

Computing Similarity between Piecewise-Linear Functions*

Pankaj K. Agarwal
Dept. of Computer Science,
Duke University

Boris Aronov
Dept. of Computer Science
and Engineering
Polytechnic Institute of NYU

Marc van Kreveld
Dept. of Information and
Computing Sciences
Utrecht University

Maarten Löffler
Computer Science Dept.
University of California, Irvine

Rodrigo I. Silveira
Dept. Matemàtica Aplicada II
Universitat Politècnica de
Catalunya

ABSTRACT

We study the problem of computing the similarity of two piecewise-linear bivariate functions defined over a common domain, where the surfaces they define in 3D—polyhedral terrains—can be transformed vertically by a linear transformation of the third coordinate (scaling and translation). We present a randomized algorithm that minimizes the maximum vertical distance between the graphs of the two functions, over all linear transformations, in $O(n^{4/3} \text{polylog } n)$ expected time, where n is the total number of vertices in the graphs of the two functions. We also study the computation of similarity of two univariate or bivariate functions by minimizing the area or volume between their graphs. For univariate functions we give a $(1 + \varepsilon)$ -approximation algorithm for minimizing the area that runs in $O(n/\sqrt{\varepsilon})$ time, for any fixed $\varepsilon > 0$. The $(1 + \varepsilon)$ -approximation algorithm for the bivariate version, where volume is minimized, runs in $O(n/\varepsilon^2)$ time, for any fixed $\varepsilon > 0$, provided the two functions are defined over the same triangulation of their domain.

*Work by P.A. is supported by NSF under grants CNS-05-40347, CCF-06-35000, IIS-07-13498, and CCF-09-40671, by ARO grants W911NF-07-1-0376 and W911NF-08-1-0452, by an NIH grant 1P50-GM-08183-01, by a DOE grant OEG-P200A070505, and by a grant from the U.S.–Israel Binational Science Foundation. Work by B.A. is supported by a grant from the U.S.–Israel Binational Science Foundation, by NSA MSP Grant H98230-06-1-0016, and NSF Grant CCF-08-30691. Work by M.v.K., M.L., and R.S. is supported by the Netherlands Organisation for Scientific Research (NWO), and M.L. is further supported by the Office of Naval Research under grant N00014-08-1-1015.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SCG'10, June 13–16, 2010, Snowbird, Utah, USA.

Copyright 2010 ACM 978-1-4503-0016-2/10/06 ...\$5.00.

Categories and Subject Descriptors

F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems—*computations on discrete structures, geometrical problems and computations*

General Terms

Algorithms, Theory

Keywords

piecewise-linear function, polyhedral terrain, similarity, approximation algorithm, randomized algorithm

1. INTRODUCTION

Many types of spatial data can be modeled mathematically as a bivariate function $f: \mathbb{D} \rightarrow \mathbb{R}$, where \mathbb{D} is a (planar) region of interest. Examples include annual precipitation, depth to ground water, soil salinity at the surface, elevation above sea level, and (steepness of) slope of the terrain. Data is usually collected by sampling at a number of points in \mathbb{D} , and the function is extended to the entire \mathbb{D} using spatial interpolation schemes [3, 11, 15]. There is much work in many disciplines, including environmental sciences, geology, and statistics on analyzing such data, computing correlations among them, and testing hypotheses and models.

The simplest type of model that may capture the correlation between two types of spatial data is a linear model. If the two types of data are collected at exactly the same set of locations, i.e., data is *isotopic*, then one can apply standard regression analysis on the pairs of values to determine whether a linear dependence of one data set on the other exists [13, 14]. Notwithstanding its popularity because of its simplicity, this approach has two serious shortcomings. First, the two types of data may have been collected at different locations, i.e., data is *heterotopic* [15], due to a difference in resolution of data acquisition techniques or cost factors, or due to inaccessibility of regions for certain types of measurement. Second, it does not account for a possible difference in sampling density in subregions of the region of interest. This will bias the result toward the linear relation that exists in the more densely sampled subregion. The first problem can be handled by using spatial interpolation on one data set to obtain pairs of values at the locations of the

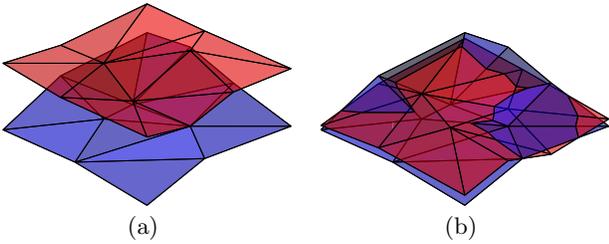


Figure 1. (a) Two triangulated terrains. (b) One terrain has been translated and scaled vertically to match the other one.

other data set. The second problem can be handled by assigning a weight to each sample location, where the weight is lower in densely sampled subregions. Alternatively, we compute piecewise-linear representations of f and g using spatial interpolation and match them to determine a linear dependence between f and g . This is likely to provide better results than using weights at the sampled locations, which in a sense form a piecewise-constant representation of each function. See [9] for a Morse-theory based method for computing the similarity of two functions.

We assume the following representation of a function. Let \mathbb{M} be a triangulation of the domain \mathbb{D} . A function value $f(v)$ is assigned to every vertex of \mathbb{M} . By linearly interpolating the function value within each triangle of \mathbb{M} , we can extend f to the entire \mathbb{D} . Thus \mathbb{M} defines a piecewise-linear function $f : \mathbb{D} \rightarrow \mathbb{R}$. The graph of f is a polyhedral *terrain*, which can be represented as an xy -monotone triangulated surface whose triangulation is induced by \mathbb{M} . Given such representations of two functions f and g , we can *match* them directly, over all possible linear transformations on these functions, to determine a linear correlation of f and g . Note that we apply a linear transformation to only *one* of the functions. Since the functions f and g are scalar-valued, a linear function is determined by two parameters: a *scaling* and a *translation*. Figure 1 shows an example of two terrains representing the functions f and g , and how one can be transformed to match the other one better.

This paper discusses the problem of computing the scaling and translation parameters that provide an optimal matching of two bivariate functions modeled as polyhedral terrains. In particular, let \mathbb{D} be a planar polygonal domain, let \mathbb{M}_f and \mathbb{M}_g be two triangulations of \mathbb{D} , and let $f, g : \mathbb{D} \rightarrow \mathbb{R}$ be two piecewise-linear functions whose linear pieces correspond to \mathbb{M}_f and \mathbb{M}_g , respectively. We assume that \mathbb{M}_f and \mathbb{M}_g together have n vertices. We consider three different measures to determine the matching (similarity, or correlation) between f and g : (i) minimizing the maximum vertical distance (L_∞ -norm) between their graphs, (ii) minimizing the volume between their graphs (L_1 -norm), and (iii) minimizing the square measure (L_2 -norm). More precisely, if s is the scaling parameter and t the translation parameter, we define functions $\mu_p : \mathbb{R}^2 \rightarrow \mathbb{R}$, $p = 1, 2, \infty$, to measure how well f and g match for a pair s and t in (1)–(3).

$$\mu_\infty(s, t) = \max_{(x,y) \in \mathbb{D}} |sf(x, y) + t - g(x, y)| \quad (1)$$

$$\mu_1(s, t) = \int_{(x,y) \in \mathbb{D}} |sf(x, y) + t - g(x, y)| \, dx dy \quad (2)$$

$$\mu_2(s, t) = \int_{(x,y) \in \mathbb{D}} (sf(x, y) + t - g(x, y))^2 \, dx dy \quad (3)$$

For $p = 1, 2, \infty$, we can now define the matching between f and g as $\sigma_p(f, g) = \min_{s,t} \mu_p(s, t)$. For each p , our goal is to compute the pair (s^*, t^*) that minimizes the function μ_p and to compute this minimum value $\sigma_p(f, g) = \mu_p(s^*, t^*)$. We note that μ_2 is a linear regression model, μ_1 is similar but uses the L_1 -norm, and μ_∞ is similar to the Hausdorff distance, a commonly used measure for shape matching [2].

