Persistent Homology under Non-uniform Error *

Paul Bendich[1,2,3], Herbert Edelsbrunner[1,2,3,4], Michael Kerber[1,2], and Amit Patel[1,2]

[1] IST Austria (Institute of Science and Technology Austria), Klosterneuburg, Austria,
[2] Dept. Comput. Sci., Duke Univ., Durham, North Carolina,
[3] Dept. Mathematics, Duke Univ. Durham, North Carolina,
[4] Geomagic, Research Triangle Park, North Carolina.

Abstract. Using ideas from persistent homology, the robustness of a level set of a real-valued function is defined in terms of the magnitude of the perturbation necessary to kill the classes. Prior work has shown that the homology and robustness information can be read off the extended persistence diagram of the function. This paper extends these results to a non-uniform error model in which perturbations vary in their magnitude across the domain.

Keywords. Topological spaces, continuous functions, level sets, perturbations, homology, extended persistence, error models, stability, robustness.

1 Introduction

Numerical errors are an inescapable by-product of scientific and other computations, and thus they have inspired the creation of entire fields of mathematical inquiry, including numerical analysis and statistics. There are many sources, such as randomness, limited resolution, and limited computational resources. There are also many coping strategies, such as improving accuracy, or finding credible substitutes for the elusive ideal. The contributions of this paper belong to the analytic approach that gives estimates on how far the result is from the ideal. More specifically, we study real-valued functions, which model many real world applications, including medical images and satellite pictures. To extract information from a function, we consider level and sublevel sets and their topology expressed in terms of homology groups. The question thus arises to what extent the homology of a level or sublevel set is sensitive to perturbations of the function. In this paper, we study the effect of perturbations which mimic the common situation in which measurements vary in their accuracy, and we call these *non-uniform perturbations*. We assume that this variation is tied to the location, and that we have complete information on how the accuracy varies across the domain. We capture this information in a (non-uniform) error model, which will be formally defined in Section 3.

On a technical level, we extend the algebraic and measure theoretic concept of persistent homology to non-uniform error models. Specifically, we define the non-uniform

* This research is partially supported by the Defense Advanced Research Projects Agency (DARPA), under grants HR0011-05-1-0057 and HR0011-09-0065, as well as the National Science Foundation (NSF), under grant DBI-0820624.

persistence diagram of a function under an error model, and we relate it to the conventional persistence diagram (defined for uniform error) and to non-uniform persistence diagrams of other functions under the same error model. Using results from prior work, we then extend these results to well diagrams, which characterize the robustness of level sets. The main technical concept is a transformation of functions that turns non-uniform into uniform error and thus extends much of the machinery of persistent and robust homology to non-uniform error. Under the additional assumption of the linearity of the error model, this includes one of the cornerstones of the theory of persistent homology, namely the stability of the diagram.

Outline. Section 2 reviews the background from persistent homology. Section 3 introduces error models and discusses their effect on the persistence diagrams of functions. Section 4 introduces dual error models and uses them to transform non-uniform to uniform error. Section 5 demonstrates that linear error models give rise to a richer theory than the more general error models. Finally, Section 6 concludes the paper.

2 Background

In this section, we review the background on homology and on persistence; see Munkres [13] and Hatcher [12] for standard texts in algebraic topology, and Edelsbrunner, Harer [10] for a recent book in computational topology. While the material in this section may seem dry and technical, we remind the reader of the connection to the fundamental problems of feature extraction, matching, and classification for images, shapes, and more general data. Indeed, persistent homology has been applied to a host of different shape and data analysis questions, including for natural images [3], trademark images [4], brain structure [5], sensor networks [8], and gene expression [9].

Persistence. The persistence of homology classes along a filtration of a topological space can be defined in a quite general context. For the purposes of this paper, we need only a particular type of filtration, one defined by the sublevel sets of a real-valued function on a compact topological space. Given such a space $\mathbb{X}$ and a function $f : \mathbb{X} \to \mathbb{R}$, we call a set $\mathbb{X}_r(f) = f^{-1}(-\infty, r]$ a *sublevel set*, and we consider the nested sequence of these sets. Whenever $r \leq s$, the inclusion $\mathbb{X}_r(f) \hookrightarrow \mathbb{X}_s(f)$ induces maps on the homology groups $\mathsf{H}(\mathbb{X}_r(f)) \to \mathsf{H}(\mathbb{X}_s(f))$. Here we use modulo 2 homology, that is, the coefficient group is $\mathbb{Z}/2\mathbb{Z}$. In addition, we simplify the notation by taking the direct sum of the homology groups over all dimensions. A real value r is called a *homological regular value* of f if there exists an $\varepsilon > 0$ such that the inclusion $\mathbb{X}_{r-\delta}(f) \hookrightarrow \mathbb{X}_{r+\delta}(f)$ induces an isomorphism between homology groups for all $\delta < \varepsilon$. Otherwise we call r a *homological critical value*. We say that f is *tame* if the homology groups of each sublevel set have finite rank and if there are only finitely many homological critical values. Assuming that f is tame, with ordered homological critical values $r_1 < r_2 < \ldots < r_n$, we select $n+1$ homological regular values s_i such that $s_0 < r_1 < s_1 < \ldots < r_n < s_n$, and set $\mathbb{X}_i = \mathbb{X}_{s_i}(f)$. Note that $\mathbb{X}_0 = \emptyset$ and $\mathbb{X}_n = \mathbb{X}$, by compactness. The inclusions $\mathbb{X}_i \hookrightarrow \mathbb{X}_j$ induce maps $\mathsf{f}^{i,j} : \mathsf{H}(\mathbb{X}_i) \to \mathsf{H}(\mathbb{X}_j)$ for $0 \leq i \leq j \leq n$ and give a *filtration* of the homology groups:

