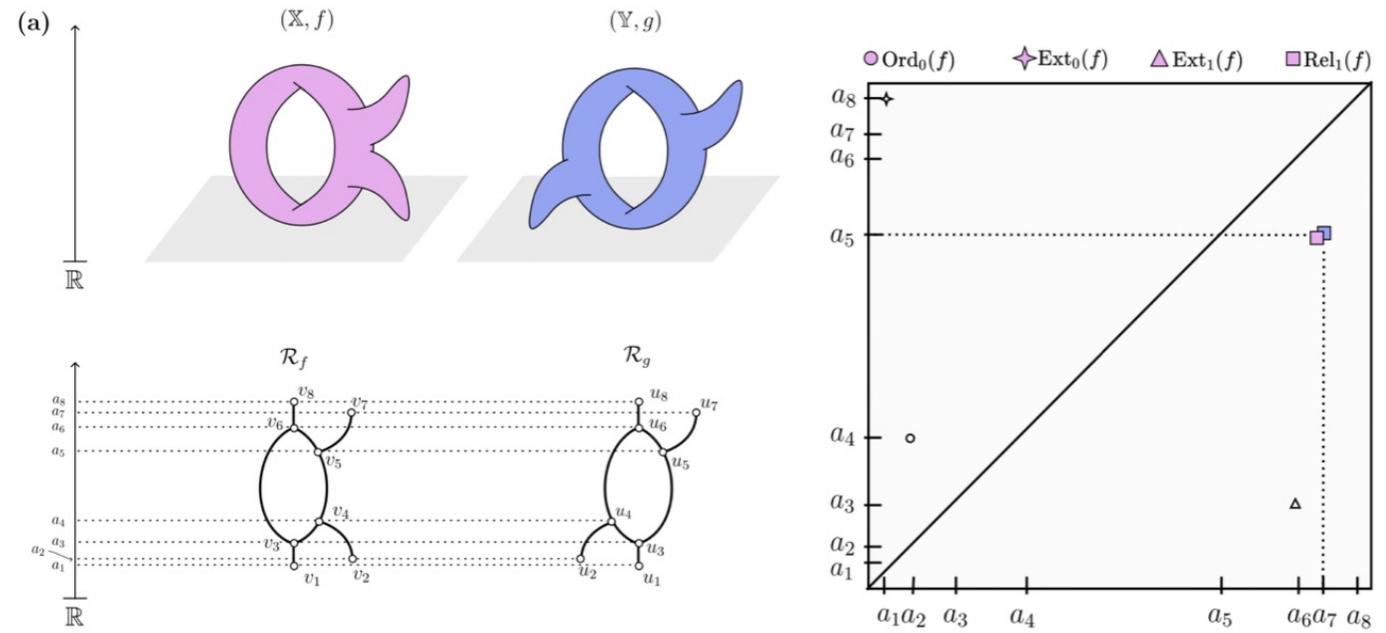
To compare two Reeb graphs, can simplify to their diagrams & compare those:



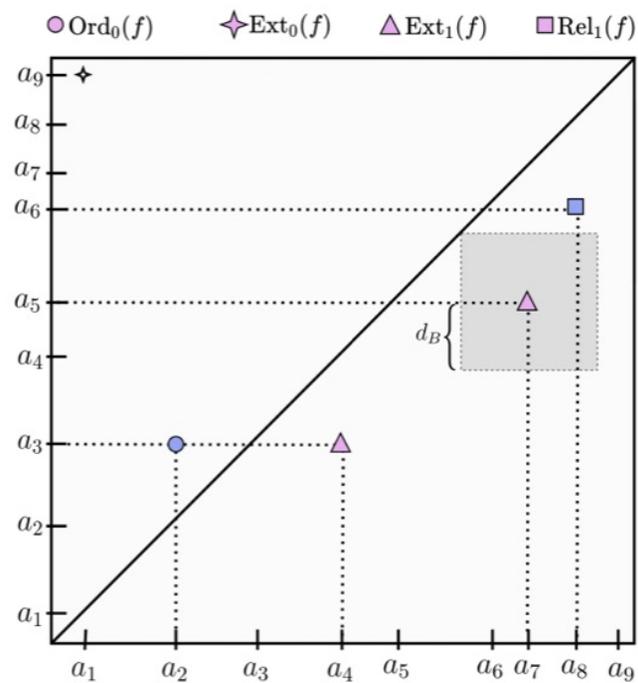
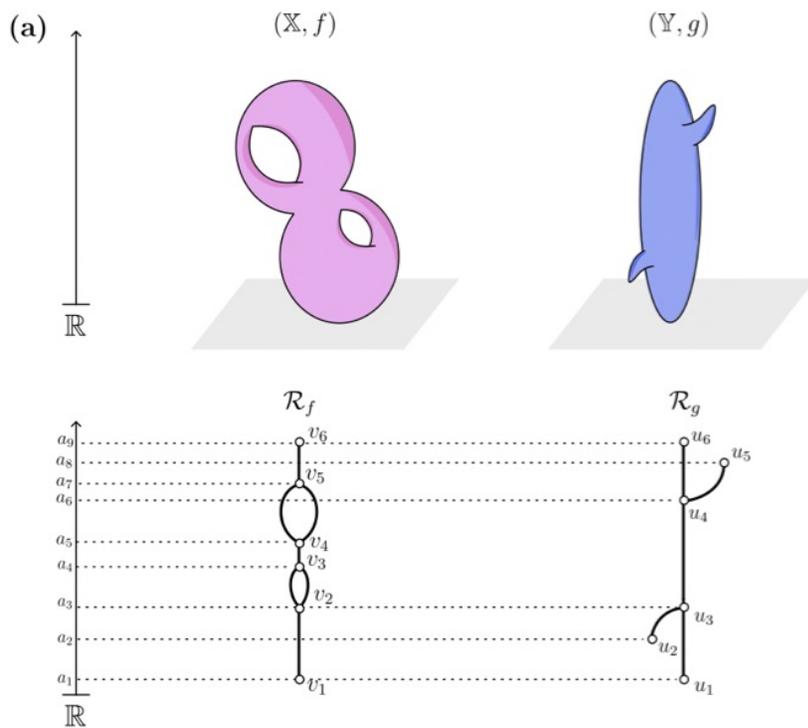
Note:

- Allowed to match any point to diagonal.
- Types can either matter (graded  $d_B$ ) or not matter (ungraded  $d_b$ ).