If the triangulations \mathbb{M}_f and \mathbb{M}_g are identical, then we say that f and g are *aligned*, otherwise they are *unaligned*. We study both aligned and unaligned versions of the problem. We can convert unaligned functions to aligned ones by computing a new triangulation \mathbb{M} that is a common refinement of \mathbb{M}_f and \mathbb{M}_g . The number of vertices in \mathbb{M} can vary between $\Theta(n)$ and $\Theta(n^2)$, depending on the complexity of the overlay of \mathbb{M}_f and \mathbb{M}_g . If the number of edge-edge intersections between the two triangulations is k , then the overlay has complexity $O(n + k)$ and can be computed in $O(n + k)$ time [10]. A triangulation of the overlay also has complexity $O(n + k)$. For triangulations that satisfy *realistic input assumptions*, one can often show that $k = O(n)$. Several models where this is the case are defined, for example, in [7]. In this case one can align a pair of unaligned functions without loss of asymptotic efficiency.

In Section 2 we develop algorithms for computing $\sigma_\infty(f, g)$. We can compute the optimal linear transformation in linear time for aligned functions, using linear programming. For unaligned terrains, we can avoid the potential quadratic running time by combining various techniques leading to an $O^*(n^{4/3})$ time algorithm (the O^* -notation omits polylogarithmic factors). In Section 4 we discuss the computation of $\sigma_1(f, g)$, but since it is rather technical, we first study the problem for univariate functions in Section 3. This simplified problem is also interesting and has been studied because of its applications in analyzing multivariate time-series data [12, 16]. Computing σ_1 even for univariate functions requires minimizing a linear-size sum of rational functions, which we cannot hope to do exactly. We show that a $(1 + \varepsilon)$ -approximation algorithm exists that takes $O(n/\sqrt{\varepsilon})$ time for univariate functions (aligned or unaligned), and $O(n/\varepsilon^2)$ time for bivariate functions in the aligned case. The latter extends immediately to $O(n^2/\varepsilon^4)$ time for the non-aligned case, or $O(n/\varepsilon^2)$ time with certain realistic input assumptions. Since the σ_2 measure can easily be computed in linear time for aligned terrains, we do not discuss it further in this paper. To simplify notation, since each section discusses a specific type of measure, we drop the subscript (1 or ∞) from μ and σ in the subsequent sections.

The main novelty and interest of this paper lie in the approach taken to determine correlation of two bivariate functions on the same domain. It gives rise to new geometric problems that can be solved using a suitable combination of advanced algorithmic and approximation techniques. The efficient algorithms that we present are based on geometric properties of the matching functions, which we provide. The convexity of $\mu_1(s, t)$ is especially surprising, as this function is a degree-3 rational function.

2. MINIMIZING THE MAX DISTANCE OF TWO BIVARIATE FUNCTIONS

Let $f, g : \mathbb{D} \rightarrow \mathbb{R}$ be two piecewise-linear bivariate func-

tions as defined above. In this section the goal is to find s^*, t^* that minimizes $\mu(s, t)$ according to (1), the maximum vertical distance between the graphs of g and $sf + t$. Let \mathbb{X} be the set of vertices in the overlay of \mathbb{M}_f and \mathbb{M}_g . If \mathbb{M}_f and \mathbb{M}_g are identical (f and g are aligned), then \mathbb{X} is the set of its vertices, otherwise \mathbb{X} is the set of vertices in \mathbb{M}_f , in \mathbb{M}_g , and intersections of their edges. As remarked, \mathbb{X} can be as large as $\Theta(n^2)$ for unaligned functions.

We first observe that the maximum of μ is always realized at a point of \mathbb{X} , implying that the problem of computing $\sigma(f, g)$ is inherently discrete. For a point $v \in \mathbb{X}$, we define the function

$$\begin{aligned} d_v(s, t) &= |sf(v) + t - g(v)| \\ &= \max\{sf(v) + t - g(v), g(v) - sf(v) - t\}, \end{aligned} \quad (4)$$

which is the vertical height between the two terrains (for a given s and t) at v . Then the above observation implies that $\mu(s, t) = \max_{v \in \mathbb{X}} d_v(s, t)$. Hence, the problem reduces to computing the minimum value on the upper envelope of $\mathcal{D} = \{d_v(s, t) \mid v \in \mathbb{X}\}$. For each v , we define two half-spaces (in the (s, t, z) -space): $\gamma_v^+ : z \geq sf(v) + t - g(v)$ and $\gamma_v^- : z \geq g(v) - sf(v) - t$. Then, by (4), the upper envelope of \mathcal{D} is the same as the boundary of the convex polyhedron $\bigcap_{v \in \mathbb{X}} \gamma_v^+ \cap \gamma_v^-$. The lowest vertex (in z -direction) of this convex polytope can be computed in $O(|\mathbb{X}|)$ expected time, using a linear-time randomized algorithm for linear programming (LP) in fixed dimensions [6].

As mentioned above, if f and g are aligned or they satisfy realistic-input assumptions, then $|\mathbb{X}| = O(n)$, but otherwise it can be $\Theta(n^2)$. In what follows we show how to reduce the potentially quadratic running time for the general case by not computing \mathbb{X} explicitly, but considering only a linear number of intersection points. Roughly speaking, we avoid computing the set \mathbb{X} (and the set of constraints induced by \mathbb{X}) explicitly, by using a random-sampling approach; this idea was first presented in [8]. Our method is similar to the LP algorithm in [6], and is sketched in Algorithm MINMAXDISTANCE. For a subset $Y \subseteq \mathbb{X}$, let $\text{OPTIMAL}(Y)$ denote the optimal solution for Y , i.e., the lowest vertex in the convex polyhedron $\bigcap_{v \in Y} \gamma_v^+ \cap \gamma_v^-$. As mentioned above, this can be computed in $O(|Y|)$ expected time.

Algorithm 1: MINMAXDISTANCE

1. $R \leftarrow$ random subset of \mathbb{X} of $\min\{|\mathbb{X}|, 9n\}$ points
2. **repeat**
3. $\xi_0 \leftarrow \text{OPTIMAL}(R)$
4. $V \leftarrow \{v \in \mathbb{X} \mid \xi_0 \notin \gamma_v^+ \cap \gamma_v^-\}$
5. **if** $|V| > 2n$ **then**
6. $R \leftarrow$ random subset of \mathbb{X} of $\min\{|\mathbb{X}|, 9n\}$ points
7. **else** $R \leftarrow R \cup V$
8. **until** $V = \emptyset$
9. **return** ξ_0

Figure 2. The algorithm for computing $\mu_\infty(s, t)$.

Since \mathbb{X} is not explicitly computed, it is not straightforward to choose R and compute V , the set of *violated* vertices. We describe these two steps in detail. Let $\xi_0 = (s_0, t_0, z_0)$ and $f_0(x, y) = s_0 f(x, y) + t_0$.

Computing the violated vertices. By definition, a point $v \in \mathbb{X}$ is violated if $d_v(s_0, t_0) > z_0$, i.e., the vertical distance between (the graphs of) f_0 and g at v is larger than z_0 .

The vertical distances between f_0 and g at the vertices of \mathbb{M}_f and \mathbb{M}_g can be computed in a total of $O(n \log n)$ time. More challenging is therefore finding the intersection points of pairs of edges at which the vertical distance is larger than z_0 —we want to report them only if there are at most $2n$ of them, without spending more time otherwise.

To report this type of violated points of \mathbb{X} , we proceed in two steps. First we use the so-called hereditary segment tree data structure to reduce the vertical-distance problem between line segments to a problem of reporting lines in 3D that are vertically more than z_0 apart. Then we solve that problem by mapping the lines in \mathbb{R}^3 to Plücker points and Plücker hyperplanes, and solving a halfspace range reporting problem in \mathbb{R}^5 . By using the trade-off techniques for halfspace range reporting we find all the violated constraints in $O^*(n^{4/3} + |V|)$ time, with the option to stop reporting if $|V| > 2n$. A more detailed description follows.

