

THE REEB GRAPH EDIT DISTANCE IS UNIVERSAL

ULRICH BAUER 

DRAGON APPLIED TOPOLOGY CONFERENCE  
SWANSEA

SEP 12, 2018

JOINT WORK WITH CLAUDIA LANDI (UNIVERSITY OF MODENA)  
AND FACUNDO MÉMOLI (OHIO STATE U)

"Here are two things that are reasonably close to each other, and I want to compare them." S.

S. WEINBERGER

# REEB GRAPHS

$$f: M \rightarrow \mathbb{R}$$



$M$

→  $R_f$

$$\tilde{f}: R_f \rightarrow \mathbb{R}$$



identify components of level sets  $f^{-1}(t)$ :

$$R_f = M / \sim_f,$$

where  $x \sim_f y \Leftrightarrow x, y$  in same component of some  $f^{-1}(t)$ ,  $t \in \mathbb{R}$ .

$$g: M \rightarrow \mathbb{R}$$



$M$

→  $R_g$

$$\tilde{g}: R_g \rightarrow \mathbb{R}$$



## FORMAL SETTING

We consider

- locally compact Hausdorff spaces (**Reeb domains**)
- proper quotient maps with connected fibers (**Reeb quotient maps**)

These maps are closed under composition, and stable under pullbacks.

Define a **Reeb graph** as

- a Reeb domain  $R_f$  with
- a function  $\tilde{f}: R_f \rightarrow \mathbb{R}$  with discrete fibers (**Reeb function**)

A Reeb graph  $R_f$  is the **Reeb graph** of a function  $f: X \rightarrow \mathbb{R}$  if

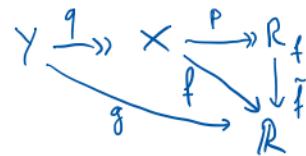
- $f = \tilde{f} \circ p$  for some Reeb quotient map  $p: X \rightarrow R_f$ .

In this case,  $R_f \cong X / \sim_f$ .

Moreover: let  $q: Y \rightarrow X$  be a Reeb quotient map.

Then  $R_f$  is also the Reeb graph of  $g = f \circ q$ .

- Reeb quotient maps **preserve Reeb graphs**.



## Goals

How to compare two Reeb graphs  $R_f, R_g$ ? ( $f, g : M \rightarrow \mathbb{R}$  are unknown)

Assign distance (extended pseudo-metric)  $d(R_f, R_g)$ .

Desirable properties:

**Stability:** For any space  $X$  and  $f, g : X \rightarrow \mathbb{R}$  yielding Reeb graphs  $R_f, R_g$ ,

$$d(R_f, R_g) \leq \|f - g\|_\infty.$$

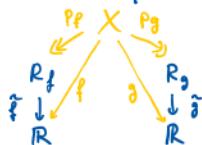
**Universality:** For any other stable distance  $d_s$ ,

$$d_s(R_f, R_g) \leq d(R_f, R_g).$$

# A CANONICAL UNIVERSAL DISTANCE

Given Reeb graphs  $R_f, R_g$  with functions  $\tilde{f}, \tilde{g}$ , define

$$d_u(R_f, R_g) = \inf \quad \|f - g\|_\infty$$

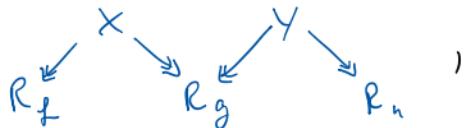


taken over all Reeb domains  $X$  and Reeb quotient maps  $p_f, p_g$ .

