Region-based approximation of probability distributions (for visibility between imprecise points among obstacles)

Region-based approximation of probability distributions (for visibility between imprecise points among obstacles)

Kevin Buchin[*] *Irina Kostitsyna*[*] *Maarten Löffler*[†] *Rodrigo I. Silveira*[‡]

Abstract

Let p and q be two imprecise points, given as probability density functions on $\mathbb{R}^2$, and let $\mathcal{R}$ be a set of n line segments (obstacles) in $\mathbb{R}^2$. We study the problem of approximating the probability that p and q can see each other; that is, that the segment connecting p and q does not cross any segment of $\mathcal{R}$. To solve this problem, we approximate each density function by a weighted set of polygons; a novel approach to dealing with probability density functions in computational geometry.

1 Introduction

Data imprecision is an important obstacle to the application of geometric algorithms to real-world problems. In the computational geometry literature, various models to deal with data imprecision have been suggested. Most generally, in this paper we describe the location of each point by a probability distribution μ_i (for instance by a Gaussian distribution). This model is often not worked with directly because of the computational difficulties arisen from its generality.

These difficulties can often be addressed by approximating the distributions by point sets. For instance, for tracking uncertain objects a particle filter uses a discrete set of locations to model uncertainty [17]. Löffler and Phillips [13] and Jørgenson *et al.* [11] discuss several geometric problems on points with probability distributions, and show how to solve them using discrete point sets (or *indecisive* points) that have guaranteed error bounds. More specifically, a 2-dimensional point set P is an ε-*quantization* of an xy-monotone function F (such as a cumulative probability density function), if for every point q in the plane the fraction of P dominated by q differs from $F(q)$ by at most ε.

Imprecise points appear naturally in many applications. They play an important role in databases [8, 2, 7, 5, 16, 1, 6], machine learning [4], and sensor networks [18], where a limited number of probes from a certain data set are gathered, each potentially representing the true location of a data point. Alternatively, data points may be obtained using imprecise measurements or are the result of inexact earlier computations.

Even though a point set may be a provably good approximation of a probability distribution, this is not good enough in all applications. Consider, for example, a situation where we wish to model visibility between imprecise points among obstacles. When both points are given by a probability distribution, naturally there is a probability that the two points see each other. However, when we discretise the distributions, the choice of points may greatly influence the resulting probability, as illustrated in Figure 1.

*Dept. of Mathematics and Computer Science, TU Eindhoven; {k.a.buchin, i.kostitsyna}@tue.nl.
†Department of Computing and Information Sciences, Utrecht University; m.loffler@uu.nl.
‡Dept. de Matemática & CIDMA, Universidade de Aveiro, and Dept. Matemàtica Aplicada II, Universitat Politècnica de Catalunya; rodrigo.silveira@ua.pt.

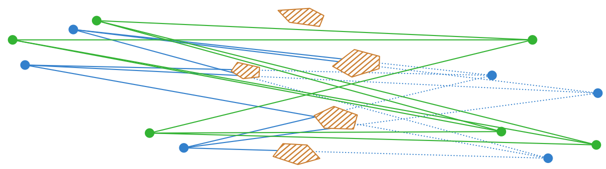
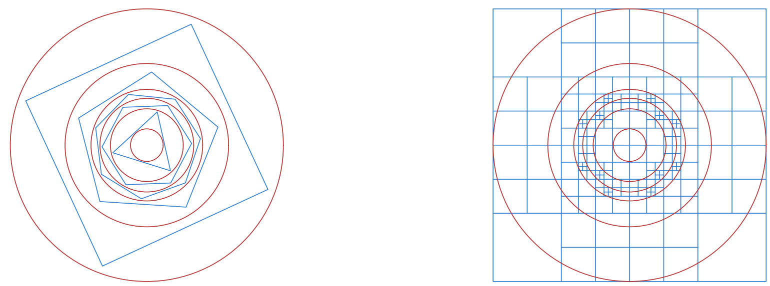


Fig. 1: Two pairs of point sets on opposite sides of a collection of obstacles. The green points can all see each other, whereas none of the blue points can.



(a)

(b)

Fig. 2: A Gaussian distribution, given by isolines at ε levels. (a) Approximation by polygons. (b) Approximation by quadtree.

Instead, we may approximate distributions by regions. The concept of describing an imprecise point by a region or shape was first introduced by Guibas *et al.* [9], motivated by finite coordinate precision, and later studied extensively in a variety of settings [10, 3, 14, 15, 12].

In this work we show how to use region-based approximation of point distributions to solve algorithmic problems on (general) imprecise points. In Section 2 we discuss several ways to do this. In Section 3, we focus on a geometric problem for which previous point-based methods do not work well: visibility computations between imprecise points.

2 Region-based approximation

Assume that an imprecise point p is given by a probability distribution μ. We wish to describe μ by a set of weighted regions $\mathcal{M}$ that provide an additive ε-approximation of the distribution: for any point in the plane, the sum of the weights of the regions containing q differs from $\mu(q)$ by at most ε.

One approach to attack this problem is to consider the *isolines* of the probability density function $f(\cdot)$ at ε levels. These are the curves where $f(\cdot)$ is exactly $k\varepsilon$, for some integer k, and they separate the plane into regions where $f(\cdot)$ has a value between $i\varepsilon$ and $(i+1)\varepsilon$. Note that if we could take the regions formed by the isolines, and give each of them a weight of ε, they would form a valid ε-approximation of f. However, the isolines are not generally polygonal. Instead, we note that if we take any polygons that stay between two consecutive isolines, and use these as polygons with weight ε, they are guaranteed to form a 2ε-approximation.

Lemma 2.1. *Let f be a function with isolines of f at ε levels. Let $\mathcal{M}$ be a set of polygons such that exactly one polygon from $\mathcal{M}$ lies between any two consecutive isolines. If every polygon in $\mathcal{M}$ has weight ε then $\mathcal{M}$ forms an additive 2ε-approximation of f.*

Proof. Let P be a polygon that lies between the ith and the $(i+1)$th isolines, and Q a polygon that lies between the $(i+1)$th and $(i+2)$th isolines. Then the region between P and Q lies completely between the ith and the $(i+2)$th isolines, hence, we know that for any point p inside P but outside Q, we have $i\varepsilon \le f(p) \le (i+2)\varepsilon$. The lemma follows. Figure 2(a) illustrates this. $\square$

Of course, the complexity of the polygons depends on the specific distribution. In the following we focus on Gaussian distributions, because they are natural and likely to occur in applications.