For the first part, we use a two-level hereditary segment tree T [5]. Let R and B be the sets of edges in \mathbb{M}_f and \mathbb{M}_g , respectively. As in [5], T can be augmented to produce a bipartite clique decomposition of the intersecting pairs of segments in $R \times B$. That is, we compute a family $\mathcal{F} = \{(R_1, B_1), \dots, (R_u, B_u)\}$, where $R_i \subseteq R, B_i \subseteq B$, such that

- (i) every segment in R_i intersects every segment in B_i ;
- (ii) the left endpoints of all segments of R_i lie below the lines supporting every segment of B_i , or all of them lie above these lines;
- (iii) for every intersecting pair $(r, b) \in R \times B$ there is an i such that $r \in R_i, b \in B_i$;
- (iv) $\sum_i (|R_i| + |B_i|) = O(n \log^2 n)$.

\mathcal{F} can be computed in $O(n \log^2 n)$ time. For each $(R_i, B_i) \in \mathcal{F}$, we “lift” every line segment $r \in R_i$ (resp. $b \in B_i$) to a line in \mathbb{R}^3 , namely, the line containing the segment $f_0(r)$ (resp. $g(b)$). Let \hat{R}_i and \hat{B}_i be the sets of resulting lines in \mathbb{R}^3 . We wish to report the pairs of lines in $\hat{R}_i \times \hat{B}_i$ that are vertically more than z_0 apart.

This second problem is transformed into a halfspace range reporting problem in \mathbb{R}^5 , as follows. Testing whether two lines ℓ_1, ℓ_2 in \mathbb{R}^3 lie at a vertical distance of more than z_0 from each other can be formulated (after translating one of the lines by z_0) as testing whether the Plücker point of ℓ_1 lies in one of the halfspaces bounded by the Plücker hyperplane of ℓ_2 . We map the lines from, say, \hat{B}_i , to Plücker points and perform a halfspace reporting query with each of the halfspaces bounded by the Plücker hyperplanes of the lines of \hat{R}_i . To solve the halfspace reporting problem we apply standard trade-off techniques for geometric range searching (see for example [1]). We build a data structure of size $O^*(n_i^{4/3})$, where $n_i = |R_i| + |B_i|$, that answers halfspace reporting queries in $O(n_i^{1/3} + p)$ time, where p is the output size of the query. In our context, each reported point corresponds to an intersecting pair of segments $(r, b) \in R \times B$ such that $d_v(s_0, t_0) > z_0$. We repeat this procedure for all bipartite cliques in \mathcal{F} . Note that in line 4 of the algorithm, we do not need to report all violated vertices of \mathbb{X} , but we can stop as soon as the number of reported pairs exceeds $2n$. Thus the total time spent in line 4 is $O^*(n^{4/3})$.

Choosing a random sample. We can use the bipartite clique decomposition \mathcal{F} to choose a random point of \mathbb{X} , as

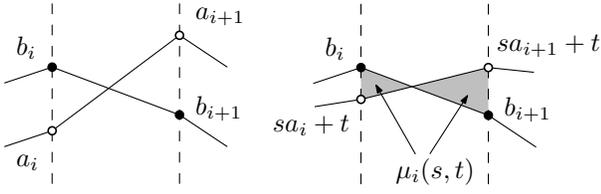


Figure 3. Edges of f and g , and the function $\mu_i(s, t)$.

follows. We first choose a random bipartite clique $(R_i, B_i) \in \mathcal{F}$; the probability of (R_i, B_i) being chosen is $|R_i| \cdot |B_i|/k$, where $k = \sum_j |R_j| \cdot |B_j|$ is the total number of intersecting pairs in $R \times B$. Next, we choose a random pair $(r, b) \in R_i \times B_i$, each pair is chosen with equal probability, and the desired point is $r \cap b$. After having computed \mathcal{F} , it takes $O(\log n)$ time to choose a random point. Hence, we can compute R in $O(n \log n)$ time.

Number of iterations. As argued in [6], it can be shown that line 7 of Algorithm MINMAXDISTANCE (i.e., $|V| \leq 2n$) will be executed at most four times. This is because every time $\xi_0 \neq \text{OPTIMAL}(\mathbb{X})$, i.e., $V \neq \emptyset$, the point of \mathbb{X} corresponding to one of the three constraints that define $\text{OPTIMAL}(\mathbb{X})$ is added to V (and thus to R in line 7), and from then on it remains in R . Therefore after executing line 7 at most three times, R will contain the three vertices that define the optimum, hence the optimum will be found in the next (fourth) iteration. Following the same argument as in [6], based on random sampling, we can prove that the expected number of iterations of the loop is $O(1)$. Hence, we can conclude:

THEOREM 1. *Given two bivariate piecewise-linear functions over a common domain, a linear transformation minimizing the maximum distance between them can be found in $O^*(n^{4/3})$ expected time, where n is the total number of vertices in the graphs of these functions. If the functions are aligned, the transformation can be found in linear time.*

3. MINIMIZING THE AREA BETWEEN TWO UNIVARIATE FUNCTIONS

In this section, we assume $\mathbb{D} \subseteq \mathbb{R}$ to be a bounded interval and $f, g : \mathbb{D} \rightarrow \mathbb{R}$ to be two univariate piecewise-linear functions. Since the overlay of two one-dimensional subdivisions has linear complexity, we can assume that f and g are aligned and they are defined by a common subdivision \mathbb{M} of \mathbb{D} . For any value of s and t , let $\mu(s, t)$ denote the area between $sf + t$ and g over \mathbb{D} . The goal is to compute $\arg \min_{s, t} \mu(s, t)$. We begin by analyzing the function μ .

Analytic form of μ . The analytic form of the bivariate function $\mu(s, t)$ depends on the set of pairs of intersecting edges of $sf + t$ and g . Let $x_1 < \dots < x_n$ be the vertices of \mathbb{M} , and let $a_i = f(x_i)$, $b_i = g(x_i)$. Consider the vertical slab $[x_i, x_{i+1}] \times \mathbb{R}$, see Figure 3. Let f_i and g_i be the functions f and g , restricted to the interval $[x_i, x_{i+1}]$, and let μ_i be the area between the graphs of these functions, over the same domain. Unless f_i is constant (we deal with this degenerate situation in the full version of this paper), there is exactly one pair s, t that makes $sf_i + t = g_i$, and $\mu_i(s, t) = 0$. This happens when $sa_i + t = b_i$ and $sa_{i+1} + t = b_{i+1}$ simultaneously.

LEMMA 1. *The region in the (s, t) -plane where $sf_i + t$ and g_i intersect is a double wedge whose bounding lines are $\ell_i: t = -a_i s + b_i$ and $\ell_{i+1}: t = -a_{i+1} s + b_{i+1}$. These lines intersect in the apex (\bar{s}_i, \bar{t}_i) , where $\bar{s}_i f_i + \bar{t}_i = g_i$ and $\mu_i(\bar{s}_i, \bar{t}_i) = 0$. Above ℓ_i and ℓ_{i+1} (in the direction of t), μ_i is linear in s and t , and the same is true below ℓ_i and ℓ_{i+1} . Below ℓ_i and above ℓ_{i+1} in the (s, t) -plane, we have*

$$\mu_i(s, t) = (x_{i+1} - x_i) \cdot \frac{(b_i - sa_i - t)^2 + (sa_{i+1} - b_{i+1} + t)^2}{2(sa_{i+1} - sa_i - b_{i+1} + b_i)}.$$

A similar expression exists for the function that is valid in the wedge above ℓ_i and below ℓ_{i+1} .

Figure 4(a) shows the function $\mu_i(s, t)$ graphically. We observe that inside the double wedge defined by ℓ_i and ℓ_{i+1} , μ_i is a fraction (i.e., rational function) that has the unknown s in the denominator. To minimize μ , we must solve $\min_{s, t} \sum_{i=1}^{n-1} \mu_i(s, t)$, which is a sum of linear and rational functions and involves solving a polynomial of linear degree, if the minimum occurs where many pairs of edges of f and g intersect. We cannot hope to find an exact combinatorial solution in this case.