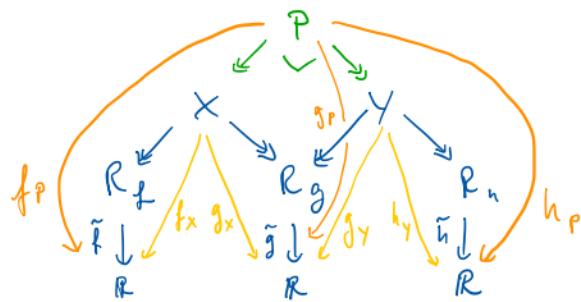
- **stability**: infimum is lower bound on all  $\|f - g\|_\infty$
- **universality**: infimum is **greatest** lower bound

# TRIANGLE INEQUALITY FOR THE CANONICAL UNIVERSAL DISTANCE

Triangle inequality: for all



consider pullbacks

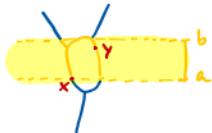


then

$$\begin{aligned}
 \|f_p - h_p\|_\infty &\leq \|f_p - g_p\|_\infty + \|g_p - h_p\|_\infty \\
 &= \|f_x - g_x\|_\infty + \|g_y - h_y\|_\infty.
 \end{aligned}$$

PREVIOUS WORK: FUNCTIONAL DISTORTION DISTANCE [B, Ge, Wang 2014]

- On a Reeb graph  $R_f$  with  $\tilde{f}: R_f \rightarrow \mathbb{R}$ , consider the metric  $d_f: (x, y) \mapsto \inf \{b-a \mid x, y \text{ in same component of } \tilde{f}^{-1}[a, b]\}$ .



- Given maps  $\phi: R_f \rightarrow R_g$ ,  $\psi: R_g \rightarrow R_f$ , consider  $G(\phi, \psi) = \{(x, \phi(x)) \mid x \in R_f\} \cup \{\psi(\psi(y), y) \mid y \in R_g\} \subseteq R_f \times R_g$ .

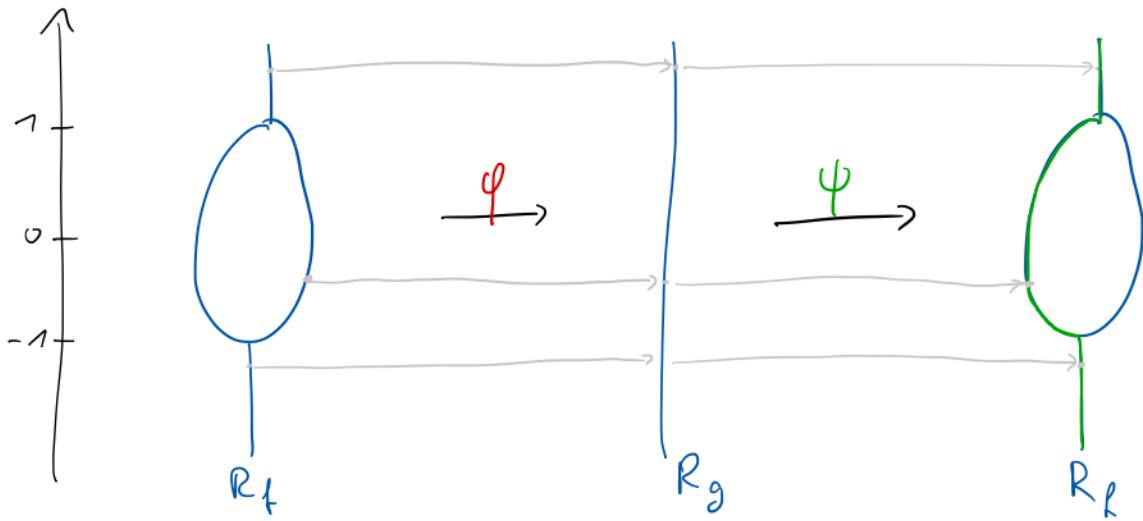
- Define the *distortion* of  $(\phi, \psi)$  as

$$D(\phi, \psi) = \sup_{(x, y), (\tilde{x}, \tilde{y}) \in G(\phi, \psi)} \frac{1}{2} |d_f(x, \tilde{x}) - d_g(y, \tilde{y})|.$$

