SIAM J. COMPUT.
Vol. 42, No. 5, pp. 1867–1887

COMPUTING CORRELATION BETWEEN PIECEWISE-LINEAR FUNCTIONS*

PANKAJ K. AGARWAL†, BORIS ARONOV‡, MARC VAN KREVELD§, MAARTEN LÖFFLER§, AND RODRIGO I. SILVEIRA¶

Abstract. We study the problem of computing correlation between two piecewise-linear bivariate functions defined over a common domain, where the surfaces they define in three dimensions—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 of one of the terrains, in $O(n^{4/3}\operatorname{polylog} n)$ expected time, where n is the total number of vertices in the graphs of the two functions. We also present approximation algorithms for minimizing the mean distance between the graphs of univariate and bivariate functions. For univariate functions we present a $(1+\varepsilon)$-approximation algorithm that runs in $O(n(1+\log^2(1/\varepsilon)))$ expected time for any fixed $\varepsilon > 0$. The $(1+\varepsilon)$-approximation algorithm for bivariate functions runs in $O(n/\varepsilon)$ time, for any fixed $\varepsilon > 0$, provided the two functions are defined over the same triangulation of their domain.

Key words. piecewise-linear function, polyhedral terrain, similarity, approximation algorithm, correlation

AMS subject classifications. 68U05, 68Q25, 68W25

DOI. 10.1137/120900708

1. Introduction.

Many types of spatial data can be modeled mathematically as a bivariate function $f\colon \mathbb{D} \to \mathbb{R}$, where $\mathbb{D}$ is a (planar) polygonal 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 are 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, 13, 21]. 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 [9]. See [11] for a Morse theory–based method for computing the similarity of two functions.

*Received by the editors November 29, 2012; accepted for publication (in revised form) July 15, 2013; published electronically September 26, 2013. A preliminary version of this paper appeared in *Proceedings of the 26th ACM Symposium on Computational Geometry* [1].

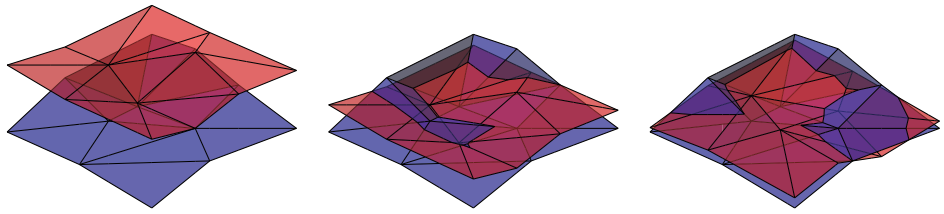
http://www.siam.org/journals/sicomp/42-5/90070.html

†Department of Computer Science, Duke University, Durham, NC 27708 (pankaj@cs.duke.edu). This author's research was supported by NSF grants CCF-09-40671, CCF-10-12254, and CCF-11-61359, by ARO grants W911NF-07-1-0376 and W911NF-08-1-0452, and by ERDC contract W9132V-11-C-0003.

‡Department of Computer Science and Engineering, Polytechnic Institute of New York University, New York, NY 11201 (aronov@poly.edu). This author's research was supported by a grant from the U.S.-Israel Binational Science Foundation, by NSA MSP grant H98230-06-1-0016, and NSF grants CCF-08-30691, CCF-11-17336, and CCF-12-18791.

§Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands (m.j.vankreveld@uu.nl, m.loffler@uu.nl). The research of these authors was supported by the Netherlands Organisation for Scientific Research (NWO) under the Free Competition project GOGO, under FOCUS/BRICKS grant 642.065.503, and under grant 639.021.123. The fourth author was further supported by the Office of Naval Research under grant N00014-08-1-1015.

¶Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (rodrigo.silveira@upc.edu). This author's research was supported by the FP7 Marie Curie Actions Individual Fellowship PIEF-GA-2009-251235, and partially supported by the ESF EUROCORES programme EuroGIGA -ComPoSe IP04-MICINN project EUI-EURC-2011-4306.



(a) (b) (c)

FIG. 1.1. *Two triangulated terrains* (a). *The top (red) terrain has been translated* (b) *and scaled* (c) *vertically to match the bottom (blue) terrain. (See electronic version of article for color figures.)*

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., if the 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 [17, 18]. 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., the data is *heterotopic* [21], 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 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.

Contribution. 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 $f\colon \mathbb{D} \to \mathbb{R}$. The graph[1] 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 1.1 shows an example of two terrains representing the functions f and g, and how one can be transformed to match the other better.

This paper discusses the problem of computing the scaling and translation parameters that minimize the vertical distance between 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 $f, g\colon \mathbb{D} \to \mathbb{R}$ be two piecewise-linear functions

[1]Slightly abusing the terminology, we will not distinguish between a function and its graph, whenever it does not cause confusion.

whose linear pieces correspond to the triangles of $\mathbb{M}_f$ and $\mathbb{M}_g$, respectively. We assume that $\mathbb{M}_f$ and $\mathbb{M}_g$ together have n vertices. We consider two different ways to measure correlation between f and g: (i) minimizing the maximum vertical distance (L_∞-norm) between their graphs, and (ii) minimizing the mean distance between their graphs (L_1-norm). More precisely, if s is the scaling parameter and t the translation parameter, we define functions $\mu_p\colon \mathbb{R}^2 \to \mathbb{R}$, $p = 1, \infty$, to measure how well f and g match for a pair s and t in (1.1)–(1.2), as follows:

$$\mu_\infty(s,t) = \max_{(x,y)\in\mathbb{D}} |sf(x,y) + t - g(x,y)|, \tag{1.1}$$

$$\mu_1(s,t) = \frac{1}{\operatorname{area}\mathbb{D}} \int_{(x,y)\in\mathbb{D}} |sf(x,y) + t - g(x,y)|\, dx\, dy. \tag{1.2}$$

We can now define the correlation between f and g as $\sigma_p(f,g) = \min_{s,t} \mu_p(s,t)$ for $p = 1, \infty$. For each p, our goal is to compute the pair (s^*, t^*) that minimizes the function μ_p and to compute this minimum value $\sigma_p(f,g) = \mu_p(s^*, t^*)$. One can also consider $p = 2$, where the goal is to minimize the root-mean-square vertical distance between the graphs of f and g; this problem was recently studied by Moroz and Aronov, who presented a near-linear algorithm for computing $\sigma_2(f,g)$ [16].

Before proceeding, we note that the functions $\mu_\infty(s,t)$ and $\mu_1(s,t)$ are convex: The expression $|sf(x,y)+t-g(x,y)|$, for fixed x and y, is a convex function of s and t. Therefore the pointwise maximum and the integral of a family of such expressions, respectively, over all suitable pairs (x,y) are also convex. In fact, a variant of this argument applies for all $p \geq 1$.

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. Using linear programming 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 [12]. 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 [10]. In this case one can align a pair of unaligned functions without loss of asymptotic efficiency.

In section 2 we develop algorithms for computing $\sigma_\infty(f,g)$. Using linear programming we can compute the optimal linear transformation in linear time for aligned functions. 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 time-series data [14, 22]. Computing σ_1 even for univariate functions requires minimizing a linear-size sum of rational functions, which we cannot hope to do exactly [19]. We show that a randomized $(1+\varepsilon)$-approximation algorithm exists that takes $O(n(1+\log^2(1/\varepsilon)))$ expected time for univariate functions (aligned or unaligned), and $O(n/\varepsilon)$ time for bivariate functions in the aligned case. The latter extends immediately to $O((n+k)/\varepsilon^2)$ time for the nonaligned case.

The main novelty of this paper lies in the algorithmic approach taken to determine correlation between 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. We prove several geometric properties of the correlation functions, which are the basis of the algorithms that we present.

2. Minimizing the maximum distance. Let $f, g\colon \mathbb{D} \to \mathbb{R}$ be two piecewise-linear bivariate functions as defined above. In this section the goal is to find s^*, t^* that minimize $\mu(s,t)$ according to (1.1), 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 $\mathbb{M}_f$, vertices in $\mathbb{M}_g$, and intersection points of the edges in $\mathbb{M}_f$ and $\mathbb{M}_g$. 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 $\sigma_\infty(f,g)$ is inherently discrete. For a point $v \in \mathbb{X}$, we define a bivariate function $d_v\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$d_v(s,t) = |sf(v) + t - g(v)| = \max\{sf(v) + t - g(v),\ g(v) - sf(v) - t\}, \tag{2.1}$$

which is the vertical distance between f and g (for a given s and t) at v. Set $\mathcal{D} = \{d_v(s,t) \mid v \in \mathbb{X}\}$. The above observation implies that $\mu(s,t) = \max_{v\in\mathbb{X}} d_v(s,t)$, i.e., μ is the upper envelope of $\mathcal{D}$ [20]. Hence, the problem of computing s^*, t^* reduces to computing the minimum value on the upper envelope of $\mathcal{D}$.

For each $v \in \mathbb{X}$, we define two halfspaces (in the stz-space): $\gamma_v^+\colon z \geq sf(v) + t - g(v)$ and $\gamma_v^-\colon z \geq g(v) - sf(v) - t$. Then, by (2.1), 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 the 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 [8]; a more involved linear-time deterministic algorithm also exists [7].

As mentioned above, if f and g are aligned or if 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. Our method is similar to the LP algorithm in [8], and is sketched in Algorithm MinMaxDistance (Figure 2.1). For a subset $Y \subseteq \mathbb{X}$, let Optimal(Y) denote the optimal solution for Y, i.e., the lowest vertex in the three-dimensional convex polyhedron $\bigcap_{v\in Y} \gamma_v^+ \cap \gamma_v^-$. As mentioned above, this can be computed in $O(|Y|)$ expected time.

Since $\mathbb{X}$ is not explicitly computed, it is not straightforward to choose R and compute V, the set of *violated* vertices in $\mathbb{X}$. 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 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 three dimensions that are vertically more than z_0 apart. Then we solve that problem by mapping

ALGORITHM 1. MINMAXDISTANCE

```
1.  V* = ∅
2.  repeat
3.     R ← random subset of 𝕏 of min{|𝕏|, 9n} points
4.     ξ0 ← OPTIMAL(R ∪ V*)
5.     V ← {v ∈ 𝕏 | ξ0 ∉ γv+ ∩ γv−}
6.     if 1 ≤ |V| ≤ 2n then
7.        V* ← V* ∪ V
8.  until V = ∅
9.  return ξ0
```

FIG. 2.1. *The algorithm for computing* $\mu_\infty(s,t)$.

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 $|V| > 2n$; recall that the O^*-notation omits polylogarithmic factors. A more detailed description of the algorithm follows.