# Examples



# Examples:



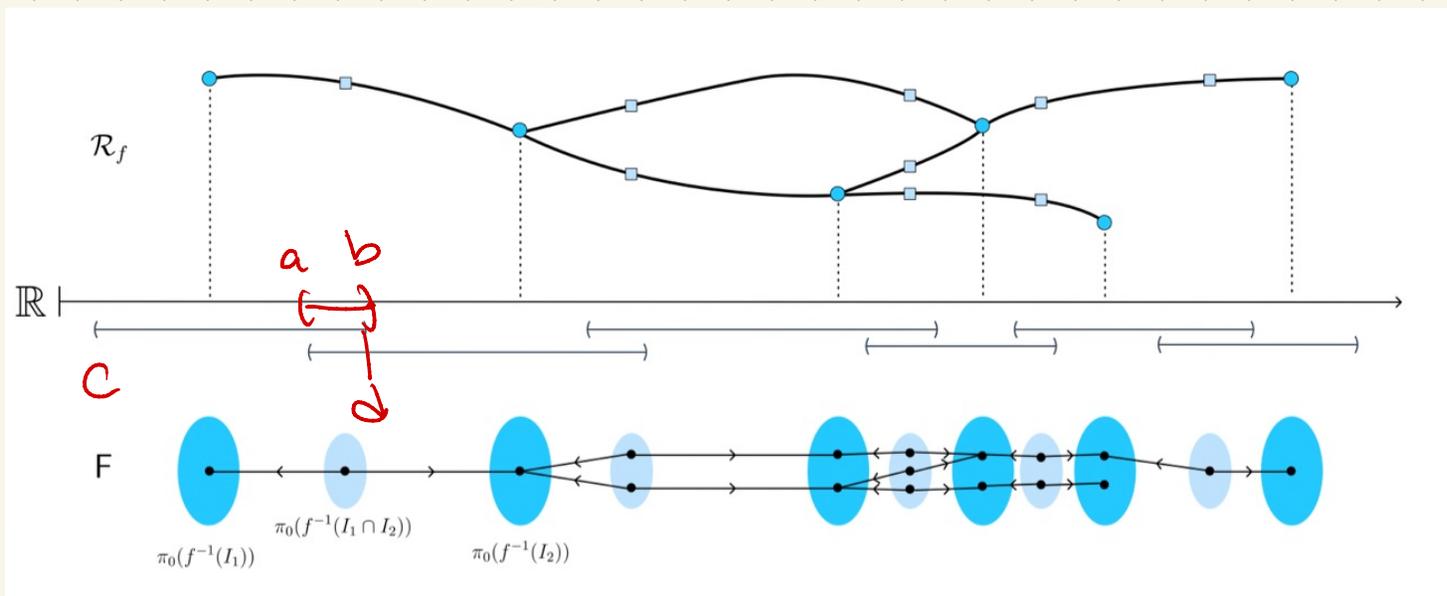
## Bottleneck pros & cons:

- Stable [Cohen-Steiner et al 2007]
- Can compute extended persistence quickly - then it's geometric matchings.  
(Subquadratic - see [Efrot et al 2001] and [Kerber et al 2016])
- It is not isomorphism invariant
- All other metrics are more discriminative [Bauer et al 2014]
- It is not path component sensitive

# Interleaving distance:

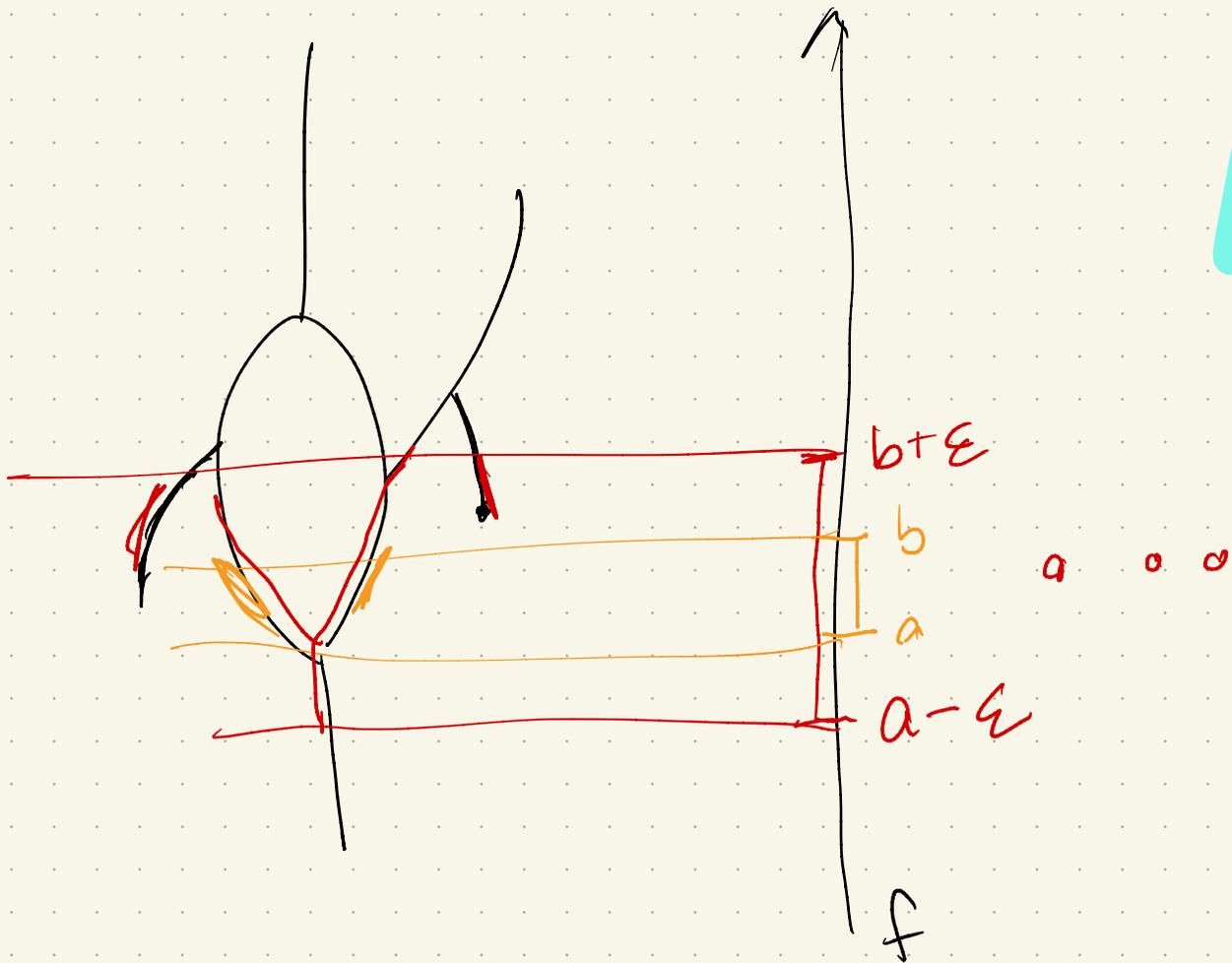
Takes a categorical point of view:  
 Reeb graph is a set valued cosheaf,  
 $F: \text{Int} \rightarrow \text{Set}$

$$\begin{array}{ccc}
 I = (a, b) & \longrightarrow & \pi_0(f^{-1}(a, b)) \\
 \downarrow & & \downarrow F[I \subseteq J] \\
 J = (c, d) & \longrightarrow & \pi_0(f^{-1}(c, d))
 \end{array}$$



