Local and Global Comparison of Continuous Functions

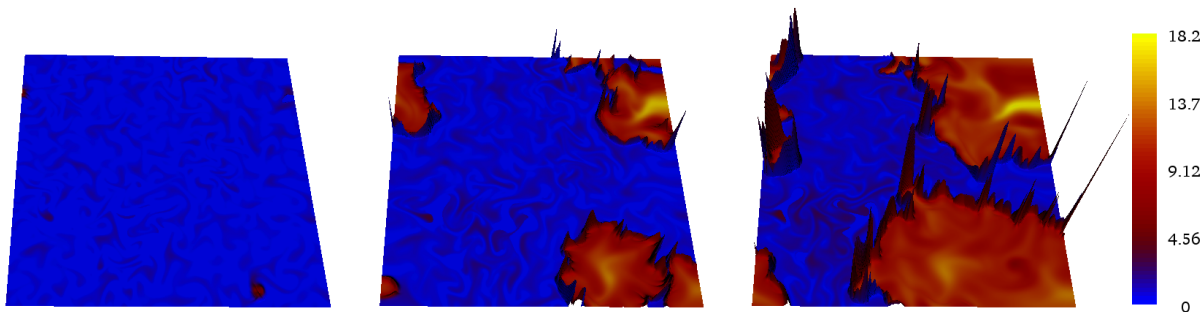Herbert Edelsbrunner[*]   John Harer[†]   Vijay Natarajan[‡]   Valerio Pascucci[§]

Figure 1: We consider the simulation of a combustion process, measure two physical quantities, and compare them. The comparison measure, κ, is mapped to the height of a terrain and the first quantity is mapped to color. From left to right: ignition phase, burning phase, and end of combustion. The front of the flames are tracked by a strip of high κ values represented by the peaks that enclose the burnt region.

ABSTRACT

We introduce local and global comparison measures for a collection of $k \leq d$ real-valued smooth functions on a common d-dimensional Riemannian manifold. For $k = d = 2$ we relate the measures to the set of critical points of one function restricted to the level sets of the other. The definition of the measures extends to piecewise linear functions for which they are easy to compute. The computation of the measures forms the centerpiece of a software tool which we use to study scientific datasets.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques.

Keywords: Visualization, Riemannian manifolds, smooth functions, time-varying data, comparison measure, differential forms.

1 INTRODUCTION

The topic of this paper is a new comparison measure for functions defined on a common Riemannian manifold. After explaining why this is an interesting topic, we state our results and compare them to prior work on the subject.

Motivation. A *scientific dataset* aimed at studying a physical phenomenon typically consists of a large number of measurements taken within a domain of interest. In recent years, such datasets have reached sizes that not only warrant but demand we think of them as continuous functions. Rather than one we usually have a number of functions, such as for example temperature, volume, pressure, and others that can be used to make predictions about the flow behavior of a fluid. Some of these functions may be redundant, which motivates the development of comparative measures that may be used to construct bases of functions sufficient to study the phenomenon.

Another application of such a measure is the study of time-varying functions. Interesting events in the evolution may be characterized by sudden changes in some of the functions or perhaps by the sudden change in the relationship between two or more of the functions. It seems natural to plot some summary measurement for individual functions but also for pairs and larger groups to search for such events. In computational simulations, knowledge about such events can be used to adjust the time-step and/or mesh size used to resolve the phenomenon.

Yet another application of comparison measures for functions are simulations of the same phenomenon performed by different pieces of software or by the same software over periods of time during which the hardware and/or software undergoes possibly non-trivial changes. The comparison measure can be used to decide whether or not these evolutionary changes have any influence on the results of the simulation.

Results. The main new concept in this paper is a comparison measure defined for $k \leq d$ functions defined on a common d-dimensional Riemannian manifold $\mathbb{M}$. Its global version is the integral of the absolute value of the wedge product of the k derivatives divided by the d-dimensional volume of the manifold. Formally,

$$\kappa_{\mathbb{M}}(F) \;\; = \;\; \int_{x \in \mathbb{M}} \|\mathrm{d}f_1 \wedge \mathrm{d}f_2 \wedge \ldots \wedge \mathrm{d}f_k\| \;/\; \mathrm{vol}(\mathbb{M}),$$

where $F = (f_1, f_2, \ldots, f_k)$ is the vector of functions. There is also a local version obtained by restricting the computation to a domain $D \subseteq \mathbb{M}$ and letting D shrink toward a point $x \in \mathbb{M}$, yielding $\kappa_x(F)$ in the limit. This gives a function $\kappa : \mathbb{M} \to \mathbb{R}$ defined by $\kappa(x) = \kappa_x(F)$, and we get the global measure as the average of the local measure. While both the local and global comparison measures are new, we believe that the local measure is a more significant development as it provides significantly more detailed information about how the functions relate to each other. For the particular case of two functions on a common 2-dimensional Riemannian manifold we