It will be useful to analyze the function μ_i closer. For a fixed scaling factor \hat{s} , consider $\mu_i(\hat{s}, t)$ as a function of t . By Lemma 1, $\mu_i(\hat{s}, t)$ is linear in t below ℓ_i and ℓ_{i+1} , or above both of them, and a quadratic function in t between ℓ_i and ℓ_{i+1} , i.e., $\mu_i(\hat{s}, t)$ consists of three pieces. Moreover, $\mu_i(\hat{s}, t)$ is symmetric and differentiable everywhere; see Figure 4(c). Despite its complicated form, we can prove the following for μ_i :

LEMMA 2. *For every i , μ_i is a convex function, and the restriction of μ_i to a ray starting at (\bar{s}_i, \bar{t}_i) is a linear function.*

PROOF. For simplicity, we assume that $x_{i+1} - x_i = 1$. The lines ℓ_i and ℓ_{i+1} partition the st -plane into four wedges. The function μ_i is linear in the wedge lying above (or below) both ℓ_i and ℓ_{i+1} , so μ_i is obviously convex within each of these two wedges and linear along any ray emanating from (\bar{s}_i, \bar{t}_i) within them. Next, we prove the convexity within each of the two other wedges. Consider the wedge W^+ lying below ℓ_i and above ℓ_{i+1} ; the other case is symmetric. Let $\ell_i(s) = b_i - a_i s$ and $\ell_{i+1}(s) = b_{i+1} - a_{i+1} s$ be the linear functions defined by the lines ℓ_i and ℓ_{i+1} . We can then write $\mu_i(s, t)$ for W^+ (cf. Lemma 1) as

$$\mu_i(s, t) = \frac{(t - (\ell_i(s) + \ell_{i+1}(s))/2)^2}{\ell_i(s) - \ell_{i+1}(s)} + \frac{\ell_i(s) - \ell_{i+1}(s)}{4}.$$

Since a convex (resp. linear) function remains convex (resp. linear) under an affine transformation of its domain, we apply the affine transform $t' = t - (\ell_i(s) + \ell_{i+1}(s))/2$ and $s' = \ell_i(s) - \ell_{i+1}(s)$ and rewrite μ_i as

$$\mu_i(s', t') = \frac{t'^2}{s'} + \frac{s'}{4}.$$

Along a ray from $(s', t') = (0, 0)$ emanating into W^+ , $t' = \alpha s'$, $\mu_i(s', \alpha s') = (\alpha^2 + 1/4)s'$ is linear. A similar argument holds in the diametrically opposite wedge W^- . Hence the graph of $\mu(s', t')$ is a cone with origin at $(0, 0)$ and to check its convexity it is sufficient to consider the univariate function $\mu(\hat{s}', t')$ along the line $\ell: s' = \hat{s}' > 0$ (a symmetric analysis handles $\hat{s}' < 0$). In W^+ , this function has the form

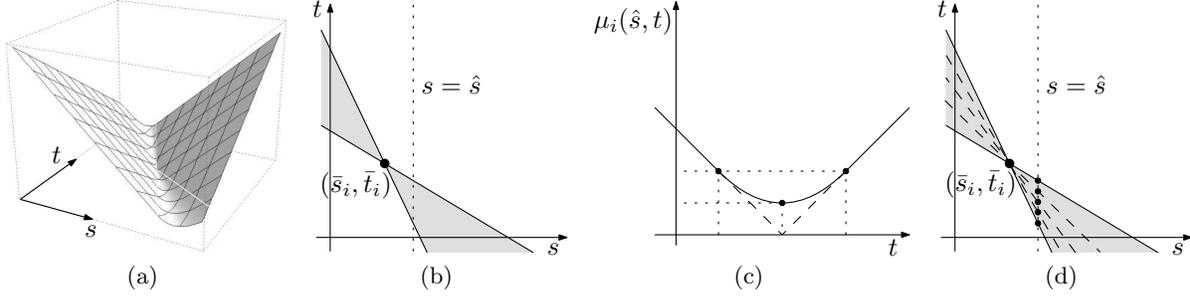


Figure 4. (a) Graph of $\mu_i(s, t)$. (b) Double wedge arising from f_i and g_i in the (s, t) -plane. (c) Cross-section in the plane $s = \hat{s}$ showing $\mu_i(\hat{s}, t)$. (d) Subdivision into double wedges in the (s, t) -plane.

$\mu_i(\hat{s}', t') = \frac{t'^2}{\hat{s}'} + \frac{\hat{s}'}{4}$ which is a convex quadratic function. Outside of W^+ , $\mu_i(\hat{s}', t') = |t'|$ as is easily checked. Hence, at the point $(\hat{s}', \pm\hat{s}'/2)$ where line ℓ leaves W^+ , the slopes of the two expressions for $\mu_i(\hat{s}', t')$ agree, so it is convex as a function of t' . From the above discussion, it follows that $\mu(s', t')$ and therefore $\mu(s, t)$ are convex over their entire range, completing the proof of the lemma. \square

COROLLARY 1. μ is a convex function.

Approximating μ . Since it is hard to compute the minimum value of $\mu(s, t)$, we define a convex piecewise-linear function $\tilde{\mu}(s, t)$, for a given $\varepsilon > 0$, such that $\mu(s, t) \leq \tilde{\mu}(s, t) \leq (1 + \varepsilon)\mu(s, t)$ for all $s, t \in \mathbb{R}$, and we compute $(\tilde{s}, \tilde{t}) = \arg \min_{s, t} \tilde{\mu}(s, t)$.

For each i , we define an approximation $\tilde{\mu}_i$ of μ_i : μ_i is already linear outside the double wedge, so we set $\tilde{\mu}_i$ to μ_i outside the double wedges and focus on approximating the quadratic part of $\mu_i(s, t)$ inside the double wedge. We first fix a value \hat{s} and approximate the univariate function $\mu_i(\hat{s}, t)$. We choose $\lceil 2/\sqrt{\varepsilon} \rceil$ points on the graph of $\mu_i(\hat{s}, t)$, equally spaced along the line $s = \hat{s}$, such that the maximum distance between the polygonal line through the chosen points and $\mu_i(\hat{s}, t)$ itself is at most $\varepsilon\mu_i(\hat{s}, t)$, see Figure 4(d). We extend this approximation to all values of s by choosing lines through the new points and $(\tilde{s}_i, \tilde{t}_i)$, giving a partition of the plane into $\Theta(1/\sqrt{\varepsilon})$ wedges. Lemma 2 ensures that $\tilde{\mu}_i(s, t) \leq (1 + \varepsilon)\mu_i(s, t)$ everywhere. We set $\tilde{\mu}(s, t) := \sum_i \tilde{\mu}_i(s, t)$. By construction and the convexity of μ , we can prove the following

LEMMA 3. $\tilde{\mu}$ is convex and $\mu(s, t) \leq \tilde{\mu}(s, t) \leq (1 + \varepsilon)\mu(s, t)$.

Approximation algorithm. We now describe the algorithm for computing the pair (\tilde{s}, \tilde{t}) that minimizes $\tilde{\mu}$. Let L be the set of lines in the (s, t) -plane, used to define the functions $\tilde{\mu}_i$. Set $n' := |L| = O(n/\sqrt{\varepsilon})$. The function $\tilde{\mu}$ is linear within each cell of the arrangement $\mathcal{A}(L)$. The minimum is achieved at a vertex of $\mathcal{A}(L)$. By traversing $\mathcal{A}(L)$ and computing the minimum value of $\tilde{\mu}$ at each vertex, (\tilde{s}, \tilde{t}) can be computed in $O(n^2/\varepsilon)$ time.