For the first part, we use a two-level hereditary segment tree T [5]. Let A 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 $A \times B$. That is, we compute a family $\mathcal{F} = \{(A_1, B_1), \ldots, (A_u, B_u)\}$, where $A_i \subseteq A$, $B_i \subseteq B$, and $u = O(n \log n)$, such that

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

$\mathcal{F}$ can be computed in $O(n \log^2 n)$ time. For each $(A_i, B_i) \in \mathcal{F}$, we "lift" every line segment $a \in A_i$ (resp., $b \in B_i$) to a line in $\mathbb{R}^3$, namely, the line containing the segment $f_0(a)$ (resp., $g(b)$). Let $\hat{A}_i$ and $\hat{B}_i$ be the sets of resulting lines in $\mathbb{R}^3$. We wish to report the pairs of lines in $\hat{A}_i \times \hat{B}_i$ 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. For a line $\ell \in \hat{A}_i$, let $\ell^+ = \ell + (0, 0, z_0)$ and $\ell^- = \ell + (0, 0, -z_0)$—translates of ℓ in the z-direction at distance z_0. Set $L_i^+ = \{\ell^+ \mid \ell \in \hat{A}_i\}$ and $L_i^- = \{\ell^- \mid \ell \in \hat{A}_i\}$. A line $\ell_2 \in \hat{B}_i$ is more than z_0 apart from $\ell_1 \in \hat{A}_i$ if ℓ_2 lies above ℓ_1^+ or below ℓ_1^-. We thus wish to report all pairs ℓ, ℓ' in $L_i^+ \times \hat{B}_i$ (resp., $L_i^- \times \hat{B}_i$) such that ℓ' lies above (resp., below) ℓ.

We now use Plücker points and hyperplanes to rephrase the question in terms of range reporting. Informally, Plücker coordinates are a way to represent oriented lines as either points or hyperplanes in oriented projective 5-space. The key property of the transformation is that a line ℓ_1 intersecting another line ℓ_2 is transformed into the Plücker point $p(\ell_1)$ of ℓ_1 lying on the Plücker hyperplane $\pi(\ell_2)$ of ℓ_2; moreover, ℓ_1 has positive orientation with respect to ℓ_2 if and only if $p(\ell_1)$ lies in the positive halfspace bounded by $\pi(\ell_2)$. We refer the reader to [6] for the technical details of the transformation.

The property (ii) of (A_i, B_i) ensures that Plücker points of all lines in $\hat{B}_i$ that lie above a line $\ell \in L_i^+$ lie in one of the halfspaces bounded by the Plücker hyperplane of ℓ [6]. Hence, the problem of reporting all pairs of $\ell, \ell' \in L_i^+ \times \hat{B}_i$ such that ℓ' lies

above ℓ reduces to halfspace range reporting in $\mathbb{R}^5$. Using a partition tree–based data structures, all such κ^+ pairs can be reported in $O^*(n_i^{4/3} + \kappa^+)$ time; see, e.g., [2]. Similarly, we can report, in $O^*(n_i^{4/3} + \kappa^-)$ time, all κ^- pairs $\ell, \ell' \in L_i^- \times \hat{B}_i$ such that ℓ' lies below ℓ. Hence, we can report all κ pairs of lines of $\hat{A}_i \times \hat{B}_i$ that are more than z_0 apart in $O^*(n_i^{4/3} + \kappa)$ time. Furthermore, if we wish to report at most κ' such pairs, the algorithm takes $O^*(n_i^{4/3} + \min\{\kappa, \kappa'\})$ time. Repeating this procedure for all bipartite cliques in $\mathcal{F}$, in $O^*(n^{4/3})$ time we either report that more than $2n$ vertices of $\mathbb{X}$ are violated or report all violated pairs of $\mathbb{X}$ (if there are at most $2n$ such points). Thus the total time spent in line 5 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 follows. We first choose a random bipartite clique $(A_i, B_i) \in \mathcal{F}$; the probability of (A_i, B_i) being chosen is $|A_i| \cdot |B_i|/k$, where $k = \sum_j |A_j| \cdot |B_j|$ is the total number of intersecting pairs in $A \times B$. Next, we choose a random pair $(a, b) \in A_i \times B_i$, each pair being chosen with equal probability, and the desired point is $a \cap b$. After having computed $\mathcal{F}$, it takes $O(\log n)$ time to choose a random point. Hence, we can choose R in $O(n \log n)$ time.

Number of iterations. It can be shown that line 7 of Algorithm MinMaxDistance (i.e., $V^* \leftarrow V^* \cup V$) will be executed at most three times using the argument in [8]. Indeed, if $\xi_0 \neq \text{Optimal}(\mathbb{X})$, then $V \neq \emptyset$ and a point of $\mathbb{X}$ corresponding to one of the three halfspaces that define $\text{Optimal}(\mathbb{X})$ belongs to V and is not currently in $V^* \cup R$. It is added to V^* in line 7, and from then on it remains in V^*. Therefore after executing line 7 at most three times, V^* will contain the three vertices of $\mathbb{X}$ that define the optimum, and hence there will not be any violated constraints in the next iteration. Following the same argument as in [8], based on random sampling, we can argue that $|V| > 2n$ with a probability at most $\frac{1}{2}$. The expected number of iterations of the loop is therefore $O(1)$. Hence, we can conclude as follows.

Theorem 2.1. *Given two bivariate piecewise-linear functions over a common domain, a linear transformation minimizing the maximum vertical 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 mean distance: Univariate functions. In this section, we assume $\mathbb{D} \subseteq \mathbb{R}$ to be a bounded interval and $f, g \colon \mathbb{D} \to \mathbb{R}$ 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 are defined by a common subdivision $\mathbb{M}$ of $\mathbb{D}$. For any value of s and t, let

$$\mu(s,t) = \int_{x \in \mathbb{D}} |sf(x) + t - g(x)| \, dx \tag{3.1}$$

be the area between the graphs of $sf + t$ and g over $\mathbb{D}$. In order to minimize the mean vertical distance between f and g, it suffices 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 $x_1 < \cdots < x_n$ 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 3.1. Let f_i and g_i be the functions f and g, restricted to the interval $[x_i, x_{i+1}]$, and let $\mu_i(s,t)$ be the area between the graphs of $sf_i + t$ and g_i in the interval $[x_i, x_{i+1}]$. Then we have

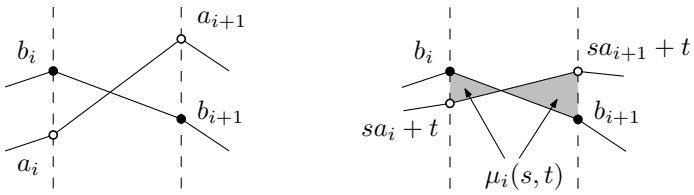


FIG. 3.1. *Edges of f and g, and the function $\mu_i(s,t)$.*

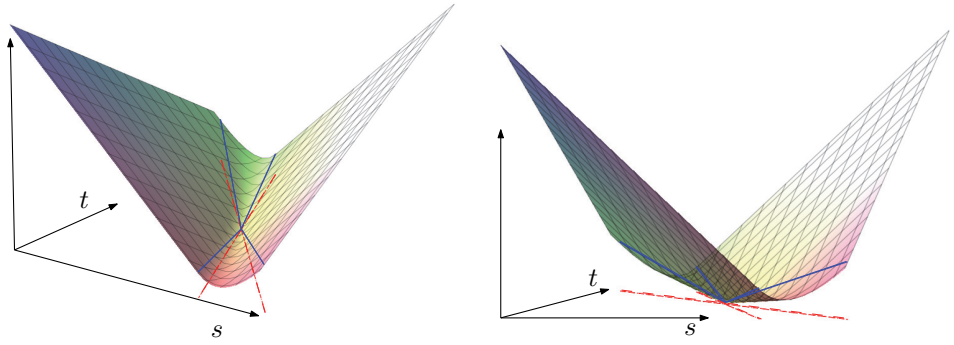


FIG. 3.2. (a) *Graph of $\mu_i(s,t)$; red (dashed) lines show ℓ_i and ℓ_{i+1}, and blue (solid) lines show the rays where the different function pieces meet.* (b) *Another view of $\mu_i(s,t)$.*

$$\mu(s,t) = \sum_{i=1}^{n-1} \mu_i(s,t).$$

Unless f_i is constant,[2] the region in the st-plane where sf_i+t and g_i intersect is a double wedge whose bounding lines are $\ell_i\colon t = -a_i s + b_i$ and $\ell_{i+1}\colon t = -a_{i+1}s + b_{i+1}$. The lines ℓ_i and ℓ_{i+1} correspond to the s,t pairs at which the functions g_i and sf_i+t intersect above x_i and x_{i+1}, respectively, and these lines intersect at the apex $(\bar{s}_i, \bar{t}_i)$ of the double wedge, where $\bar{s}_i f_i + \bar{t}_i = g_i$ and $\mu_i(\bar{s}_i, \bar{t}_i) = 0$. Slightly abusing the notation, we also view ℓ_i as the graph of the linear function $\ell_i(s) = b_i - a_i s$ and regard ℓ_{i+1} similarly. The following lemma describes the form of μ_i illustrated in Figures 3.1 and 3.2.

LEMMA 3.1. *The function μ_i is linear in s and t above ℓ_i and ℓ_{i+1} (in the direction of t), and below ℓ_i and ℓ_{i+1}. For points lying below ℓ_i and above ℓ_{i+1}, in the situation depicted in Figure* 3.1, *where $t < \ell_i(s)$ and $t > \ell_{i+1}(s)$, we have*

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

A symmetric expression holds for points where $t > \ell_i(s)$ and $t < \ell_{i+1}(s)$.

Figure 3.2 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 variable s in the denominator. It will be useful to analyze the function μ_i closer.

[2] If f_i is a constant, the expressions that follow simplify.

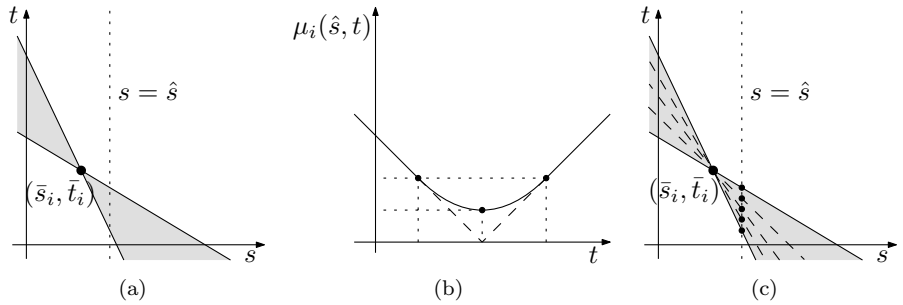


FIG. 3.3. (a) *Double wedge arising from f_i and g_i in the st-plane.* (b) *Cross-section in the plane $s = \hat{s}$ showing $\mu_i(\hat{s}, t)$.* (c) *Subdivision into double wedges in the st-plane.*

For a fixed scaling factor $\hat{s}$, consider $\mu_i(\hat{s}, t)$ as a function of t. By Lemma 3.1, $\mu_i(\hat{s}, t)$ is linear in t below ℓ_i and ℓ_{i+1}, or above both of them, and a convex 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 3.3(b). Despite its complicated form, we can prove the following for μ_i.