---

[*]Departments of Computer Science and Mathematics, Duke University, Durham, North Carolina, and Raindrop Geomagic, Research Triangle Park, North Carolina. e-mail: edels@cs.duke.edu

[†]Department of Mathematics, Duke University, Durham, North Carolina. e-mail: john.harer@duke.edu

[‡]Department of Computer Science, Duke University, Durham, North Carolina. e-mail: vijay@cs.duke.edu

[§]Lawrence Livermore National Laboratory, Livermore, California. e-mail: pascucci1@llnl.gov

have formulas that express a relationship between $\kappa(F) = \kappa_{\mathbb{M}}(F)$ and the Jacobi set recently introduced in [5].

We have developed a visualization tool for comparing scientific datasets, including options geared towards time-varying functions. We illustrate the local and global versions of our comparison measure by applying this tool to a suite of synthetic and scientific datasets.

Prior work. We compare our measure to two concepts: the correlation coefficient from statistics and the Earth mover's distance from computer vision. The *correlation coefficient* of two sets of values x_i and y_i is formally defined as

$$\rho \quad = \quad \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{[\sum_{i=1}^{n}(x_i - \bar{x})^2] \cdot [\sum_{i=1}^{n}(y_i - \bar{y})^2]}},$$

where $\bar{x} = \frac{1}{n}\sum x_i$ and $\bar{y} = \frac{1}{n}\sum y_i$ are the mean values of the two sets [9, Chapter 8]. If applied to functions, we would sample the domain at a set of points and use the function values at these points to compute ρ. The correlation coefficient is commonly used to check whether the two sets are independent ($\rho = 0$) or whether one is an affine function of the other ($\rho = \pm 1$). It is fairly straightforward to extend the definition to three or more sets of values or functions. One major difference between our comparison measure and the correlation coefficient is that the former is sensitive to connectivity information among the sample points (vertices of the triangulation) and the latter is not. This is the reason why ρ is exclusively used as a global measure without meaningful local analog. The correlation coefficient can determine whether the sets or functions covary positively or negatively, which is not possible with our comparison measure. As a consequence, we may get $\rho = 0$ even if there is strong co-variation, namely strongly positive in one half and strongly negative in the other half of the domain. In contrast, $\kappa(F)$ is more sensitive to the local alignment of features of the functions, in particular the critical points and the gradient flow between them. It is oblivious to the orientation of the flow and measures the similarity between two functions f and g the same as between $-f$ and g.

To describe the Earth mover's distance we consider two smooth functions and the corresponding sets of critical points. Each critical point has a weight, which may be its function value. We think of one set of critical points as a supply of dirt and the other as a collection of holes. The goal is to determine the least amount of work necessary to move the dirt around to fill the holes. Formally, the *Earth mover's distance* is the weight of the minimum weight matching between the two sets of points. Other than the weights of the points, we also need a notion of edge length to compute the cost for the transportation of the dirt, and this notion depends on the application at hand. Originally introduced for comparing color patterns by Rubner, Tomasi and Guibas [12], this measure has since been adapted to solve a variety of problems, including contour matching, object tracking, polyhedral shape matching, and vector function comparison. Similar to the correlation coefficient, the Earth mover's distance is a global measure without any meaningful local analog. Furthermore, nearness does not require similar features within the same region of the domain. It therefore does not capture local similarity the way our comparison measure does.

Outline. Section 2 introduces the comparison measure. Section 3 offers interpretations based on the Jacobi set of the functions. Section 4 illustrates the measure using results from computational experiments performed on synthetic and scientific datasets. Section 5 mentions open questions and concludes the paper.

2 THE MEASURE

In this section, we introduce the comparison measure for smooth functions on a common manifold. After defining the local and global versions of the measure, we study some of its properties and explain how to compute it for piecewise linear functions. A central concept is the *k*-form of a collection of *k* smooth functions and we refer to Darling [3] and Weintraub [13] for background on the topic.

The definition. Let $\mathbb{M}$ be a smooth compact manifold with a Riemannian metric. When $\mathbb{M} \subseteq \mathbb{R}^d$, the manifold inherits a natural metric by identifying the tangent space of $\mathbb{M}$ at a point x with a subspace of $\mathbb{R}^d$ endowed with the Euclidean metric. Let $F = (f_1, f_2, \ldots, f_k) : \mathbb{M} \to \mathbb{R}$ be a vector of $k \leq d$ smooth functions. For each f_i, the *(differential) 1-form* of f_i is written

$$\mathrm{d}f_i \quad = \quad \frac{\partial f_i}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f_i}{\partial x_2}\mathrm{d}x_2 + \cdots + \frac{\partial f_i}{\partial x_d}\mathrm{d}x_d,$$