A more efficient solution can be obtained by exploiting the convexity of $\tilde{\mu}$ and using a prune-and-search approach. We compute a $(1/2)$ -cutting Ξ of L of size $O(1)$ in $O(n')$ time [4]. Ξ is a tiling of the plane by a constant number of triangles with the property that at most half of the lines of L cross each triangle of Ξ and each vertex of Ξ is a vertex

of $\mathcal{A}(L)$. Along each of the $O(1)$ edges e of Ξ , we compute $\tilde{\mu}$ over e and find its minimum and gradient at the minimum in $O(n')$ time, which will reveal the triangle $\Delta \in \Xi$ containing (\tilde{s}, \tilde{t}) . We recurse on the lines of L that cross Δ ; their number is at most $n'/2$ by the definition of $(1/2)$ -cuttings. The contributions from the remaining $\tilde{\mu}_i$ sum up to a single linear function over Δ and can be easily kept track of in constant additional time. After $O(\log n')$ phases of divide-and-conquer, which together take $O(n')$ time, we find a triangle that lies in a single cell of $\mathcal{A}(L)$ and (\tilde{s}, \tilde{t}) is a vertex of the triangle, from which we can easily compute (\tilde{s}, \tilde{t}) .

THEOREM 2. Let f and g be two univariate piecewise-linear functions with a total of n pieces. For any fixed $\varepsilon > 0$, a pair (\tilde{s}, \tilde{t}) can be computed in $O(n/\sqrt{\varepsilon})$ time such that $\sigma(f, g) \leq \mu(\tilde{s}, \tilde{t}) \leq (1 + \varepsilon)\sigma(f, g)$.

4. MINIMIZING THE VOLUME BETWEEN TWO BIVARIATE FUNCTIONS

We extend the approach of the previous section to the case where f and g are two aligned bivariate functions, defined by a common triangulation \mathbb{M} of \mathbb{D} . The main difficulty in generalizing the above approach is that the analogue of function μ_i , though convex, is much less well behaved. In the previous case, the bivariate function $\mu_i(s, t)$ behaves essentially as a univariate function, in the sense that its graph is a surface of a convex cone. As we require a different method to approximate the volume over a single triangle of \mathbb{M} , we obtain an $O(n/\varepsilon^2)$ time algorithm to compute a pair \tilde{s}, \tilde{t} such that $\sigma(f, g) \leq \mu(\tilde{s}, \tilde{t}) \leq (1 + \varepsilon)\sigma(f, g)$.

Analysis of μ over a triangle. Let $\Delta = \Delta v_i v_j v_k$ be a triangle of \mathbb{M} , and let Δ_f and Δ_g be the triangles in the graph of f and g , respectively, defined by Δ (i.e., Δ is their xy -projection). Let a_i, a_j , and a_k be the vertices of Δ_f , and let b_i, b_j , and b_k be the corresponding vertices of Δ_g . With slight abuse of notation, we use a_i, a_j , and a_k , and b_i, b_j , and b_k to denote the function values of these vertices as well (i.e., $a_i = f(v_i)$, $b_i = g(v_i)$, etc.). Without loss of generality, we assume that Δ is an equilateral triangle with side length 1. Depending on the scaling s and translation t , triangle $s \cdot \Delta_f + t$ may intersect, be above, or be below Δ_g . The volume between the triangles is therefore composed of one or two parts.

In the (s, t) -plane, let $\ell_i : t = b_i - sa_i$ represent the set of points (s, t) at which vertex v_i of Δ_g and $s\Delta_f + t$ coincide;

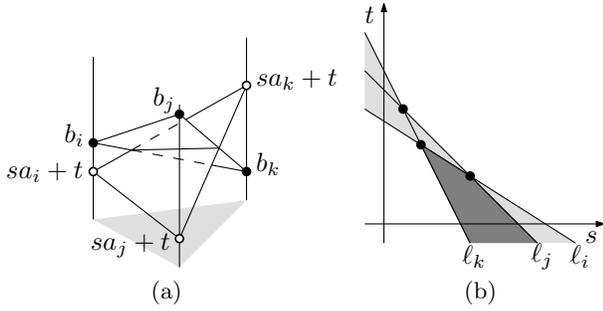


Figure 5. (a) A pair of intersecting triangles with common (x, y) -projection. (b) The situation of (a) corresponds to values of (s, t) in the dark cell of the parameter space.

l_j and l_k are defined analogously. We will also view l_i as a univariate function $l_i(s) = b_i - sa_i$ and use a similar convention for l_j and l_k . Generally, these three lines partition the (s, t) -plane into seven cells, see Figure 5(b). Let \mathcal{A}_3 denote this planar decomposition. Inside a cell of \mathcal{A}_3 , for every choice of s and t , the same subset of vertices of $s \cdot \Delta_f + t$ lies above Δ_g , and the volume between these triangles can be expressed by a single analytic function. For example, for two triangles in the configuration shown in Figure 5(a), the expression for the volume between the triangles (assuming their projection is equilateral with unit side length) is

$$\begin{aligned} \mu_{ijk}(s, t) &= \frac{\sqrt{3}}{12} (l_i(s) + l_j(s) + l_k(s) - 3t) \\ &+ \frac{\sqrt{3}}{6} \cdot \frac{(t - l_k(s))^3}{(l_i(s) - l_k(s)) \cdot (l_j(s) - l_k(s))}. \end{aligned} \quad (5)$$

More precisely, the formula is valid whenever $sa_i + t$ lies below b_i , $sa_j + t$ lies below b_j , and $sa_k + t$ is above b_k . Let us examine a value \hat{s} of s at which the line $s = \hat{s}$ intersects the lines l_k, l_i, l_j , in this order, bottom to top. The form of the function μ_{ijk} in the cells of \mathcal{A}_3 met by this line are, in bottom to top order: $\frac{\sqrt{3}}{12} (l_i(s) + l_j(s) + l_k(s) - 3t)$, the form given by (5),

$$\begin{aligned} &\frac{\sqrt{3}}{12} (3t - (l_i(s) + l_j(s) + l_k(s))) \\ &+ \frac{\sqrt{3}}{6} \frac{(l_j(s) - t)^3}{(l_j(s) - l_i(s)) \cdot (l_j(s) - l_k(s))}, \end{aligned} \quad (6)$$

and finally $\frac{\sqrt{3}}{12} (3t - l_i(s) - l_j(s) - l_k(s))$. The equations correspond, respectively, to $s \cdot \Delta_f + t$ lying below all vertices of Δ_g , above only b_k , above b_k and b_i , and above all of them. In the remaining three faces of this arrangement, the form of the function is given by similar formulas, obtained from (5) or (6) by a permutation of indices.

LEMMA 4. $\mu_{ijk}(s, t)$ is a convex function.

PROOF. Since μ_{ijk} measures the volume lying vertically between two surfaces in xyz -space, its convexity is not affected by an affine transformation of the (x, y) -plane. In particular, we have already taken the liberty of assuming that Δ is an equilateral triangle with side length 1 in the above expressions for μ_{ijk} .

Our proof consists of two steps: We first verify that μ_{ijk} is convex over every face of \mathcal{A}_3 and then check that the convexity is not locally violated along the ‘‘seams,’’ i.e., along the three lines separating the cells of \mathcal{A}_3 , which form the

locus of points where the analytic expression defining μ_{ijk} changes.

For the first step, there is nothing to prove in the top and bottom cells of \mathcal{A}_3 , as the function is linear there. In the other faces, up to a permutation of i, j, k , it has a form given in (5) or (6).

Without loss of generality, suppose the function has the form of (5). For the purpose of determining convexity, since the first term is linear, the function is convex if and only if the second term is convex. By applying an appropriate transform to simplify the second term and computing its Hessian directly (and recalling that, by construction, in the relevant face of \mathcal{A}_3 , we have $t > l_k(s)$, $l_i(s) > l_k(s)$, and $l_j(s) > l_k(s)$), we can verify that the Hessian is non-negative in the desired region. We omit the tedious details from this abstract.