- Define the *functional distortion distance* as

$$d_{FD}(R_f, R_g) = \inf_{\phi, \psi} (\max \{ D(\phi, \psi), \|f - g \circ \phi\|_\infty, \|g - f \circ \psi\|_\infty \}).$$

## EXAMPLE: FUNCTIONAL DISTORTION DISTANCE



$$D(\varphi, \psi) = \sup \frac{1}{2} (d(x, \tilde{x}) - d(y, \tilde{y})) = \frac{1}{2}$$

where  $x, \tilde{x} \in R_f$ ,  $y, \tilde{y} \in R_g$

with  $\varphi(x) = y$  or  $x = \psi(y)$ ,

$\varphi(\tilde{x}) = \tilde{y}$  or  $\tilde{x} = \psi(\tilde{y})$

PREVIOUS WORK: INTERLEAVING DISTANCE [Bubenik & al. 2015; deSilva & al. 2016]

- Interpret Reeb graph  $R_f$  as a functor (in fact, a *co*leaf)

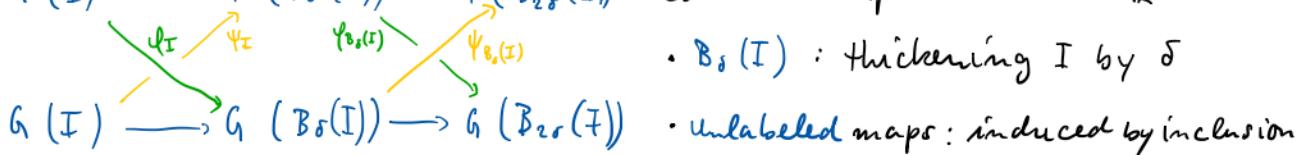
$$F : \text{Int}_{\mathbb{R}} \rightarrow \text{Set}, I \mapsto \pi_0(\tilde{f}^{-1}(I))$$

( $\text{Int}_{\mathbb{R}}$ : open intervals, as a poset wrt.  $\subseteq$ )

- A  $\delta$ -interleaving between  $F$  and  $G$  is a pair of natural transformations

$\varphi, \psi$  (with components  $\varphi_I : F(I) \rightarrow G(B_\delta(I)), \dots$ ) such that

$$F(I) \longrightarrow F(B_\delta(I)) \longrightarrow F(B_{2\delta}(I)) \quad \text{commutes for all } I \in \text{Int}_{\mathbb{R}}.$$