# Smoothing a Reeb co-sheaf:

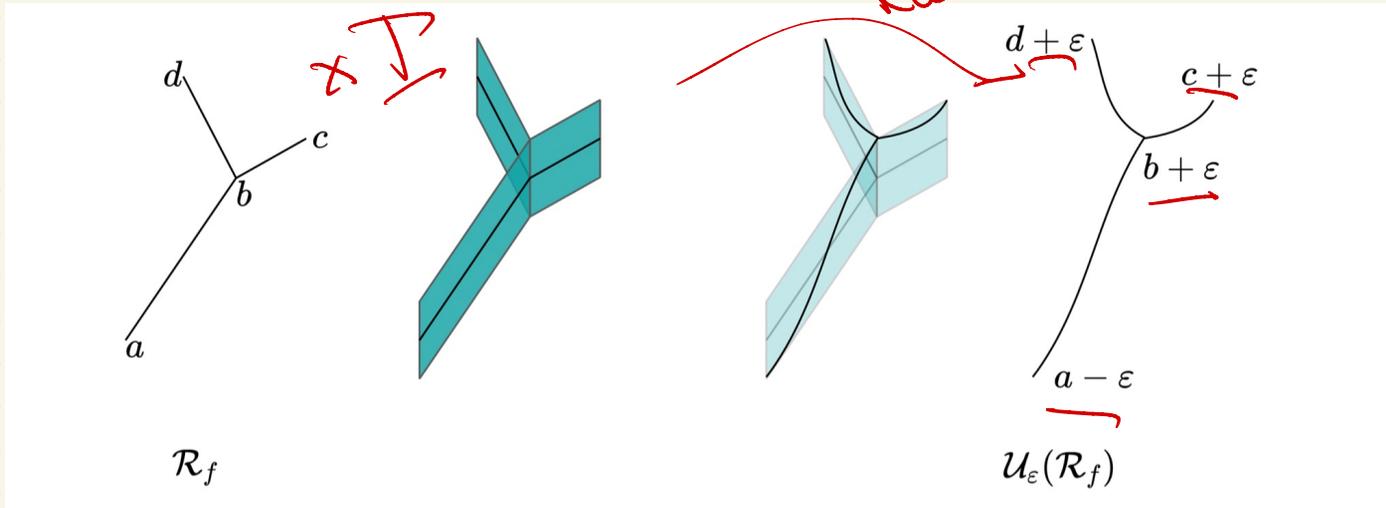
Given  $F: \text{Int} \rightarrow \text{Set}$   
return  $S_\epsilon(F): \text{Int} \rightarrow \text{Set}$   
 $(a, b) \longrightarrow f^{-1}(\pi_0(a-\epsilon, b+\epsilon))$



## Result:

a different  
Reeb co-sheaf,  
related to  
original

Result: In [de Silva et al 2016], they connect this to the thickening function:



Result: Nice connection to a geometric notion of similarity.

# Interleaving distance: Why do we care??

Given 2 Reeb cosheaves  
 $F, G: \text{Int} \rightarrow \text{Set}$

an  $\epsilon$ -interleaving is  $\varphi$  and  $\psi$  such  
that the diagram commutes:

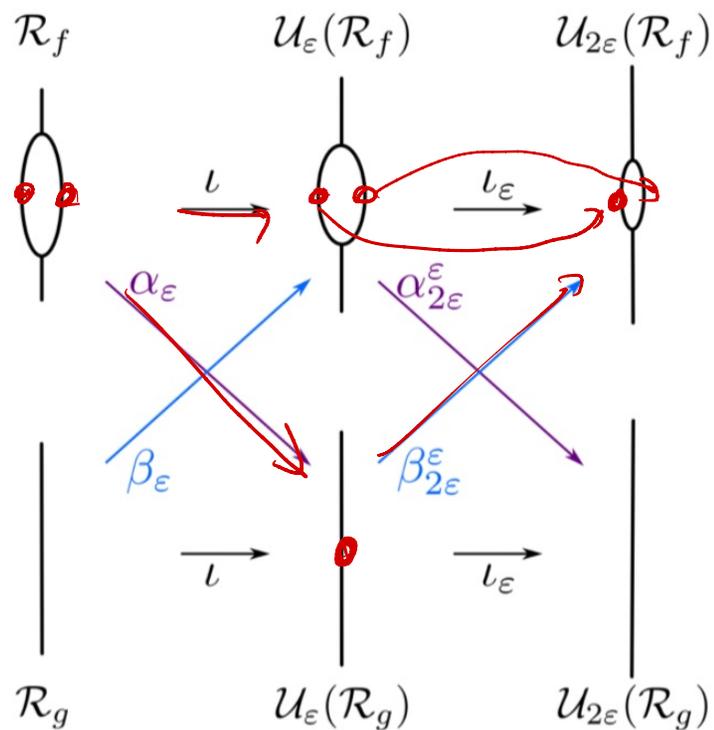
$$\begin{array}{ccccc} F & \longrightarrow & S_\epsilon F & \longrightarrow & S_{2\epsilon} F \\ & \searrow \text{red} & \nearrow \text{blue} & & \\ & & G & \longrightarrow & S_\epsilon G \\ & & & \searrow \text{red} & \nearrow \text{blue} \\ & & & & S_{2\epsilon} G \end{array}$$

Reeb Interleaving distance is

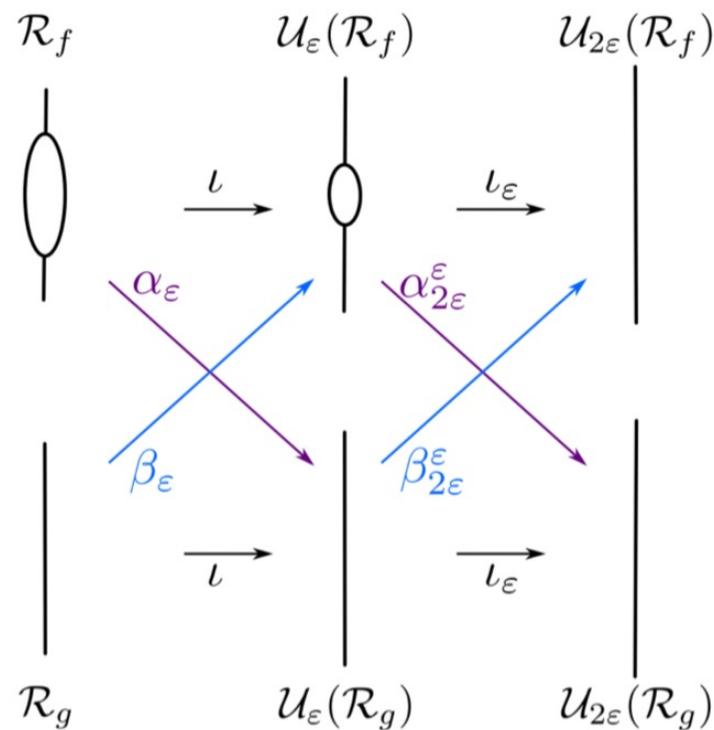
$$d_I(F, G) = \inf \{ \epsilon \geq 0 \mid \exists \epsilon\text{-interleaving} \}$$

# Example $\epsilon$

Bad  $\epsilon$ :



Good  $\epsilon$ :



# Interleaving distance pros & cons:

It is: - discriminative [Bauer et al 2015]

- stable

- path component sensitive

- isomorphism invariant

[de Silva et al, 2016]

Unfortunately, it's Graph-Isomorphism  
hard to compute.