For the second step of the proof, let $M = \{(s, t, \mu_{ijk}(s, t)) : (s, t) \in \mathbb{R}^2\}$ be the graph of μ_{ijk} . We will argue that M has a well-defined tangent plane at every point (projecting to a point) on the bounding lines l_i, l_j, l_k , by determining the tangent plane at an interior point (s, t) of a cell of \mathcal{A}_3 and computing its limit as (s, t) approaches a bounding line. Since the analytical expressions for the function valid on opposite sides of an arrangement edge agree along the edge (the function is continuous), the ‘‘slopes’’ of the two expressions along the edge agree, so it is sufficient to check that their slopes in some other direction, not parallel to the line, agree as well, to conclude that the tangent planes given by the expressions on both sides of the edge coincide and therefore M has a well-defined tangent plane at the point.

We pick the direction of a line $s = \hat{s}$, for some fixed \hat{s} . Our plan is to check that the function $\mu_{ijk}(\hat{s}, t)$ of t has a well-defined tangent at the three points where this line crosses the bounding lines l_i, l_j, l_k ; this is done by computing the one-sided derivatives of this univariate function on both sides of each line and verifying that they agree.

For the situation described in Figure 5, $\mu_{ijk}(s, t)$ is given by (5). The following calculation assumes that the line $\gamma : s = \hat{s}$ crosses the lines l_k, l_i, l_j , in this order, bottom to top. (5) applies to the second face from the bottom in \mathcal{A}_3 crossed by γ . The one-sided derivative at $\gamma \cap l_i$ is given by

$$\begin{aligned} \left. \frac{\partial}{\partial t} \mu_{ijk}(\hat{s}, t) \right|_{t \uparrow l_i(\hat{s})} &= \frac{\partial}{\partial t} \left[\frac{\sqrt{3}}{12} (l_i(\hat{s}) + l_j(\hat{s}) + l_k(\hat{s}) - 3t) \right. \\ &\quad \left. + \frac{\sqrt{3}}{6} \frac{(t - l_k(\hat{s}))^3}{(l_i(\hat{s}) - l_k(\hat{s})) \cdot (l_j(\hat{s}) - l_k(\hat{s}))} \right]_{t \uparrow l_i(\hat{s})} \\ &= \left[-\frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{6} \frac{3(t - l_k(\hat{s}))^2}{(l_i(\hat{s}) - l_k(\hat{s})) \cdot (l_j(\hat{s}) - l_k(\hat{s}))} \right]_{t \uparrow l_i(\hat{s})} \\ &= -\frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{2} \cdot \frac{(l_i(\hat{s}) - l_k(\hat{s}))^2}{(l_i(\hat{s}) - l_k(\hat{s})) \cdot (l_j(\hat{s}) - l_k(\hat{s}))} \\ &= -\frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{2} \cdot \frac{l_i(\hat{s}) - l_k(\hat{s})}{l_j(\hat{s}) - l_k(\hat{s})} \\ &= \frac{\sqrt{3}}{4} \cdot \frac{(l_i(\hat{s}) - l_k(\hat{s})) + (l_i(\hat{s}) - l_j(\hat{s}))}{l_j(\hat{s}) - l_k(\hat{s})}. \end{aligned}$$

We now repeat the calculation for $\mu_{ijk}(s, t)$ above the line l_i , as given by (6), obtaining

$$\left. \frac{\partial}{\partial t} \mu_{ijk}(\hat{s}, t) \right|_{t \downarrow l_i(\hat{s})} = \frac{\sqrt{3}}{4} \cdot \frac{(l_i(\hat{s}) - l_k(\hat{s})) + (l_i(\hat{s}) - l_j(\hat{s}))}{l_j(\hat{s}) - l_k(\hat{s})},$$

showing that indeed $\mu_{ijk}(\hat{s}, t)$, as a function of t , is smooth

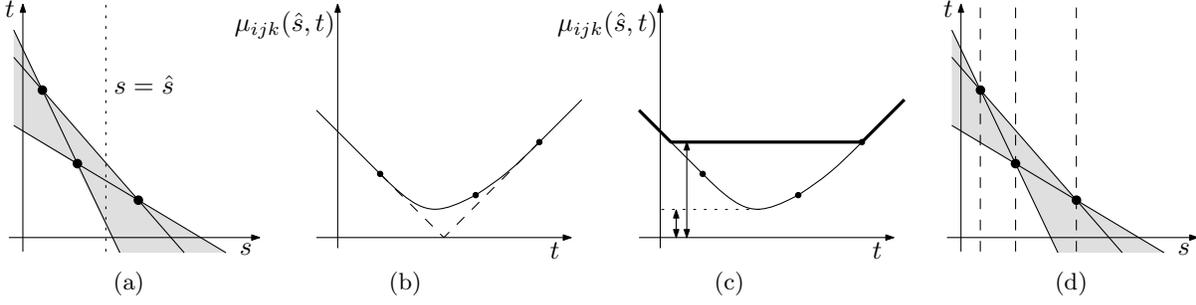


Figure 6. Three lines in the (s, t) -plane and its vertical cross-section showing $\mu_{ijk}(\hat{s}, t)$.
BORIS SAYS: this figure needs more elaborate captions, one for each subfigure

at $t = \ell_i(\hat{s})$. (We have considered the hard case, where two non-linear expressions for μ_{ijk} meet across a line, here ℓ_i . A similar argument can be made where the linear and non-linear portions of μ_{ijk} meet; here at the intersections of γ with ℓ_k and ℓ_j . The details are omitted in this version.) Therefore along the line $s = \hat{s}$, the function $\mu_{ijk}(\hat{s}, t)$ is smooth everywhere. By the above argument, $\mu_{ijk}(s, t)$ must have a well-defined tangent plane at every point of its domain. This, together with the fact that it is convex over every cell of \mathcal{A}_3 , implies that $\mu_{ijk}(s, t)$ is convex everywhere. \square

Considering the situation along a line $s = \hat{s}$ again, we see that when $\hat{s} \cdot \Delta_f + t$ and Δ_g do not intersect, $\mu_{ijk}(\hat{s}, t)$ is a linear function in t . When they do intersect, the function is piecewise-cubic. More precisely, as can be seen from (5) and (6) by setting $s = \hat{s}$, $\mu_{ijk}(\hat{s}, t)$ has (at most) three break points, it is a cubic function between consecutive break points and linear outside the break points, see Figure 6(b). The proof implies that it is differentiable everywhere including at the break points.

A constant-factor approximation for μ_{ijk} . We continue by showing that $\mu_{ijk}(s, t)$ can be approximated within a factor 16 by a constant-complexity piecewise-linear function.

LEMMA 5. *For any fixed \hat{s} , the value of $\mu_{ijk}(\hat{s}, t)$ at its break points is at most 16 times the value of $\mu_{ijk}(\hat{s}, t)$ at its minimum.*

PROOF. By rescaling the triangle, we will assume for the duration of this proof that the xy -projections of Δ_f and Δ_g have unit area. If, for $s = \hat{s}$ and some t , the triangles do not intersect, then the volume between $\hat{s} \cdot \Delta_f + t$ and Δ_g is simply the average distance between the triangles at the three vertices. At the break points, one distance is zero: $a_i \hat{s} + t = b_i$, $a_j \hat{s} + t = b_j$, or $a_k \hat{s} + t = b_k$. Assume without loss of generality that $\hat{s} \cdot \Delta_f + t$ and Δ_g do not properly intersect when $a_j \hat{s} + t = b_j$, nor when $a_k \hat{s} + t = b_k$; they define the extreme break points in Figure 6(b).

Let $\alpha = (a_j \hat{s} + t) - b_j + b_k - (a_k \hat{s} + t) = a_j \hat{s} - b_j + b_k - a_k \hat{s}$ be the maximum vertical distance between any two points on Δ_f and Δ_g if they intersect. Let $\beta = (a_i \hat{s} + t) - b_i + b_k - (a_k \hat{s} + t) = a_i \hat{s} - b_i + b_k - a_k \hat{s}$, and assume $\beta \geq \alpha/2$ (otherwise, exchange the roles of j and k and we are in the desired situation). Now β is the maximum height difference at the vertices with index i when $\hat{s} \cdot \Delta_f + t$ and Δ_g intersect. See Figure 7.