LEMMA 3.2. *For every $i = 1, \ldots, n-1$, 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. The convexity of μ_i follows from the same arguments as the convexity of μ, as explained in the introduction. As for the rest, we assume, without loss of generality, that $x_{i+1} - x_i = 1$. The lines ℓ_i and ℓ_{i+1} partition the st-plane into four wedges. We assume that we are in the situation depicted in Figure 3.1, where $sa_i + t < b_i$ and $sa_{i+1} + t > b_{i+1}$; the other situation is symmetric. 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 linear along any ray emanating from $(\bar{s}_i, \bar{t}_i)$ within them. Next, we consider the two other wedges. Consider the wedge W^+ lying below ℓ_i and above ℓ_{i+1}; the other case is symmetric. We can then write $\mu_i(s, t)$ for W^+ (see Lemma 3.1) as

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

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^+, $t' = \alpha s'$, so $\mu_i(s', \alpha s') = (\alpha^2 + 1/4)s'$ is linear. Hence the graph of $\mu(s', t')$ restricted to W^+ is a cone with apex at the origin. A similar argument holds in the diametrically opposite wedge W^-, completing the proof of the lemma. ∎

Approximating μ. Since it seems difficult to compute the minimum value of $\mu(s, t)$ exactly, we define a convex piecewise-linear function $\tilde{\mu}(s, t)$, for a given $\varepsilon > 0$, such that $\mu(s, t) \le \tilde{\mu}(s, t) \le (1 + \varepsilon)\mu(s, t)$ 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 there and focus on approximating the part of $\mu_i(s, t)$

inside the double wedge. We first fix a value $\hat{s}$ and approximate the univariate function $\mu(\hat{s}, t)$. 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 3.3(c). We extend this approximation to all values of s by choosing lines through the new points and $(\bar{s}_i, \bar{t}_i)$, giving a partition of the plane into $\Theta(1/\sqrt{\varepsilon})$ wedges over each of which $\tilde{\mu}_i$ is linear; note that the graph of $\tilde{\mu}_i$ is once again a convex cone with its apex at the origin. Lemma 3.2 and the discussion above ensure that $\mu_i(s,t) \le \tilde{\mu}_i(s,t) \le (1+\varepsilon)\mu_i(s,t)$ everywhere. We set $\tilde{\mu}(s,t) := \sum_i \tilde{\mu}_i(s,t)$.

LEMMA 3.3. *The function $\tilde{\mu}$ is convex, and*

$$\mu(s,t) \le \tilde{\mu}(s,t) \le (1+\varepsilon)\mu(s,t).$$

Approximation algorithm. We now describe the algorithm for computing the point $\pi^* = (\tilde{s}, \tilde{t})$ that minimizes $\tilde{\mu}$. For each $i = 1, \ldots, n-1$, let L_i denote the set of lines, including ℓ_i, ℓ_{i+1}, that define $\tilde{\mu}_i$; $|L_i| = O(1/\sqrt{\varepsilon})$. We assume that the lines in L_i are ordered from ℓ_i to ℓ_{i+1}. Since the lines in L_i are regularly distributed, the wedge of the arrangement $\mathcal{A}(L_i)$ containing a given point $x \in \mathbb{R}^2$ can be determined in $O(1)$ time, assuming that the $\lfloor \cdot \rfloor$ operation takes $O(1)$ time. Set $L = \bigcup_{i=1}^{n-1} L_i$ and $|L| = m$. Note that the lines in L are not in general position, as all lines within each L_i pass through the point $\ell_i \cap \ell_{i+1}$. For simplicity, assume that these are the only high-degree vertices of the arrangement $\mathcal{A}(L)$; every other vertex of $\mathcal{A}(L)$ is incident on two lines of L.

The function $\tilde{\mu}$ is linear within each cell of $\mathcal{A}(L)$, and therefore the minimum of $\tilde{\mu}$ is achieved at a vertex of $\mathcal{A}(L)$. By traversing $\mathcal{A}(L)$ and computing the value of $\tilde{\mu}$ at each vertex, π^* can be computed in $O(n^2/\varepsilon)$ time. We next describe a more efficient procedure by exploiting the convexity of $\tilde{\mu}$ and using a prune-and-search approach.

Let γ be a zero-, one-, or two-dimensional simplex (point, segment, or triangle) in the *st*-plane; we treat them as closed. We use $L_{|\gamma} \subseteq L$ to denote the set of lines that *cross* the interior of γ, i.e., they intersect the interior of γ but do not contain γ; set $|L_{|\gamma}| = m_\gamma$. (If γ is a point, then $L_{|\gamma} = \emptyset$.) Next, we use $D_\gamma \subseteq [1 : n-1]$ (where $[a : b]$ stands for $\{a, a+1, \ldots, b\}$) to denote the set of indices j for which a line of L_j intersects γ but does not cross its interior. By the general position assumption above, $|D_\gamma| \le 6$; notice that there might be many such lines; we only claim that they come from at most six families L_j. Let

$$\pi^*_\gamma = \arg\min_{x \in \gamma} \tilde{\mu}(x).$$

In order to compute $\tilde{\mu}(x)$ efficiently for any point $x \in \gamma$, we represent $\tilde{\mu}$ compactly, as follows. We observe that if γ is not crossed by any line of L_i, then $\tilde{\mu}_i$ restricted to γ is a linear function, which we denote by $h_{i,\gamma}$. We represent $\tilde{\mu}$ as a pair $\Psi_\gamma = (J_\gamma, h_\gamma)$, where $J_\gamma \subseteq [1 : n-1]$ is the set of indices j for which at least one of the lines of L_j crosses the interior of γ, i.e., $J_\gamma = \{1 \le j \le n-1 \mid L_j \cap L_{|\gamma} \ne \emptyset\}$, and $h_\gamma = \sum_{j \notin J_\gamma} h_{j,\gamma}$. Put $|J_\gamma| = n_\gamma$. For any point $x \in \gamma$,

$$\tilde{\mu}(x) = \sum_{j \in J_\gamma} \tilde{\mu}_j(x) + h_\gamma(x). \tag{3.2}$$

Let γ be a segment or a triangle. Given Ψ_γ and D_γ, each of the following operations can be performed in $O(n_\gamma)$ time. A low-level description of some of these operations is somewhat tedious, but they are nevertheless straightforward and the details are omitted.

(A1) For any $x \in \gamma$, we can compute $\tilde{\mu}(x)$ using (3.2). Furthermore, if γ is a segment, we can also determine whether $x = \pi^*_\gamma$, and, if the answer is NO, whether π^*_γ lies to the left of x or to its right.

(A2) Let γ be a segment. Given π^*_γ, we can determine whether $\pi^* = \pi^*_\gamma$, and, if the answer is NO, whether π^* lies in γ^- or γ^+, where γ^- (resp., γ^+) is the halfplane lying below (resp., above) the line containing γ.

(A3) We can compute the value of m_γ, the number of lines of L crossing the interior of γ. Indeed, for every $i \in J_\gamma$, by locating the vertices of γ in $\mathcal{A}(L_i)$, we can compute $m_{i,e} = |L_{|\gamma} \cap L_i|$ in $O(1)$ time. By repeating this for all $i \in J_\gamma$, the value of m_γ can be computed in $O(n_\gamma)$ time.

(A4) We can choose a random line from $L_{|\gamma}$. First, we compute $m_{i,\gamma}$ for every $i \in J_\gamma$, as in (A3). Next, we choose an index $j \in J_\gamma$ randomly, with the probability of j being chosen equal to $m_{j,\gamma}/m_j$. Then we choose a random line of $L_j \cap L_{|\gamma}$; each line has equal probability of being chosen. Since $L_j \cap L_{|\gamma}$ is composed of at most two contiguous subsequences of L_j, this step can be done in $O(1)$ time.

We now describe two procedures. The first one computes π^*_e for a segment e, and the second procedure computes $\pi^*_\triangle$ for a triangle $\triangle$.

Computing the minimum along a segment. Let e be a segment in the plane. Assuming that Ψ_e and D_e are given, we describe how to compute π^*_e. We first check in $O(n_e)$ time whether π^*_e is one of its endpoints. If so, we return that endpoint and stop. So assume that π^*_e lies in the interior of e. We describe a recursive procedure for computing π^*_e.

Fix a sufficiently large integer constant $r > 1$. If $m_e \le r$, then we set $R = L_{|e}$. Otherwise (i.e., $m_e > r$), $R \subseteq L_{|e}$ is a subset of r randomly and independently chosen lines (if the same line is picked more than once, we keep only one copy). Using (A4), R can be computed in $O(n_e)$ time.

For each line $\ell \in R$, we compute $\pi = e \cap \ell$ and $\tilde{\mu}(e \cap \ell)$, and we determine whether $\pi^*_e = \pi$ using (A1). If the answer is YES, then we return π and stop; otherwise we also know whether π^*_e lies to the left or right of π. If the algorithm has not stopped in this step, i.e., $\pi^*_e \neq e \cap \ell$ for any $\ell \in R$, then we find the two consecutive intersection points π_1, π_2 of R with e such that π^*_e lies in the interior of the segment $\hat{e} = \pi_1\pi_2 \subset e$. We compute $\Psi_{\hat{e}}$ from Ψ_e: we begin by setting $J_{\hat{e}} = \emptyset$ and $h_{\hat{e}} = h_e$. Next, for each $j \in J_e$, we check whether $\hat{e}$ intersects any line of L_j. If yes, then we add j to $J_{\hat{e}}$; otherwise we compute $h_{j,\hat{e}}$ in $O(1)$ time and set $h_{\hat{e}} = h_{\hat{e}} + h_{j,\hat{e}}$. We also compute $D_{\hat{e}}$ from D_e. Using (A3), we compute the value of $m_{\hat{e}}$. If $|R| < r/2$ or $m_{\hat{e}} > m_e/4$, then we choose another sample R and repeat the above steps; otherwise we recursively solve the problem for $\hat{e}$.