$$0 = \mathsf{H}(\mathbb{X}_0) \to \mathsf{H}(\mathbb{X}_1) \to \ldots \to \mathsf{H}(\mathbb{X}_n) = \mathsf{H}(\mathbb{X}). \qquad (1)$$

Given a class $\alpha \in \mathsf{H}(\mathbb{X}_i)$, we say that α is *born* at $\mathbb{X}_i$ if $\alpha \notin \operatorname{im} \mathsf{f}^{i-1,i}$. A class α born at $\mathbb{X}_i$ is said to *die entering* $\mathbb{X}_j$ if $\mathsf{f}^{i,j}(\alpha) \in \operatorname{im} \mathsf{f}^{i-1,j}$ but $\mathsf{f}^{i,j-1}(\alpha) \notin \operatorname{im} \mathsf{f}^{i-1,j-1}$. We remark that if a class α is born at $\mathbb{X}_i$, then every class in the coset $[\alpha] = \alpha + \operatorname{im} \mathsf{f}^{i-1,i}$ is born at the same time. Of course, whenever such an α dies entering $\mathbb{X}_j$, the entire coset $[\alpha]$ also dies with it.

Extended persistence. The filtration in (1) begins with the zero group but ends with a potentially nonzero group. Hence, it is possible to have classes that are born but never die. We call these *essential* classes, as they represent the actual homology of the space $\mathbb{X}$. To measure the persistence of the essential classes, we follow [7] and extend (1) using relative homology groups. More precisely, we consider for each i the *superlevel set* $\mathbb{X}^i = f^{-1}[s_{n-i}, \infty)$. By compactness, we have $\mathbb{X}^0 = \emptyset$ and $\mathbb{X}^n = \mathbb{X}$ and therefore $\mathsf{H}(\mathbb{X}, \mathbb{X}^0) = \mathsf{H}(\mathbb{X})$ and $\mathsf{H}(\mathbb{X}, \mathbb{X}^n) = 0$. For $0 \leq i \leq j \leq n$, the inclusions $\mathbb{X}^i \hookrightarrow \mathbb{X}^j$ induce maps on relative homology. We then consider the extended filtration:

$$\begin{aligned} 0 = \mathsf{H}(\mathbb{X}_0) &\to \mathsf{H}(\mathbb{X}_1) \to \ldots \to \mathsf{H}(\mathbb{X}_n) = \mathsf{H}(\mathbb{X}) \\ &= \mathsf{H}(\mathbb{X}, \mathbb{X}^0) \to \mathsf{H}(\mathbb{X}, \mathbb{X}^1) \ldots \to \mathsf{H}(\mathbb{X}, \mathbb{X}^n) = 0, \end{aligned} \tag{2}$$

and extend the notions of birth and death in the obvious way. Now all classes will eventually die, as this filtration begins and ends with the zero group. The information contained within the extended filtration (2) can be compactly represented by *persistence diagrams* $\operatorname{Dgm}_p(f)$, one for each dimension p in homology. These diagrams are multisets of points drawn in three copies of the extended plane, shrunk to finite size and arranged side by side, as shown in Figure 1. For technical reasons, we always consider the diagram to contain infinitely many copies of each point on the baseline, where the birth and death coordinates coincide. By $\operatorname{Dgm}(f)$, we mean the points of all diagrams in all dimensions, overlaid as one multiset of points.

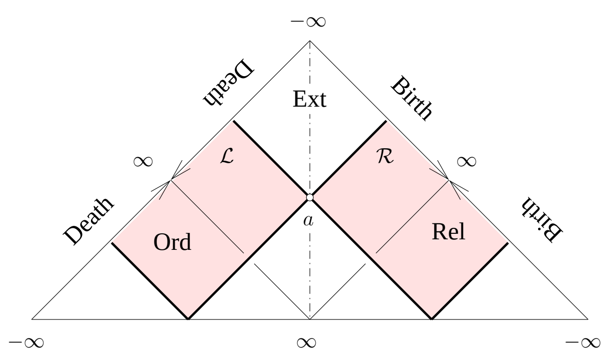

Fig. 1: The persistence diagram of a function consists of three subdiagrams, Ord, Ext, and Rel, arranged in a triangle as shown. The points in the two shaded rectangles, $\mathcal{L}$ and $\mathcal{R}$, represent the homology of the level set defined by a.