Theorem 2.2. *A Gaussian distribution with standard deviation σ can be ε-approximated by $O(\sigma^{-2}\varepsilon^{-1})$ polygons of total complexity $O(\sigma^{-4}\varepsilon^{-2})$.*

Proof. The isolines of a bivariate Gaussian distribution are concentric circles that subdivide $\mathbb{R}^2$ into annuli, and we wish to compute a polygon that stays within each annulus. We observe that the complexity of such a polygon depends only on the relative width of its annulus; that is, given an annulus with inner radius r and outer radius r', we can fit a regular $\lceil \pi / \arccos \frac{r}{r'} \rceil$-gon. Refer to Figure 2(a) for some examples.

The probability density function with standard deviation σ is given by the equation

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

The number of annuli depends on the height of the peak of the function we wish to approximate, which is at $\frac{1}{2\pi\sigma^2}$, so we need $k = \frac{1}{2\pi\sigma^2\varepsilon}$ isolines.

If we solve $f(x,0) = i\varepsilon$ for x, we get

$$x = \sqrt{-2\sigma^2 \log(2\pi\sigma^2 i\varepsilon)}$$

so the ith annulus has relative width

$$\frac{r}{r'} = \sqrt{\frac{\log(2\pi\sigma^2 i\varepsilon)}{\log(2\pi\sigma^2 (i+1)\varepsilon)}}.$$

Hence, the total complexity of all polygons is

$$\sum_{i=1}^{k} \left\lceil \pi / \arccos \sqrt{\frac{\log(2\pi\sigma^2 i\varepsilon)}{\log(2\pi\sigma^2 (i+1)\varepsilon)}} \right\rceil,$$

which we rather coarsely bound by k times the maximum of these terms,

$$k \max_i \left\lceil \pi / \arccos \sqrt{\frac{\log(2\pi\sigma^2 i\varepsilon)}{\log(2\pi\sigma^2 (i+1)\varepsilon)}} \right\rceil.$$

attained at $i = k/2$. We obtain:

$$\frac{1}{2\pi\sigma^2\varepsilon} \left\lceil \pi / \arccos \sqrt{\frac{\log 1/2}{\log(1/2 + 2\pi\sigma^2\varepsilon)}} \right\rceil.$$

As the argument of the arccos approaches 1, the value approaches 0 as the square of the argument, leading to a $O(\frac{1}{\sigma^2\varepsilon})$ growth rate. The lemma follows. $\square$

Alternatively, we may subdivide space into grid cells and give each cell a weight depending on the value of f. The advantage of a grid-based approach is that the subdivision of the plane does not depend on the actual distributions, and that squares are particularly nice polygons. A problem with this approach is that the resolution of the grid depends on the steepest part of f: when the value of f varies by more than ε in a cell, the approximation is not valid. Instead, we may also choose to compute a non-uniform grid, for example based on a quadtree. If we use a quadtree to subdivide $\mathbb{R}^2$ until no cell is crossed by more than one isoline, and we weigh a cell crossed by the ith isoline by $i\varepsilon$, we again obtain a 2ε-approximation. Figure 2(b) illustrates this.

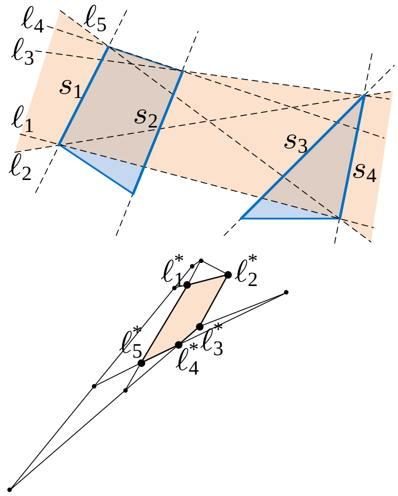


Fig. 3: Left: Two polygons P_1 and P_2 in primary space. The orange region is the set of lines intersecting P_1 and P_2 through s_1, s_2, s_3, s_4. Right: Partition L^* in dual space. The orange cell corresponds to all lines in the primary space intersecting the same four segments s_1, s_2, s_3, s_4.

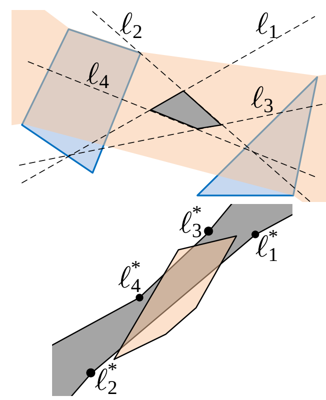


Fig. 4: Left: Polygons P_1 and P_2 and an obstacle in between in the primary space. Right: The "hour-glass" shape H^* in the dual space that corresponds to a set H of all lines in the primary space that intersect the obstacle.

3 Visibility between two regions

Consider a set of obstacles $\mathcal{R}$ in the plane. We assume that the obstacles are disjoint simple convex polygons with m vertices in total. For two imprecise points with probability distributions μ_1 and μ_2 we approximate them with two sets of weighted regions $\mathcal{M}_1$ and $\mathcal{M}_2$, each consisting of convex polygons. For every pair of polygons $P_1 \subset \mathcal{M}_1$ and $P_2 \subset \mathcal{M}_2$, we compute the probability that a point p_1 chosen uniformly at random from P_1 can see a point p_2 chosen uniformly at random from P_2. We say that two points can "see" each other if and only if the straight line segment connecting them does not intersect any obstacle from $\mathcal{R}$. The probability of two points $p_1 = (x_1, y_1) \in P_1$ and $p_2 = (x_2, y_2) \in P_2$ seeing each other can be computed by the formula:

$$prob = \frac{\iiiint v(x_1, y_1, x_2, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}x_2\mathrm{d}y_2}{\iiiint \mathrm{d}x_1\mathrm{d}y_1\mathrm{d}x_2\mathrm{d}y_2}, \tag{1}$$

where $v(x_1, y_1, x_2, y_2)$ is 1 if the points see each other, and 0 otherwise.