Putting everything together, the running time of each iteration is $O(n_e)$. Notice that, in the recursive case, $|R| \ge r/2$ with probability at least $\frac{1}{2}$ and, if r is chosen sufficiently large, then a random sampling argument [15] implies that the probability that $m_{\hat{e}} > m_e/4$ is at most $\frac{1}{4}$. In particular, the probability of restarting the procedure is at most $\frac{3}{4}$, so the number of restarts required to reach the next level of recursion is expected to be constant. Let $\varphi(n_e, m_e)$ be the expected running time of the overall algorithm with $|J_e| = n_e$ and $|L_{|e}| = m_e$.

We obtain the following recurrence for $\varphi(\cdot, \cdot)$:

$$\varphi(n_e, m_e) \le \begin{cases} \varphi(n_{\hat{e}}, m_{\hat{e}}) + c_1 n_e & \text{if } m_e > r, \\ c_2 & \text{if } m_e \le r, \end{cases}$$

where $c_1, c_2 > 0$ are constants, $m_{\hat{e}} \le m_e/4$, and $n_{\hat{e}} \le \min\{n_e, m_{\hat{e}}\}$. We claim that the solution to the above recurrence is

$$\varphi(n_e, m_e) \le An_e \left(1 + \log_2 \frac{m_e}{n_e}\right),$$

where $A \ge \max\{2c_1, c_2\}$ is a constant. The claim is obviously true for $m_e \le r$, so assume that $m_e > r$. By the induction hypothesis,

$$\varphi(n_e, m_e) \le An_{\hat{e}} \left(1 + \log_2 \frac{m_{\hat{e}}}{n_{\hat{e}}}\right) + c_1 n_e \le An_{\hat{e}} \left(1 + \log_2 \frac{m_e}{4n_{\hat{e}}}\right) + c_1 n_e.$$

First, assume that $n_{\hat{e}} \ge n_e/2$. Then

$$\varphi(n_e, m_e) \le An_e \left(1 + \log_2 \frac{m_e}{2n_e}\right) + c_1 n_e \le An_e \left(1 + \log_2 \frac{m_e}{n_e}\right),$$

because $A \ge 2c_1$. Next, assume $n_{\hat{e}} < n_e/2$. Then we obtain

$$\begin{aligned}
\varphi(n_e, m_e) &\le An_{\hat{e}} \left(1 + \log_2 \frac{m_e}{4n_{\hat{e}}}\right) + c_1 n_e \\
&= A\frac{n_e}{2} \left(1 + \log_2 \frac{m_e}{n_e}\right) - An_{\hat{e}} \left(\frac{n_e}{2n_{\hat{e}}} \left(1 + \log_2 \frac{m_e}{n_e}\right) - 1 - \log_2 \frac{m_e}{4n_{\hat{e}}}\right) + c_1 n_e \\
&\le An_e \left(1 + \log_2 \frac{m_e}{n_e}\right) - An_{\hat{e}} \left(\left(\frac{n_e}{2n_{\hat{e}}} - 1\right) \left(1 + \log_2 \frac{m_e}{n_e}\right) - \log_2 \frac{n_e}{4n_{\hat{e}}}\right) \\
&\le An_e \left(1 + \log_2 \frac{m_e}{n_e}\right) - An_{\hat{e}} \left(\frac{n_e}{2n_{\hat{e}}} - \log_2 \frac{n_e}{2n_{\hat{e}}}\right) \\
&\qquad \text{(as } n_e \ge 2n_{\hat{e}} \text{ and } \log_2(m_e/n_e) \ge 0\text{)} \\
&\le An_e \left(1 + \log_2 \frac{m_e}{n_e}\right).
\end{aligned}$$

The last inequality follows because $x - \log_2 x \ge 1$ for $x \ge 1$. Hence, the expected running time of the algorithm is $O(n_e(1 + \log(m_e/n_e)))$.

Computing the minimum in a triangle. Let $\triangle$ be a triangle in the st-plane. Assuming that we are given $\Psi_\triangle$ and $D_\triangle$, we describe how to compute $\pi^*_\triangle$. We assume that $\pi^* = \pi^*_\triangle$ and that it lies in the interior of $\triangle$, because otherwise $\pi^*_\triangle$ lies on the boundary of $\triangle$ and can be computed using the previous procedure (see below how we detect whether π^* lies in the interior of $\triangle$). We describe a recursive procedure, an extension of the one described above, for computing $\pi^*_\triangle = \pi^*$.

Fix a sufficiently large integer constant $r > 1$. If $m_\triangle \le r$, then we set $R = L_{|\triangle}$; otherwise $R \subseteq L_{|\triangle}$ is a subset of r randomly and independently chosen lines, with duplicates removed. By (A4), R can be computed in $O(n_\triangle)$ time. We construct $\mathcal{A}(R)$, the arrangement of R clipped within $\triangle$, and its bottom-vertex triangulation $\mathcal{A}^\nabla(R)$; see [15] for the definition of bottom-vertex triangulation. For each edge $e \in \mathcal{A}^\nabla(R)$, we first compute Ψ_e from $\Psi_\triangle$ in $O(n_\triangle)$ time and then compute π^*_e in $O(n_e(1 + \log(m_e/n_e)))$ expected time. We also compute $D_{\pi^*_e}$. Using (A2), we now determine whether $\pi^*_e = \pi^*$. If the answer is YES, we return π^* and stop. Otherwise (A2) also determines on which side of (the line supporting) e the point π^* lies.

Suppose $\pi^*_e \ne \pi^*$ for any edge e of $\mathcal{A}^\nabla(R)$. Since $\tilde{\mu}$ is convex, π^* lies in the interior of a unique triangle of $\mathcal{A}^\nabla(R)$. Fix a triangle $\tau \in \mathcal{A}^\nabla(R)$, and let e_1, e_2, e_3 be

the edges of τ. Then π^* lies in the interior of τ if and only if π^* lies on the same side of e_i as τ for every $i = 1, 2, 3$. For each edge $e \in \mathcal{A}^\nabla(R)$ we know whether π^* lies in e^- or e^+, and therefore we identify in $O(1)$ time the triangle $\tau \in \mathcal{A}^\nabla(R)$ that contains π^*. We first compute D_τ, Ψ_τ from $D_\triangle, \Psi_\triangle$ and then m_τ, using (A3), in $O(n_\tau)$ time. If $m_\tau > m_\triangle/4$ or $|R| < r/2$, we choose another sample R and repeat the above steps. Otherwise, we recursively search in τ.

Again, $|R| > r/2$ with probability at least $\frac{1}{2}$, and if r is chosen sufficiently large, then the theory of random sampling [15, Chapter 4] implies that $m_\tau \le m_\triangle/4$ with probability at least $\frac{3}{4}$; so we expect a constant number of retrials. Let $\psi(n_\triangle, m_\triangle)$ be the expected running time of the above algorithm. Then

$$\psi(n_\triangle, m_\triangle) \le \begin{cases} \psi(n_\tau, m_\tau) + c_3 n_\triangle \left(1 + \log_2 \frac{m_\triangle}{n_\triangle}\right) & \text{if } m_\triangle > r, \\ c_4 & \text{if } m_\triangle \le r, \end{cases}$$

where c_3, c_4 are constants, $n_\tau \le \min\{n_\triangle, m_\tau\}$, and $m_\tau \le m_\triangle/4$. Following the same reasoning as above, the solution to the above recurrence can be shown to be

$$\psi(n_\triangle, m_\triangle) = O(n_\triangle(1 + \log_2^2(m_\triangle/n_\triangle))).$$

Finally, to compute π^*, we call this procedure with $\triangle = \mathbb{R}^2$, and we set $J_\triangle = [1 : n-1]$, $D_\triangle = \emptyset$, and $h_\triangle = 0$. Since $|L| = O(n/\sqrt{\varepsilon})$, the expected running time is $O(n(1 + \log^2(1/\varepsilon)))$.

THEOREM 3.4. *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 $O(n(1 + \log^2(1/\varepsilon)))$ expected time such that $\sigma_1(f, g) \le \mu(\tilde{s}, \tilde{t}) \le (1 + \varepsilon)\sigma_1(f, g)$.*

4. Minimizing the mean distance: 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 $\mathbb{D}$. Extending the definition from the last section, for any values of s and t, let $\mu(s, t)$ denote the volume between the graphs of $sf + t$ and g over $\mathbb{D}$, that is,

$$\mu(s, t) = \int_{(x,y)\in\mathbb{D}} |sf(x, y) + t - g(x, y)|\, dx\, dy. \tag{4.1}$$

The main difficulty in generalizing the previous approach is that the analogue of function μ_i, though also 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 the boundary of a convex cone. Now μ_i is a "truly bivariate" function and we are forced to use a different method to approximate it, resulting in an $O(n/\varepsilon)$ time algorithm to compute a pair $\tilde{s}, \tilde{t}$ such that

$$\sigma_1(f, g) \le \frac{1}{\operatorname{area}\mathbb{D}} \mu(\tilde{s}, \tilde{t}) \le (1 + \varepsilon)\sigma_1(f, g).$$

Analysis of μ over a triangle. Let $\Delta = \triangle v_i v_j v_k$ 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 common 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.). Depending on the scaling s and translation

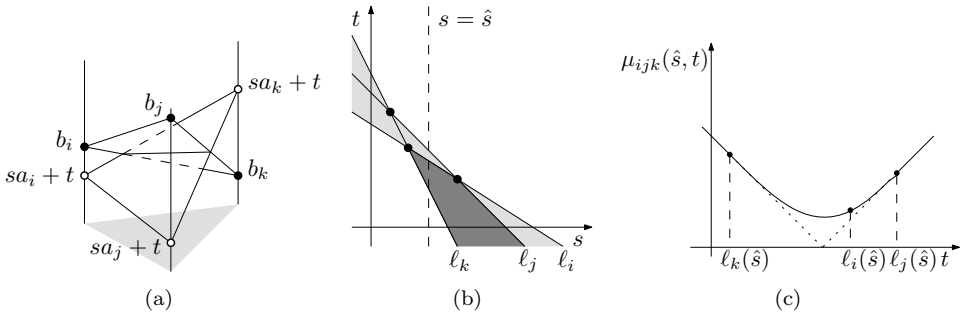


FIG. 4.1. (a) *A pair of intersecting triangles with common xy-projection.* (b) *The situation of* (a) *corresponds to values of* (s,t) *in the dark cell of the parameter space.* (c) *The vertical cross-section of the surface* $\mu_{ijk}(s,t)$ *for a fixed value* $s = \hat{s}$ *consists of four pieces (two linear and two cubic), separated by three breakpoints that correspond to the three vertices of the triangle.*

t, triangle $s \cdot \Delta_f + t$ may intersect, lie above, or lie below Δ_g. The volume between the two triangles is therefore composed of one or two parts.

