

# Euler and Betti curves are stable under Wasserstein deformations of distributions of stochastic processes

Daniel Perez<sup>\*1,2,3</sup>

<sup>1</sup>Département de mathématiques et applications, École normale supérieure, CNRS, PSL University, 75005 Paris, France

<sup>2</sup>Laboratoire de mathématiques d'Orsay, Université Paris-Saclay, CNRS, 91405 Orsay, France

<sup>3</sup>DataShape, Centre Inria Saclay, 91120 Palaiseau, France

November 24, 2022

## Abstract

Euler and Betti curves of stochastic processes defined on a  $d$ -dimensional compact Riemannian manifold which are almost surely in a Sobolev space  $W^{n,s}(X, \mathbb{R})$  (with  $d < n$ ) are stable under perturbations of the distributions of said processes in a Wasserstein metric. Moreover, Wasserstein stability is shown to hold for all  $p > \frac{d}{n}$  for persistence diagrams stemming from functions in  $W^{n,s}(X, \mathbb{R})$ .

## Contents

<b>1 Introduction</b>	<b>1</b>
1.1 Glossary, definitions and conventions . . . . .	2
<b>2 Stability results</b>	<b>3</b>
2.1 Preliminary results . . . . .	3
2.2 Stability theorems . . . . .	4
<b>3 Euler and Betti curves</b>	<b>5</b>

## 1 Introduction

The study of random fields by Adler and Taylor using geometric and topological methods gave rise to the following celebrated and interesting result.

**Theorem 1.1** (Adler, Taylor, Theorem 12.4.1 [2]). *Let  $f$  be a centered unit variance Gaussian field on a  $d$ -dimensional  $C^2$ -submanifold  $X$  of a  $C^3$  manifold. Suppose  $f$  satisfies the conditions of [2, Corollary 11.3.5] (henceforth the GKF conditions), then,*

$$\mathbb{E}[\chi(\{f \geq x\})] = \sum_{j=0}^d \mathcal{L}_j(X) \rho_j(x), \quad (1.1)$$

where  $\chi$  denotes the Euler characteristic and  $\rho_j(x)$  denotes some known functions of  $x$  and  $\mathcal{L}_j$  denote the Lipschitz-Killing curvatures of  $X$ .

In recent years, much work has been done to weaken the Gaussian assumption of the above theorem. A notable step in this direction is the following remarkable result by Bobrowski and Borman.

---

\*Email: [daniel.perez@ens.fr](mailto:daniel.perez@ens.fr)

**Theorem 1.2** (Bobrowski, Borman, Theorem 4.1, [3]). *Let  $X$  be a compact  $d$ -dimensional stratified space, and let  $f : X \rightarrow \mathbb{R}^k$  be a  $k$ -dimensional Gaussian random field satisfying the GKF conditions. For a piecewise  $C^2$ -function  $G : \mathbb{R}^k \rightarrow \mathbb{R}$  and let  $g = G \circ f$ . Setting  $D_u = G^{-1} ] - \infty, u]$ , we have*

$$\mathbb{E} \left[ \int_X g d\chi \right] = \chi(X) \mathbb{E}[g] - \sum_{j=1}^d (2\pi)^{j/2} \mathcal{L}_j(X) \int_{\mathbb{R}} \mathcal{M}_j(D_u) du, \quad (1.2)$$

where  $\mathbb{E}[g]$  has constant mean.

As pointed out by the authors of the above theorem in their original paper [3], the difficulty in evaluating the expression above lies in computing the Minkowski functionals  $\mathcal{M}_j(D_u)$ . In this paper, we aim to somewhat compensate this difficulty by providing a stability result regarding the above expectations, and moreover, the distributions of the Euler curves (*i.e.* the curves  $x \mapsto \chi(\{f \geq x\})$ ). More precisely, our main result is the following.

**Theorem 1.3.** *Let  $X$  be a compact  $d$ -dimensional manifold and let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. Consider two stochastic processes  $f, g : \Omega \rightarrow W^{n,s}(X, \mathbb{R})$  with  $d < n$ . Then, the Euler curves of  $f$  and  $g$  are well-defined as elements of  $L^1(\mathbb{R})$  and as distributions, and for every  $\frac{d}{n} < \alpha < 1$ , there exists a finite constant  $C_{X,n,s}$  such that*

$$\|\mathbb{E}[\chi(\{f \geq x\})] - \mathbb{E}[\chi(\{g \geq x\})]\|_{L^1_x(\mathbb{R})} \leq C_{X,n,s} \left[ \|f\|_{L^1(\Omega, W^{n,s})}^\alpha + \|g\|_{L^1(\Omega, W^{n,s})}^\alpha \right] W_{1,L^\infty}^{1-\alpha}(f_\# \mathbb{P}, g_\# \mathbb{P}), \quad (1.3)$$

where  $W_{1,L^\infty}$  denotes the Wasserstein distance on the space of probability measures on the space  $(L^\infty(X, \mathbb{R}) \cap W^{n,s}(X, \mathbb{R}), \|\cdot\|_{L^\infty})$ . Moreover, the same statement holds for all Betti curves.