$\bullet B_\delta(I)$  : thickening  $I$  by  $\delta$

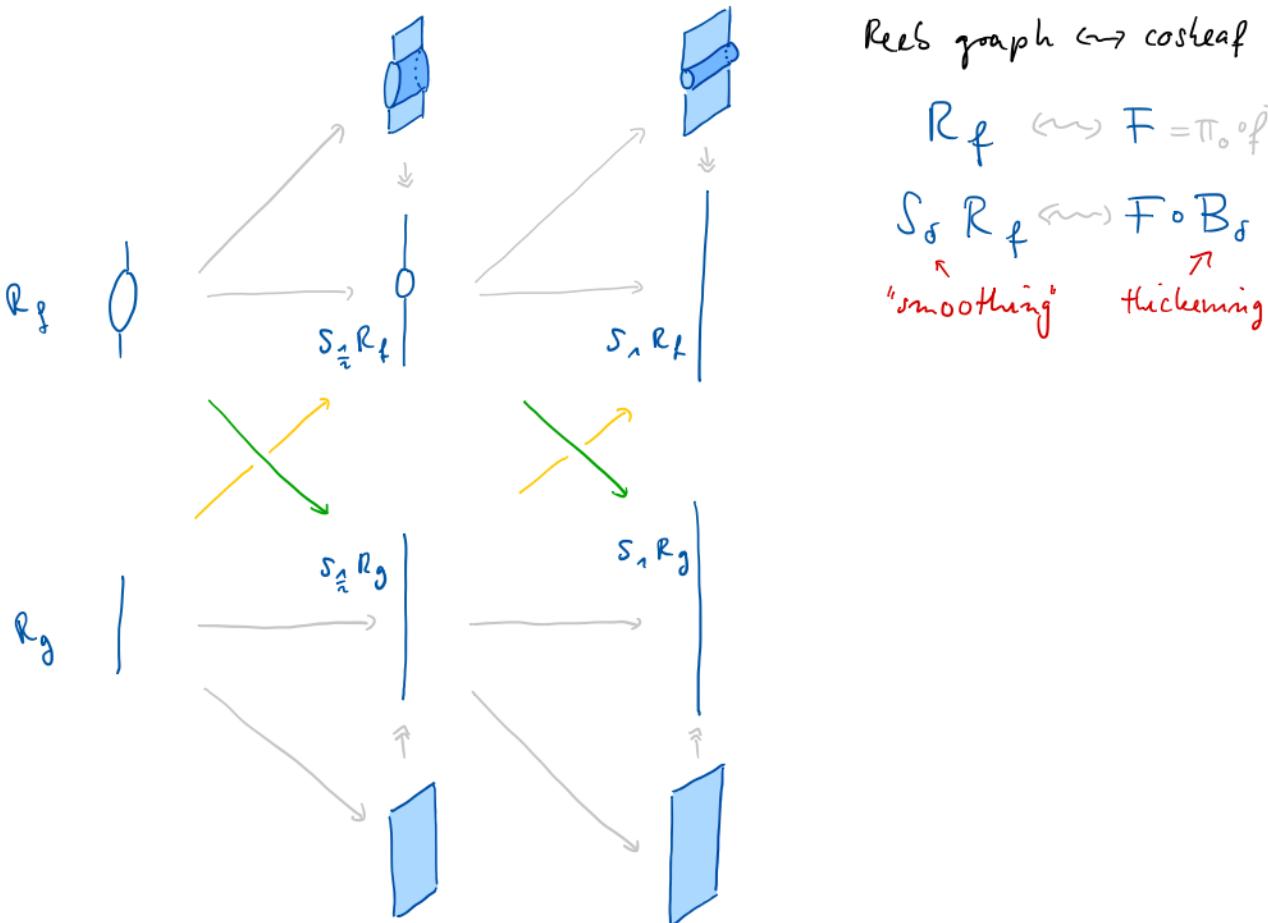
$\bullet$  unlabeled maps: induced by inclusion

- The interleaving distance is

$$d_I(R_f, R_g) := \inf \{ \delta \mid \exists \text{ a } \delta\text{-interleaving between } F \text{ and } G \}$$

Thm [B., Munch, Wang 2015]  $\frac{1}{3} d_{FD} \leq d_I \leq d_{FD}$ .

# ABSTRACT AND TOPOLOGICAL INTERLEAVINGS [desilva & al. 2016]



## LEVEL SET PERSISTENT HOMOLOGY

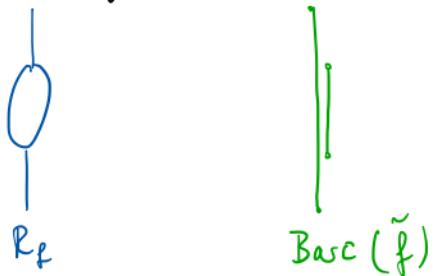
Thm [Carlsson, de Silva, Morozov 2009]

Given  $f: X \rightarrow \mathbb{R}$  (PL, with  $X$  compact):