where x_1 to x_d is a local coordinate system. Using the wedge product, we combine the 1-forms of the f_i to make the *k*-form $\mathrm{d}f_1 \wedge \mathrm{d}f_2 \wedge \ldots \wedge \mathrm{d}f_k$. We may think of it as a vector in a $\binom{d}{k}$-dimensional vector space. For a domain $D \subseteq \mathbb{M}$ we define the comparison measure over D as the normalized integral of the value of the *k*-form,

$$\kappa_D(F) \quad = \quad \int_{x \in D} \|\mathrm{d}f_1 \wedge \mathrm{d}f_2 \wedge \ldots \wedge \mathrm{d}f_k\| \, / \, \mathrm{vol}(D),$$

where the metric on $\mathbb{M}$ is used to define the norm of the wedge product as well as the volume. We obtain the *global comparison measure* as $\kappa(F) = \kappa_{\mathbb{M}}(F)$ and note that $0 \leq \kappa(F) < \infty$. We may also shrink D toward a point $x \in \mathbb{M}$ and obtain the value $\kappa_x(F)$ in the limit. This furnishes the *local comparison measure*, which is the function $\kappa : \mathbb{M} \to \mathbb{R}$ defined by $\kappa(x) = \kappa_x(F)$. Note that the global measure is the average local measure:

$$\kappa(F) \quad = \quad \int_{x \in \mathbb{M}} \kappa(x) \, \mathrm{d}x \, / \, \mathrm{vol}(\mathbb{M}).$$

Using this relationship we can deduce properties of the global from properties of the local measure.

Evaluation. A *k*-form can be written in canonical form as an element of the $\binom{d}{k}$-dimensional vector space spanned by the *k*-forms $\mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \ldots \wedge \mathrm{d}x_{i_k}$ with $1 \leq i_1 < i_2 < \ldots < i_k \leq d$. Specifically,

$$\mathrm{d}f_1 \wedge \ldots \wedge \mathrm{d}f_k \quad = \quad \sum_{1 \leq i_1 < \cdots < i_k \leq d} \omega_{i_1,\ldots,i_k} \cdot (\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_k}),$$

where x_1 to x_d form a local coordinate chart,

$$\omega_{i_1,\ldots,i_k} \quad = \quad \det \begin{bmatrix} \frac{\partial f_1}{\partial x_{i_1}} & \frac{\partial f_2}{\partial x_{i_1}} & \cdots & \frac{\partial f_k}{\partial x_{i_1}} \\ \frac{\partial f_1}{\partial x_{i_2}} & \frac{\partial f_2}{\partial x_{i_2}} & \cdots & \frac{\partial f_k}{\partial x_{i_2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_{i_k}} & \frac{\partial f_2}{\partial x_{i_k}} & \cdots & \frac{\partial f_k}{\partial x_{i_k}} \end{bmatrix}.$$

If we choose the coordinates $x_1, x_2, \ldots, x_d$ to be orthonormal at x (but we may not be able to do this uniformly throughout D) then the local measure at x is

$$\kappa(x) \quad = \quad \sqrt{\sum_{1 \leq i_1 < \cdots < i_k \leq d} \omega_{i_1,\ldots,i_k}^2}.$$

For example, if we have $k = 2$ functions on a 2-manifold embedded in $\mathbb{R}^3$ with the standard Euclidean metric, then $\kappa(x)$ is the length

of the cross-product of the two gradients at x. More generally, $\kappa(x)$ is the k-dimensional volume of the parallelepiped spanned by the k gradients [8, Chapter V].

Properties. The comparison measure enjoys a number of useful algebraic properties, but the triangle inequality is not one of them. To see this consider $k = 2$ and let g be a constant function on $\mathbb{M}$. Then $\kappa(f_1, g) = \kappa(g, f_2) = 0$ but $\kappa(f_1, f_2)$ may take any arbitrary positive value. The first three properties follow immediately from the definition:

Symmetry: $\kappa(\ldots, f_i, \ldots, f_j, \ldots) = \kappa(\ldots, f_j, \ldots, f_i, \ldots)$ for every pair of indices $i \neq j$.

Degeneracy: $\kappa(F) = 0$ if there is a subset of linearly dependent 1-forms.

Scaling: $\kappa(\alpha f_1 + \beta, f_2, \ldots, f_k) = |\alpha| \cdot \kappa(f_1, f_2, \ldots, f_k)$, with $\alpha, \beta \in \mathbb{R}$.

The next property follows from simple vector algebra. Call the $(k - 1)$-dimensional parallelepiped spanned by the gradients of f_2 to f_k the *base* of the k-dimensional parallelepiped spanned by the same gradients together with the gradient of $f_1 + g_1$. The property says that the volume of this parallelepiped is at most the volume of the two parallelepipeds obtained by expanding the same base with the gradients of f_1 and of g_1.