The key will be to adopt the point of view of persistent homology, for which we refer the unfamiliar reader to the now classical references on the subject [5, 10] for an introduction. It is noteworthy that the relevance of the study of persistence diagrams was already understood by Bobrowski and Borman, and was later also elaborated upon by different authors [1, 9]. The results of this paper can be seen as direct consequences of the theory developed in [11] and of a recent result by Buhovsky *et al.* [4].

## 1.1 Glossary, definitions and conventions

**Notation 1.4.** We will interchangeably denote the persistent homology in degree  $k$  on  $X$  with respect to a filtration function  $f : X \rightarrow \mathbb{R}$  by  $H_k(X, f)$  or  $\text{Dgm}_k(f)$ . We use persistence barcodes, diagrams, modules and their decompositions interchangeably via the usual identifications [5]. We reserve the notation  $H_*(X)$  for the homology groups of  $X$  exclusively.

**Convention 1.5.** Unless otherwise specified, we will always consider the filtration by **superlevel** sets induced by a function  $f : X \rightarrow \mathbb{R}$ .

**Convention 1.6.** Unless otherwise specified, we shall always consider the (infinite) bars of the barcode to be truncated. That is,  $b = b \cap [\inf f, \sup f]$  for every  $b \in \text{Dgm}_k(f)$ .

**Notation 1.7.** The length of a bar  $b \in \text{Dgm}_k(f)$  is denoted by  $\ell(b)$ . With this definition, we define

$$\text{Pers}_p^p(H_k(X, f)) := \sum_{b \in \text{Dgm}_k(f)} \ell(b)^p. \quad (1.4)$$

**Notation 1.8.** We equip  $\mathbb{R}^2$  with a metric  $d_{\mathbb{R}^2, \infty}$  defined by

$$d_{\mathbb{R}^2, \infty}((x, y), (x', y')) = \max\{|x - x'|, |y - y'|\}. \quad (1.5)$$

and define the distance from a point  $(x, y)$  to the diagonal  $\Delta$  as  $d_{\mathbb{R}^2, \infty}((x, y), \Delta) = \frac{1}{2}|y - x|$ . We will use the following notation to design different sets of  $\mathbb{R}^2$ . For **superlevel** set filtrations we denote

- $\mathcal{X} = \{(x, y) \in \mathbb{R}^2 \mid y < x\}$  and  $\overline{\mathcal{X}}$  its closure.
- $R_x := [x, \infty[ \times ]-\infty, x] \subset \overline{\mathcal{X}}$ .

The reader can adjust the statements by modifying the definitions in the following way if he or she is considering **sublevel** set filtrations, as is customary in the topological data analysis community.

- $\mathcal{X} = \{(x, y) \in \mathbb{R}^2 \mid x < y\}$  and  $\overline{\mathcal{X}}$  its closure.
- $R_x := ]-\infty, x] \times [x, \infty[ \subset \overline{\mathcal{X}}$ .

**Definition 1.9.** Define  $\mathcal{D}_p$  to be the space of Radon measures on  $\overline{\mathcal{X}}$  (called **persistence measures**) with finite  $\text{Pers}_p$  [8, Equations 3 and 4]. We equip  $\mathcal{D}_p$  with the topology of the optimal *partial* transport distance  $d_p$ . For persistence diagrams

$$d_p(D, D') := \inf_{\gamma \in \Gamma(D, D')} \left[ \sum_{b \in D} d_{\mathbb{R}^2, \infty}(b, \gamma(b))^p \right]^{1/p}, \quad (1.6)$$

where the infimum is taken over  $\Gamma(D, D')$ , the set of all bijections between  $D \cup \partial\mathcal{X}$  and  $D' \cup \partial\mathcal{X}$ . We refer the reader to [8, Definition 2.1] for the definition of  $d_p$  on general persistence measures.