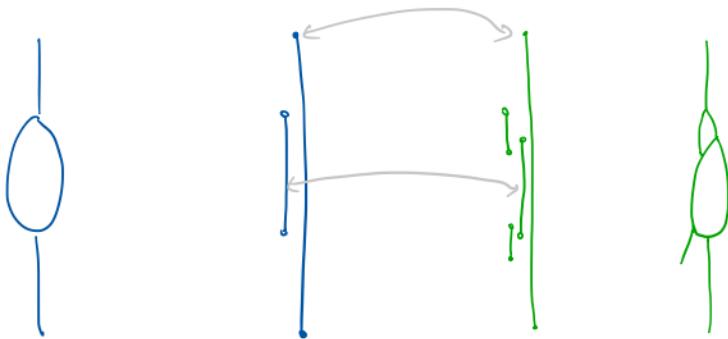
Homology of level sets  $H_*(f^{-1}(t); \mathbb{F})$  (and more generally,  
of inclusions  $f^{-1}(I) \hookrightarrow f^{-1}(J)$  for intervals  $I \subseteq J$ )

is encoded (up to isomorphism) by a unique  
collection of intervals (level set persistence barcode).

Example for a Reeb graph:



# THE BOTTLENECK DISTANCE BETWEEN PERSISTENCE BARCODES



A  $\delta$ -matching between two barcodes  $\text{Barc}(f)$ ,  $\text{Barc}(g)$  satisfies:

- matched intervals  $(I, J)$  have distance  $d_H(I, J) \leq \delta$
- unmatched intervals have length  $\leq 2\delta$

The bottleneck distance  $d_B(f, g)$  is

$$\inf \delta : \exists \delta\text{-matching between } \text{Barc}(f), \text{Barc}(g)$$

# A ZOO OF DISTANCES AND INEQUALITIES

[Carlsson, de Silva, Morozou 2009]

$$d_B(R_f, R_g) \leq \|f - g\|_\infty$$

Open problem:

is the lower bound tight?

[B., Ye, Wang 2014]

$$\frac{1}{3} d_B(R_f, R_g) \leq d_{FD}(R_f, R_g) \leq \|f - g\|_\infty$$

[B., Munch, Wang 2015]

$$\frac{1}{3} d_{FD}(R_f, R_g) \leq d_I(R_f, R_g) \leq d_{FD}(R_f, R_g)$$

[Botnan, Lesnick 2016]

$$\frac{1}{5} d_B(R_f, R_g) \leq d_I(R_f, R_g)$$

[Björkevik 2016]

$$\frac{1}{2} d_B(R_f, R_g) \leq d_I(R_f, R_g)$$

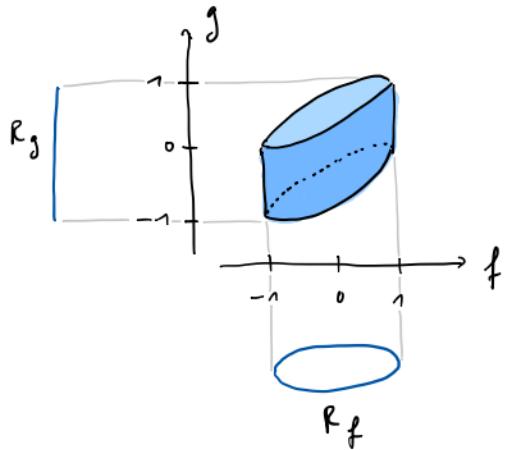
•  $d_B$ : bottleneck distance  
(of level set  $H_0$  barcode)

•  $d_{FD}$ : functional distortion distance

•  $d_I$ : interleaving distance

# FUNCTIONAL DISTORTION & INTERLEAVING DISTANCES ARE NOT UNIVERSAL

Consider a cylinder with two functions f, g:



- $d_u(R_f, R_g) = \|f - g\|_\infty = 1$
- $d_I(R_f, R_g) \leq d_{FD}(R_f, R_g) \leq \frac{1}{2} < d_u(R_f, R_g)$ :