Sub-additivity: $\kappa(f_1 + g_1, f_2, \ldots, f_k) \leq \kappa(f_1, f_2, \ldots, f_k) + \kappa(g_1, f_2, \ldots, f_k)$.

The last property is easier to state for the local measure. It says that the k-dimensional volume of the parallelepiped spanned by k vectors is at most the i-dimensional volume of the parallelepiped spanned by $1 \leq i < k$ of the vectors times the $(k - i)$-dimensional volume of the parallelepiped spanned by the remaining $k - i$ vectors. We have equality iff the i and $k - i$ vectors span orthogonal linear subspaces of dimension i and $k - i$.

Sub-multiplicativity: $\kappa_x(f_1, \ldots, f_i, f_{i+1}, \ldots, f_k) \leq \kappa_x(f_1, \ldots, f_i) \cdot \kappa_x(f_{i+1}, \ldots, f_k)$, for every point $x \in \mathbb{M}$ and each index i.

The corresponding inequality of the global measure is

$$\kappa(f_1, \ldots, f_k) \quad \leq \quad \kappa(f_1, \ldots, f_i) \cdot \kappa(f_{i+1}, \ldots, f_k) \cdot \mathrm{vol}(\mathbb{M}).$$

Computation. We are interested in computing the local and global comparison measures for functions obtained by piecewise linear interpolation from measurements at a finite set of points. We let these points be the vertices of a triangulation K of the manifold $\mathbb{M}$. All k functions are linear on a d-simplex of K and their gradients are therefore constant. We may evaluate $\kappa(x)$ at a point x in the interior of the d-simplex as described above. This gives the local version of the measure. To get the global measure, we compute

$$\kappa(F) \quad = \quad \sum_{\sigma} \kappa(x) \cdot \mathrm{vol}(\sigma) \, / \, \mathrm{vol}(K),$$

where σ ranges over all d-simplices in K and x is a point in the interior of σ.

3 JACOBI SET INTERPRETATION

In this section, we develop an alternate interpretation of the comparison measure in terms of critical points. We have a result only for the case of two functions on a 2-manifold. We begin by introducing Morse functions [10, 11] and Jacobi sets [5].

Morse functions and Jacobi sets. Let $\mathbb{M}$ be a smooth 2-manifold embedded in $\mathbb{R}^3$ with its inherited Riemannian metric and let $f : \mathbb{M} \to \mathbb{R}$ be a smooth function. Given local coordinates in a neighborhood of a point $x \in \mathbb{M}$, we define the *gradient*, $\nabla f(x)$, as the vector of partial derivatives. The point x is *critical* if the gradient is the zero-vector. Non-critical points are *regular*. The *Hessian* of the point x is the matrix of second-order partial derivatives. The critical point x is *non-degenerate* if the Hessian at x is invertible. Finally, f is a *Morse function* if

   I. every critical point is non-degenerate;

   II. every two critical points have different function values.

Intuitively, Morse functions are generic smooth functions. Their critical points are isolated and if $\mathbb{M}$ is compact then there are only finitely many such points.

Suppose now $f, g : \mathbb{M} \to \mathbb{R}$ are two Morse functions. We are interested in how the two functions interact and define the *Jacobi set* as the set of points at which the Jacobian matrix of partial derivatives does not have full rank,

$$\mathbb{J}(f, g) \quad = \quad \{x \in \mathbb{M} \mid \mathrm{rank}\, J(x) < 2\}.$$

This is the set of points where the gradients of f and g are linearly dependent. Equivalently, it is the set of points x with $\kappa_x(f, g) = 0$. Yet another characterization is obtained by taking level sets of one function and collecting the critical points of the other function restricted to these level sets; see Figure 2. Generically, the Jacobi set of two functions is a 1-manifold smoothly embedded in $\mathbb{M}$.
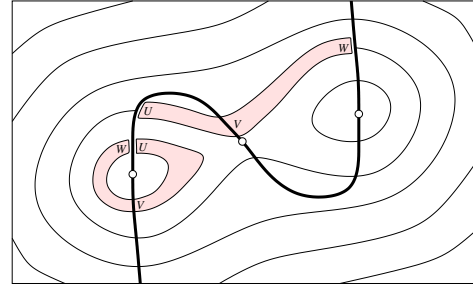

Figure 2: The function f increases from bottom to top and has horizontal level lines that are not shown. The function g has two maxima and one saddle and is indicated by its solid level lines. The boldface Jacobi set is an S-shaped curve that passes through the three critical points. The two shaded strips illustrate the argument used to reinterpret κ in terms of critical points of f restricted to level sets of g.