**Notation 1.10.** Let  $D$  denote a persistence diagram, then  $N_D^\varepsilon$  denotes the number of bars in the diagram of length  $\geq \varepsilon$ .

**Notation 1.11.** We note  $W^{n,s}(X, \mathbb{R})$  the Sobolev space with exponents  $(n, s)$ , which is not to be confused with the  $p$ -Wasserstein distance on the set of probability measures on a Polish metric space  $(\mathcal{M}, d)$ , which we denote  $W_{p,d}$ .

## 2 Stability results

### 2.1 Preliminary results

The following important result was recently shown by Buhovsky *et al.*

**Theorem 2.1** (Buhovsky, Payette, Polterovich, Polterovich, Shelukhin, Stojisavljević, [4]). *Let  $X$  be a compact Riemannian manifold of dimension  $d$  and let  $f \in W^{n,s}(X, \mathbb{R})$  and  $n > \frac{d}{s}$ . Then,  $f$  is continuous and for all  $\varepsilon > 0$*

$$N_{H_k(X, f)}^\varepsilon \leq C_{X, n, s} \|f\|_{W^{n, s}}^{d/n} \varepsilon^{-d/n} + \dim H_k(X). \quad (2.7)$$

This theorem gives us a bound of the  $\text{Pers}_p^p$ -functional via the following lemma.

**Lemma 2.2.**

$$\text{Pers}_p^p(H_k(X, f)) = p \int_0^\infty \varepsilon^{p-1} N_{H_k(X, f)}^\varepsilon d\varepsilon. \quad (2.8)$$

*Proof.* The length of a bar  $b \in H_k(X, f)$  can be written as

$$\ell(b)^p = p \int_0^\infty 1_{[0, \ell(b)]}(\varepsilon) \varepsilon^{p-1} d\varepsilon. \quad (2.9)$$

The result follows by summing over all bars  $b$  in the barcode of  $H_k(X, f)$  and applying the Fubini theorem. ■

**Corollary 2.3.** *Let  $f \in W^{n,s}(X, \mathbb{R})$ , then there exists a finite constant  $C_{X, n, s}$  such that for all  $p > \frac{d}{n}$ ,*

$$\text{Pers}_p^p(H_k(X, f)) \leq C_{X, n, s} \frac{pn}{d} \|f\|_{W^{n, s}}^p + (2 \|f\|_\infty)^p \dim H_k(X). \quad (2.10)$$

*Remark 2.4.* Corollary 2.3 means

$$f \in W^{n,s}(X, \mathbb{R}) \implies \text{Dgm}_k(f) \in \bigcap_{p > d/n} \mathcal{D}_p. \quad (2.11)$$

*Remark 2.5.* Examining the limiting case where  $n \rightarrow \infty$  and  $s \rightarrow \infty$ , one can wonder whether  $C^\infty$ -functions have finite  $\text{Pers}_0$  (in other words, a finite number of bars). However, this is not the case as is shown by the example  $e^{-1/x^2} \sin(1/x)$  on  $[0, 1]$ .

Finding the optimal degree of regularity for which  $\text{Pers}_0$  is finite is an interesting question. As shown in [7], a sufficient condition is that the functions be subanalytic in an  $\mathcal{o}$ -minimal structure. Once  $\text{Pers}_0$  is finite, it follows that  $\text{Pers}_p^p$  is a holomorphic function everywhere on  $\mathbb{C}$ .

On the other side of the story, we can ask whether the  $C^\alpha$  (up to reparametrization by a homeomorphism) condition is necessary to establish that  $\text{Pers}_p^p$  is finite for some  $p$ . This question has been positively answered in [11] in the 1D case (in fact, the correct notion of regularity is the  $p$ -variation, but is equivalent to asking that the function be  $C^\alpha$  for some  $\alpha$ , up to reparametrization).

One can ask whether the bound  $p > \frac{d}{n}$  for the finiteness of  $\text{Pers}_p^p$  is sharp within the degree of regularity considered. This question amounts to asking whether the asymptotics provided by theorem 2.1 are in some sense sharp. Indeed, lemma 2.2 entails

**Proposition 2.6.**

$$\limsup_{\varepsilon \rightarrow 0} \frac{\log N_{H_k(X,f)}^\varepsilon}{\log(1/\varepsilon)} = \inf\{p \mid \text{Pers}_p(H_k(X, f)) < \infty\}. \quad (2.12)$$

The sharpness of the asymptotics of  $N_{H_k(X,f)}^\varepsilon$  was partially addressed in [4, §1.5] where a positive answer was given (more precisely, explicit examples saturating the bound were given). For  $\alpha$ -Hölder functions, this sharpness was also addressed in [11], where it was shown that

$$\inf\{p \mid \text{Pers}_p(H_k(X, f)) < \infty\} = \frac{d}{\alpha} \text{ generically in } C^\alpha. \quad (2.13)$$

This parallel suggests the following conjecture, which might be addressed in a later work.