To compute $prob$ we consider a dual space $\mathbb{L}$ where a point with coordinates (α, β) corresponds to a line $y = \alpha x - \beta$ in the primary space. We construct a region $\mathbb{L}^*$ in the dual space that corresponds to the set of lines that stab both P_1 and P_2. This region can be partitioned into cells, each corresponding to a set of lines that cross the same four segments of P_1 and P_2 (refer to Figure 3). The following follows from the fact that each vertex of $\mathbb{L}^*$ corresponds to a line in primary space through two vertices of P_1 and P_2.

Lemma 3.1. *Given two convex polygons P_1 and P_2 of total size n, the complexity of partition L^* in the dual space that corresponds to a set of lines that stab P_1 and P_2 is $O(n^2)$.*

Proof. Each vertex of L^* corresponds to a line in the primary space that goes through a pair of vertices of P_1 and P_2. f $\square$

For each obstacle $h \subset \mathcal{R}$ we construct a region H^* in the dual space, that corresponds to the set of lines that intersect h. H^* has an "hour-glass" shape (refer to Figure 4). We now compute the

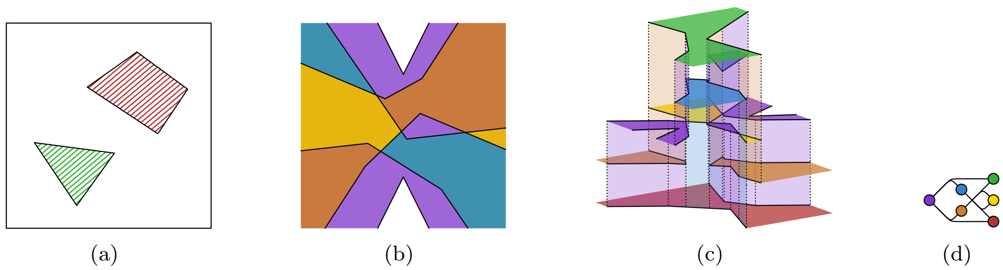


Fig. 5: (a) A square polygon P with two holes (obstacles). (b) The dual space Ł (cropped to a square). The coloured area corresponds to all lines in Ł that intersect the square domain. The orange/yellow area corresponds to the lines that intersect the red obstacle; the blue/yellow area corresponds to the lines that intersect the green obstacle. (c) The space of maximal line segments $\mathfrak{S}$. The purple layer are the segments that miss both obstacles (they extend from one end of the square to the other). The yellow layer are the segments that touch both obstacles. The red and orange layers are the segments that touch the red obstacle, but miss the green obstacle. The blue and green layers are the segments that touch the green obstacle, but miss the red obstacles. Vertical panels indicate which edges of layers are connected. (d) Schematic view of how layers are connected to each other.

subdivision Ł of the dual plane resulting from overlaying the partition Ł* and the regions H^*. Since the objects involved are bounded by a total of $O(m+n)$ line segments in the primal space, Ł has complexity $O((m+n)^2)$.

First consider the case that P_1, P_2 and the obstacles are disjoint. We can assume that all obstacles lie in the convex hull of P_1 and P_2. Then a pair of points from P_1 and P_2 see each other exactly if the line through the points does not intersect an obstacle. Thus, we only need to identify the cells in Ł not intersecting any of the regions H^*, and integrate over these cells. Details on evaluating the integral for one cell is given in Section 4. Overall this case can be handled in $O((m+n)^2)$ time. Next, consider the case that P_1 and P_2 are disjoint but might intersect obstacles. Now we need to consider the length of each line segment from the last obstacle in P_1 to the boundary and from the boundary of P_2 to the first obstacle. We can annotate the cells of Ł with this information by a traversal of Ł. Between neighboring cells this information can be updated in constant time. Thus, this case can be handled with the same asymptotic running time as the previous case. As a third case, consider P_1 overlapping P_2 but with no obstacles in the overlap area. The computations needed remain the same as in the case of non-overlapping P_1 and P_2. provided we actually evaluated this integral, we should now be able to compute the value in $O(n^2)$ time.

Finally, we consider the general case, in which obstacles might also lie in the overlap of P_1 and P_2. In the cells of Ł that correspond to the overlap of P_1 and P_2 we now need to consider the sum of the lengths of each line segment between boundaries of obstacles. If we simply traverse Ł, maintaining the ordered list of intersected obstacle boundaries, then computing the sum of lengths in one cell requires $O(m)$ time, leading to a total running time of $O(m(m+n)^2)$. Instead, we investigate the structure of the problem a little more closely.

Lemma 3.2. *Let P be a polygon, possibly with holes or multiple components, of total complexity n. Let $\mathfrak{S}$ be the space of all maximal line segments, that is, segments which lie in the (closed) interior of P but which are not contained in larger line segments that also lie in the interior of P. Then $\mathfrak{S}$ has complexity $O(n^2)$.*

Proof. Line segments have four degrees of freedom, but the condition that they must be locally maximal removes two of them, so $\mathfrak{S}$ is intrinsically two-dimensional. We may project $\mathfrak{S}$ onto the set Ł of all lines (by extending each segment to a line), but this way we may map multiple segments onto the same line. However, we only map finitely many segments to a line. We can visualize this as a finite set of "copies" of (patches of) Ł above each other. Then, as we move (translate or rotate) our segment through P, it may split into two segments when we hit a vertex; this corresponds to one of the copies of Ł splitting into two copies. The "seams" along which the copies of patches of Ł are sewn together in $\mathfrak{S}$ are one-dimensional curves, which correspond to the segment in P rotating around (and touching) a vertex. The endpoints of these seams are points which correspond to segments in P that connect two vertices. Figure 5 illustrates P, Ł and $\mathfrak{S}$ for a small example.

Clearly, there can be at most $O(n^2)$ segments that connect two vertices in P, thus, there are only $O(n^2)$ vertices in $\mathfrak{S}$. This does not immediately give the bound, though, since $\mathfrak{S}$ is not planar. However, each vertex in $\mathfrak{S}$ (corresponding to a pair of vertices in P) can be incident to at most two seams in $\mathfrak{S}$: one that corresponds to a segment rotating around either vertex in P. So, the total number of seams can also be at most $O(n^2)$. Since a seam always connects exactly three patches, the total complexity of $\mathfrak{S}$ is $O(n^2)$. □

If we apply Lemma 3.2 to our setting, then P is the imprecise point with obstacles as holes, of total complexity $(n + m)$. We arrive at the following intermediate result. In the next section, we show how to compute the probability for a given combinatorial configuration.