In the st-plane, let $\ell_i \colon t = b_i - sa_i$ represent the set of points (s,t) at which the vertices of Δ_g and $s\Delta_f + t$ above v_i coincide; ℓ_j and ℓ_k are defined analogously. We also view ℓ_i as a univariate linear function $\ell_i(s) = b_i - sa_i$ and use a similar convention for ℓ_j and ℓ_k. Generally, these three lines partition the st-plane into seven cells;[3] see Figure 4.1(b). Let $\mathcal{A}_3$ denote this arrangement. 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 4.1(a), the expression for the volume between the triangles is

$$\mu_{ijk}(s,t) = \frac{A}{3}\left[(\ell_i(s) + \ell_j(s) + \ell_k(s) - 3t) + \frac{2(t - \ell_k(s))^3}{(\ell_i(s) - \ell_k(s)) \cdot (\ell_j(s) - \ell_k(s))}\right], \tag{4.2}$$

where $A = \operatorname{area}\Delta$, the area of the common projection of Δ_g and $s\Delta_f + t$; $\mu_{ijk}(s,t)$ is the signed volume between Δ_g and $s\Delta_f + t$ plus twice the volume of the tetrahedron formed by the intersection segment of the two triangles and the vertices b_k and $sa_k + t$ (see Figure 4.1(a)). The above formula is valid whenever $t \geq \ell_k(s)$, $t \leq \ell_i(s)$, and $t \leq \ell_j(s)$. Consider 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 (see Figure 4.1(c)). The form of the function μ_{ijk} in the cells of $\mathcal{A}_3$ met by this line are, in bottom to top order,

$$\frac{A}{3}(\ell_i(s) + \ell_j(s) + \ell_k(s) - 3t), \tag{4.3}$$

the form given by (4.2),

$$\frac{A}{3}\left[(3t - (\ell_i(s) + \ell_j(s) + \ell_k(s))) + \frac{2(\ell_j(s) - t)^3}{(\ell_j(s) - \ell_i(s)) \cdot (\ell_j(s) - \ell_k(s))}\right], \tag{4.4}$$

and finally

$$\frac{A}{3}(3t - \ell_i(s) - \ell_j(s) - \ell_k(s)). \tag{4.5}$$

[3] In the degenerate case where some choice of s,t makes $\Delta_g = s \cdot \Delta_f + t$, the st-plane is divided into only six cells, and the expressions that follow simplify.

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 (4.2) or (4.4) by a permutation of indices.

A constant-factor approximation for μ_{ijk}. We show that $\mu_{ijk}(s,t)$ can be approximated within a factor of $\sqrt{2}+2$ by a constant-complexity piecewise-linear function. Recall that $\mu_{ijk}(s,t)$ is convex, due to the same arguments that hold for μ and μ_i. We first consider the situation along a line $s = \hat{s}$, i.e., fix the scaling parameter to $\hat{s}$. As evident from (4.2) and (4.4), the univariate function $\mu_{ijk}(\hat{s},t)$ has (at most) three breakpoints—it is a cubic function between consecutive breakpoints and linear outside the breakpoints; see Figure 4.1(c). The next lemma suggests how to approximate $\mu_{ijk}(\hat{s},t)$ by a piecewise-linear convex function.

LEMMA 4.1. *For any fixed $\hat{s}$, the value of $\mu_{ijk}(\hat{s},t)$ at its breakpoints is at most $(\sqrt{2}+2)\min_{t\in\mathbb{R}} \mu_{ijk}(\hat{s},t)$.*

Proof. Without loss of generality, we assume that $\operatorname{area}\Delta = 1$. For any $t \in \mathbb{R}$, let $\bar{\Delta}_t$ denote $\hat{s}\Delta_f + t$. Suppose the line $s = \hat{s}$ intersects ℓ_j, ℓ_i, and ℓ_k from top to bottom in this order. Then Δ_g and $\bar{\Delta}_t$ intersect if t lies in the interval $[\ell_k(\hat{s}), \ell_j(\hat{s})]$, which we refer to as the *intersection interval*. The breakpoints of $\mu_{ijk}(\hat{s},t)$ occur at $t = \ell_k(\hat{s})$, $t = \ell_i(\hat{s})$, and $t = \ell_j(\hat{s})$. We set $\alpha = \ell_j(\hat{s}) - \ell_k(\hat{s})$ and $\beta = \ell_i(\hat{s}) - \ell_k(\hat{s})$, $0 \le \beta \le \alpha$; see Figure 4.2(a), (c).

Among the three breakpoints of $\tilde{\mu}_{ijk}(\hat{s},\cdot)$, the volume is maximum at $t = \ell_k(\hat{s})$ or $t = \ell_j(\hat{s})$ (Figure 4.2(c)), and by (4.3) and (4.5) the values at these two breakpoints are $\frac{1}{3}(\alpha+\beta)$ and $\frac{1}{3}(2\alpha-\beta)$, respectively. For simplicity, we assume $\beta \ge \alpha/2$, i.e., $\ell_i(\hat{s}) \ge (\ell_j(\hat{s}) + \ell_k(\hat{s}))/2$; then the volume is maximum at $t = \ell_k(\hat{s})$, i.e., when Δ_g and $\bar{\Delta}_t$ touch each other above the vertex v_k of Δ (see Figure 4.2(a)), and its value is

$$V_0 := \mu_{ijk}(\hat{s}, \ell_k(\hat{s})) = \frac{1}{3}(\alpha+\beta).$$

(If $\beta < \alpha/2$, we reverse the direction of the t-axis and use a symmetric argument.)

Next, consider the t-interval $[\ell_k(\hat{s}), \ell_i(\hat{s})]$. In this interval, Δ_g and $\bar{\Delta}_t$ intersect in the configuration shown in Figures 4.1(a) and 4.2(b), so using (4.2), $\mu_{ijk}(\hat{s},t)$ can be written as

$$V(t) := \mu_{ijk}(\hat{s},t) = \frac{1}{3}(\alpha+\beta-3(t-\ell_k(\hat{s})) + \frac{2(t-\ell_k(\hat{s}))^3}{3\alpha\beta}.$$

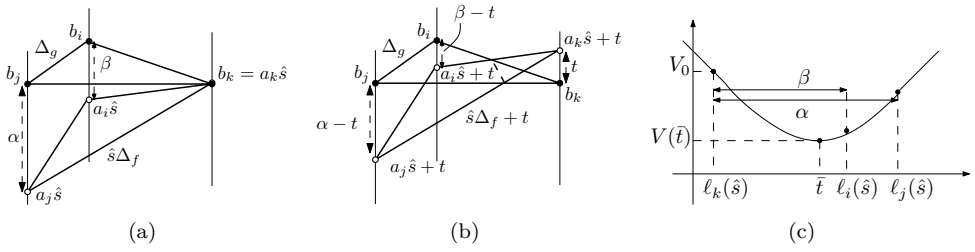


FIG. 4.2. *Illustration of the proof of Lemma* 4.1*:* (a) $t = \ell_k(\hat{s})$, (b) $t \in [\ell_k(\hat{s}), \ell_i(\hat{s})$, (c) *function* $\mu_{ijk}(\hat{s},t)$.

$V(t)$ is minimized at $t = \bar{t} := \sqrt{\alpha\beta/2} + \ell_k(\hat{s})$, and

$$V(\bar{t}) = \frac{1}{3}(\alpha + \beta - 3\sqrt{\alpha\beta/2}) + \frac{2}{3\alpha\beta}(\alpha\beta/2)^{3/2} = \frac{1}{3}(\alpha + \beta - \sqrt{2\alpha\beta}).$$

Since $\beta \geq \alpha/2$, $\bar{t}$ lies in the interval $[\ell_k(\hat{s}), \ell_i(\hat{s})]$. The convexity of μ_{ijk} implies that $\min_{t\in\mathbb{R}} \mu_{ijk}(\hat{s}, t) = V(\bar{t})$ (Figure 4.2). Hence,

$$\frac{\min_{t\in\mathbb{R}} \mu_{ijk}(\hat{s}, t)}{V_0} = \frac{V(\bar{t})}{V_0} = 1 - \frac{\sqrt{2\alpha\beta}}{\alpha + \beta}.$$

For $0 \leq \beta \leq \alpha$, the above expression is always at least $1 - 1/\sqrt{2}$, for the minimum is attained when $\beta = \alpha$. Hence,

$$V_0 \leq \frac{\sqrt{2}}{\sqrt{2} - 1} V(\bar{t}) = (\sqrt{2} + 2) \min_{t\in\mathbb{R}} \mu_{ijk}(\hat{s}, t),$$

as claimed. □

It follows that a piecewise-linear $(\sqrt{2}+2)$-approximation of $\mu_{ijk}(\hat{s}, t)$ can be created by taking its highest breakpoint, 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 4.3. Let $\chi_{ijk}(\hat{s})$ denote the t-interval on which the function is constant; we refer to this interval as the *valley* of $\tilde{\mu}_{ijk}(\hat{s}, \cdot)$.

We now return to the entire st-plane and extend the above function to a piecewise-linear bivariate function $\tilde{\mu}_{ijk} \colon \mathbb{R}^2 \to \mathbb{R}$. For an arbitrary value of s, we denote the valley of $\tilde{\mu}_{ijk}$ as $\chi_{ijk}(s) = [\gamma^-(s), \gamma^+(s)]$; the precise form is given below in (4.7). Assume $\ell_j(s) \geq \ell_i(s) \geq \ell_k(s)$. (For other cases, the expressions for $\tilde{\mu}_{ijk}$ and χ_{ijk} can be written by permuting the indices i, j, k.) Let $A = \operatorname{area}\Delta$ and $\bar{\ell}(s) = \frac{1}{2}(\ell_j(s) + \ell_k(s))$. Then (4.3), (4.5), and the proof of Lemma 4.1 imply that

$$\tilde{\mu}_{ijk}(s,t) = \begin{cases} \frac{A}{3}(\ell_i(s) + \ell_j(s) + \ell_k(s) - 3t) & \text{for } t < \gamma^-(s), \\ \frac{A}{3}(\ell_i(s) + \ell_j(s) - 2\ell_k(s)) & \text{for } t \in \chi_{ijk}(s), \ell_i(s) \geq \bar{\ell}(s), \\ \frac{A}{3}(2\ell_j(s) - \ell_i(s) - \ell_k(s)) & \text{for } t \in \chi_{ijk}(s), \ell_i(s) < \bar{\ell}(s), \\ \frac{A}{3}(3t - \ell_i(s) - \ell_j(s) - \ell_k(s)) & \text{for } t > \gamma^+(s). \end{cases} \tag{4.6}$$

Recall from the proof of Lemma 4.1 that if $\ell_i(s) \geq \bar{\ell}(s)$, then the highest breakpoint of $\mu_{ijk}(s)$ is $\ell_k(s)$, so in this case $\gamma^-(s) = \ell_k(s)$ and $\gamma^+(s)$ is the value of t for which

$$\frac{A}{3}(\ell_i(s) + \ell_j(s) - 2\ell_k(s)) = \frac{A}{3}(3t - \ell_i(s) - \ell_j(s) - \ell_k(s)),$$