**Conjecture 2.7.** Let  $X$  be a compact Riemannian manifold of dimension  $d$ , then

$$\inf\{p \mid \text{Pers}_p(H_k(X, f)) < \infty\} \leq \frac{d}{r + \alpha} \quad (2.14)$$

and this inequality is saturated generically in the sense of Baire in  $C^{r+\alpha}$ .

*Remark 2.8.* By the work performed in [11, §3.2], it suffices to exhibit a function  $g$  on the cube (or the ball), such that the bound is saturated. We obtain a dense family by perturbing any dense family to make it constant on a small open ball  $B \subset X$  and adding  $g$  to  $f$  on this ball. For  $H_0$ , such a function may be constructed by considering packings and smoothed versions of distances to point clouds. In higher degrees of homology, exhibiting explicit examples might be more difficult. Experience shows that it might be easier to find a random field which saturates the bound, as done in [11].

## 2.2 Stability theorems

Having bounded  $\text{Pers}_p^p$ , a stability result follows from the standard argument considered in [6].

**Theorem 2.9** (Wasserstein stability on compact Riemannian manifolds). *Let  $X$  be a compact Riemannian manifold of dimension  $d$  and let  $f, g \in W^{n,s}(X, \mathbb{R})$ . Then, for all  $p > q > \frac{d}{n}$ ,*

$$d_p^p(\text{Dgm}_k(f), \text{Dgm}_k(g)) \leq C_{X,n,s} (\|f\|_{W^{n,s}}^q \vee \|g\|_{W^{n,s}}^q) \|f - g\|_\infty^{p-q}. \quad (2.15)$$

*Proof.* Following the proof of the stability theorem shown in [6], it is sufficient to show that  $\text{Pers}_q(\text{Dgm}_k(f))$  and  $\text{Pers}_q(\text{Dgm}_k(g))$  can be bounded. This bound is provided by corollary 2.3, yielding the result. ■

Similarly, using [11, §5.2], this bound immediately entails a stochastic stability result.

**Theorem 2.10** (Stochastic Wasserstein Stability). *Let  $f$  and  $g$  be two a.s.  $W^{n,s}(X, \mathbb{R})$  stochastic processes on a  $d$ -dimensional compact Riemannian manifold  $X$ , defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Then, for any  $0 \leq k < d$ , every  $\frac{d}{n} < q < p < \infty$  and any  $r, s \in ]1, \infty[$  satisfying  $\frac{1}{r} + \frac{1}{s} = 1$  and  $(p - q)s \geq 1$ , there exists a constant  $C = C_{X,n,s}$  such that*

$$W_{p,d_p}((\text{Dgm}_k \circ f)_{\#}\mathbb{P}, (\text{Dgm}_k \circ g)_{\#}\mathbb{P}) \leq C \left[ \mathbb{E}[\|f\|_{W^{n,s}}^{qr}]^{\frac{1}{r}} + \mathbb{E}[\|g\|_{W^{n,s}}^{qr}]^{\frac{1}{r}} \right]^{\frac{1}{p}} W_{(p-q)s,\infty}^{1-\frac{q}{p}}(f_{\#}\mathbb{P}, g_{\#}\mathbb{P}) \quad (2.16)$$

$$\leq C \left[ \mathbb{E}[\|f\|_{W^{n,s}}^{qr}]^{\frac{1}{r}} + \mathbb{E}[\|g\|_{W^{n,s}}^{qr}]^{\frac{1}{r}} \right]^{\frac{1}{p}} \|f - g\|_{L^{(p-q)s}(\Omega, L^\infty(X, \mathbb{R}))}^{1-\frac{q}{p}}. \quad (2.17)$$

*Proof.* Here, we follow the proof of [11, Theorem 5.9]. The proof is the same, replacing the  $C^\alpha$ -seminorms by the Sobolev norms. ■