Lemma 3.3. *Given two polygons P_1 and P_2 of total size n and obstacles of total complexity m, we can compute the probability that a pair of points drawn uniformly at random from $P_1 \times P_2$ can see each other in $O((m + n)^2)$ time, assuming we can compute the necessary information within each cell.*

4 Computing the probability for a fixed combinatorial configuration

For simplicity of presentation, we assume that P_1 and P_2 are separable by a vertical line, and P_1 and P_2 are disjoint from $\mathcal{R}$. This will allow us to write the solution in a more concise way without loss of generality.

Consider line ℓ, given by the formula $y = \alpha x - \beta$, that goes through two points $p_1(x_1, y_1) \in P_1$ and $p_2(x_2, y_2) \in P_2$. In the dual space, point ℓ^*, corresponding to line ℓ, has coordinates (α, β). Substitute variables y_1 and y_2 in Formula 1 with α and β: $(x_1, y_1, x_2, y_2) \leftarrow (x_1, \alpha, x_2, \beta)$, where $\alpha(x_1, y_1, x_2, y_2) = y_2 - y_1/x_2 - x_1$ and $\beta(x_1, y_1, x_2, y_2) = (x_1y_2 - x_2y_1)/(x_2 - x_1)$. We can express the probability of two points, distributed uniformly at random in P_1 and P_2, seeing each other as

$$prob = \frac{\iiiint v(x_1, \alpha, x_2, \beta)|J|\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}\alpha\mathrm{d}\beta}{\iiiint |J|\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}\alpha\mathrm{d}\beta}, \tag{2}$$

where

$$J = \det\begin{bmatrix} \frac{dy_1}{\mathrm{d}\alpha} & \frac{\mathrm{d}y_1}{\mathrm{d}\beta} \\ \frac{dy_2}{\mathrm{d}\alpha} & \frac{\mathrm{d}y_2}{\mathrm{d}\beta} \end{bmatrix} = \frac{1}{\det\begin{bmatrix} \frac{\mathrm{d}\alpha}{\mathrm{d}y_1} & \frac{\mathrm{d}\beta}{\mathrm{d}y_1} \\ \frac{\mathrm{d}\alpha}{\mathrm{d}y_2} & \frac{\mathrm{d}\beta}{\mathrm{d}y_2} \end{bmatrix}} = x_2 - x_1\,.$$

The denominator of (2) can be written as a sum of integrals over all cells of partition L^* in the dual space:

$$\sum_{C \subset L^*} \iint\limits_C \left(\int\limits_{X_1(\alpha,\beta)}^{X_2(\alpha,\beta)} \int\limits_{X_3(\alpha,\beta)}^{X_4(\alpha,\beta)} (x_2 - x_1)\mathrm{d}x_2\mathrm{d}x_1 \right) \mathrm{d}\alpha\mathrm{d}\beta\,,$$

where $X_1(\alpha, \beta)$, $X_2(\alpha, \beta)$, $X_3(\alpha, \beta)$, and $X_4(\alpha, \beta)$ are the x-coordinates of intersections of line $y = \alpha x - \beta$ with the boundary segments of P_1 and P_2.

The numerator of (2) can be written as a sum of integrals over all cells of partition $L^* \setminus \cup_h H^*$ in the dual:

$$\sum_{C \subset L^* \setminus \cup_h H^*} \iint_C \left(\int_{X_1(\alpha,\beta)}^{X_2(\alpha,\beta)} \int_{X_3(\alpha,\beta)}^{X_4(\alpha,\beta)} (x_2 - x_1) \mathrm{d}x_2 \mathrm{d}x_1 \right) \mathrm{d}\alpha \mathrm{d}\beta.$$

In Appendix A we give a detailed case-by-case closed-form evaluation of the integrals. Since we integrate over constant-size subproblems, we obtain:

Theorem 4.1. *Given two convex polygons P_1 and P_2 of total size n and a set of obstacles of total size m, we can compute the probability that a point p_1 chosen uniformly at random in P_1 sees a point p_2 chosen uniformly at random in P_2 in $O((m+n)^2)$ time.*

5 Main result

Combining Theorems 2.2 and 4.1, our main result follows:

Theorem 5.1. *Given two imprecise points, modelled as Gaussian distributions μ_1 and μ_2 with standard deviations σ_1 and σ_2, and n obstacles, we can ε-approximate the probability that p and q see each other in $O(\sigma_1^{-2}\sigma_2^{-2}\varepsilon^{-2}((\sigma_1^{-2}+\sigma_2^{-2})\varepsilon^{-1}+n)^2)$ time.*

Proof. According to Theorem 2.2, we need to solve $O(\sigma_1^{-2}\sigma_2^{-2}\varepsilon^{-2})$ individual problems. For each, we have $m = O((\sigma_1^{-2}+\sigma_2^{-2})\varepsilon^{-1})$, so using Theorem 4.1 we solve them in $O(((\sigma_1^{-2}+\sigma_2^{-2})\varepsilon^{-1}+n)^2)$ time. This leads to $O(\sigma_1^{-2}\sigma_2^{-2}\varepsilon^{-2}((\sigma_1^{-2}+\sigma_2^{-2})\varepsilon^{-1}+n)^2)$ running time . □

Acknowledgments. K.B., I.K., and M.L. are supported by the Netherlands Organisation for Scientific Research (NWO) under grant no. 612.001.207, 612.001.106, and 639.021.123, respectively. R.S. was funded by Portuguese funds through CIDMA and FCT, within project PEst-OE/MAT/UI4106/2014, and by FCT grant SFRH/BPD/88455/2012. In addition, R.S. was partially supported by projects MINECO MTM2012-30951/FEDER, Gen. Cat. DGR2009SGR1040, and by ESF EUROCORES program EuroGIGA-ComPoSe IP04-MICINN project EUI-EURC-2011-4306.

References

[1] P. K. Agarwal, S.-W. Cheng, Y. Tao, and K. Yi. Indexing uncertain data. In *PODS*, pages 137–146, 2009.

[2] P. Agrawal, O. Benjelloun, A. D. Sarma, C. Hayworth, S. Nabar, T. Sugihara, and J. Widom. Trio: A system for data, uncertainty, and lineage. In *PODS*, 2006.

[3] D. Bandyopadhyay and J. Snoeyink. Almost-Delaunay simplices: Nearest neighbor relations for imprecise points. In *SODA*, pages 410–419, 2004.

[4] J. Bi and T. Zhang. Support vector classification with input data uncertainty. In *NIPS*, 2004.