Contained within each $\mathrm{Dgm}_p(f)$ are three subdiagrams, corresponding to three different combinations of birth and death location. The *ordinary subdiagram*, $\mathrm{Ord}_p(f)$, contains the point (r_i, r_j) for each coset of classes that are born at $\mathbb{X}_i$ and die entering $\mathbb{X}_j$. Here, birth and death both happen during the first half of (2). The *extended subdiagram*, $\mathrm{Ext}_p(f)$, contains (r_i, r_j) for each coset of classes that is born at $\mathbb{X}_i$ and dies entering $(\mathbb{X}, \mathbb{X}^{n-j+1})$. Finally, the *relative subdiagram*, $\mathrm{Rel}_p(f)$, contains (r_i, r_j) for each coset of classes that is born at $(\mathbb{X}, \mathbb{X}^{n-i+1})$ and dies entering $(\mathbb{X}, \mathbb{X}^{n-j+1})$. We arrange the three subdiagrams side by side, while reversing the birth axis of the extended subdiagram and both axes of the relative subdiagram. We do so to simplify the interpretation of the diagram, as will be explained later.

Stability. An essential property of the persistence diagrams is their stability under small changes of the function. To make this precise, we need to define a distance between functions and a distance between diagrams. Given two functions $f, h : \mathbb{X} \to \mathbb{R}$, and a real number $r \geq 0$, we call h an *r-perturbation* of f if $|f(x) - h(x)| \leq r$, for every $x \in \mathbb{X}$. This relation is symmetric and can be used to define a metric on the space of real-valued functions on $\mathbb{X}$, setting $\|f - h\|_\infty$ equal to the minimum r such that f and h are r-perturbations of each other. This is of course the standard L_∞-distance between functions. Given any two persistence diagrams, $\mathcal{D}$ and $\mathcal{D}'$, we define the *bottleneck distance* between them as the largest distance between matched points (in maximum norm) under the best possible matching between the diagrams. More formally,

$$W_\infty(\mathcal{D}, \mathcal{D}') \;=\; \inf_\gamma \sup_u \|u - \gamma(u)\|_\infty, \tag{3}$$

where u ranges over all points of the diagram $\mathcal{D}$, and γ ranges over all bijections from $\mathcal{D}$ to $\mathcal{D}'$. We then have:

1 (Stability Theorem [6]) *Given continuous and tame functions $f, h : \mathbb{X} \to \mathbb{R}$ on a compact topological space, we have $W_\infty(\mathrm{Dgm}_p(f), \mathrm{Dgm}_p(h)) \leq \|f - h\|_\infty$ for each homological dimension p.*

This result seems natural as we can construct a homotopy between f and h in which the values change continuously, each by at most $\|f - h\|_\infty$. However, consider that critical points may appear and disappear and global rearrangements may cause the pairing between critical values change during the homotopy.

Robustness. The persistence diagram of a real-valued function f carries a wealth of information. For example, it allows us to measure the robustness of the homology of level sets to perturbations of f. We now make this precise.

Fixing some value $a \in \mathbb{R}$ and a real number $r \geq 0$, we consider the preimage of the interval: $\mathbb{X}_{[a-r,a+r]}(f) = f^{-1}[a - r, a + r]$. For every r-perturbation h of f, the level set $h^{-1}(a)$ will be a subset of this preimage, and hence there is an induced map on homology $\mathsf{j}_h : \mathsf{H}(h^{-1}(a)) \to \mathsf{H}(\mathbb{X}_{[a-r,a+r]}(f))$. Following [11], we say that a class $\alpha \in \mathsf{H}(\mathbb{X}_{[a-r,a+r]}(f))$ is *supported* by h if it belongs to the image of j_h; in other words, if the set $h^{-1}(a)$ carries a chain representative of α. The *well group* of f and $r \geq 0$ is then defined to consist of those classes that are supported by all r-perturbations of f. It is a subgroup of $\mathsf{H}(\mathbb{X}_{[a-r,a+r]}(f))$. The sequence of well groups no longer forms a filtration but a more general zigzag module, introduced recently in [2]. These modules can still be characterized, albeit less directly, by their persistence diagrams. In the case of well groups, all births happen at the beginning, so the diagram simplifies to a multiset of points that mark deaths on the real line. We refer to this multiset as the *well diagram* of the function and the value defining the level set. It expresses what we call the *robustness* of the homology classes, that is, their resilience to perturbations of the function. In [1], the authors demonstrate a simple relationship between the persistence diagrams of f and the well diagrams of f and a, for every $a \in \mathbb{R}$. To describe this relationship, we first define for each homological dimension p two multisets of points:

$$\begin{aligned} \mathcal{L}_p[a] &= \{(x,y) \in \mathrm{Ord}_p(f) \mid x < a, y > a\} \sqcup \{(x,y) \in \mathrm{Ext}_p(f) \mid x < a, y > a\}, \\ \mathcal{R}_p[a] &= \{(x,y) \in \mathrm{Ext}_p(f) \mid x > a, y < a\} \sqcup \{(x,y) \in \mathrm{Rel}_p(f) \mid x > a, y < a\}, \end{aligned}$$

where x refers of course to the birth coordinate and y to the death coordinate of the point; see Figure 1. Then the p-dimensional homology of $f^{-1}(a)$ is characterized by points (x,y) in $\mathcal{L}_p[a] \cup \mathcal{R}_{p+1}[a]$. If the point belongs to $\mathcal{L}_p[a]$, then its robustness is equal to $\min\{a-x, y-a\}$, while if the points belongs to $\mathcal{R}_{p+1}[a]$, then its robustness is $\min\{x-a, a-y\}$. To get the well diagram of f and a, we then just plot the robustness value for every point in $\mathcal{L}_p[a] \cup \mathcal{R}_{p+1}[a]$ on the real line.

3 Non-uniform Error

In this section, we extend the concepts of persistence and robustness from a uniform to a non-uniform notion of error. We begin by introducing the error model as a 1-parameter family of functions.

Error model. It is convenient to substitute the extended real line, $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, for $\mathbb{R}$ as the range of our functions, and therefore also of the error model.

2 (Definition) *An* error model *on a compact topological space* $\mathbb{X}$ *is a continuous mapping* $E : \mathbb{X} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ *that satisfies the following two properties:*

monotonicity: $E(x,r) < E(x,s)$ *for all* $x \in \mathbb{X}$ *and all* $r < s$*;*
normalization: $E(x,0) = 0, E(x,\infty) = \infty, E(x,-\infty) = -\infty$*, for all* $x \in \mathbb{X}$*.*

The error model is uniform *if* $E(x,r) = r$ *for all* $(x,r) \in \mathbb{X} \times \bar{\mathbb{R}}$*, and it is* non-uniform *otherwise.*

Fixing either a point x or a radius r, we get restricted functions $e_x : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ and $e_r : \mathbb{X} \to \bar{\mathbb{R}}$ defined by $e_x(r) = e_r(x) = E(x,r)$. Note that the two conditions above guarantee that e_x is invertible, a property we make use of later. Intuitively, one might imagine r to be a global noise parameter which leads, via the error model E, to a variable amount $e_x(r)$ of noise across $\mathbb{X}$. The continuity of E models our assumption that errors are not independent and indeed are closely related for nearby points. In Section 5, we will add the further assumption that E is linear, meaning $E(x,r) = r \cdot E(x,1)$ for all $(x,r) \in \mathbb{X} \times \bar{\mathbb{R}}$. For the moment we make no such requirement of our model.

Non-uniform filtrations and persistence. Given a function $f : \mathbb{X} \to \bar{\mathbb{R}}$ and any two values $r \leq s$, the standard notion of an *interlevel set* is the preimage of the interval: $\mathbb{X}_{[r,s]}(f) = f^{-1}[r,s]$. In applications in which $\mathbb{X}$ is 3-dimensional, this construct is often referred to as an *interval volume*. Extending it to our framework, we define the *non-uniform interlevel set* as the set of points with image between the bounds specified by the error model:

$$\mathbb{X}_{[r,s]}(f,E) = \{x \in \mathbb{X} \mid E(x,r) \leq f(x) \leq E(x,s)\}.$$

As illustrated in Figure 2, we can construct the non-uniform interlevel set by intersecting the graph of f with the strip of points between the graphs of e_r and e_s, and projecting the intersection to $\mathbb{X}$. In the special case in which $r = -\infty$, we write

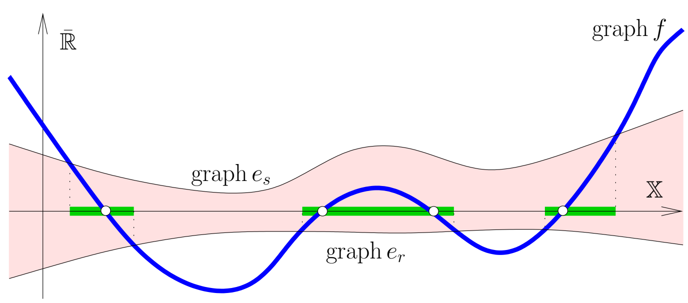

Fig. 2: The graph of a function f, the shaded strip bounded from below and above by the graphs of restrictions of the error model, and the non-uniform interlevel set obtained by projecting the intersection. For $r \leq 0 \leq s$, the non-uniform interlevel set contains the zeroset of f.

$\mathbb{X}_s(f,E) = \mathbb{X}_{[-\infty,s]}(f,E)$ and call it a *non-uniform sublevel set*. Similarly, if $s = \infty$, we write $\mathbb{X}^r(f,E) = \mathbb{X}_{[r,\infty]}(f,E)$ and call it a *non-uniform superlevel set*.