**Theorem 2.11** (Stochastic stability of representations, [11]). *Let  $\mathcal{B}$  be a Banach space and  $\Psi : \mathcal{D}_p \rightarrow \mathcal{B}$  be an  $\alpha$ -Hölder continuous functional. Let  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_{\alpha q}(\mathcal{D}_p)$ , then*

$$\|\mathbb{E}_{\mathbb{P}}[\Psi] - \mathbb{E}_{\mathbb{Q}}[\Psi]\|_{\mathcal{B}} \leq W_{q,\|\cdot\|_{\mathcal{B}}}(\Psi_{\#}\mathbb{P}, \Psi_{\#}\mathbb{Q}) \leq \|\Psi\|_{C^\alpha(\mathcal{D}_p, \mathcal{B})} W_{q\alpha, d_p}^\alpha(\mathbb{P}, \mathbb{Q}). \quad (2.18)$$

### 3 Euler and Betti curves

Degree-by-degree, we define the so-called Betti and Euler curves.

**Definition 3.1.** Let  $f \in W^{n,s}(X, \mathbb{R})$  with  $n > \frac{d}{s}$ . The  $k$ th degree Betti curve of  $f$ , denoted  $\beta_k(x, f)$  is defined as

$$\beta_k(f)(x) := \beta_k(f, x) := \text{Dgm}_k(f)(R_x), \quad (3.19)$$

where  $\text{Dgm}_k(f)$  is seen as a persistence measure. The Euler curve of  $f$  is defined as

$$\chi(f)(x) := \chi(f, x) := \sum_{k=0}^d (-1)^k \text{Dgm}_k(f)(R_x) \quad (3.20)$$

*Remark 3.2.* We extend this definition to persistence measures in the obvious way.

Considered pointwise, showing a stability result for  $\beta_k(f, x)$  is hopeless. This follows from the fact that we may perturb  $f$  so that we displace the points in the diagram pushing them over the boundary of  $R_x$ , thereby introducing a perturbation in  $\beta_k(f, x)$  of a potentially infinite number of points.

The correct point of view is to allow ourselves to consider the  $\beta_k(f, x)$  as an element of  $L^1(\mathbb{R})$ , or more generally as a distribution.

**Proposition 3.3.** Let  $X$  be a  $d$ -dimensional Riemannian manifold and let  $f \in W^{n,s}(X, \mathbb{R})$  for  $d < n$ . Then, for every  $k$ ,  $\beta_k(f), \chi(f) \in L^1(\mathbb{R})$  and thus also in  $\mathcal{D}'(\mathbb{R})$ . Moreover,

$$\|\beta_k(f)\|_{L^1(\mathbb{R})} = \text{Pers}_1(\text{Dgm}_k(f)). \quad (3.21)$$

*Remark 3.4.*  $\beta_k(f)$  can be infinite for some values of  $x$ . Nonetheless, the degree of regularity considered ensures that we can still make sense of average values of  $\beta_k$  around small neighbourhoods of a level  $x$ . This justifies that we look at the Betti and Euler curves as distributions.

*Proof.* We do the proof for sublevel sets, noticing that the proof for superlevel sets is completely analogous. Since  $\beta_k(f)$  is a positive function,

$$\begin{aligned} \|\beta_k(f)\|_{L^1(\mathbb{R})} &= \int_{\mathbb{R}} \beta_k(f, x) dx = \int_{\mathbb{R}} \text{Dgm}_k(f)(R_x) dx = \int_{\mathbb{R}} \left[ \int_{\overline{\mathcal{X}}} 1_{R_x}(z) d\text{Dgm}(f)(z) \right] dx \\ &= \int_{\overline{\mathcal{X}}} \left[ \int_{\mathbb{R}} 1_{R_x}(z) dx \right] d\text{Dgm}(f)(z). \end{aligned}$$

For any  $z = (z_1, z_2) \in \overline{\mathcal{X}}$ ,

$$1_{R_x}(z_1, z_2) = 1_{[0, \infty[}(z_2 - x) 1_{[0, \infty[}(x - z_1). \quad (3.22)$$

Using the translation invariance of the Lebesgue measure on  $\mathbb{R}$ , we get

$$\int_{\mathbb{R}} 1_{R_x}(z_1, z_2) dx = z_2 - z_1. \quad (3.23)$$

This is nothing other than the distance  $d_{\mathbb{R}^2, \infty}$  from  $(z_1, z_2)$  to the diagonal, so

$$\int_{\overline{\mathcal{X}}} (z_2 - z_1) d\text{Dgm}(f)(z_1, z_2) = \text{Pers}_1(\text{Dgm}_k(f)) < \infty, \quad (3.24)$$

as soon as  $\frac{d}{n} < 1$  by virtue of corollary 2.3. ■

**Corollary 3.5** (No stability for  $d_p$ -stability for  $p > 1$ ). *Let  $p > 1$  and consider  $\beta : \mathcal{D}_p \rightarrow L^1(\mathbb{R})$  the map associating  $D \mapsto [x \mapsto D(R_x)]$ . Then,  $\beta$  is discontinuous.*

*Remark 3.6.* This suggests that  $\chi : \mathcal{D}_p \rightarrow L^1(\mathbb{R})$  is discontinuous as well.