Reformulation of comparison measure. For two functions f and g on a 2-manifold embedded in $\mathbb{R}^3$, the limit of the comparison measure at a point $x \in \mathbb{M}$ is the length of the cross-product of the two gradients, $\kappa_x(f, g) = \|\nabla f(x) \times \nabla g(x)\|$. Writing $\mathrm{area}(\mathbb{M})$ for the area of the manifold we therefore have

$$\kappa(f, g) \quad = \quad \int_{x \in \mathbb{M}} \|\nabla f(x) \times \nabla g(x)\| \, \mathrm{d}x \, / \, \mathrm{area}(\mathbb{M}).$$

The length of the cross-product remains the same if we replace $\nabla f(x)$ be its orthogonal projection to the plane normal to $\nabla g(x)$. This projection is the gradient of f restricted to the level set $g^{-1}(g(x))$ that passes through x. Writing $t = g(x)$ and f_t for the restriction of f to $g^{-1}(t)$, the length of the cross-product further simplifies to the product of lengths of gradients, $\|\nabla f_t(x)\| \cdot \|\nabla g(x)\|$.

We now rewrite $\kappa(f, g)$ as an integral over the Jacobi set of f and g. Recall that a point $v \in \mathbb{J} = \mathbb{J}(f, g)$ is a critical point of f restricted to $g^{-1}(g(x))$. Let u and w be the preceding and succeeding

critical points along the same level set. Let V be a short interval along $\mathbb{J}$ that contains v and let U and W be the corresponding intervals containing u and w, as illustrated in Figure 2. Integrating the comparison measure over the strip D of paths connecting points in U with points in W we get

$$\kappa_D(f,g) = \int_{x \in D} \|\nabla f(x) \times \nabla g(x)\| \, dx \, / \, \text{area}(D)$$
$$= \int_{v \in V} \left(\int_{x \in uw} \|\nabla f_t(x)\| dx \right) dg \, / \, \text{area}(D),$$

where uw is the path on $g^{-1}(g(x))$ that connects u to w and passes through v and dg is the derivative of g restricted to V times dv. The crucial step from the first to the second line in the above expression for $\kappa_D(f,g)$ is the observation that $\int \|\nabla g(y)\| dy$ along an integral line of g crossing the strip D is the same everywhere, namely equal to the difference in g-value along the two sides of the strip. It is also equal to $\int_{v \in V} dg$. The restriction of f is monotonic from u to v and from v to w. The inner integral is therefore $|f(v) - f(u)| + |f(v) - f(w)|$, giving $\kappa_D(f,g) = \int_{v \in V} |2f(v) - f(u) - f(w)| dg \, / \, \text{area}(D)$. Integrating over all strips of this form is the same as integrating over $\mathbb{M}$ twice. Hence

$$\kappa(f,g) = \int_{v \in \mathbb{J}} |2f(v) - f(u) - f(w)| \, dg \, / \, 2\text{area}(\mathbb{M}). \quad (1)$$

According to Equation (1), the comparison measure can be interpreted as an average difference between functions values at neighboring critical points of f restricted to level sets of g. A point $x \in \mathbb{J}$ that is a maximum of f restricted to $g^{-1}(g(x))$ contributes four times its function value to the above integral: $2f(x)$ when $x = v$, $f(x)$ when $x = u$ and $f(x)$ when $x = w$. Similarly, x contributes minus four times its function value if it is a minimum of f restricted to $g^{-1}(g(x))$. Hence

$$\kappa(f,g) = 2 \int_{v \in \mathbb{J}} \text{sgn}(v) f(v) \, dg \, / \, \text{area}(\mathbb{M}), \quad (2)$$

where $\text{sgn}(v) = 1$ if v is a maximum and -1 if v is a minimum. In case the Jacobi curve is given, Equation (2) gives a fast algorithm for computing the global comparison measure between f and g.

Algorithm using persistence. We may think of Equations (1) and (2) as alternative algorithms that compute $\kappa(f,g)$ by integrating contributions of the points in the Jacobi set. We further develop the integral in Equation (1) to find a formulation in which the contributions express more directly the role of the points in the restriction of f to the level sets of g.

Each generic level set $g^{-1}(t)$ is a collection of topological circles and f restricted to this level set has equally many minima and maxima. Letting k be this common number, we form a pairing $\{(u_i, v_i) \mid 1 \le i \le k\}$ between the minima u_i and the maxima v_i such that $f(v_i) - f(u_i) > 0$ for all i. Writing $\text{pers}(u_i) = \text{pers}(v_i) = f(v_i) - f(u_i)$, we get

$$\kappa(f,g) = \int_{v \in \mathbb{J}} \text{pers}(v) \, dg \, / \, \text{area}(\mathbb{M}). \quad (3)$$

Indeed, Equation (1) is a special case in which the integration is done over two pairings of the points in the Jacobi set. A more meaningful (single) pairing is obtained using the concept of extended persistent homology [1, 7]. It is easy to explain for a Morse function $f_t : \mathbb{S}^1 \to \mathbb{R}$, which has equally many minima and maxima. Sweeping the circle in the direction of increasing function value, we get a new component whenever we pass a minimum and we merge two components whenever we pass a maximum, except that we complete the circle when we pass the last maximum. Each component is represented by its oldest minimum (the one with smallest function value).