Whenever $r \leq s$, the monotonicity requirement guarantees the inclusion of $\mathbb{X}_r(f,E)$ in $\mathbb{X}_s(f,E)$, and the inclusion of $\mathbb{X}^s(f,E)$ in $\mathbb{X}^r(f,E)$. Hence, just as in Section 2, the non-uniform sublevel and superlevel sets give an extended filtration of $\mathbb{X}$. As a result, we have in each homological dimension p a *non-uniform persistence diagram*, denoted by $\mathrm{Ngm}_p(f,E)$. As before, we write $\mathrm{Ngm}(f,E)$ for the overlay of the diagrams in all dimensions. In Section 5, we will see that, under the assumption of a linear error model, these non-uniform diagrams are stable.

Non-uniform perturbations and robustness. Now suppose that we have a function $f : \mathbb{X} \to \bar{\mathbb{R}}$ as well as an error model $E : \mathbb{X} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$. As promised, we create a theoretical language to quantify the robustness of the homology of a level set $f^{-1}(0)$ under non-uniform perturbation. Given another function $h : \mathbb{X} \to \bar{\mathbb{R}}$ and a value $r \geq 0$, we say that h is a *non-uniform r-perturbation* of f, with respect to E, if

$$E(x,-r) \leq f(x) - h(x) \leq E(x,r), \tag{4}$$

for all $x \in \mathbb{X}$. For example, every function of the form $f - e_s$, with $s \in [-r, r]$, is a non-uniform r-perturbation of f. If E is linear, or indeed if each e_x is odd, meaning $e_x(-r) = -e_x(r)$ for every r, then f will also be a non-uniform r-perturbation of h, but this need not be true in the general case. It is useful to understand the connection between non-uniform perturbations and interlevel sets.

3 (Non-uniform Perturbation Lemma) *A function $h : \mathbb{X} \to \bar{\mathbb{R}}$ is a non-uniform r-perturbation of $f : \mathbb{X} \to \bar{\mathbb{R}}$, under the error model $E : \mathbb{X} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$, only if $h^{-1}(0) \subseteq \mathbb{X}_{[-r,r]}(f, E)$.*

Proof. Starting with the definition of a non-uniform r-perturbation, we get

$$E(x, -r) \ \leq \ f(x) \ \leq \ E(x, r),$$

for all points x with $h(x) = 0$. The two inequalities define the non-uniform interlevel set defined by $-r \leq r$, which implies the claimed containment. ◻

We note that $\mathbb{X}_{[-r,r]}(f, E)$ is the smallest interlevel set that contains the zeroset of every non-uniform r-perturbation. In other words, it is the union of all these zerosets. Compare this with the fact that $\mathbb{X}_{[-r,r]}(f, E)$ is also the union of the zerosets of the functions $f - e_s$, for all $s \in [-r, r]$. Adopting the terminology from Section 2, we can now define the *non-uniform well group* of f, E, and a value $r \geq 0$ to consist of those classes in $\mathsf{H}(\mathbb{X}_{[-r,r]}(f, E))$ that are supported by all non-uniform r-perturbations of f. Correspondingly, we get the *non-uniform well diagram* that characterizes the *non-uniform robustness* of the homology of $f^{-1}(a)$ under the error model E.

4 Transformation to Uniform Error

In this section, we show that the non-uniform persistence and well diagrams of a function f and an error model E are really just the uniform diagrams of another function. To do so, we create a dual error model, E^*, which enables us to transform non-uniform interlevel sets and perturbations into uniform interlevel sets and perturbations.

Dual error model. Given an error model $E : \mathbb{X} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$, we recall that the function $e_x : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ defined by $e_x(r) = E(x, r)$ is invertible for every $x \in \mathbb{X}$. We thus have the following definition.

4 (Definition) *The* dual error model *of E is the unique mapping $E^* : \mathbb{X} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ that satisfies $E^*(x, E(x, r)) = r$ for every (x, r) in $\mathbb{X} \times \bar{\mathbb{R}}$.*

Considering the restrictions of E and E^* obtained by fixing a point x, we note that $e_x^* = e_x^{-1}$. The name of E^* is justified by the following result. For technical reasons, it needs the additional assumption that $\mathbb{X}$ be first-countable [14]. This assumption is rather mild; for example, it is satisfied whenever $\mathbb{X}$ can be embedded in finite-dimensional Euclidean space.

5 (Duality Lemma) *Given an error model $E : \mathbb{X} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ on a compact and first-countable topological space $\mathbb{X}$, then (i) E^* is an error model, (ii) $(E^*)^* = E$, (iii) E^* is linear iff E is linear.*

Proof. Claims (ii) and (iii) are obvious. For Claim (i), we note that monotonicity and normalization for E immediately imply the same properties for E^*. So we must only prove that E^* is a continuous function from $\mathbb{X} \times \bar{\mathbb{R}}$ to $\bar{\mathbb{R}}$. To do this, we make use of the following lemma from point-set topology: Assuming a first-countable space $\mathbb{W}$ and a compact space $\mathbb{Y}$, a mapping $H : \mathbb{W} \to \mathbb{Y}$ is continuous iff the graph of H is a closed subset of $\mathbb{W} \times \mathbb{Y}$; see eg. [14]. In our context, we have $\mathbb{W} = \mathbb{X} \times \bar{\mathbb{R}}$ and $\mathbb{Y} = \bar{\mathbb{R}}$.