*Proof.* We give two different arguments which prove the statement. First, consider a diagram  $D \notin \mathcal{D}_1$  such that  $D \in \mathcal{D}_p$  and any other diagram  $D' \in \mathcal{D}_1 \cap \mathcal{D}_p$ . Then,  $d_p(D, D') < \infty$ , but the reverse triangle inequality entails

$$\left| \|\beta(D)\| - \|\beta(D')\| \right|_{L^1} \leq \|\beta(D) - \beta(D')\|_{L^1} = \infty, \quad (3.25)$$

since the lower bound is infinite. Second, suppose that  $\beta$  is continuous. Then the map  $\mathcal{D}_p \rightarrow \mathbb{R}$ , given by  $D \mapsto \|\beta(D)\|_{L^1}$  must be continuous by continuity of the norm. However, by [8, Proposition 5.1] this map cannot be continuous, since

$$\int_{\mathbb{R}} 1_{R_x}(z_1, z_2) dx = z_2 - z_1 \quad (3.26)$$

is linear in the distance to the diagonal, and in particular not  $O(|z_2 - z_1|^p)$  as would be required for continuity with respect to  $d_p$ . ■

*Remark 3.7.* Divol and Lacombe's work [8, Proposition 5.1] also explain why it is necessary – even within  $\mathcal{D}_1$  – to regard the Betti and Euler curves as distributions. Indeed, notice that pointwise in  $x$ ,  $1_{R_x}$  doesn't decrease to 0 at the required rate.

**Corollary 3.8** ( $d_1$ -stability of Betti and Euler curves). *Let  $D, D' \in \mathcal{D}_1$ , then*

$$\|\beta(D) - \beta(D')\|_{L^1(\mathbb{R})} \leq 2 d_1(D, D'). \quad (3.27)$$

*Remark 3.9.* This shows that  $\beta : \mathcal{D}_1 \rightarrow L^1(\mathbb{R})$  is a 2-Lipschitz representation.

*Proof.* Consider a transport map  $\pi \in \Gamma(D, D')$ . Then,

$$\|\beta(D) - \beta(D')\|_{L^1(\mathbb{R})} \leq \int_{\overline{\mathcal{X}}^2} \left[ \int_{\mathbb{R}} |1_{R_x}(z) - 1_{R_x}(w)| dx \right] d\pi(z, w).$$

We carefully check that

$$\int_{\mathbb{R}} |1_{R_x}(z) - 1_{R_x}(w)| dx \leq 2 d_{\mathbb{R}^2, \infty}(z, w), \quad (3.28)$$

which entails the result by taking  $\pi$  to be an optimal transport for  $d_1$ . ■

**Proposition 3.10** (Interpolation for optimal transport). Let  $0 < p < q \leq \infty$  and  $\theta \in ]0, 1[$ . Define  $p_\theta$  by

$$\frac{1}{p_\theta} = \frac{\theta}{p} + \frac{1-\theta}{q}. \quad (3.29)$$

Then, for  $\mu, \nu \in \mathcal{D}_p \cap \mathcal{D}_q$

$$d_{p_\theta}(\mu, \nu) \leq 2^{1-\theta} d_p^\theta(\mu, \nu) (\text{Pers}_q(\mu) + \text{Pers}_q(\nu))^{1-\theta} \quad (3.30)$$

$$d_{p_\theta}(\mu, \nu) \leq 2^\theta d_q(\mu, \nu)^{1-\theta} (\text{Pers}_p(\mu) + \text{Pers}_p(\nu))^\theta. \quad (3.31)$$

Consequently, if  $p \leq r \leq q$ , then  $\mathcal{D}_p \cap \mathcal{D}_q \subset \mathcal{D}_r$ .