[5] G. Cormode and M. Garafalakis. Histograms and wavelets of probabilitic data. In *ICDE*, 2009.

[6] G. Cormode, F. Li, and K. Yi. Semantics of ranking queries for probabilistic data and expected ranks. In *ICDE*, 2009.

[7] G. Cormode and A. McGregor. Approximation algorithms for clustering uncertain data. In *PODS*, 2008.

[8] N. Dalvi and D. Suciu. Efficient query evaluation on probabilitic databases. *The VLDB Journal*, 16:523–544, 2007.

[9] L. J. Guibas, D. Salesin, and J. Stolfi. Epsilon geometry: building robust algorithms from imprecise computations. In *SoCG*, pages 208–217, 1989.

[10] L. J. Guibas, D. Salesin, and J. Stolfi. Constructing strongly convex approximate hulls with inaccurate primitives. *Algorithmica*, 9:534–560, 1993.

[11] A. Jørgensen, M. Löffler, and J. Phillips. Geometric computations on indecisive points. In *WADS*, pages 536–547, 2011.

[12] M. Löffler. *Data Imprecision in Computational Geometry.* PhD thesis, Utrecht University, 2009.

[13] M. Löffler and J. Phillips. Shape fitting on point sets with probability distributions. In *ESA*, pages 313–324, 2009.

[14] T. Nagai and N. Tokura. Tight error bounds of geometric problems on convex objects with imprecise coordinates. In *Jap. Conf. on Discrete and Comput. Geom.*, pages 252–263, 2000.

[15] Y. Ostrovsky-Berman and L. Joskowicz. Uncertainty envelopes. In *EuroCG*, pages 175–178, 2005.

[16] Y. Tao, R. Cheng, X. Xiao, W. K. Ngai, B. Kao, and S. Prabhakar. Indexing multi-dimensional uncertain data with arbitrary probability density functions. In *VLDB*, 2005.

[17] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan. The unscented particle filter. In *Adv. Neural Inf. Process. Syst.*, volume 8, pages 351–357, 2000.

[18] Y. Zou and K. Chakrabarty. Uncertainty-aware and coverage-oriented deployment of sensor networks. *J. Parallel Distrib. Comput*, pages 788–798, 2004.

A Integral

Here we'll show how to calculate the following integral for a cell C of the partition L^* of in the dual space:

$$I = \iint\limits_C \left(\int\limits_{X_1(\alpha,\beta)}^{X_2(\alpha,\beta)} \int\limits_{X_3(\alpha,\beta)}^{X_4(\alpha,\beta)} (x_2 - x_1) \mathrm{d}x_2 \mathrm{d}x_1 \right) \mathrm{d}\alpha \mathrm{d}\beta \, .$$

Suppose lines corresponding to C intersect four segments s_1, s_2, s_3, and s_4 that belong to the lines with the following formulas:

$$\begin{aligned} a_1 x + b_1 y + c_1 = 0 \, , \quad a_2 x + b_2 y + c_2 = 0 \, , \\ a_3 x + b_3 y + c_3 = 0 \, , \quad a_4 x + b_4 y + c_4 = 0 \, . \end{aligned}$$

Then, the limits of integration can be expressed as:

$$\begin{aligned} X_1(\alpha,\beta) = \frac{b_1\beta - c_1}{b_1\alpha + a_1}, \quad X_2(\alpha,\beta) = \frac{b_2\beta - c_2}{b_2\alpha + a_2} \, , \\ X_3(\alpha,\beta) = \frac{b_3\beta - c_3}{b_3\alpha + a_3}, \quad X_4(\alpha,\beta) = \frac{b_4\beta - c_4}{b_4\alpha + a_4} \, . \end{aligned}$$

After solving the inner two integrals we get:

$$\begin{aligned} I &= \iint\limits_C \frac{(X_2 - X_1)(X_4 - X_3)(X_3 + X_4 - X_1 - X_2)}{2} \mathrm{d}\alpha \mathrm{d}\beta = \\ &= \frac{1}{2} \iint\limits_C \big(- X_1^2 X_3 + X_2^2 X_3 + X_1 X_3^2 - X_2 X_3^2 \\ &\qquad + X_1^2 X_4 - X_2^2 X_4 - X_1 X_4^2 + X_2 X_4^2 \big) \mathrm{d}\alpha \mathrm{d}\beta \, . \end{aligned}$$

For some i and j:

$$X_i X_j^2 = \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} \, .$$

Denote I_{ij} to be:

$$\begin{aligned} I_{ij} &= \iint\limits_C X_i X_j^2 \mathrm{d}\alpha \mathrm{d}\beta = \\ &= \iint\limits_C \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} \mathrm{d}\alpha \mathrm{d}\beta = \\ &= \sum_{C_v \subset C} \int\limits_{\alpha_1}^{\alpha_2} \left(\int\limits_{A_1\alpha + B_1}^{A_2\alpha + B_2} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} \mathrm{d}\beta \right) \mathrm{d}\alpha \, , \end{aligned}$$

where cell C is split into vertical splines C_v, with each C_v bounded by left and right vertical segments with α-coordinates equal to α_1 and α_2, and bottom and top segments defined by formulas $\beta = A_1\alpha + B_1$ and $\beta = A_2\alpha + B_2$. Then,

$$I = \frac{1}{2} \left(I_{13} - I_{31} - I_{23} + I_{32} - I_{14} + I_{41} + I_{24} - I_{42} \right) .$$

Denote $F_{ij}(\alpha)$ to be an indefinite integral with additive constant equal to 0:

$$F_{ij}(\alpha) = \int \left(\int\limits_{A_1\alpha + B_1}^{A_2\alpha + B_2} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} \mathrm{d}\beta \right) \mathrm{d}\alpha \, . \tag{3}$$