By assumption, E is continuous, and thus the graph of E is a closed subset of $\mathbb{X} \times \bar{\mathbb{R}} \times \bar{\mathbb{R}}$. On the other hand, the graphs of E and E^* are homeomorphic. To see this, recall that $\mathrm{graph}\, E = \{(x, r, a) \in \mathbb{X} \times \bar{\mathbb{R}} \times \bar{\mathbb{R}} \mid E(x, r) = a\}$. Switching the last two arguments gives a homeomorphism $\psi : \mathbb{X} \times \bar{\mathbb{R}} \times \bar{\mathbb{R}} \to \mathbb{X} \times \bar{\mathbb{R}} \times \bar{\mathbb{R}}$ defined by $\psi(x, r, a) = (x, a, r)$. Then $(x, r, a) \in \mathrm{graph}\, E$ iff $(x, a, r) \in \mathrm{graph}\, E^*$. In other words, the restriction of ψ to the graph of E provides a homeomorphism between $\mathrm{graph}\, E$ and $\mathrm{graph}\, E^*$. In particular, the graph of E^* is closed. Applying the lemma once again, we conclude that E^* is continuous. ◻

Transforming functions. The dual error model E^* of E allows us to associate to each function f a transformation $\mathit{\Gamma}_f$ that turns non-uniform interlevel sets into uniform ones, as we now explain. Given such f, we create a new function $\mathit{\Gamma}_f(f)$ defined by the rule $\mathit{\Gamma}_f(f)(x) = E^*(x, f(x))$ for each $x \in \mathbb{X}$. This definition is illustrated in Figure 3. To construct $\mathit{\Gamma}_f(f)$ geometrically, we use the subset of the graph of E that projects

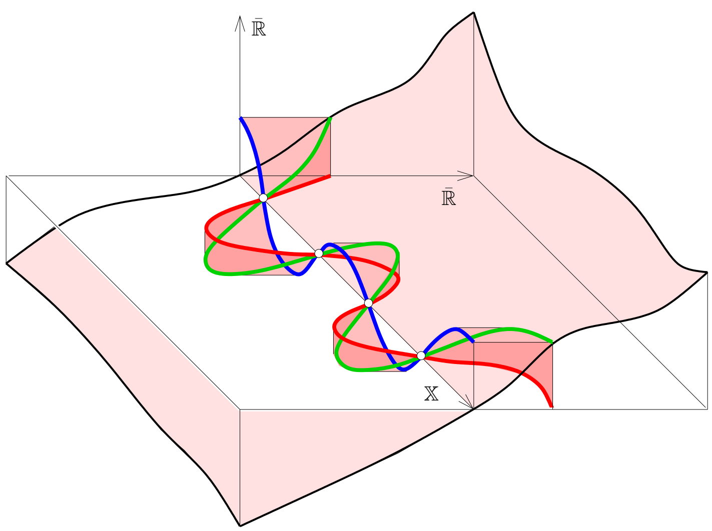

Fig. 3: The surface is the graph of the error model, E. We also see the graph of f in the vertical plane, the graph of $\mathit{\Gamma}_f(f)$ in the horizontal plane, and the curve in the surface that projects to both. We note that f and $\mathit{\Gamma}_f(f)$ have the same zeroset.

to the graph of f in $\mathbb{X} \times \bar{\mathbb{R}}$. The projection of the same subset to $\mathbb{X}$ times the other copy of $\bar{\mathbb{R}}$ gives the graph of $\Gamma_f(f)$. As mentioned earlier, this transformation forms a correspondence between non-uniform and uniform interlevel sets.

6 (Interlevel Set Transformation Lemma) *For every $r \leq s$, the non-uniform interlevel set of f and E is the uniform interlevel set of $\Gamma_f(f)$; that is, $\mathbb{X}_{[r,s]}(f,E) = \mathbb{X}_{[r,s]}(\Gamma_f(f))$.*

Proof. Assume first that $x \in \mathbb{X}_{[r,s]}(f,E)$. Applying the strictly increasing function e_x^* to the chain of inequalities $e_x(r) \leq f(x) \leq e_x(s)$ gives

$$r \;\leq\; e_x^*(f(x)) \;=\; E^*(x, f(x)) \;=\; \Gamma_f(f)(x) \;\leq\; s,$$

and so $x \in \mathbb{X}_{[r,s]}(\Gamma_f(f))$. Reversing the argument proves the claim. ◻

In particular, the non-uniform sublevel and superlevel sets of f and E are the uniform sublevel and superlevel sets of $\Gamma_f(f)$. This implies that the extended filtrations defined by f and E and by $\Gamma_f(f)$ are the same. Hence, they define the same sequence of homology groups and maps between them, and therefore also the same persistence diagrams.