↳ Essentially, can use  $O$ -interleaved  
to check for graph isomorphism.

## Edit distances:

Edit distances are well studied for strings and abstract graphs.

↳ See [Bille 2005] for the many variants on graphs.

In a graph, usually have:

- vertex insertion / deletion
- edge insertion / deletion

and some cost associated with each operation.

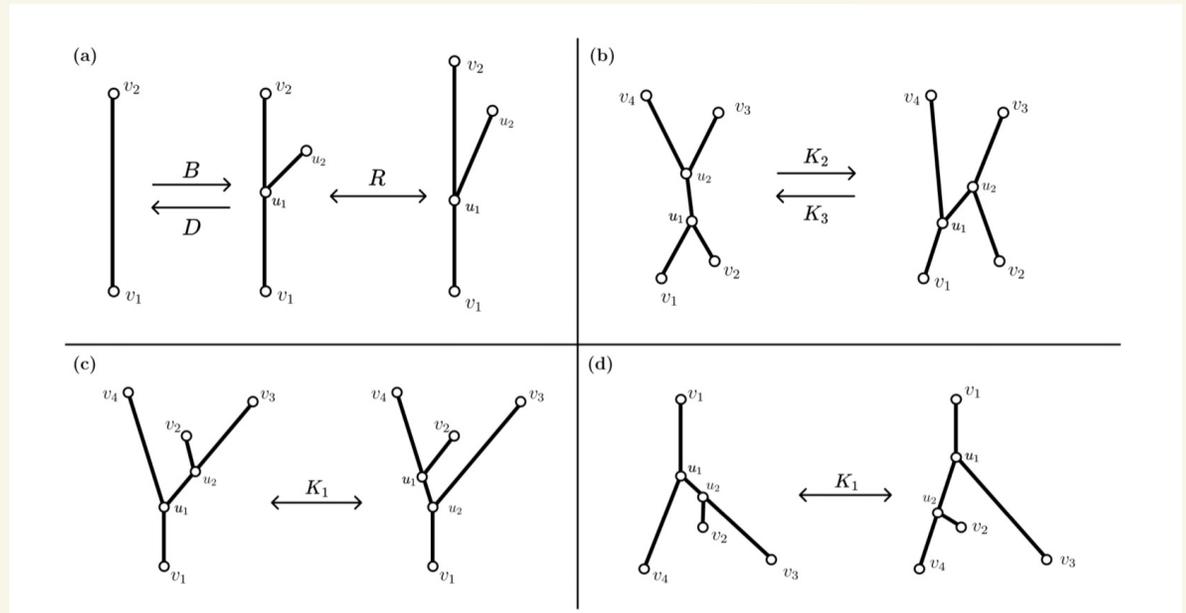
$$\text{Total Edit distance} = \min \left\{ \sum (\text{edit costs}) \right\}$$

# Edits:

In Reeb graphs, only certain edits will correspond to topological deformations on original space.

[Di Fabio & Landi, 2012 and 2016]

- insert
- delete
- relabel
- K-type

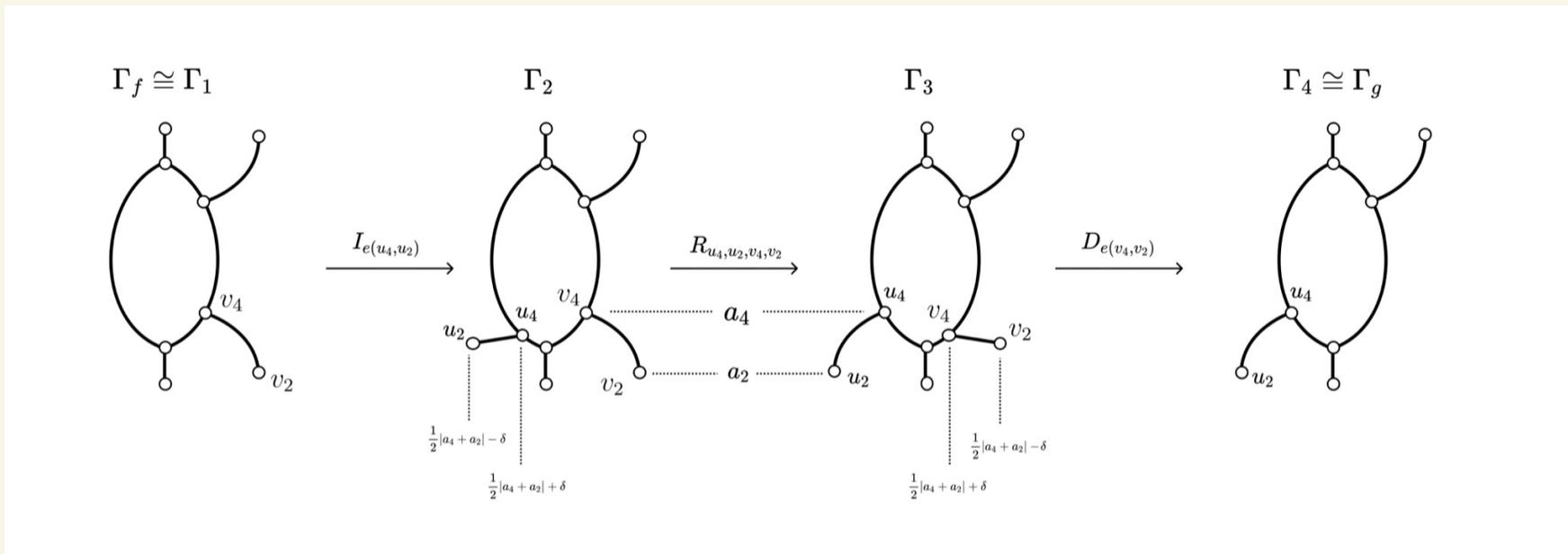


(Reason: Morse theory)

# Downside

Could only compare homeomorphic spaces, originally.