Rule 1. If a maximum merges two components we pair it with the younger of the two minima representing the two components. The older minimum stays on to represent the merged component.

Rule 2. The last maximum is paired with the first minimum.

The persistence is a notion of importance of a critical point that has found use in a number of applications involving smooth functions, see for example [1, 2, 6]. It is fairly straightforward to implement Equation (3) for piecewise linear functions f and g on a triangulation K of $\mathbb{M}$. First, we use the algorithm of [5] to identify the Jacobi set as a collection of edges that form closed cycles in K. Second, we sweep K in the direction of increasing value of g, maintaining the level set, $g^{-1}(t)$, as another collection of closed cycles. We then use the persistence algorithm to compute $\text{pers}(v)$ for all critical points v of $g^{-1}(t)$, which are the intersection points between the level set and the Jacobi set. We get a local picture of the computation by coloring each edge of the Jacobi set with the persistence of its points, as in Figure 3.

4 COMPUTATIONAL EXPERIMENTS

We illustrate the new comparison measure with results from several preliminary computational experiments. We begin by comparing a suite of simple synthetic functions and proceed to studying a series of two-dimensional datasets describing a combustion process and several three-dimensional electrostatic potentials describing a protein complex.

Synthetic functions. We use a suite of five simple synthetic functions to get a first feel for the local and global versions of our comparison measure. All five functions are defined on the torus modeled by the square $[-2\pi, 2\pi] \times [-2\pi, 2\pi]$ with opposite sides identified:

$$\text{cup}(x,y) = x^2 + y^2; \quad \sin(x,y) = \sin x + \sin y;$$
$$\text{sad}(x,y) = x^2 - y^2; \quad \cos(x,y) = \cos x + \cos y;$$
$$\text{abs}(x,y) = |x|.$$

We triangulate the torus, evaluate each function at the vertices, and approximate each function by piecewise linear interpolation. We have $\kappa(\text{cup}) = \kappa(\text{sad}) = 9.61$, $\kappa(\sin) = \kappa(\cos) = 0.96$, and $\kappa(\text{abs}) = 1.0$. Table 1 shows the global comparison measure, $\kappa(f,g)$, for every pair of the five functions. We note that κ is not scale invariant and can therefore not be used to decide whether a given pair of functions is more or less similar to another pair. We

		cup	sad	sin	cos	abs
cup	κ	0.00	78.96	5.76	5.24	6.28
	1%	0.00	0.28	0.03	0.07	0.07
	5%	0.00	8.31	0.98	1.11	2.36
sad	κ	78.96	0.00	5.76	6.28	6.28
	1%	0.28	0.00	0.01	0.02	0.07
	5%	8.31	0.00	0.34	0.35	1.85
sin	κ	5.76	5.76	0.00	0.63	0.64
	1%	0.03	0.01	0.00	0.00	0.00
	5%	0.98	0.34	0.00	0.02	0.11
cos	κ	5.24	6.28	0.63	0.00	0.63
	1%	0.07	0.02	0.00	0.00	0.01
	5%	1.11	0.35	0.02	0.00	0.01
abs	κ	6.28	6.28	0.64	0.63	0.00
	1%	0.07	0.07	0.00	0.01	0.00
	5%	2.36	1.85	0.11	0.12	0.00

Table 1: Compare κ (top rows) with the error introduced when 1% (middle rows) and 5% random noise (bottom rows) is added to the synthetic functions.

use the same five synthetic functions to get a feeling for how sensitive the comparison measure is to small changes in function values. Specifically, we select a percentage R and for each vertex u we pick a random number $r \in [-R, R]$ and change the function value at u by adding $\frac{r}{100}$ times the range of the function. The results of this experiment are shown in Table 1.