In the general case, when $b_i \neq 0$, $b_j \neq 0$, and $a_i/b_i \neq a_j/b_j$,

$$
\begin{aligned}
F_{ij}(\alpha) =& \frac{\log\left(a_i+\alpha b_i\right)}{12 b_i^2\left(a_j b_i-a_i b_j\right)^2}\Big[3\left(A_2^4-A_1^4\right) a_i^4 b_j^2-12 a_i^3\left(A_2^3 B_2-A_1^3 B_1\right) b_i b_j^2 \\
&\quad +18 a_i^2\left(A_2^2 B_2^2-A_1^2 B_1^2\right) b_i^2 b_j^2-12 a_i\left(A_2 B_2^3-A_1 B_1^3\right) b_i^3 b_j^2+3\left(B_2^4-B_1^4\right) b_i^4 b_j^2 \\
&\quad +\left(b_j c_i+2 b_i c_j\right)\Big(4\left(A_2^3-A_1^3\right) a_i^3 b_j-12 a_i^2\left(A_2^2 B_2-A_1^2 B_1\right) b_i b_j \\
&\qquad +12 a_i\left(A_2 B_2^2-A_1 B_1^2\right) b_i^2 b_j-4\left(B_2^3-B_1^3\right) b_i^3 b_j\Big) \\
&\quad +\left(2 b_j c_i+b_i c_j\right)\Big(6\left(A_2^2-A_1^2\right) a_i^2 b_i c_j-12 a_i\left(A_2 B_2-A_1 B_1\right) b_i^2 c_j+6\left(B_2^2-B_1^2\right) b_i^3 c_j\Big) \\
&\quad +12\left(A_2-A_1\right) a_i b_i^2 c_i c_j^2-12\left(B_2-B_1\right) b_i^3 c_i c_j^2\Big] \\
&+\frac{\log\left(a_j+\alpha b_j\right)}{12 b_j^2\left(a_j b_i-a_i b_j\right)^2}\Big[3\left(A_2^4-A_1^4\right) a_j^3 b_i\left(3 a_j b_i-4 a_i b_j\right)+12 a_j^2\left(A_2^3 B_2-A_1^3 B_1\right) b_i b_j\left(3 a_i b_j-2 a_j b_i\right) \\
&\quad +18 a_j\left(A_2^2 B_2^2-A_1^2 B_1^2\right) b_i b_j^2\left(a_j b_i-2 a_i b_j\right)+12 a_i\left(A_2 B_2^3-A_1 B_1^3\right) b_i b_j^4-3\left(B_2^4-B_1^4\right) b_i^2 b_j^4 \\
&\quad +\left(b_j c_i+2 b_i c_j\right)\Big(4\left(A_2^3-A_1^3\right) a_j^2\left(2 a_j b_i-3 a_i b_j\right)-12 a_j\left(A_2^2 B_2-A_1^2 B_1\right) b_j\left(a_j b_i-2 a_i b_j\right) \\
&\qquad -12 a_i\left(A_2 B_2^2-A_1 B_1^2\right) b_j^3+4\left(B_2^3-B_1^3\right) b_i b_j^3\Big) \\
&\quad +\left(2 b_j c_i+b_i c_j\right)\Big(6\left(A_2^2-A_1^2\right) a_j c_j\left(a_j b_i-2 a_i b_j\right)+12 a_i\left(A_2 B_2-A_1 B_1\right) b_j^2 c_j-6\left(B_2^2-B_1^2\right) b_i b_j^2 c_j\Big) \\
&\quad -12 a_i b_j^2 c_i c_j^2\left(A_2-A_1\right)-12 b_i b_j^2 c_i c_j^2\left(B_2-B_1\right)\Big] \\
&+\frac{1}{24 b_i b_j^2\left(a_j b_i-a_i b_j\right)\left(a_j+\alpha b_j\right)}\Big[3 \alpha^3 b_i b_j^3\left(a_j b_i-a_i b_j\right)\left(A_2^4-A_1^4\right) \\
&\quad +\alpha^2\left(a_i b_j-a_j b_i\right)\Big(3 b_j^2\left(3 a_j b_i+2 a_i b_j\right)\left(A_2^4-A_1^4\right)-24 b_i b_j^3\left(A_2^3 B_2-A_1^3 B_1\right) \\
&\qquad +8 b_j^2\left(b_j c_i+2 b_i c_j\right)\left(A_2^3-A_1^3\right)\Big) \\
&\quad -\alpha\left(a_j b_i-a_i b_j\right)\Big(6 a_j b_j\left(2 a_j b_i+a_i b_j\right)\left(A_2^4-A_1^4\right)-24 a_j b_i b_j^2\left(A_2^3 B_2-A_1^3 B_1\right) \\
&\qquad +8 a_j b_j\left(b_j c_i+2 b_i c_j\right)\left(A_2^3-A_1^3\right)\Big) \\
&\quad +6\left(A_2^4-A_1^4\right) a_j^4 b_i^2-24 a_j^3\left(A_2^3 B_2-A_1^3 B_1\right) b_i^2 b_j \\
&\quad +36 a_j^2\left(A_2^2 B_2^2-A_1^2 B_1^2\right) b_i^2 b_j^2-24 a_j\left(A_2 B_2^3-A_1 B_1^3\right) b_i^2 b_j^3+6\left(B_2^4-B_1^4\right) b_i^2 b_j^4 \\
&\quad +8\left(A_2^3-A_1^3\right) a_j^3 b_i\left(b_j c_i+2 b_i c_j\right)-24 a_j^2\left(A_2^2 B_2-A_1^2 B_1\right) b_i b_j\left(b_j c_i+2 b_i c_j\right) \\
&\quad +24 a_j\left(A_2 B_2^2-A_1 B_1^2\right) b_i b_j^2\left(b_j c_i+2 b_i c_j\right)-8\left(B_2^3-B_1^3\right) b_i b_j^3\left(b_j c_i+2 b_i c_j\right) \\
&\quad +12\left(A_2^2-A_1^2\right) a_j^2 b_i c_j\left(2 b_j c_i+b_i c_j\right)-24 a_j\left(A_2 B_2-A_1 B_1\right) b_i b_j c_j\left(2 b_j c_i+b_i c_j\right) \\
&\quad +12\left(B_2^2-B_1^2\right) b_i b_j^2 c_j\left(2 b_j c_i+b_i c_j\right)+24\left(A_2-A_1\right) a_j b_i b_j c_i c_j^2-24\left(B_2-B_1\right) b_i b_j^2 c_i c_j^2\Big].
\end{aligned}
$$