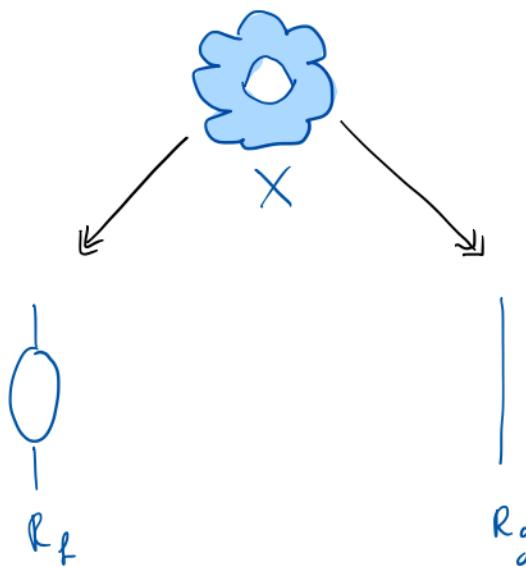
$$R_f \xrightarrow{\phi} |R_g| \quad R_g \xrightarrow{\psi} {}^{\text{im } \phi} R_f$$

# FROM CLOSE REEB GRAPHS TO CLOSE FUNCTIONS

Open problem

Given two Reeb graphs  $R_f, R_g$  with  $d_I(R_f, R_g) = \delta$ .

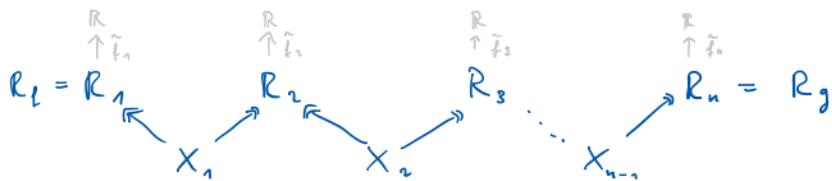
Is there a space  $X$  with  $f, g : X \rightarrow \mathbb{R}$ ,  $\|f - g\| \leq C \cdot \delta$ ,  
yielding Reeb graphs  $R_f, R_g$ , for some fixed constant  $C$ ?



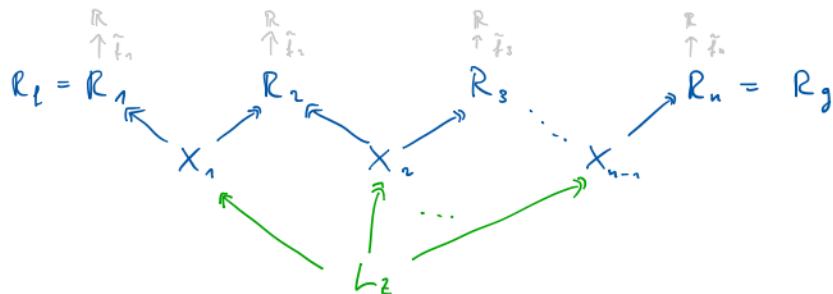
By the previous example:  
if yes, then  $C \geq 2$ .

## TOPOLOGICAL EDIT ZIG-ZAGS

- Consider zig-zag diagrams  $\mathcal{Z}$  of Reeb quotient maps



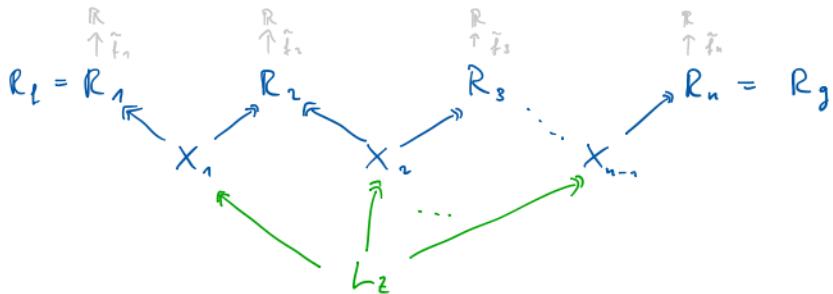
and take the limit  $L_z$



(note : all maps are Reeb quotient maps).