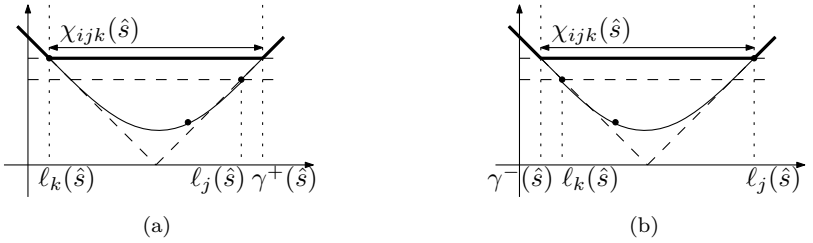


FIG. 4.3. *Function $\mu_{ijk}(\hat{s}, t)$ and its approximation $\tilde{\mu}(\hat{s}, t)$ (thick lines):* (a) $\beta \geq \alpha/2$, (b) $\beta < \alpha/2$. *The dashed lines denote the four linear functions defined in* (4.6)*; $\tilde{\mu}(\hat{s}, t)$ is the upper envelope of these four functions.*

i.e., $t = \frac{1}{3}(2\ell_i(s)+2\ell_j(s)-\ell_k(s))$. Similarly we can compute the values of $\gamma^-(s), \gamma^+(s)$ when $\ell_i(s) < \bar{\ell}(s)$. We can thus write χ_{ijk} as

$$\chi_{ijk}(s) = \begin{cases} \left[\ell_k(s), \frac{1}{3}(2\ell_i(s) + 2\ell_j(s) - \ell_k(s))\right] & \text{for } \ell_i(s) \geq \bar{\ell}(s), \\ \left[\frac{1}{3}(2\ell_i(s) + 2\ell_k(s) - \ell_j(s)), \ell_j(s)\right] & \text{for } \ell_i(s) < \bar{\ell}(s). \end{cases} \tag{4.7}$$

The following lemma is straightforward from the construction; see Figure 4.3.

LEMMA 4.2. *For any fixed value of s, say, $\hat{s}$, $\tilde{\mu}_{ijk}(\hat{s}, t)$ is the upper envelope of the four linear functions on the right-hand side of* (4.6).

The convexity of $\tilde{\mu}_{ijk}$ over the entire st-plane follows from the following lemma.

LEMMA 4.3. *The function $\tilde{\mu}_{ijk}$ can be represented as the upper envelope of at most six linear functions.*

Proof. We will describe a set of six linear functions such that $\tilde{\mu}_{ijk}$ is their upper envelope. Consider the arrangement $\mathcal{A}_3$ of ℓ_i, ℓ_j, ℓ_k. Without loss of generality, assume that area $\Delta = 1$, that ℓ_i, ℓ_j, ℓ_k appear from top to bottom in this order at $s = -\infty$, that the three lines are in general position, and that two vertices of $\mathcal{A}_3$ appear on the upper envelope λ^+ and one on the lower envelope λ^- of the three lines; refer to Figure 4.4. Let λ_1 be the median level of $\mathcal{A}_3$. We view $\lambda^+, \lambda^-, \lambda_1$ as polygonal lines as well as piecewise-linear functions of s. Note that γ^-, the lower boundary of χ_{ijk}, lies on or below λ^-, and γ^+, the upper boundary of χ_{ijk}, lies on or above λ^+. By (4.6), for points (s,t) lying above γ^+, $\tilde{\mu}_{ijk}(s,t)$ is defined by the linear function $t - \frac{1}{3}(\ell_i(s) + \ell_j(s) + \ell_k(s))$, and for points lying below γ^-, it is defined by the linear function $\frac{1}{3}(\ell_i(s) + \ell_j(s) + \ell_k(s)) - t$.

We next focus on points lying between γ^- and γ^+ (shaded area in Figure 4.4(b)). Let $\bar{\lambda}(s) = (\lambda^+(s) + \lambda^-(s))/2$, which is also a polygonal chain. Let s_1, s_2, s_3 be the s-coordinates of the vertices of $\mathcal{A}_3$. The lines $s = s_i$, $i = 1, 2, 3$, partition the st-plane into four strips $\Sigma_0, \ldots, \Sigma_3$. Note that the vertices of $\bar{\lambda}$ lie on the boundaries of these strips, so $\bar{\lambda}$ is a linear function within each strip.

We claim that λ_1 intersects $\bar{\lambda}$ exactly once in each of Σ_1 and Σ_2, and at most once in $\Sigma_0 \cup \Sigma_3$. Indeed, $\lambda_1(s_1) = \lambda^+(s_1) > \bar{\lambda}(s_1)$ and $\lambda_1(s_2) = \lambda^+(s_2) < \bar{\lambda}(s_2)$, implying that λ_1 and $\bar{\lambda}$ intersect in Σ_1. A similar argument shows that they intersect once in Σ_2. Finally, $\bar{\lambda}(s) = (\ell_i(s) + \ell_k(s))/2$ and $\lambda_1 = \ell_j$ in $\Sigma_0 \cup \Sigma_3$, and therefore they intersect at most once in $\Sigma_0 \cup \Sigma_3$. This proves the claim.

Let $\bar{s}_1, \bar{s}_2, \bar{s}_3$ be the s-coordinates of the intersection points of λ_1 and $\bar{\lambda}$ (we assume that they intersect at three points; fewer intersections are easier to handle). If an

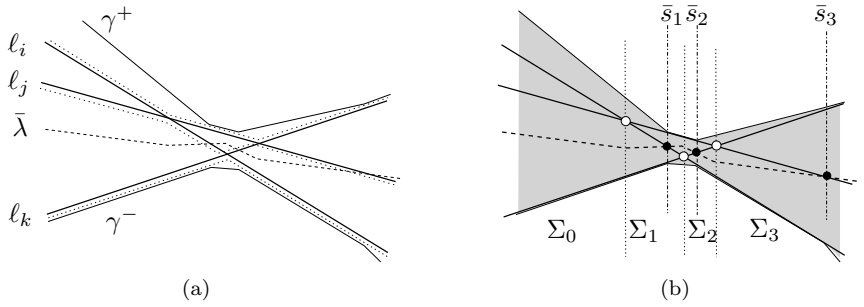


FIG. 4.4. (a) *Arrangement $\mathcal{A}$, λ^+, λ_1, λ^- (dotted lines), $\bar{\lambda}$ (dashed line), and γ^-, γ^+;* (b) *$\mathcal{A}$, valley (shaded region) of $\tilde{\mu}_{ijk}$, and strips $\Sigma_0, \ldots, \Sigma_3$; black circles denote the intersection points of $\bar{\lambda}$ and λ_1, the s-values where $\tilde{\mu}_{ijk}$ changes inside the valley.*

s-interval I does not contain any s_a or $\bar{s}_b$ for $1 \le a, b \le 3$, then clearly $\tilde{\mu}_{ijk}$ is defined by a single linear function for all points (s,t) with $s \in I$ and $t \in \chi_{ijk}(s)$. Next, we claim that the functional form of $\tilde{\mu}_{ijk}$ does not change even at s_a for $a = 1, 2, 3$. Indeed, suppose s_a is the s-coordinate of a vertex on the upper envelope λ^+ formed by $\ell_i \cap \ell_j$. Since $\lambda_1(s_a) = \lambda^+(s_a) > \bar{\lambda}(s_a)$, by (4.6), we observe that $\tilde{\mu}_{ijk}(s,t) = \frac{1}{3}(\ell_i(s) + \ell_j(s) - 2\ell_k(s))$ for $s \uparrow s_a$ as well as for $s \downarrow s_a$ (and $t \in \chi_{ijk}(s)$). That is, the functional form of $\tilde{\mu}_{ijk}$ does not change at s_a, as λ_1 and λ^+ switch between ℓ_i and ℓ_j that vertex. A similar argument holds if s_a is a vertex of the lower envelope. Hence, inside the valley, the functional form of $\tilde{\mu}_{ijk}$ only changes at $\bar{s}_1, \bar{s}_2, \bar{s}_3$, implying that it is defined by at most four linear functions inside the valley. Altogether, $\tilde{\mu}_{ijk}$ is defined by at most six linear functions.

Finally, we prove that $\tilde{\mu}_{ijk}$ is the upper envelope of these six functions. Fix a value $s = \hat{s}$. Assume that $\ell_j(\hat{s}) \ge \ell_i(\hat{s}) \ge \ell_k(\hat{s})$. By Lemma 4.2, $\tilde{\mu}_{ijk}(\hat{s}, t)$ is the upper envelope of the four functions

$$f_1(\hat{s},t) = \frac{1}{3}(\ell_i(\hat{s}) + \ell_j(\hat{s}) + \ell_k(\hat{s})) - t,$$
$$f_2(\hat{s},t) = t - \frac{1}{3}(\ell_i(\hat{s}) + \ell_j(\hat{s}) + \ell_k(\hat{s})),$$
$$f_3(\hat{s},t) = \frac{1}{3}(\ell_i(\hat{s}) + \ell_j(\hat{s}) - 2\ell_k(\hat{s})),$$
$$f_4(\hat{s},t) = \frac{1}{3}(2\ell_j(\hat{s}) - \ell_i(\hat{s}) - \ell_k(\hat{s})).$$

The remaining two functions $f_5(s,t), f_6(s,t)$ are of the form f_3 or f_4, obtained by permuting the indices i, j, k. Suppose f_5 is obtained from f_3 by permuting the indices. Since $\ell_k(\hat{s}) \le \ell_i(\hat{s}), \ell_j(\hat{s})$, we have $f_3(\hat{s},t) \ge f_5(\hat{s},t)$. Similarly if f_5 is obtained from f_4 by permuting the indices, then $f_4(\hat{s},t) \ge f_5(\hat{s},t)$ because $\ell_j(\hat{s}) \ge \ell_k(\hat{s}), \ell_i(\hat{s})$. Hence, for every $(s,t) \in \mathbb{R}^2$, $\tilde{\mu}_{ijk}(s,t) = \max_{1 \le h \le 6} f_h(s,t)$. This completes the proof of the lemma. □

Lemmas 4.1 and 4.3 together imply the following.

COROLLARY 4.4.