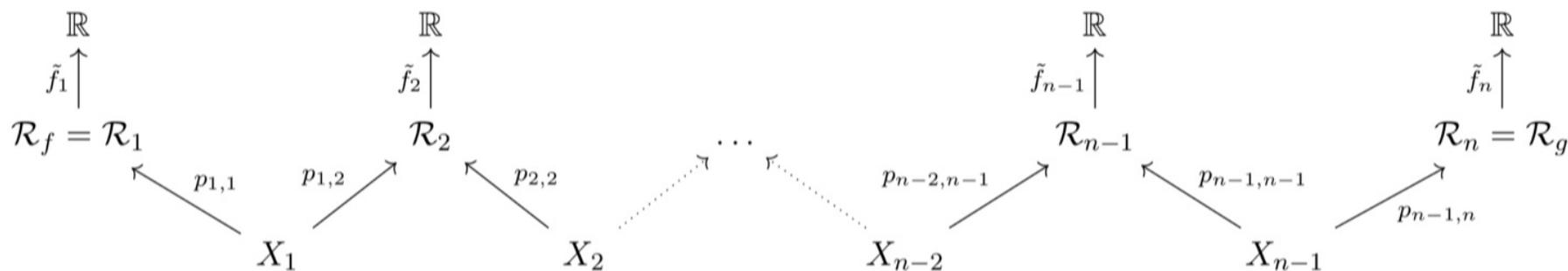
Also a more subtle flaw: can pay less by infinitesimally shifting vertices, since relabels cost max of operations.



# Categorical edit distance

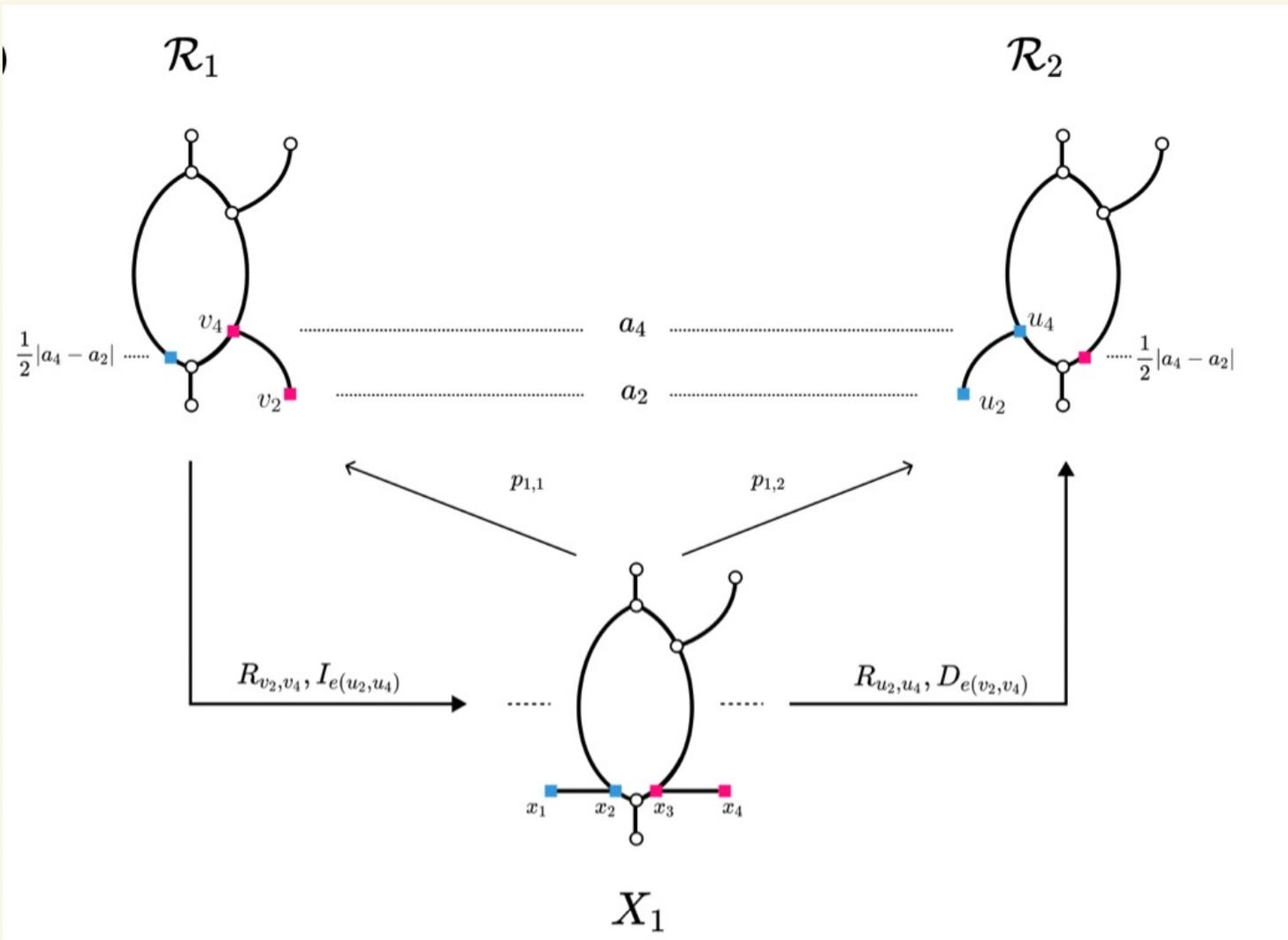
[Bauer et al 2021]

Instead of "direct" edits to Reeb graphs, goal is to find a sequence of topological spaces making this diagram commute.