If segment i is pointing towards segment j (line drawn through i intersects j), and one of the corners of the integration spline corresponds to the line going through i, then the following equalities hold

$$
\begin{aligned}
a_i+\alpha' b_i &= 0\,, \\
A_1 a_i-B_1 b_i+c_i &= 0\,, \\
A_2 a_i-B_2 b_i+c_i &= 0\,,
\end{aligned}
$$

where α' corresponds to the corner of the spline. In that case,

$$
\begin{aligned}
F_{ij}(\alpha) = & \frac{\log\left(a_j + \alpha b_j\right)}{12 b_i^2 b_j^2} \Big[9\left(A_2^4 - A_1^4\right)\left(a_j b_i - a_i b_j\right)^2 + 16\left(A_2^3 - A_1^3\right)\left(a_j b_i - a_i b_j\right)\left(b_i c_j - b_j c_i\right) \\
& \qquad + 6\left(A_2^2 - A_1^2\right)\left(b_j c_i - b_i c_j\right)^2 \Big] \\
& + \frac{1}{24 b_i^3 b_j^2 \left(a_j + \alpha b_j\right)} \Big[3\alpha^4 b_i^3 b_j^3 \left(A_2^4 - A_1^4\right) \\
& \qquad - \alpha^2 b_i^2 b_j^2 \Big(9\left(a_j b_i - 2 a_i b_j\right)\left(A_2^4 - A_1^4\right) + 16\left(b_i c_j - b_j c_i\right)\left(A_2^3 - A_1^3\right) \Big) \\
& \qquad + 2 a_j b_i^2 b_j \alpha \Big(3\left(3 a_i b_j - 2 a_j b_i\right)\left(A_2^4 - A_1^4\right) - 8\left(b_i c_j - b_j c_i\right)\left(A_2^3 - A_1^3\right) \Big) \\
& \qquad + \left(6 a_j^2 b_i^2 - 9 a_i^2 b_j^2\right)\left(A_2^4 - A_1^4\right) + 16 a_j b_i \left(b_i c_j - b_j c_i\right)\left(A_2^3 - A_1^3\right) + 12\left(b_j c_i - b_i c_j\right)^2\left(A_2^2 - A_1^2\right) \Big] .
\end{aligned}
$$

If segment j is pointing towards segment i (line drawn through j intersects i), and one of the corners of the integration spline corresponds to the line going through j, then the following equalities hold

$$
\begin{aligned}
a_j + \alpha' b_j &= 0 , \\
A_1 a_j - B_1 b_j + c_j &= 0 , \\
A_2 a_j - B_2 b_j + c_j &= 0 ,
\end{aligned}
$$

where α' corresponds to the corner of the spline. In that case,

$$
\begin{aligned}
F_{ij}(\alpha) = & \frac{\log\left(a_i + \alpha b_i\right)}{12 b_i^2 b_j^2} \Big[4\left(A_2^3 - A_1^3\right)\left(a_j b_i - a_i b_j\right)\left(b_i c_j - b_j c_i\right) + 3\left(A_2^4 - A_1^4\right)\left(a_j b_i - a_i b_j\right)^2 \Big] \\
& + \frac{1}{24 b_i^2 b_j^2} \Big[3\alpha^2 \left(A_2^4 - A_1^4\right) b_i b_j + \alpha \Big(\left(A_2^4 - A_1^4\right)\left(12 a_j b_i - 6 a_i b_j\right) - 8\left(A_2^3 - A_1^3\right)\left(b_j c_i - b_i c_j\right) \Big) \Big] .
\end{aligned}
$$

In case when $b_i \neq 0$, $b_j \neq 0$, but lines are parallel ($a_i/b_i = a_j/b_j$),

$$
\begin{aligned}
F_{ij}(\alpha) =& \frac{\log(a_i+\alpha b_i)}{2b_i^2 b_j^2}\Big[3\left(A_2^4 - A_1^4\right) a_i^2 b_j^2 - 6a_i\left(A_2^3 B_2 - A_1^3 B_1\right) b_i b_j^2 + 3\left(A_2^2 B_2^2 - A_1^2 B_1^2\right) b_i^2 b_j^2 \\
&\quad + 2\left(A_2^3 - A_1^3\right) a_i b_j\left(b_j c_i + 2b_i c_j\right) - 2\left(A_2^2 B_2 - A_1^2 B_1\right) b_i b_j\left(b_j c_i + 2b_i c_j\right) \\
&\quad + \left(A_2^2 - A_1^2\right) b_i c_j\left(2b_j c_i + b_i c_j\right)\Big] \\
&+ \frac{1}{24 b_i^2 b_j^2 (a_i+\alpha b_i)^2}\Big[3\alpha^4\left(A_1^4 - A_2^4\right) b_i^4 b_j^2 \\
&\quad + \alpha^3\Big(12 a_i b_i^3 b_j^2 (A_2^4 - A_1^4) - 24 b_i^4 b_j^2 (A_2^3 B_2 - A_1^3 B_1) + 8 b_i^3 b_j\left(b_j c_i + 2b_i c_j\right)\left(A_2^3 - A_1^3\right)\Big) \\
&\quad + \alpha^2\Big(33 a_i^2 b_i^2 b_j^2 (A_2^4 - A_1^4) - 48 a_i b_i^3 b_j^2 (A_2^3 B_2 - A_1^3 B_1) + 16 a_i b_i^2 b_j\left(b_j c_i + 2b_i c_j\right)\left(A_2^3 - A_1^3\right)\Big) \\
&\quad - \alpha\Big(6 a_i^3 b_i b_j^2 (A_2^4 - A_1^4) - 48 a_i^2 b_i^2 b_j^2 (A_2^3 B_2 - A_1^3 B_1) + 72 a_i b_i^3 b_j^2 (A_2^2 B_2^2 - A_1^2 B_1^2) \\
&\qquad - 24 b_i^4 b_j^2 (A_2 B_2^3 - A_1 B_1^3) + 16 a_i^2 b_i b_j\left(b_j c_i + 2b_i c_j\right)\left(A_2^3 - A_1^3\right) \\
&\qquad - 48 a_i b_i^2 b_j\left(b_j c_i + 2b_i c_j\right)\left(A_2^2 B_2 - A_1^2 B_1\right) + 24 b_i^3 b_j\left(b_j c_i + 2b_i c_j\right)\left(A_2 B_2^2 - A_1 B_1^2\right) \\
&\qquad + 24 a_i b_i^2 c_j\left(2b_j c_i + b_i c_j\right)\left(A_2^2 - A_1^2\right) - 24 b_i^3 c_j\left(2b_j c_i + b_i c_j\right)\left(A_2 B_2 - A_1 B_1\right) \\
&\qquad + 24 b_i^3 c_i c_j^2\left(A_2 - A_1\right)\Big) \\
&\quad - 21 a_i^4 b_j^2\left(A_2^4 - A_1^4\right) + 60 a_i^3 b_i b_j^2\left(A_2^3 B_2 - A_1^3 B_1\right) - 54 a_i^2 b_i^2 b_j^2\left(A_2^2 B_2^2 - A_1^2 B_1^2\right) \\
&\quad + 12 a_i b_i^3 b_j^2\left(A_2 B_2^3 - A_1 B_1^3\right) + 3 b_i^4 b_j^2\left(B_2^4 - B_1^4\right) \\
&\quad - \left(b_j c_i + 2b_i c_j\right)\Big(20 a_i^3 b_j\left(A_2^3 - A_1^3\right) - 36 a_i^2 b_i b_j\left(A_2^2 B_2 - A_1^2 B_1\right) \\
&\qquad + 12 a_i b_i^2 b_j\left(A_2 B_2^2 - A_1 B_1^2\right) + 4 b_i^3 b_j\left(B_2^3 - B_1^3\right)\Big) \\
&\quad - \left(2b_j c_i + b_i c_j\right)\Big(18 a_i^2 b_i c_j\left(A_2^2 - A_1^2\right) - 12 a_i b_i^2 c_j\left(A_2 B_2 - A_1 B_1\right) - 6 b_i^3 c_j\left(B_2^2 - B_1^2\right)\Big) \\
&\quad - 12 a_i b_i^2 c_i c_j^2\left(A_2 - A_1\right) - 12 b_i^3 c_i c_j^2\left(B_2 - B_1\right)\Big].
\end{aligned}
$$