7 (Persistence Diagram Lemma) *The transformation Γ_f turns non-uniform persistence diagrams into uniform ones; that is,* $\mathrm{Ngm}(f,E) = \mathrm{Dgm}(\Gamma_f(f))$.

Transforming perturbations. We now generalize the above construction, applying Γ_f to any other function $h : \mathbb{X} \to \mathbb{R}$ by setting

$$\Gamma_f(h)(x) \;=\; \Gamma_f(f)(x) - E^*(x, f(x) - h(x)),$$

for each $x \in \mathbb{X}$. Of course, we could perform a similar operation using Γ_h, but $\Gamma_f(h)$ and $\Gamma_h(h)$ are not necessarily the same. In Section 5, we will see that a linear error model guarantees the equality of $\Gamma_f(h)$ and $\Gamma_h(h)$. Even without this assumption, we have the following:

8 (Perturbation Transformation Lemma) *The transformation Γ_f turns non-uniform r-perturbations into uniform ones; that is, h is a non-uniform r-perturbation of f iff $\Gamma_f(h)$ is a uniform r-perturbation of $\Gamma_f(f)$.*

Proof. Assume that h is a non-uniform r-perturbation of f. By definition, we have $e_x(-r) \leq f(x) - h(x) \leq e_x(r)$ for every $x \in \mathbb{X}$. When we apply the strictly increasing function e_x^* to this chain of inequalities, we get

$$-r \;\leq\; e_x^*(f(x) - h(x)) \;=\; E^*(x, f(x) - h(x)) \;=\; \Gamma_f(f)(x) - \Gamma_f(h)(x) \;\leq\; r,$$

which implies that $\Gamma_f(h)$ is a uniform r-perturbation of $\Gamma_f(f)$. Reversing the steps proves the claim. ◻

The two Transformation Lemmas together imply that the non-uniform well diagram of f and E is identical to the uniform well diagram of Γ_f. In other words, the robustness of each level set to non-uniform perturbation can be read directly off $\mathrm{Ngm}(f,E) = \mathrm{Dgm}(\Gamma_f(f))$ in the manner discussed in Section 2.

5 Linear Error

In this section, we demonstrate that the imposition of linearity on our error model E leads to a richer theory of persistence. More specifically, we assume that E is linear and use this fact to define a metric on the space of all functions. Then we show that, under this metric, non-uniform persistence diagrams are stable.

Metric. Recall that an error model E is *linear* if $E(x, r) = r \cdot E(x, 1)$, for all $x \in \mathbb{X}$ and all $r \in \bar{\mathbb{R}}$. Put another way, E is linear if all of the functions e_x are linear. Whenever we have such a model and two functions f and h, we may rewrite the inequalities in (4) as $|f(x) - h(x)| \leq E(x, r)$. In other words, f is a non-uniform r-perturbation of h iff h is a non-uniform r-perturbation of f. This leads us to define the following notion of *non-uniform distance* between h and f:

$$d_E(f, h) \;=\; \min\{r \mid |f(x) - h(x)| \leq E(x, r), \text{ for all } x \in \mathbb{X}\}.$$

9 (Metric Lemma) *Assuming a linear error model E, d_E is a metric on the vector space of all $\bar{\mathbb{R}}$-valued functions on $\mathbb{X}$.*

Proof. Since $E(x, 0) = 0$, we have $f = h$ iff $d_E(f, h) = 0$. Furthermore, $d_E(f, h) = d_E(h, f)$, as discussed above. It remains to prove the triangle inequality: $d_E(f, h) \leq d_E(f, g) + d_E(g, h)$. Put $R = d_E(f, g)$ and $S = d_E(g, h)$ and fix a point $x \in \mathbb{X}$. Then using the triangle inequality for the standard metric in $\bar{\mathbb{R}}$, we find

$$\begin{aligned} |f(x) - h(x)| \;&\leq\; |f(x) - g(x)| + |g(x) - h(x)| \\ &\leq\; E(x, R) + E(x, S) \\ &\leq\; E(x, R + S). \end{aligned}$$

This implies $d_E(f, h) \leq R + S$ and the claim follows. ◻

We note that the Metric Lemma still holds under a weaker assumption on E. Namely, we need only assume that for each $x \in \mathbb{X}$, $e_x(-r) = -e_x(r)$ and $e_x(r) + e_x(s) \leq e_x(r + s)$, for all $r, s \geq 0$. In words, each e_x is odd and also convex on the non-negative half of the extended real line.

Stability. We now compare the non-uniform persistence diagrams of two functions f and h on $\mathbb{X}$, and prove that their bottleneck distance is bounded from above by the non-uniform distance between the two functions. First we need to show that for a linear error model, the transformations defined by different functions are the same.