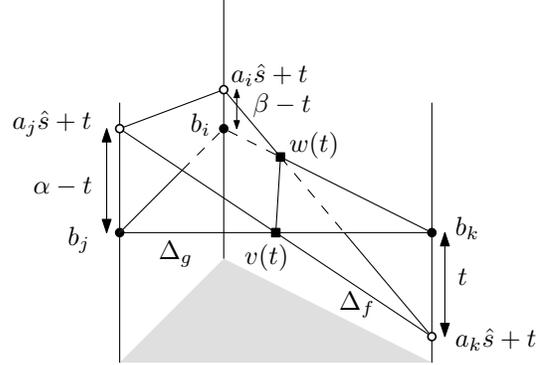


Figure 7. Illustration of the proof of Lemma 5.

The maximum volume at a break point then is $(\alpha + \beta + 0)/3 \leq 2\alpha/3$, which occurs when t is such that $a_k \hat{s} + t = b_k$. Reparameterize t such that $t = 0$ in this case, and t is positive when $\hat{s} \cdot \Delta_f + t$ and Δ_g intersect. The minimum volume in between occurs for some $0 \leq t \leq \alpha$. We will distinguish two cases depending on the value of t , and show that a relatively big tetrahedron exists between $\hat{s} \cdot \Delta_f + t$ and Δ_g whose volume is at least $\alpha/24$.

Let $e(t)$ be the intersection edge of $\hat{s} \cdot \Delta_f + t$ and Δ_g . Let $v(t)$ be the endpoint of $e(t)$ that lies on the side $b_j b_k$ of Δ_g , and let $w(t)$ be the other endpoint, on $b_i b_k$ or $b_i b_j$.

First, assume that $t \geq \alpha/2$. If $w(t) \in b_i b_k$, then the projected area of $\Delta b_k v(t) w(t)$ is at least $1/4$, and hence the volume of the tetrahedron $b_k v(t) w(t) (a_k \hat{s} + t)$ is at least $\alpha/24$ because $b_k - (a_k \hat{s} + t) = t \geq \alpha/2$. And if $w(t) \in b_i b_j$, then the projected area of $\Delta b_k v(t) b_i$ is at least $1/2$, and hence the volume of the tetrahedron $b_k v(t) b_i (a_k \hat{s} + t)$ is at least $\alpha/12$.

Second, assume that $t < \alpha/2$. Then $w(t) \in b_i b_k$ by the assumption $\beta \geq \alpha/2$. The projected area of $\Delta b_j b_i v(t)$ is at least $1/2$, and the volume of the tetrahedron $b_j b_i v(t) (a_j \hat{s} + t)$ is at least $\alpha/12$ because $(a_j \hat{s} + t) - b_j = t > \alpha/2$.

Hence, the volume at a break point is at most 16 times the volume at the minimum in all cases. \square

It follows that we can create a piecewise-linear 16-approximation of $\mu_{ijk}(\hat{s}, t)$ by taking the highest break point, choosing the unique other point on $\mu_{ijk}(\hat{s}, t)$ that has the same height, using a constant function between them, and the linear functions of $\mu_{ijk}(\hat{s}, t)$ outside, see Figure 6(c).

Let $[\tau_1, \tau_2]$ be the interval of t -values for which $\hat{s} \cdot \Delta_f + t$ and Δ_g intersect. We call $[\tau_1, \tau_2]$ the *intersection range* for Δ_f , Δ_g , and \hat{s} , and let $r = \tau_2 - \tau_1$.

LEMMA 6. For a fixed $s = \hat{s}$, the piecewise-linear 16-approximation is exact for all $t \leq \tau_1 - r$ and also for all $t \geq \tau_2 + r$.

PROOF. First, observe that the extensions of the linear parts of $\mu_{ijk}(\hat{s}, t)$ intersect the s -axis of the (s, t) -plane in the same point, see also Figure 6(b). Since the slopes of the linear parts of $\mu_{ijk}(\hat{s}, t)$ are also the same, the constant part of the 16-approximation has a range for t that cannot be more than twice the intersection range of Δ_f , Δ_g , and \hat{s} . The lemma follows. \square

Now we go back to the whole (s, t) -plane to extend the 16-approximation to a piecewise-linear bivariate function. If we consider $\mu_{ijk}(\hat{s}, t)$ for changing values of \hat{s} , the outer break points translate linearly in \hat{s} , because $\mu_{ijk}(s, t)$ is linear in the bottom and top cells of \mathcal{A}_3 . Hence, extending the 16-approximation of $\mu_{ijk}(\hat{s}, t)$ to a piecewise-linear function over the (s, t) -plane is natural, as long as the same line among ℓ_i , ℓ_j , and ℓ_k defines the highest break point. If we draw vertical (i.e., parallel to the t -axis) lines through the three intersections of ℓ_i , ℓ_j , and ℓ_k , the (s, t) -plane is partitioned into four strips, and inside each strip, we have the same outer two lines, see Figure 6(d). Which one gives the highest $\mu_{ijk}(\hat{s}, \cdot)$ value, for increasing values of \hat{s} , can change only once in each strip. We take another vertical line at such a location, giving at most eight strips. Inside each strip, we use three linear functions, and hence the graph of the function consists of no more than 24 patches. Finally, it can **BORIS SAYS: do we remember why? does it need to be put in this version? finally, is it true?** be proved that the resulting function is convex. Hence, we obtain:

LEMMA 7. A piecewise-linear convex factor-16 approximation $\tilde{\mu}_{ijk}(s, t)$ of $\mu_{ijk}(s, t)$ exists with $O(1)$ complexity. Moreover, for any \hat{s} , $\tilde{\mu}_{ijk}(\hat{s}, t) = \mu_{ijk}(\hat{s}, t)$ for all t at which $\mu_{ijk}(\hat{s}, t) \geq \mu_{ijk}(\hat{s}, \hat{t})$, where (\hat{s}, \hat{t}) is the highest break point of the function $\mu_{ijk}(\hat{s}, t)$.

A $(1 + \varepsilon)$ -approximation for μ_{ijk} . To obtain a $(1 + \varepsilon)$ -approximation for μ_{ijk} , we will use a different approach than the one in the previous section. Let $\varepsilon > 0$ be given, and assume $\varepsilon \leq \frac{1}{2}$. We will partition Δ into a family Ξ of $\Theta(1/\varepsilon^2)$ equal triangles by partitioning the three edges of Δ using m equally-spaced points on each edge, and choosing edges parallel to the sides of Δ . The parameter m is chosen to be $\lceil c/\varepsilon \rceil$ for a suitable constant c . For each triangle $\rho \in \Xi$, we define the functions μ_ρ and $\tilde{\mu}_\rho$ analogous to μ_{ijk} and $\tilde{\mu}_{ijk}$. By construction, $\sum_{\rho \in \Xi} \mu_\rho(s, t) = \mu_{ijk}(s, t)$.

We now apply Lemma 6 to the subtriangles. Assume that $s \cdot \Delta_f + t$ and Δ_g intersect along an edge e . If we fix s but vary t , the edge $e(t)$ moves parallel to itself in, say, Δ_g . For any subtriangle ρ , the intersection range $[\tau_1, \tau_2]$ is exactly the interval of t where $e(t)$ intersects ρ . By Lemma 6, the 16-approximation of the volume at ρ is exact when $t \leq \tau_1 - r$ or $t \geq \tau_2 + r$. This is true when the supporting line of $e(t)$ misses ρ by a sufficient distance, namely, exactly the same distance as the extent of ρ in the direction normal to $e(t)$. In Figure 8, this is the case for subtriangles that do not intersect the gray strip, which we denote Ψ . Note that if

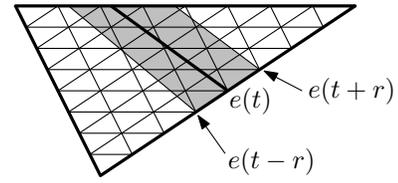


Figure 8. The subtriangles that do not intersect the gray strip Ψ give the exact volume in the 16-approximation.

we increase m , the subtriangles shrink, and so does Ψ . Also note that Ψ depends on s , t , m , and the particular triangles Δ_f and Δ_g .