*Proof.* We prove the first inequality, the second one following by analogy, by interverting the roles of  $p$  and  $q$ . Let  $\pi$  be an optimal transport for  $d_p$ . Applying Hölder's inequality,

$$\begin{aligned} d_{p_\theta}(\mu, \nu) &\leq \|d_{\mathbb{R}^2, \infty}\|_{L^{p_\theta}(\pi)} \leq \|d_{\mathbb{R}^2, \infty}\|_{L^p(\pi)}^\theta \|d_{\mathbb{R}^2, \infty}\|_{L^q(\pi)}^{1-\theta} \\ &= d_p(\mu, \nu)^\theta \left[ \int_{\bar{\mathcal{X}}^2} d_{\mathbb{R}^2, \infty}^q(z, z') d\pi(z, z') \right]^{\frac{1-\theta}{q}} \\ &\leq d_p(\mu, \nu)^\theta \left[ 2^q \int_{\bar{\mathcal{X}}^2} d_{\mathbb{R}^2, \infty}^q(z, \Delta) + d_{\mathbb{R}^2, \infty}^q(\Delta, z') d\pi(z, z') \right]^{\frac{1-\theta}{q}}, \end{aligned}$$

where the inequality on the last line holds everywhere on the support of  $\pi$ . This can be shown by defining

$$S = \{(z, z') \in \bar{\mathcal{X}}^2 \cap \text{supp}(\pi) \mid d_{\mathbb{R}^2, \infty}(z, z') > d_{\mathbb{R}^2, \infty}(z, \Delta) + d_{\mathbb{R}^2, \infty}(z', \Delta)\}. \quad (3.32)$$

This set  $S$  either has null or positive measure. If it has positive measure, then we can modify the transport plan  $\pi$  by sending the projections of  $S$  to the diagonal, thereby producing a transport plan of strictly inferior cost to that of  $\pi$ , which is a contradiction. Hence,  $S$  is of null measure, so the equality holds over the support of the measure. This entails

$$d_{p_\theta}(\mu, \nu) \leq 2^{1-\theta} d_p(\mu, \nu)^\theta (\text{Pers}_q(\mu) + \text{Pers}_q(\nu))^{1-\theta}.$$

If  $q = \infty$ , since  $\pi$  is an optimal transport between  $\mu$  and  $\nu$  and  $\mu, \nu \in \mathcal{D}_\infty$ ,  $\pi$  itself must have compact support and the diameter of the support is bounded above by  $\text{Pers}_\infty(\mu) \vee \text{Pers}_\infty(\nu)$ , so the inequality of the proposition follows. ■

*Remark 3.11.* Taking  $0 < p < 1 \leq q$ , proposition 3.10 allows us to say that, despite  $\beta$  being discontinuous on  $\mathcal{D}_q$ , this functional is continuous on  $(\mathcal{D}_p \cap \mathcal{D}_q, d_q)$ . As shown in this paper, we may always do this for diagrams stemming from functions in  $W^{n,s}(X, \mathbb{R})$ , provided that  $n > d$ . This reconciles our results with the ones found in [7], where the case of subanalytic functions was studied (in which case  $\text{Dgm}(f) \in \mathcal{D}_0 \cap \mathcal{D}_\infty$ ).

**Theorem 3.12** (Stability of Euler and Betti curves under perturbations in distribution). *Let  $f$  and  $g$  be two stochastic processes on a  $d$ -dimensional compact Riemannian manifold  $X$ , defined on a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that almost surely  $f, g \in W^{n,s}(X, \mathbb{R})$  with  $\frac{d}{n} < 1$ . Then, for every  $\frac{d}{n} < \alpha < 1$ , there exists a constant  $C = C_{X,n,s}$  such that,*

$$W_{1,L^1}(\beta_k(f)_\# \mathbb{P}, \beta_k(g)_\# \mathbb{P}) \leq C \left[ \|f\|_{L^1(\Omega, W^{n,s})}^\alpha + \|g\|_{L^1(\Omega, W^{n,s})}^\alpha \right] W_{1,L^\infty}^{1-\alpha}(f_\# \mathbb{P}, g_\# \mathbb{P}), \quad (3.33)$$

and a similar inequality holds by replacing the Betti curves with the Euler curves  $\chi$ .