If $b_i = 0$ (segment i is vertical), and $b_j \neq 0$, then

$$
\begin{aligned}
F_{ij}(\alpha) =& \frac{\log(a_j+\alpha b_j)}{a_i b_j^2}\Big[-a_j^2 c_i\left(A_2^3 - A_1^3\right) + 2a_j b_j c_i\left(A_2^2 B_2 - A_1^2 B_1\right) - b_j^2 c_i\left(A_2 B_2^2 - A_1 B_1^2\right) \\
&\quad - 2a_j c_i c_j\left(A_2^2 - A_1^2\right) + 2b_j c_i c_j\left(A_2 B_2 - A_1 B_1\right) - c_i c_j^2\left(A_2 - A_1\right)\Big] \\
&+ \frac{1}{6 a_i b_j^2 (a_j+\alpha b_j)}\Big[-\alpha^3 b_j^3 c_i\left(A_2^3 - A_1^3\right) \\
&\quad + 3\alpha b_j^2 c_i\Big(a_j\left(A_2^3 - A_1^3\right) + 2b_j\left(A_1^2 B_1 - A_2^2 B_2\right) + 2c_j\left(A_2^2 - A_1^2\right)\Big) \\
&\quad + 2\alpha a_j b_j c_i\Big(2a_j\left(A_2^3 - A_1^3\right) + 3b_j\left(A_1^2 B_1 - A_2^2 B_2\right) + 3c_j\left(A_2^2 - A_1^2\right)\Big) \\
&\quad - 2a_j^3 c_i\left(A_2^3 - A_1^3\right) + 6a_j^2 b_j c_i\left(A_2^2 B_2 - A_1^2 B_1\right) - 6a_j b_j^2 c_i\left(A_2 B_2^2 - A_1 B_1^2\right) + 2b_j^3 c_i\left(B_2^3 - B_1^3\right) \\
&\quad - 6a_j^2 c_i c_j\left(A_2^2 - A_1^2\right) + 12 a_j b_j c_i c_j\left(A_2 B_2 - A_1 B_1\right) - 6 b_j^2 c_i c_j\left(B_2^2 - B_1^2\right) \\
&\quad - 6a_j c_i c_j^2\left(A_2 - A_1\right) + 6 b_j c_i c_j^2\left(B_2 - B_1\right)\Big].
\end{aligned}
$$

If segment j is pointing towards segment i:

$$\begin{aligned} a_j + \alpha' b_j &= 0\,, \\ A_1 a_j - B_1 b_j + c_j &= 0\,, \\ A_2 a_j - B_2 b_j + c_j &= 0\,, \end{aligned}$$

where α' corresponds to the corner of the spline. In that case,

$$F_{ij}(\alpha) = \frac{\alpha c_i \left(2a_j + \alpha b_j\right) \left(A_1^3 - A_2^3\right)}{6 a_i b_j}\,.$$

If $b_i \neq 0$, and $b_j = 0$ (segment j is vertical), then

$$\begin{aligned} F_{ij}(\alpha) =& \frac{\log\left(a_i + \alpha b_i\right)}{2a_j^2 b_i^2} \Big[a_i^2 c_j^2 \left(A_2^2 - A_1^2\right) - 2a_i b_i c_j^2 \left(A_2 B_2 - A_1 B_1\right) + b_i^2 c_j^2 \left(B_2^2 - B_1^2\right) \\ & \qquad + 2a_i c_i c_j^2 \left(A_2 - A_1\right) - 2b_i c_i c_j^2 \left(B_2 - B_1\right) \Big] \\ & + \frac{1}{6 a_i b_j^2 \left(a_j + \alpha b_j\right)} \Big[\alpha^2 \left(A_2^2 - A_1^2\right) b_i c_j^2 \\ & \qquad - 2\alpha c_j^2 \Big(a_i \left(A_2^2 - A_1^2\right) - 2b_i \left(A_2 B_2 - A_1 B_1\right) + 2c_i \left(A_2 - A_1\right) \Big) \Big]\,. \end{aligned}$$

If segment i is pointing towards segment j:

$$\begin{aligned} a_i + \alpha' b_i &= 0\,, \\ A_1 a_i - B_1 b_i + c_i &= 0\,, \\ A_2 a_i - B_2 b_i + c_i &= 0\,, \end{aligned}$$

where α' corresponds to the corner of the spline. In that case,

$$F_{ij}(\alpha) = \frac{\alpha c_j^2 \left(2a_i + \alpha b_i\right) \left(A_2^2 - A_1^2\right)}{4 a_j^2 b_i}\,.$$

If $b_i = 0$ and $b_j = 0$ (both segments are vertical), then

$$F_{ij}(\alpha) = -\frac{c_i c_j^2 \left(\alpha^2 (A_2 - A_1) + 2\alpha (B_2 - B_1)\right)}{2 a_i a_j^2}\,.$$