To visualize the local comparison measure, we have developed a tool for functions on a 2-dimensional domain. It displays individual functions as terrains, with color, or through level curves. The first two options are combined to get the pictures in Figure 1. The last two options are used to display the `sin` and `cos` functions as well as their comparison in Figure 3. We note that κ is small in
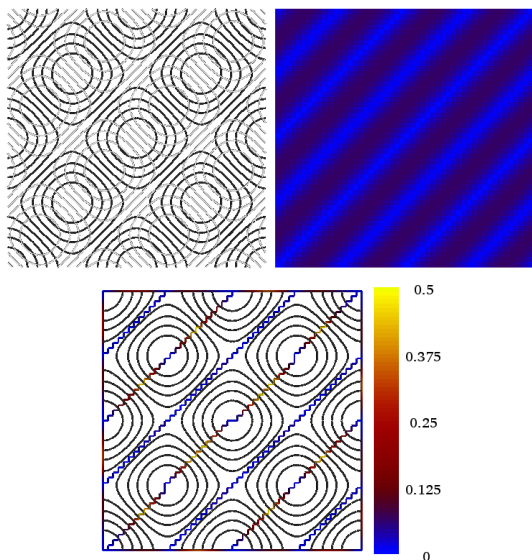
Figure 3: Upper left: level sets extracted from the `sin` (gray) and the `cos` (black) datasets. Upper right: visualization of $\kappa : [-2\pi, \pi]^2 \to \mathbb{R}$ using color. Lower left: the Jacobi set of `sin` and `cos` with color indicating the contribution of individual points to the global comparison measure. Lower right: color key used in the upper right and lower left pictures.

regions where the level curves of `sin` and `cos` are nearly parallel and it is large where the level curves are nearly orthogonal. The two functions are periodic as is obvious from the visualization.

Time series data. We illustrate the use of our tool in visualizing time-varying data obtained by simulating the process inside a combustion engine. The simulation keeps track of various quantities in an attempt to understand the influence of turbulence on ignition, flame propagation, and burnout [4]. Upon compression, the inhomogeneity in the air-fuel mixture causes ignition at multiple spots. Depending on the air-fuel ratio, the flame propagates from these spots outwards or it burns out. We look at two quantities,

- `progress`, a measure of completion of the combustion;
- H_2 or hydrogen, the fuel in the process.

These quantities are known at the vertices of a 600-by-600 grid and at 67 time steps. As usual, we represent the domain by a triangulation and the functions by piecewise linear maps obtained by interpolating the values at the vertices. To get a rough picture of how `progress` and H_2 relate during the process, we compute the global comparison measure, $\kappa(\text{progress}, H_2)$, at each time step and display the result as a function over time; see Figure 4. Initially, both `progress` and H_2 are turbulent but have low concentrations throughout the domain resulting in values of κ close to zero. Something interesting happens around time step 28 after which κ takes on
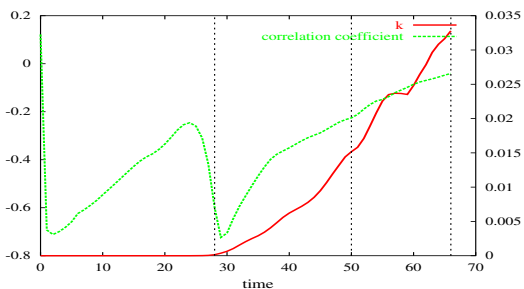

Figure 4: The comparison measure κ (scale on right vertical axis) and the correlation coefficient (scale on left vertical axis), both as functions of time as `progress` and H_2 from the combustion dataset change. The vertical markers indicate the time steps at which the pictures in Figure 1 are extracted. The plot of the correlation coefficient shows that `progress` and H_2 are negatively correlated, indicating the obvious, namely that progress depletes fuel.

progressively larger values. The function is not informative enough to detect the phases of the combustion process, but we get additional insights by visualizing the two quantities using the local comparison measure at the time steps indicated by the three vertical dotted bars which roughly correspond to ignition, burning, and end of combustion; see Figure 1. Note that the flame front is tracked by a strip of large contributions to κ. In passing through this strip both functions undergo sharp changes although at vastly different rates.

Electrostatic potentials. We consider a problem in structural biology to study our local and global comparison measures for functions defined on a three-dimensional domain. A protein complex consists of two or more proteins docked in a stable conformation. For example, the barnase-barstar complex (1BRS) consists of two

	N	S	1BRS	N,S	N,1BRS	S,1BRS	N,S,1BRS
κ	4.01	3.22	7.22	2.30	6.83	5.17	18.66

Table 2: The global measure κ computed for all combinations of the three electrostatic datasets.

proteins. The electrostatic potential defined by barnase (N) and barstar (S) individually in their docked conformation and the potential defined by the complex are available to us as functions sampled over the space. We triangulate the space and linearly interpolate to
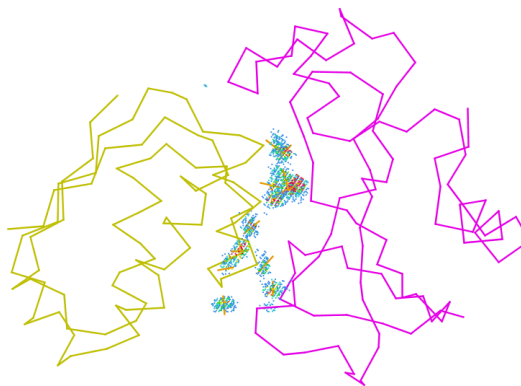