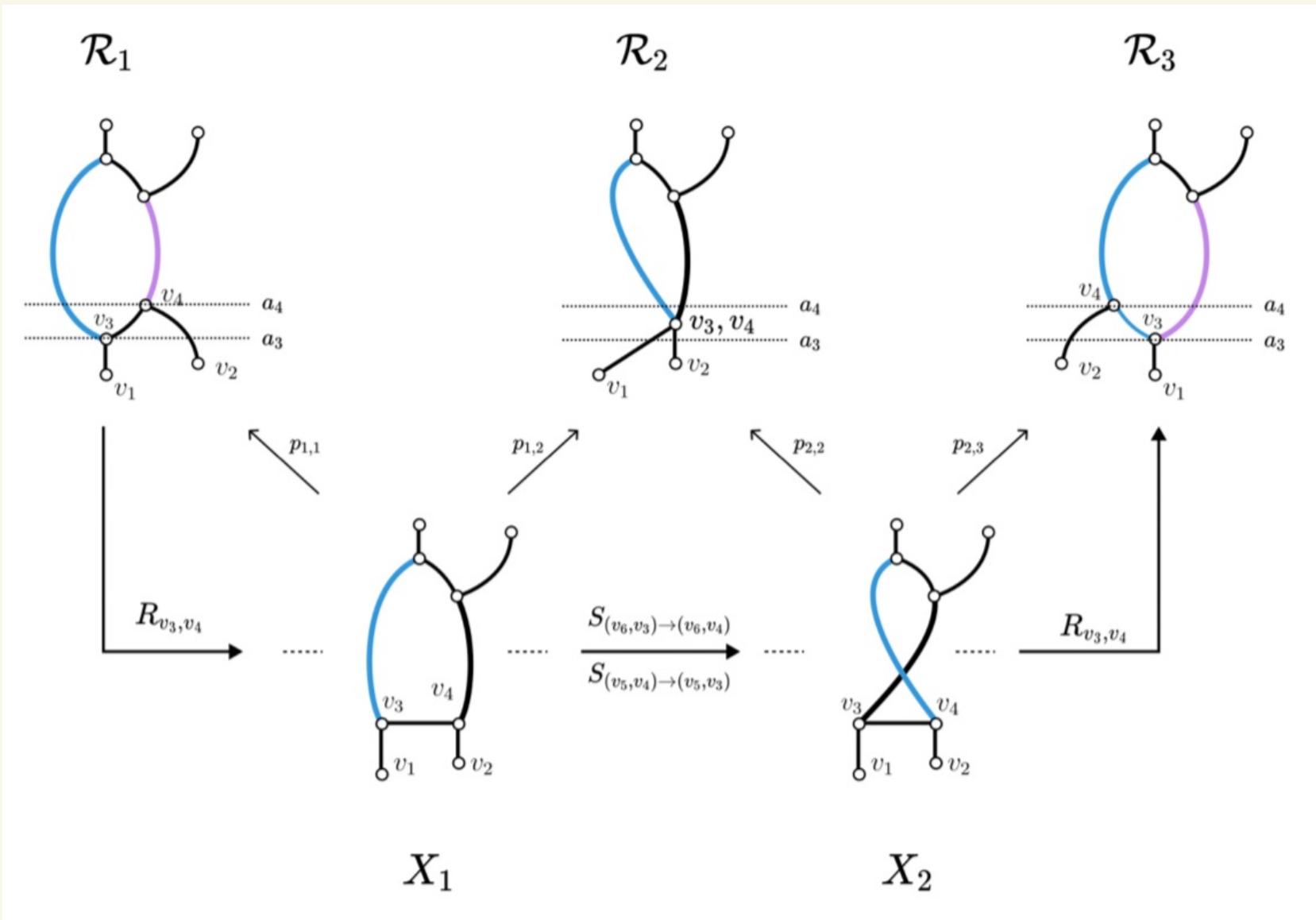


Then, pull back (or cone) of this diagram must exist, & can measure how much things move across the maps.

# Example



# Alternative:



# Categorical edit distance pros + cons:

It is:

- stable
- discriminative
- isomorphism invariant
- path component sensitive

and universal:

for any stable distance  $S$ ,

$$d_S(R_f, R_g) \leq d_E(R_f, R_g)$$

[Bauer et al 2020]

Unfortunately, no idea how to  
compute it!

# Functional distortion

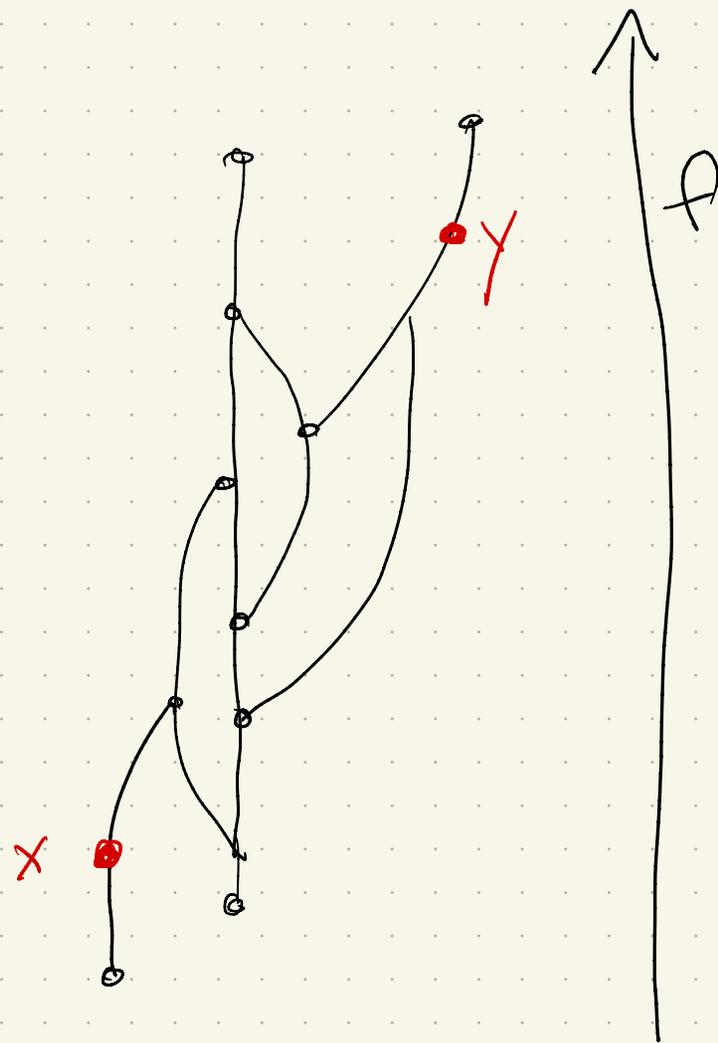
Based on Gromov-Hausdorff idea:

Height of a path:

given path  $\Pi$ ,

$$\max_{x \in \Pi} f(x) - \min_{x \in \Pi} f(x)$$

$d(x, y) =$  height of  
smallest such  $\Pi$



## Point distortion:

Fix maps  $\Psi: R_f \rightarrow R_g$

$\Phi: R_g \rightarrow R_f$

(continuous, but not function preserving)