*Proof.* Recognizing that  $\beta_k$  are Lipschitz representations, it is a simple application of the stochastic stability theorem and its analog for representations, where we set  $q = \alpha$ ,  $p = 1$  and  $s = \frac{1}{1-\alpha}$  and  $r = \frac{1}{\alpha}$ . ■

*Remark 3.13.* Using the Kantorovich-Rubenstein duality, a lower bound of  $W_{1,L^1}$  can be easily established. Recall that [12, Particular case 5.16]

$$\sup_{\psi \in \text{Lip}_1(L^1(\mathbb{R}), \mathbb{R})} \mathbb{E}[\psi(\chi(f)) - \psi(\chi(g))] \leq W_{1,L^1}(\chi(f)_{\#}\mathbb{P}, \chi(g)_{\#}\mathbb{P}). \quad (3.34)$$

This point of view further justifies our view of  $\beta_k$  and  $\chi$  as distributions. Indeed, for  $\varphi \in \mathcal{D}(\mathbb{R})$ , the map

$$\psi : \chi \mapsto \int_{\mathbb{R}} \chi(x) \varphi(x) dx, \quad (3.35)$$

is clearly Lipschitz. This distributional view of the Euler and Betti curves is (unsurprisingly) also stable with respect to perturbations in the distributions of the random fields from which they stem from.

A particularly interesting family of Lipschitz functionals (in light of recent developments [9]) to consider is

$$\psi_{\theta} : \chi \mapsto \int_{\mathbb{R}} e^{-ix\theta} \chi(x) dx. \quad (3.36)$$

Using the Kantorovich-Rubenstein duality on the real and imaginary components of the above integral separately, and taking the supremum over all  $\theta$  we arrive that the conclusion that

$$\|\mathbb{E}[\mathcal{F}\chi(f) - \mathcal{F}\chi(g)]\|_{L^{\infty}(\mathbb{R})} \leq \sqrt{2} W_{1,L^1}(\chi(f)_{\#}\mathbb{P}, \chi(g)_{\#}\mathbb{P}), \quad (3.37)$$

where  $\mathcal{F}$  denotes the Fourier transform. Interestingly, one can apply the same reasoning for different integral transforms, provided that their kernel is such that the transform is Lipschitz on  $L^1$ . Of course, these arguments are all also valid for the Betti curves as well. Applying the stability of the Euler and Betti curves under perturbations in the distributions of  $f$  and  $g$ , we arrive at the conclusion that these transforms are stable with respect to changes in distribution.

*Remark 3.14.* Looking at Euler curves, it is possible to recover the Euler characteristic of level sets of  $f$ . Since for every  $\varepsilon > 0$ ,  $\{f > x - \varepsilon\} \cap \{f < x + \varepsilon\} = \{x - \varepsilon < f < x + \varepsilon\}$  and that  $\{f > x - \varepsilon\} \cup \{f < x + \varepsilon\} = X$ , by the valuation property of the Euler characteristic, we recover

$$\chi(\{x - \varepsilon < f < x + \varepsilon\}) + \chi(X) = \chi(f, x - \varepsilon) + \chi(-f, -x - \varepsilon), \quad (3.38)$$

by taking  $\varepsilon \rightarrow 0$ , we obtain  $\chi(f = x)$ , whenever it is defined.

## References

- [1] R. J. Adler, O. Bobrowski, M. S. Borman, E. Subag, and S. Weinberger. Persistent homology for random fields and complexes. In *Institute of Mathematical Statistics Collections*, pages 124–143. Institute of Mathematical Statistics, 2010.
- [2] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer New York, 2007.
- [3] O. Bobrowski and M. S. Borman. Euler integration of Gaussian random fields and persistent homology. *Journal of Topology and Analysis*, 04(01):49–70, mar 2012.
- [4] L. Buhovsky, J. Payette, I. Polterovich, L. Polterovich, E. Shelukhin, and V. Stojisavljević. Coarse nodal count and topological persistence, 06 2022.
- [5] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. *The Structure and Stability of Persistence Modules*. Springer International Publishing, 2016.
- [6] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have  $L^p$ -stable persistence. *Foundations of Computational Mathematics*, 10(2):127–139, Jan 2010.
- [7] J. Curry, S. Mukherjee, and K. Turner. How many directions determine a shape and other sufficiency results for two topological transforms, 05 2018.



- [8] V. Divol and T. Lacombe. Understanding the topology and the geometry of the persistence diagram space via optimal partial transport. *CoRR*, abs/1901.03048, 2019.
- [9] V. Lebovici. Hybrid transforms of constructible functions. *Foundations of Computational Mathematics 2022*, 11 2021.
- [10] S. Y. Oudot. *Persistence Theory - From Quiver Representations to Data Analysis*, volume 209 of *Mathematical surveys and monographs*. American Mathematical Society, 2015.
- [11] D. Perez. On  $C^0$ -persistent homology and trees. <https://arxiv.org/abs/2012.02634>, Dec. 2020.
- [12] C. Villani. *Optimal Transport*. Springer Berlin Heidelberg, 2009.