10 (Linear Transformation Lemma) *Let $f, h : \mathbb{X} \to \bar{\mathbb{R}}$ be two functions. Assuming a linear error model E, we have $\Gamma_f(h) = \Gamma_h(h)$.*

Proof. Fix a point $x \in \mathbb{X}$. Starting with the definition, we get

$$\begin{aligned} \Gamma_f(h)(x) \;&=\; \Gamma_f(f)(x) - E^*(x, f(x) - h(x)) \\ &=\; E^*(x, f(x)) - E^*(x, f(x) - h(x)) \\ &=\; E^*(x, h(x)), \end{aligned}$$

where we use linearity to get from the second to the third line. But the third line is equal to $\Gamma_h(h)(x)$, and the claim follows. ◻

The Linear Transformation Lemma justifies the notation $\Gamma(h)$ to refer to the common function $\Gamma_f(h)$ for any other function f. As an immediate consequence of the Perturbation Transformation Lemma, we then find $d_E(f,h) = \|\Gamma(f) - \Gamma(h)\|_\infty$. Finally, we get the stability of the non-uniform persistence diagrams.

11 (Non-uniform Stability Theorem) *For any two $f, h : \mathbb{X} \to \bar{\mathbb{R}}$, and a linear error model E, we have:* $W_\infty(\mathrm{Ngm}(f,E), \mathrm{Ngm}(h,E)) \leq d_E(f,h)$.

Proof. First we transform f and h into $\Gamma(f)$ and $\Gamma(h)$, recalling from the Persistence Diagram Lemma that $\mathrm{Ngm}(f,E) = \mathrm{Dgm}(\Gamma(f))$ and $\mathrm{Ngm}(h,E) = \mathrm{Dgm}(\Gamma(h))$. Applying the Stability Theorem to the two uniform persistence diagrams, we find

$$W_\infty(\mathrm{Dgm}(\Gamma(f)), \mathrm{Dgm}(\Gamma(h))) \;\leq\; \|\Gamma(f) - \Gamma(h)\|_\infty \;=\; d_E(f,h),$$

and the result follows. ◻

6 Discussion

The main contribution of this paper is the extension of the machinery of persistent homology to non-uniform error models. This extension is not complete and many questions remain yet unanswered.

On the technical level, it would be interesting to gain a more detailed understanding on how the difference between $\mathrm{Dgm}(f)$ and $\mathrm{Ngm}(f,E)$ relates to the error model, E. Similarly, can we extend the Non-uniform Stability Theorem from linear to non-linear error models? A more challenging question is the extension of robustness under non-uniform error to a full-blown probability theory of the homology of level sets.

Acknowledgments

The authors thank Dmitriy Morozov for insightful discussions on the topic of this paper.

References

1. P. BENDICH, H. EDELSBRUNNER, D. MOROZOV AND A. PATEL. Robustness of level and interlevel sets. Manuscript, IST Austria, Klosterneuburg, Austria, 2010.
2. G. CARLSSON AND V. DE SILVA. Zigzag persistence. *Found. Comput. Math.* (2010).
3. G. CARLSSON, T. ISHKHANOV, V. DE SILVA AND A. ZOMORODIAN. On the local behavior of spaces of local images. *Internat. J. Comput. Vision* **76** (2008), 1–12.
4. A. CERRI, M. FERRI AND D. GIORGI. Retrieval of trademark images by means of size functions. *Graphical Models* **68** (2006), 451–471.
5. M. K. CHUNG, P. BUBENIK AND P. T. KIM. Persistence diagrams of cortical surface data. In *Information Processing in Medical Imaging*, Springer-Verlag, LNCS **5636**, 2009, 386–397.

6. D. Cohen-Steiner, H. Edelsbrunner and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.* **28** (2007), 511-533.
7. D. Cohen-Steiner, H. Edelsbrunner and J. Harer. Extending persistence using Poincaré and Lefschetz duality. *Found. Comput. Math.* **9** (2009), 79-103.
8. V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Alg. Geom. Topology* **7** (2007), 339–358.
9. M.-L. Dequèant, S. Ahnert, H. Edelsbrunner, T. M. A. Fink, E. F. Glynn, G. Hattem, A. Kudlicki, Y. Mileyko, J. Morton, A. R. Mushegian, L. Pachter, M. Rowicka, A. Shiu, B. Sturmfels and O. Pourquié. Comparison of pattern detection methods in microarray time series of the segmentation clock. *PLoS ONE* **3** (2008), e2856, doi:10.1371/journal.pone.0002856.
10. H. Edelsbrunner and J. L. Harer. *Computational Topology. An Introduction.* Amer. Math. Soc., Providence, Rhode Island, 2009.
11. H. Edelsbrunner, D. Morozov and A. Patel. Quantifying transversality by measuring the robustness of intersections. Manuscript, Dept. Comput. Sci., Duke Univ., Durham, North Carolina, 2009.
12. A. Hatcher. *Algebraic Topology.* Cambridge Univ. Press, Cambridge, England, 2002.
13. J. R. Munkres. *Elements of Algebraic Topology.* Perseus, Cambridge, Massachusetts, 1984.
14. W. J. Thron. *Topological Structures.* Holt, Rinehart and Winston, New York, 1966.