(i) *The function $\tilde{\mu}_{ijk}$ is a piecewise-linear convex function, consisting of at most six pieces, and can be constructed in $O(1)$ time.*

(ii) *For all $s,t \in \mathbb{R}^2$ with $t \notin \chi_{ijk}(s)$, $\tilde{\mu}_{ijk}(s,t) = \mu_{ijk}(s,t)$.*

(iii) *For all $s,t \in \mathbb{R}^2$ with $t \in \chi_{ijk}(s)$, $\mu_{ijk}(s,t) \le \tilde{\mu}_{ijk}(s,t) \le (\sqrt{2}+2)\mu_{ijk}(s,t)$.*

A $(1+\varepsilon)$-approximation for μ_{ijk}. Using $\tilde{\mu}_{ijk}$ we now construct a new function $\bar{\mu}_{ijk}$, which is a $(1+\varepsilon)$-approximation of μ_{ijk}. Let $\varepsilon > 0$ be a parameter. Set $\delta = \sqrt{\varepsilon}/c_1$, for a sufficiently large constant $c_1 > 0$, chosen so that $1/\delta$ is an integer. We partition $\Delta = \triangle v_i v_j v_k$ into a family Ξ of $\Theta(1/\delta^2) = \Theta(1/\varepsilon)$ congruent triangles by placing $\frac{1}{\delta} - 1$ equally spaced points on each edge of Δ and forming a triangular grid; refer to Figure 4.5(a). Each triangle $\rho \in \Xi$ is of the form $\rho = \delta\Delta + a$ or $\rho = -\delta\Delta + a$ for some $a \in \mathbb{R}^2$. 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 define

$$\bar{\mu}_{ijk}(s,t) = \sum_{\rho \in \Xi} \tilde{\mu}_\rho(s,t) \qquad \text{for all } s,t \in \mathbb{R}^2.$$

LEMMA 4.5. *The function $\bar{\mu}_{ijk}$ is convex, and for all $(s,t) \in \mathbb{R}^2$,*

$$\mu_{ijk}(s,t) \le \bar{\mu}_{ijk}(s,t) \le (1+\varepsilon)\mu_{ijk}(s,t).$$

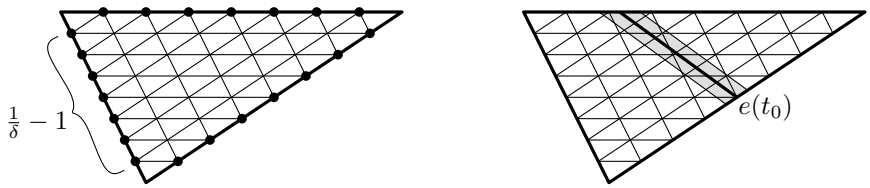


FIG. 4.5. (a) *Partitioning of a triangle by segments parallel to the sides;* (b) $\tilde{\mu}_\rho(s,t) = \mu_\rho(s,t)$ *for subtriangles* ρ *that do not intersect the gray strip* Ψ.

Proof. The convexity of $\bar{\mu}_{ijk}$ follows from that of $\tilde{\mu}_\rho$ for each $\rho \in \Xi$, and the first inequality follows since $\tilde{\mu}_\rho$ is an upper bound on μ_ρ. It thus suffices to prove the second inequality.

Fix a point (s_0, t_0). Without loss of generality, assume that $\operatorname{area}\Delta = 1$ and $\ell_k(s_0) \le \ell_i(s_0) \le \ell_j(s_0)$. Following the notation from the proof of Lemma 4.1, set $\alpha = \ell_j(s_0) - \ell_k(s_0)$ and $\beta = \ell_i(s_0) - \ell_k(s_0)$. Assume $\beta \ge \alpha/2$, so the value of $\mu_{ijk}(s_0, \cdot)$ at its highest breakpoint (which corresponds to the pair s_0, t_1 for $t_1 \in \mathbb{R}$, at which Δ_g and $s_0\Delta_f + t_1$ touch each other above the vertex v_k; refer to Figure 4.2) is $V_0 = \frac{1}{3}(\alpha + \beta)$.

Let $\Xi_1 = \{\rho \in \Xi \mid t_0 \in \chi_\rho(s_0)\}$. Since $\tilde{\mu}_\rho(s_0, t_0) = \mu_\rho(s_0, t_0)$ for $\rho \in \Xi \setminus \Xi_1$,

$$\bar{\mu}_{ijk}(s_0, t_0) = \sum_{\rho \in \Xi_1} \tilde{\mu}_\rho(s_0, t_0) + \sum_{\rho \in \Xi \setminus \Xi_1} \mu_\rho(s_0, t_0).$$

Hence,

$$\bar{\mu}_{ijk}(s_0, t_0) - \mu_{ijk}(s_0, t_0) \le \sum_{\rho \in \Xi_1} \tilde{\mu}_\rho(s_0, t_0). \tag{4.8}$$

We prove the following two claims below.

CLAIM 1. *For any* $\rho \in \Xi_1$, $\tilde{\mu}_\rho(s_0, t_0) = \delta^3 V_0$.

CLAIM 2. *There exists a constant* $c_2 > 0$ *such that* $|\Xi_1| \le c_2/\delta$.

Plugging the bounds from Claims 1 and 2 into (4.8), we obtain

$$\bar{\mu}_{ijk}(s_0, t_0) - \mu_{ijk}(s_0, t_0) \le \delta^3 V_0 |\Xi_1| \le c_2 \delta^2 V_0 \le \frac{c_2}{c_1^2} \varepsilon V_0.$$

By Lemma 4.1, $V_0 \le (\sqrt{2}+2)\mu_{ijk}(s_0, t_0)$, so by choosing $c_1 \ge \sqrt{(\sqrt{2}+2)c_2}$, we have $\bar{\mu}_{ijk}(s_0, t_0) \le (1+\varepsilon)\mu_{ijk}(s_0, t_0)$, as claimed. ∎

Proof of Claim 1. Let ρ be a triangle in Ξ. We assume $\rho = \delta\Delta + a$ for some $a \in \mathbb{R}^2$; a symmetric argument holds if $\rho = -\delta\Delta + a$.

Let $\bar{v}_k$ be the vertex of ρ corresponding to the vertex v_k of Δ, and let ρ_f and ρ_g be the triangles in the graphs of f and g, respectively, corresponding to ρ. If $\rho \in \Xi_1$, then by definition, $\tilde{\mu}_\rho(s_0, t_0)$ is the value of $\mu_\rho(s_0, \cdot)$ at its highest breakpoint. Since $\rho = \delta\Delta + a$ and Δ_g and $s_0\Delta_f + t_1$ touch each other above v_k at the highest breakpoint (s_0, t_1) of $\tilde{\mu}_{ijk}(s_0, \cdot)$, the highest breakpoint of $\mu_\rho(s_0, \cdot)$ occurs at a pair (s_0, t_2) for some $t_2 \in \mathbb{R}$, at which ρ_g and $s_0\rho_f + t_2$ touch each other above the vertex $\bar{v}_k$. Furthermore, the volume between ρ_g and $s_0\rho_f + t_2$ is δ^3 times the volume between Δ_g and $s_0\Delta_f + t_1$ (because the tetrahedron formed by ρ_g and $s_0\rho_f + t_2$ is a homothetic copy of that formed by Δ_g and $s_0\Delta_f + t_1$, with a scaling factor of δ), which is V_0. Hence, $\tilde{\mu}_\rho(s_0, t_0) = \mu_\rho(s_0, t_2) = \delta^3 V_0$. ∎

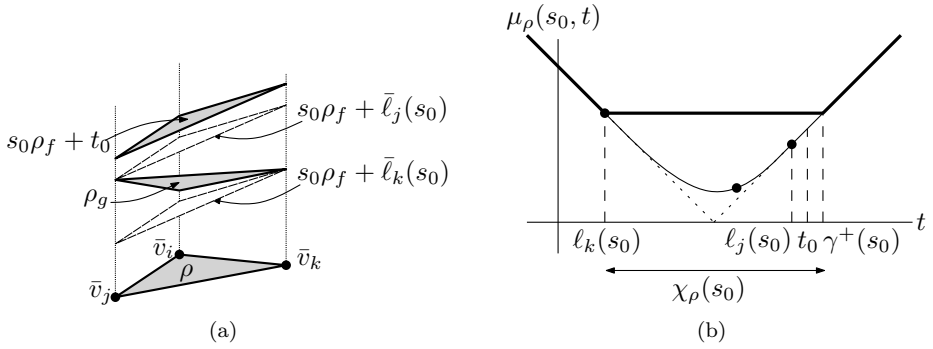


FIG. 4.6. *Illustration of the proof of Claim* 2: (a) ρ, ρ_g, *and* $s_0\rho_f + t_0$; (b) *function* $\mu_\rho(s_0, t)$.

Proof of Claim 2. We keep s fixed at s_0 and vary t. For any $t \in \mathbb{R}$, let $e(t)$ be the xy-projection of the intersection line of the planes supporting Δ_g and $s_0\Delta + t$; see Figure 4.5(b). The slope of $e(t)$ is fixed and $e(t)$ translates as t varies. Let u be the direction (in the xy-plane) normal to $e(t)$, and let ω be the width of Δ in direction u, i.e., the distance between the two lines parallel to $e(t)$ that support Δ. Recall that $s_0\Delta_f + t$ and Δ_g intersect for $t \in [\ell_k(s_0), \ell_k(s_0) + \alpha]$; i.e., α is the vertical distance between the two planes parallel to $s_0\Delta_f$ that support Δ_g. Therefore,

($\star$) for any $l > 0$, the distance between $e(t)$ and $e(t + l\alpha)$ is $l\omega$.

Let ρ be a triangle in Ξ_1. Assume ρ is of the form $\delta\Delta + a$. Extending the notation from the proof of Claim 1, let $\bar{v}_i, \bar{v}_j, \bar{v}_k$ be the vertices of ρ corresponding to the vertices v_i, v_j, v_k of Δ, let ρ_f and ρ_g be the triangles in the graphs of f and g, respectively, corresponding to ρ, and let $\bar{\ell}_i, \bar{\ell}_j, \bar{\ell}_k$ be the lines for ρ analogous to ℓ_i, ℓ_j, ℓ_k for Δ; see Figure 4.6. Since $\rho = \delta\Delta + a$, $\bar{\ell}_k(s_0) \le \bar{\ell}_i(s_0) \le \bar{\ell}_j(s_0)$, $\bar{\alpha} := \bar{\ell}_j(s_0) - \bar{\ell}_k(s_0) = \delta\alpha$, $\bar{\beta} := \bar{\ell}_i(s_0) - \bar{\ell}_k(s_0) = \delta\beta$, and $\bar{\beta} \ge \bar{\alpha}/2$. Since $\bar{\beta} \ge \bar{\alpha}/2$, by (4.7), $\chi_\rho(s_0) = [\bar{\ell}_k(s_0), \frac{1}{3}(2\bar{\ell}_j(s_0) + 2\bar{\ell}_i(s_0) - \bar{\ell}_k(s_0))]$. Note that ρ_g and $s_0\rho_f + t$ intersect for $t \in [\bar{\ell}_k(s_0), \bar{\ell}_j(s_0)]$, and $e(t_0)$ intersects ρ in this case. If the two triangles do not intersect and $\rho \in \Xi_1$, then $s_0\rho_f + t_0$ lies above ρ_g, $\ell_j(s_0) < t_0 \le \frac{1}{3}(2\bar{\ell}_j(s_0) + 2\bar{\ell}_i(s_0) - \bar{\ell}_k(s_0))$, and the height difference between ρ_g and $s_0\rho_f + t_0$ above the vertex $\bar{v}_j$ of ρ is at most $r := \frac{1}{3}(\bar{\ell}_j(s_0) - \bar{\ell}_k(s_0)) = \bar{\alpha}/3 = \delta\alpha/3$; the last claim follows from the fact that ρ_g and $s_0\rho_f + \bar{\ell}_j(s_0)$ touch each other above the vertex $\bar{v}_j$. In either case, ρ intersects $e(t_0 - r')$ for some $0 \le r' \le r$. Similarly we can argue that if $\bar{\beta} < \bar{\alpha}/2$, then ρ intersects $e(t_0 + r')$ for some $0 \le r' \le r$.