$$G(\Psi, \Phi) = \{ (x, \Psi(x)) \mid x \in R_f \}$$

$$\cup \{ (\Phi(y), y) \mid y \in R_g \}$$

## Point distortion

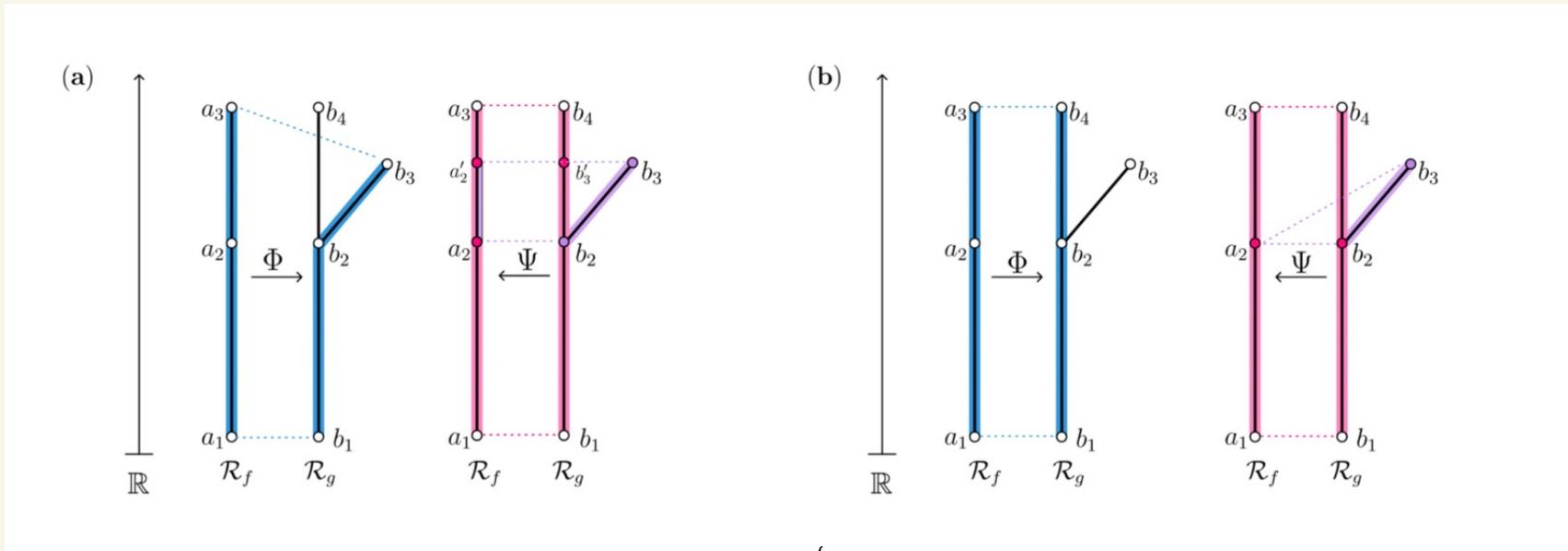
$$d((x, y), (x', y')) = \frac{1}{2} |d_f(x, x') - d_g(y, y')|$$

## Map distortion:

supremum over all pairs of points  
of the point distortion

# Example

Two possible  $\Phi, \Psi$  maps:



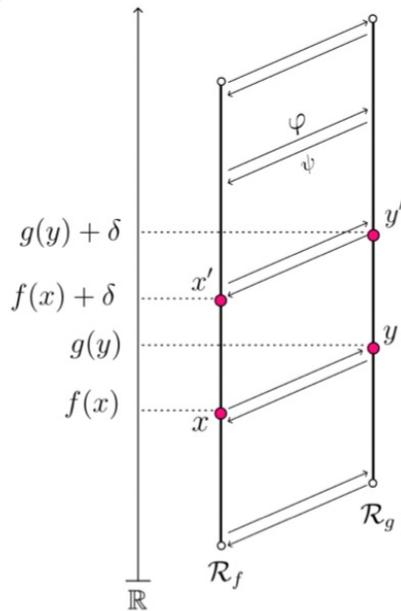
worst pair:  
 $(a_3, b_3)$  &  $(a_2, b_2)$

worst pair:  
 $(a_2, b_3)$  and  $(a_2, b_2)$

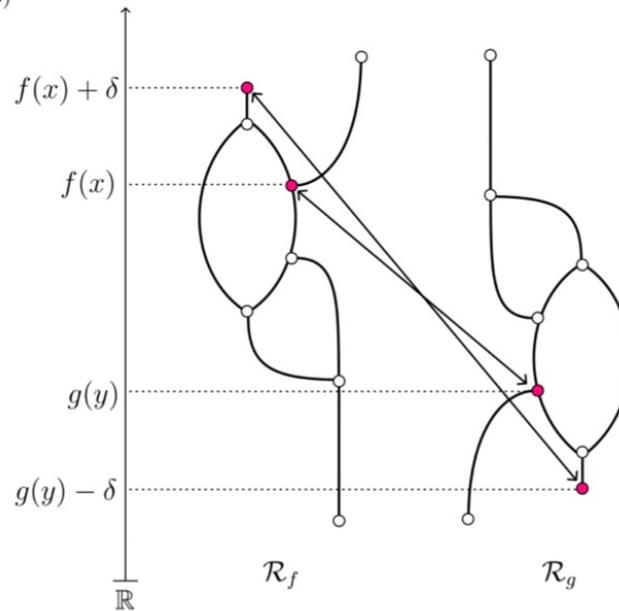
# Functional distortion

$$d_{FD}(\mathcal{R}_f, \mathcal{R}_g) = \inf_{\Phi, \Psi} \max \{ D(\Phi, \Psi), \|f - g \circ \Phi\|_{\infty}, \|f \circ \Psi - g\|_{\infty} \}$$

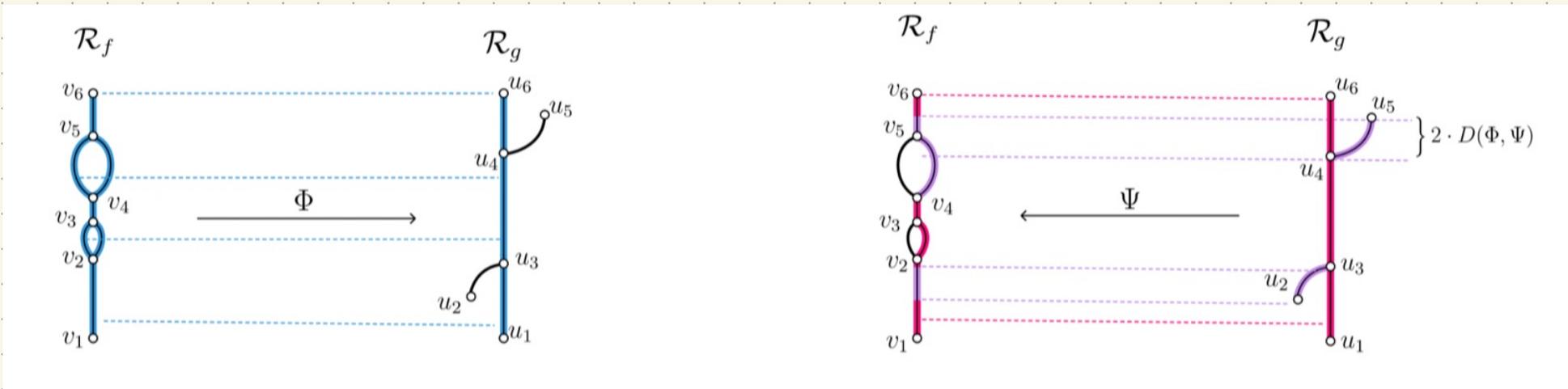
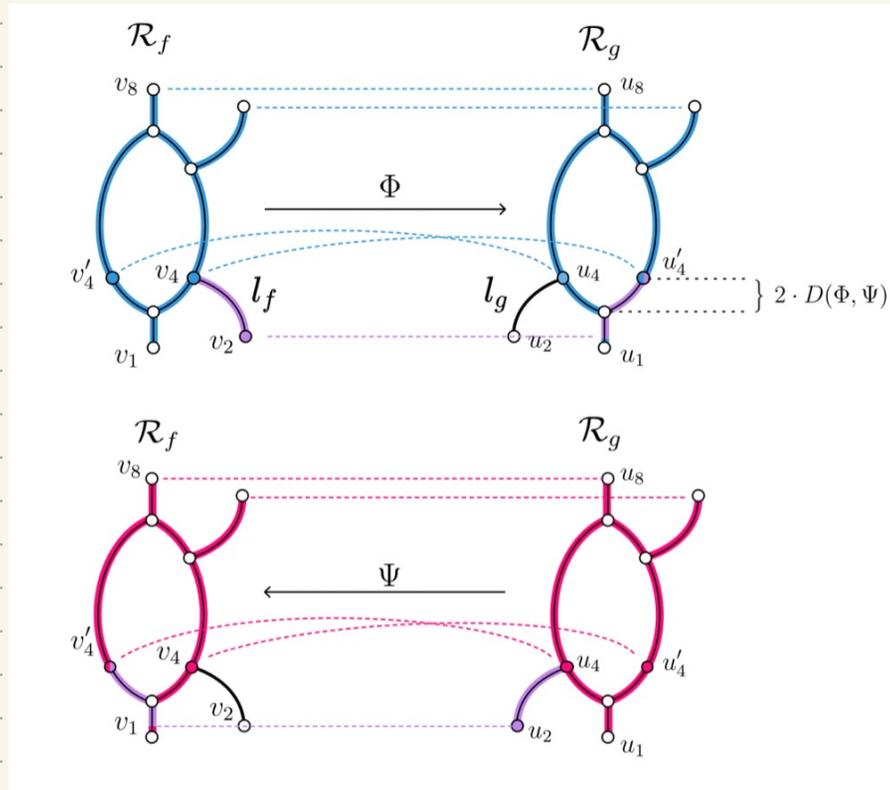
(a)