## THE TOPOLOGICAL EDIT DISTANCE

- Consider zig-zag limit  $L_z$



- Define the *spread* of the functions  $f_i: L_z \rightarrow R_i \rightarrow \mathbb{R}$  as

$$s_z : L_z \rightarrow \mathbb{R}, \quad x \mapsto \max_i f_i(x) - \min_j f_j(x).$$

(Note:  $\|f - g\|_\infty \leq \|s_z\|_\infty$ , with equality for  $n = 2$ )

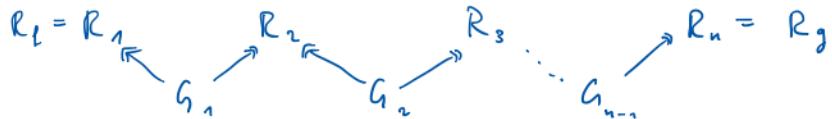
- Define the *(topological) edit distance* as

$$d_{\text{top}}(R_f, R_g) = \inf_z \|s_z\|_\infty.$$

Prop.  $d_{\text{top}}$  is stable and universal.

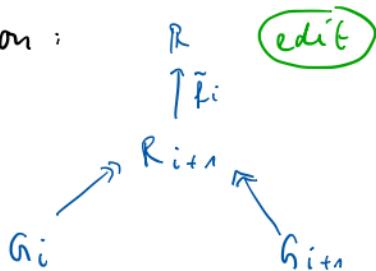
# THE REEB GRAPH EDIT DISTANCE

- Consider zig-zag diagrams  $\mathcal{Z}$  of Reeb quotient maps

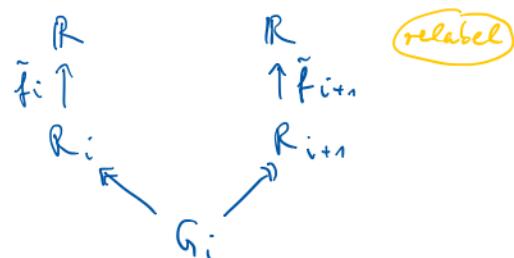


as before, but **restrict**  $R_i, G_j$  in  $\mathcal{Z}$  to be **finite graphs**.

Interpretation :



modify  $G_i$  to  $G_{i+1}$ ,  
maintaining the Reeb graph  $R_{i+1}$



modify  $f_i$  to  $f_{i+1} : G_i \rightarrow R_i$ ,  
maintaining the domain  $G_i$

- Define the **Reeb graph edit distance** analogously as

$$d_{\text{Graph}}(R_f, R_g) = \inf_{\mathcal{Z}} \|s_{\mathcal{Z}}\|_{\infty}.$$

## MAIN RESULT

Then [B., Landi, Mémoli] The Reeb graph edit distance is stable & universal.

- We restrict to the (compact) PL category here.
- The hard part is stability:  
given  $f, g : X \rightarrow \mathbb{R}$  (PL, for triangulation  $X = |K|$ ),  
construct edit zigzag between  $R_f$  and  $R_g$  with spread  $\leq \|f - g\|_\infty$ ?
- Idea:
  - Consider straight-line homotopy  $f_\lambda = \lambda f + (1-\lambda)g$
  - The structure of  $R_\lambda = R_{f_\lambda}$  changes only finitely often  
(say, at parameters  $0 = \lambda_0 < \dots < \lambda_n = 1$ ). Choose  $\rho_i \in (\lambda_i, \lambda_{i+1})$ .
  - Construct zigzag  $R_f = R_{\lambda_0} \xrightarrow{\quad} R_{\lambda_1} \xrightarrow{\quad} \dots \xrightarrow{\quad} R_{\lambda_i} \xrightarrow{\quad} R_{\lambda_{i+1}} \xrightarrow{\quad} \dots \xrightarrow{\quad} R_{\lambda_n} = R_g$
  - How to get the Reeb quotient maps in this zigzag?