In other words, ρ intersects the strip Ψ bounded by the lines $e(t_0 - r)$ and $e(t_0 + r)$. A symmetric argument shows that ρ intersects Ψ even if it is of the form $-\delta\Delta + a$. Since $r \le \frac{\delta}{3}$, by ($\star$), the width of Ψ is $\frac{2}{3}\delta\omega$, while the width of ρ in direction u is $\delta\omega$. Therefore $\rho \not\subseteq \Psi$, and if ρ intersects Ψ, then it also intersects $e(t_0 - r)$ or $e(t_0 + r)$. Each of these two lines intersects $O(1/\delta)$ triangles of Ξ. Hence, $|\Xi_1| = O(1/\delta)$. ∎

A $(1 + \varepsilon)$-approximation algorithm. We have all ingredients for a $(1 + \varepsilon)$-approximation algorithm to compute $\sigma_1(f, g)$. Given ε, we choose the parameter δ as above. We partition each triangle $\Delta = \triangle v_i v_j v_k \in \mathbb{M}$ into a set Ξ_Δ of $O(1/\delta^2) = O(1/\varepsilon)$ tiny triangles and construct the function $\tilde{\mu}_\rho$ for each $\rho \in \Xi_\Delta$, as described above. Let L_ρ be the set of $O(1)$ lines containing the edges of the st-projection of the graph of $\tilde{\mu}_\rho$. Set $\Xi = \bigcup_{\Delta \in \mathbb{M}} \Xi_\Delta$, $L = \bigcup_{\rho \in \Xi} L_\rho$, and $m := |L| = O(n/\varepsilon)$. We define

$$\bar{\mu}(s,t) = \sum_{\rho \in \Xi} \tilde{\mu}_\rho(s,t).$$

Obviously $\tilde{\mu}$ is a piecewise-linear convex function, and it is linear within a cell of the arrangement $\mathcal{A}(L)$. We compute the minimum value of $\bar{\mu}$ by using a prune-and-search method similar to the one in section 3.

For a simplex γ in the st-plane, let $L_{|\gamma} \subset L$ denote the set of lines that cross the interior of γ; set $m_\gamma = |L_{|\gamma}|$. If, for a triangle $\rho \in \Xi$, the lines of L_ρ do not intersect γ, then $\tilde{\mu}_\rho$ restricted to γ is a linear function, so let $h_\gamma(s,t) = \sum_{\rho \in \Xi : L_\rho \cap L_{|\gamma} = \emptyset} \tilde{\mu}_\rho(s,t)$. Then

$$\bar{\mu}(s,t) = \sum_{\rho \in \Xi : L_\rho \cap L_{|\gamma} \neq \emptyset} \tilde{\mu}_\rho(s,t) + h_\gamma(s,t). \tag{4.9}$$

Let $\pi^*_\gamma = \min_{p \in \gamma} \tilde{\mu}(p)$. Given a segment γ, $L_{|\gamma}$, and h_γ, each of the following two operations, which are analogous to (A1) and (A2) in section 3, can be performed in $O(m_\gamma)$ time.

(B1) For any $x \in \gamma$, we can compute $\bar{\mu}(x)$ using (4.9). Furthermore, we can also determine whether $x = \pi^*_\gamma$, and, if the answer is NO, whether π^*_γ lies to the left of x or to its right.

(B2) Given π^*_γ, we can determine whether $\pi^* = \pi^*_\gamma$, and, if the answer is NO, whether π^* lies in γ^- or γ^+, where γ^- (resp., γ^+) is the halfplane lying below (resp., above) the line containing γ.

Let τ be a triangle in the st-plane. Given $L_{|\tau}$ and h_τ, we use the same recursive procedure as in section 3 to compute π^*_τ. Initially τ is the entire st-plane, $L_{|\tau} = L$, and $h_\tau = 0$. Using (B1) and (B2), the procedure computes, in expected $O(m_\tau)$ time, either π^*_τ or a subtriangle $\tau' \subset \tau$ that contains π^*_τ and $m_{\tau'} \leq m_\tau / r$ for some constant $r \geq 2$. In the latter case, the procedure also computes $L_{|\tau'}$ and $h_{\tau'}$. Hence, the overall expected running time spent in computing π^*_τ is $O(m_\tau)$. Initially $m_\tau = O(n/\varepsilon)$, so the minimum value of $\bar{\mu}$ can be computed in $O(n/\varepsilon)$ expected time.

Putting everything together, we obtain the main result of this section.

THEOREM 4.6. *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. A pair $\bar{s}, \bar{t}$ and the value $\mu(\bar{s}, \bar{t})$ can be computed in $O(n/\varepsilon)$ expected time such that $\sigma_1(f,g) \leq \mu(\bar{s}, \bar{t}) \leq (1+\varepsilon)\sigma_1(f,g)$.*

Remark. (i) The bound on the running time can be made worst-case by using a deterministic algorithm for computing a $(1/r)$-cutting at each step of the recursive procedure [4].

(ii) It is not clear if the expected running time can be improved to $n \operatorname{polylog}(1/\varepsilon)$ as in Theorem 3.4, for we do not know how to represent $\tilde{\mu}_{ijk}$ compactly. However, using Claim 2 and some geometric data structures, it is likely that the expected running time can be improved to roughly $n/\sqrt{\varepsilon}$, but we prefer simplicity over this improvement.

(iii) For unaligned functions, the factor n in the time bound is replaced by $n + k$, where k is the number of edge intersections in the overlay, 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 mean distance measure,

a strong linear-time approximation scheme exists even for the bivariate case, i.e., an algorithm with running time of the form $O(n+q(\varepsilon))$ for some function q. It may also be possible to improve on the worst-case quadratic running time results for the bivariate minimum mean distance measure, for nonaligned terrains. Another natural extension would be to increase the dimension of the domain (e.g., $\mathbb{D} = \mathbb{R}^3$); this paper focused on the cases $\mathbb{D} = \mathbb{R}$ and $\mathbb{D} = \mathbb{R}^2$. Finally, optimizing over more complex transformations than linear, or finding a linear model for more than two bivariate functions, leads to problems with more than two unknowns that are worth studying further.

Acknowledgment. The authors thank the anonymous referees for their useful comments.

REFERENCES

[1] P. K. Agarwal, B. Aronov, M. J. van Kreveld, M. Löffler, and R. I. Silveira, *Computing similarity between piecewise-linear functions*, in Proceedings of the 26th Annual ACM Symposium on Computational Geometry, 2010, pp. 375–383.

[2] P. K. Agarwal and J. Erickson, *Geometric range searching and its relatives*, in Advances in Discrete and Computational Geometry, B. Chazelle, J. E. Goodman, and R. Pollack, eds., Contemp. Math. 23, AMS, Providence, RI, 1999, pp. 1–56.

[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 Comput. Geom., 9 (1993), pp. 145–158.

[5] B. Chazelle, H. Edelsbrunner, L. J. Guibas, and M. Sharir, *Algorithms for bichromatic line segment problems and polyhedral terrains*, Algorithmica, 11 (1994), pp. 116–132.

[6] B. Chazelle, H. Edelsbrunner, L. J. Guibas, M. Sharir, and J. Stolfi, *Lines in space: Combinatorics and algorithms*, Algorithmica, 15 (1996), pp. 428–447.

[7] B. Chazelle and J. Matousek, *On linear-time deterministic algorithms for optimization problems in fixed dimension*, J. Algorithms, 21 (1996), pp. 579–597.

[8] K. L. Clarkson, *Las Vegas algorithms for linear and integer programming when the dimension is small*, J. ACM, 42 (1995), pp. 488–499.

[9] N. Cressie, *Statistics for Spatial Data*, John Wiley, Chichester, UK, 1991.

[10] M. de Berg, H. J. Haverkort, S. Thite, and L. Toma, *I/O-efficient map overlay and point location in low-density subdivisions*, in Proceedings of the 18th Annual International Symposium on Algorithms and Computation, Lecture Notes in Comput. Sci. 4835, Springer, Berlin, 2007, pp. 500–511.

[11] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, *Local and global comparison of continuous functions*, in IEEE Visualization, IEEE Computer Society Press, 2004, pp. 275–280.

[12] U. Finke and K. Hinrichs, *Overlaying simply connected planar subdivisions in linear time*, in Proceedings of the 11th Annual ACM Symposium on Computational Geometry, 1995, pp. 119–126.

[13] E. Isaaks and R. Srivastava, *Applied Geostatistics*, Oxford University Press, 1989.

[14] H. Li and R. Tasay, *A unified approach to identifying multivariate time series models*, J. Amer. Statist. Assoc., 93 (1998), pp. 770–782.

[15] J. Matoušek, *Lectures on Discrete Geometry*, Springer, Berlin, 2002.

[16] G. Moroz and B. Aronov, *Computing the distance between piecewise-linear bivariate functions*, in Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, 2012, pp. 288–293.

[17] L. O'Brian, *Introducing Quantitative Geography*, Routledge, 1992.

[18] D. O'Sullivan and D. Unwin, *Geographic Information Analysis*, John Wiley, 2003.

[19] S. Schaible and J. Shi, *Fractional programming: The sum-of-ratios case*, Optim. Methods Software, 18 (2003), pp. 219–229.

[20] M. Sharir and P. K. Agarwal, *Davenport-Schinzel Sequences and Their Geometric Applications*, Cambridge University Press, New York, 1995.

[21] H. Wackernagel, *Multivariate Geostatistics: An Introduction with Applications*, 3rd ed., Springer, 2003.

[22] W. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, Addison-Wesley, Redwood City, CA, 1990.