(b)



# Examples:



# Functional distortion pros + cons

It is:

- stable
- discriminative

[Bauer et al 2014]

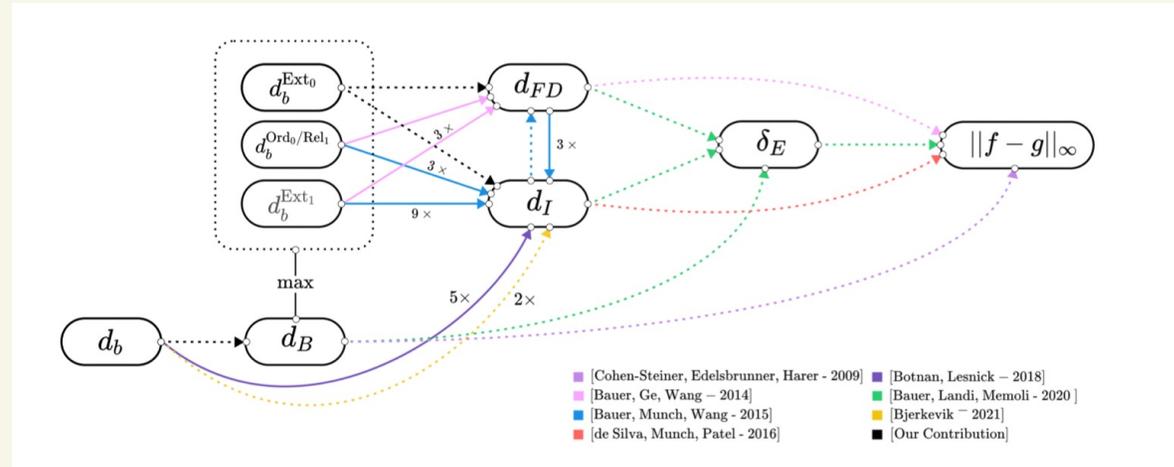
- isomorphism invariant
- path component sensitive

[Bauer et al 2016]

No idea if it's computable in any efficient way.  
(or even at all, except on trees!)

# Take-aways:

① This are very connected:



In fact, we conclude our survey with a conjecture:

**Conjecture 9.3.** *The functional distortion distance, interleaving distance, and universal edit distance are equivalent on the space of PL Reeb graphs where the domain  $\mathbb{X}$  is simply connected. That is*

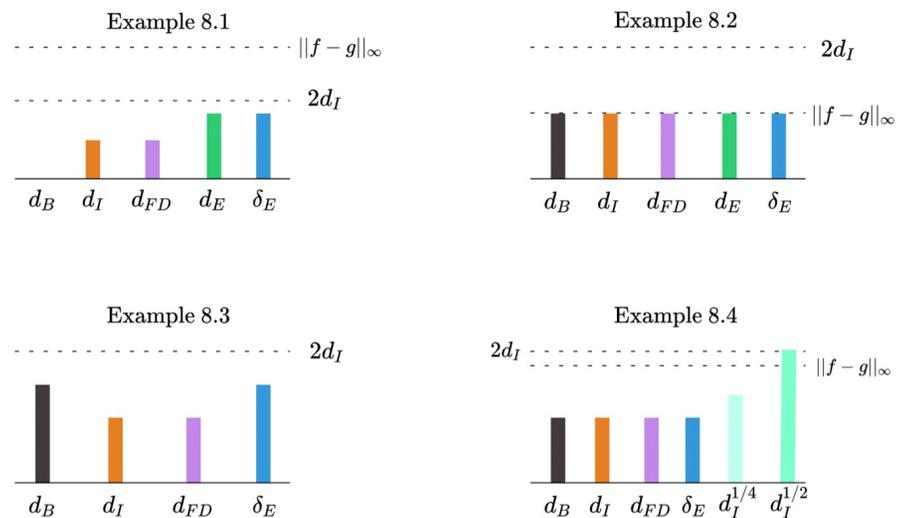
$$d_B \leq d_I = d_{FD} = \delta_E.$$

(up to constant factors)

② They do capture different types of features, even though there are connections!

	$d_B$	$d_I$	$d_I^m$	$d_{FD}$	$d_E$	$\delta_E$
Stable	[27]	[62]	-	[4]	[35]	[5]
Discriminative	-	[6], 7.24	[22], C.3	[4], 7.20	[35]	[5]
Isomorphism Invariant	-	[62]	[22]	[6], 7.10	[35]	[5]
Path Component Sensitive	-	[62]	[22]	[6]	-	[5], 7.39
Universal	-	-	-	-	-	[5]

**Table 1:** Table of distance properties. Entry corresponds to a citation where the distance was proved. We supply additional references to statements we contributed in this work which solidify these properties. We denote disproven properties or properties that are not applicable to the given distance with “-”.



**Figure 16:** Visual summary of the distance values attained for each example. For Example 4, we show two different choices of  $m$  for the truncated interleaving distance to illustrate the affect that  $m$  plays on the distance values.

### ③ Computation of metrics:

The real negative:  
all are "Hard" on general Reeb  
graphs.