# CRITICAL INSTANTS OF A PL STRAIGHT-LINE HOMOTOPY

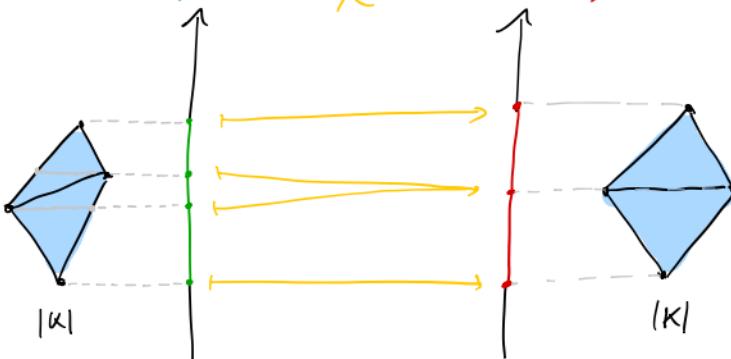
w.l.o.g.:

$$f = f_{P_i}$$

(regular)

$$fx_i = g$$

(critical)



order-preserving surjection

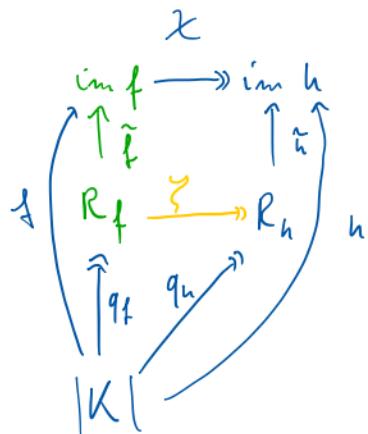
$$\chi: \text{im } f \rightarrow \text{im } g \quad (\text{PL})$$

- We have  $\chi \circ f(v) = g(v) \quad \forall v \in \text{Vert}(K)$
- But  $\chi \circ f \neq g$  !
- However:  $\chi \circ f$  and  $g$  have the same Reeb graph ...

## LIFTING REPARAMETRIZATIONS

**Lemma** Let  $\chi: \text{im } f \rightarrow \text{im } h$ ,  $h = \chi \circ f$ .

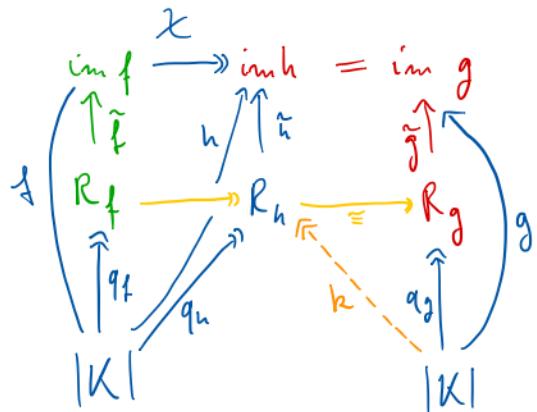
Then  $\chi$  lifts to a Reeb quotient map  $\tilde{\chi}: R_f \rightarrow R_h$ .



# THE WORKHORSE : REEB QUOTIENT MAPS FROM INTERPOLATION

$\chi : \text{im } f \rightarrow \text{im } g$  (order-preserving PL surjection),

$$g(v) = \underbrace{\chi \circ f(v)}_h \quad \forall v \in \text{Vert } K.$$



**Lemma** The relation

$$k = q_h \circ ((h^{-1} \circ g) \cap \text{st}_K)$$

is a Reeb quotient map.

**Corollary**  $R_h \cong R_g$ , and  
 $\chi$  lifts to a Reeb quotient map

$$R_f \rightarrow R_g.$$

This provides the maps

$$\begin{array}{ccc} R_{\chi_i} & \xleftarrow{R_{\pi_i}} & R_{\chi_{i+1}} \\ & \nearrow & \searrow \\ & R_{\pi_i} & \end{array}$$

for our interpolation zigzag.

## CONCLUSION

- A universal distance is the most discriminative stable distance between Reeb graphs
- There is a simple construction of a universal distance
- Interleaving and functional distortion distances are not universal
- A universal distance in PL can be constructed using graph edit zigzags

## Questions:

- What is the complexity of computing the distance?
- Is  $d_u \leq C \cdot d_I$  for some constant  $C$ ?