We define the approximation $\tilde{\mu}_{ijk}(s, t) = \sum_{\rho \in \Xi} \tilde{\mu}_\rho(s, t)$. We now prove that $\tilde{\mu}_{ijk}$ is a piecewise-linear convex function and that $\tilde{\mu}_{ijk}(s, t) \leq (1 + \varepsilon)\mu_{ijk}(s, t)$. The convexity of $\tilde{\mu}_{ijk}$ follows directly from Lemma 7. It remains to prove the approximation factor.

The intuition behind the approximation factor is the following. Regardless of the value of s and t , $s \cdot \Delta_f + t$ and Δ_g intersect in at most one segment. The subtriangles that intersect the strip Ψ will give (at worst) a 16-approximation of the volume. All subtriangles that miss Ψ give the exact value of the volume. Since there are many more (by a factor of $\Omega(1/\varepsilon)$) subtriangles that give the exact volume, and the other ones stay within a constant factor, we can argue that our approximation $\tilde{\mu}_{ijk}(s, t)$ stays within a factor $1 + \varepsilon$ of the real $\mu_{ijk}(s, t)$, for the right choice of c .

BORIS SAYS: I have a mental block about this lemma. Can someone explain what says, and why is it true?

LEMMA 8. For any s, t , any subtriangle ρ that intersects $(s \cdot \Delta_f + t) \cap \Delta_g$ defines a volume that is at most twice as large as the volume of a subtriangle that misses $(s \cdot \Delta_f + t) \cap \Delta_g$.

PROOF. Follows from elementary geometry. \square

LEMMA 9. At most $6m+3$ subtriangles intersect the strip Ψ , and at least $m^2 - 4m - 3$ subtriangles miss it.

PROOF. A triangle is partitioned into $m^2 + 2m$ subtriangles. Any line intersects at most $2m + 1$ of them. All subtriangles meeting Ψ meet three lines parallel to $e(t)$. \square

Consider the optimal s and t , and consider two triangles $s \cdot \Delta_f + t$ and Δ_g that intersect. Let V be the largest volume that two intersecting subtriangles of $s \cdot \Delta_f + t$ and Δ_g can give. Then all subtriangles that miss Ψ give a total volume in between of at least $(m^2 - 4m - 3)V/2$ by Lemmas 8 and 9, and for them the “approximation” is the exact volume by Lemma 7. The total volume of the triangles intersecting Ψ that our approximation returns is at most $16(6m + 3)V$ by Lemmas 7 and 9. Hence the relative error is at most $16(6m + 3) \cdot V / ((m^2 - 4m - 3) \cdot V/2) = \Theta(1/m)$. Since $m = \lceil c/\varepsilon \rceil$ and $\varepsilon < \frac{1}{2}$, we can choose c large enough but still constant to obtain a relative error of at most ε .

A $(1 + \varepsilon)$ -approximation algorithm. We have all ingredients to give a $(1 + \varepsilon)$ -approximation algorithm for computing $\sigma(f, g)$. For each triangle $\Delta = \Delta_{v_i v_j v_k}$ in \mathbb{M} , we compute the piecewise-linear factor- $(1 + \varepsilon)$ approximation $\tilde{\mu}_{ijk}$ of μ_{ijk} . We set $\tilde{\mu} = \sum_{\Delta_{v_i v_j v_k} \in \mathbb{M}} \tilde{\mu}_{ijk}$ and observe that $\tilde{\mu}$ is convex as well. We compute the minimum value of $\tilde{\mu}$, using the same prune-and-search procedure as in the previous section. Since each $\tilde{\mu}_{ijk}$ is defined as the sum of $O(1/\varepsilon^2)$

piecewise-linear convex functions, it follows that the total time spent is $O(n/\varepsilon^2)$. Omitting the details, we conclude with the following result.

THEOREM 3. *Let f and g be two aligned piecewise-linear bivariate functions, defined by a triangulation composed of n triangles, and let $\varepsilon > 0$ be a parameter. We can compute in $O(n/\varepsilon^2)$ time a pair \tilde{s}, \tilde{t} and the value $\mu(\tilde{s}, \tilde{t})$ such that $\sigma(f, g) \leq \mu(\tilde{s}, \tilde{t}) \leq (1 + \varepsilon)\sigma(f, g)$.*

Remark. For unaligned functions, the n in the time bound is replaced by k , the overlay complexity, as described in the introduction. In the worst case, $k = \Theta(n^2)$, but in practice, one can expect it to be close to linear.

5. CONCLUSIONS

We studied the algorithmic aspects of determining the best linear model for two bivariate functions, modeled as polyhedral terrains. A number of open problems remain. It is possible that for the minimum area and volume measures, a strong LTAS exists, where the running time dependence on n and ε is in separate terms. It may also be possible to improve on the worst-case quadratic running time results for the volume and squared measures, in case the two terrains are not aligned. Finally, optimizing over more complex transformations than linear, or finding a linear model for more than two bivariate functions, leads to problems where there are more than two unknowns. These extensions are worth studying further.

6. REFERENCES

- [1] P. K. Agarwal and J. Erickson. Geometric range searching and its relatives. In B. Chazelle, J. E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*, pages 1–56. American Mathematical Society, Providence, RI, 1999.
- [2] H. Alt and L. J. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 121–153. Elsevier, 2000.
- [3] P. Burrough and R. McDonnell. *Principles of Geographical Information Systems*. Oxford University Press, 1998.
- [4] B. Chazelle. Cutting hyperplanes for divide-and-conquer. *Discrete & Computational Geometry*, 9:145–158, 1993.
- [5] B. Chazelle, H. Edelsbrunner, L. J. Guibas, and M. Sharir. Algorithms for bichromatic line segment problems and polyhedral terrains. *Algorithmica*, 11:116–132, 1994.
- [6] K. L. Clarkson. Las Vegas algorithms for linear and integer programming when the dimension is small. *Journal of the ACM*, 42:488–499, 1995.
- [7] M. de Berg, H. J. Haverkort, S. Thite, and L. Toma. I/O-efficient map overlay and point location in low-density subdivisions. In *Proc. 18th Annu. Int. Symp. on Algorithms and Computation*, volume 4835 of *Lecture Notes in Computer Science*, pages 500–511, 2007.
- [8] H. Edelsbrunner, L. J. Guibas, J. Hershberger, R. Seidel, M. Sharir, J. Snoeyink, and E. Welzl. Implicitly representing arrangements of lines or segments. *Discrete & Computational Geometry*, 4:433–466, 1989.
- [9] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci. Local and global comparison of continuous functions. In *IEEE Visualization*, pages 275–280, 2004.
- [10] U. Finke and K. Hinrichs. Overlaying simply connected planar subdivisions in linear time. In *Proc. 11th Annu. Symp. on Computational Geometry*, pages 119–126, 1995.
- [11] E. Isaaks and R. Srivastava. *Applied Geostatistics*. Oxford University Press, 1989.
- [12] H. Li and R. Tasay. A unified approach to identifying multivariate time series models. *Journal of the American Statistical Association*, 93:770–782, 1998.
- [13] L. O’Brian. *Introducing Quantitative Geography*. Routledge, 1992.
- [14] D. O’Sullivan and D. Unwin. *Geographic Information Analysis*. Wiley, 2003.
- [15] H. Wackernagel. *Multivariate Geostatistics: An Introduction with Applications*. Springer, 3rd edition, 2003.
- [16] W. Wei. *Time Series Analysis: Univariate and Multivariate Methods*. Addison-Wesley, 1990.