Figure 5: The regions with high value of $\kappa(f_N, f_S)$ are shaded between the two proteins, which are shown as alpha-carbon traces, with barnase in magenta and barstar in yellow.

obtain three piecewise-linear functions, f_N, f_S and f_{1BRS}. Table 2

lists the values of our global comparison measure for the individual functions, the three pairs, and the triplet. Initial observations show that regions where our local comparison measure between f_N and f_S is high correspond to salt bridges or strong hydrogen bonds. Figures 5 and 6 visualize the local comparison measure of the electrostatic potentials f_N and f_S defined by barnase and barstar in the complex 1BRS. The colored dots indicate higher values of κ, namely those in the range $[0.002, 0.0207]$ and are mapped from blue to red. Dots with values lower than 0.002 are not displayed. The gold lines indicate the hydrogen bonds corresponding to the regions with high κ value.
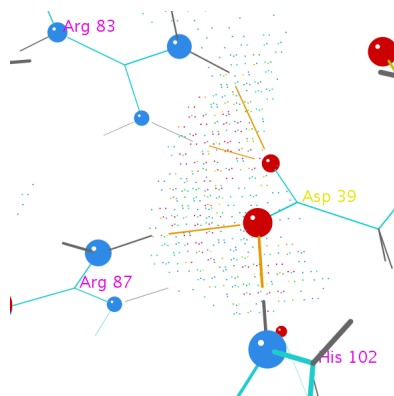
Figure 6: A close-up view of a region with high value of $\kappa(f_N, f_S)$ and the corresponding hydrogen bond cluster. `Asp 39` of barstar forms hydrogen bonds with `Arg 87`, `Arg 83`, and `His 102` of barnase. All four residues are important in the interaction between barnase and barstar.

5 DISCUSSION

Various questions related to the extension of our comparison measures remain open. We mention three:

- Our definition restricts the number of functions to at most the dimension of the manifold. It would be interesting to extend it to the case $k > d$.
- For the particular case $k = d = 2$, we give alternative interpretations of κ using Jacobi sets. These interpretations generalize to the case $k = d > 2$, but what about $k < d$?
- What is the sensitivity of our measure to the triangulation of the manifold? A detailed understanding of this question is useful in situations where functions are given on different triangulations of the same manifold.

REFERENCES

[1] P. K. AGARWAL, H. EDELSBRUNNER, J. HARER AND Y. WANG. Extreme elevation on a 2-manifold. *In* "Proc. 20th Ann. Sympos. Comput. Geom., 2004", 357–365.

[2] G. CARLSSON, A. ZOMORODIAN, A. COLLINS AND L. GUIBAS. Persistence barcodes for shapes. *In* "Proc. Eurographics Sympos. Geom. Process., 2004".

[3] R. W. R. DARLING. *Differential Forms and Connections.* Cambridge Univ. Press, New York, 1994.

[4] T. ECHEKKI AND J. H. CHEN. Direct numerical simulation of auto-ignition in inhomogeneous hydrogen-air mixtures. *In* "Proc. 2nd Joint Meeting U.S. Sections Combustion Institute, 2001".

[5] H. EDELSBRUNNER AND J. HARER. Jacobi sets of multiple Morse functions. *Foundations of Computational Mathematics, Minneapolis 2002*, 37–57, eds. F. Cucker, R. DeVore, P. Olver, E. Süli, Cambridge Univ. Press, England, 2004.

[6] H. EDELSBRUNNER, J. HARER AND A. ZOMORODIAN. Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds. *Discrete Comput. Geom.* **30** (2003), 87–107.

[7] H. EDELSBRUNNER, D. LETSCHER AND A. ZOMORODIAN. Topological persistence and simplification. *Discrete Comput. Geom.* **28** (2002), 511–533.

[8] C. H. EDWARDS, JR. *Advanced Calculus of Several Variables.* Dover, New York, 1973.

[9] W. FELLER. *An Introduction to Probability Theory and Its Applications. Volume I.* Third edition, John Wiley & Sons, New York, 1968.

[10] Y. MATSUMOTO. *An Introduction to Morse Theory.* Translated from Japanese by K. Hudson and M. Saito, Amer. Math. Soc., 2002.

[11] J. MILNOR. *Morse Theory.* Princeton Univ. Press, New Jersey, 1963.

[12] Y. RUBNER, C. TOMASI AND L. J. GUIBAS. The earth mover's distance as a metric for image retrieval. *Internat. J. Comput. Vision* **40** (2000), 99–121.

[13] S. H. WEINTRAUB. *Differential Forms: A Complement to Vector Calculus.* Academic Press, 1996.