



Region-Based Approximation of Probability Distributions (for Visibility Between Imprecise Points Among Obstacles)

Kevin Buchin¹ · Irina Kostitsyna¹  · Maarten Löffler² · Rodrigo I. Silveira³

Received: 28 December 2016 / Accepted: 21 January 2019 / Published online: 16 February 2019
© The Author(s) 2019

Abstract

Let p and q be two imprecise points, given as probability density functions on \mathbb{R}^2 , and let \mathcal{O} be a set of disjoint polygonal obstacles in \mathbb{R}^2 . We study the problem of approximating the probability that p and q can see each other; i.e., that the segment connecting p and q does not cross any obstacle in \mathcal{O} . To solve this problem, we first approximate each density function by a weighted set of polygons. Then we focus on computing the visibility between two points inside two of such polygons, where we can assume that the points are drawn uniformly at random. We show how this problem can be solved exactly in $O((n + m)^2)$ time, where n and m are the total complexities of the two polygons and the set of obstacles, respectively. Using this as a subroutine, we show that the probability that p and q can see each other amidst a set of obstacles of total complexity m can be approximated within error ε in $O(1/\varepsilon^3 + m^2/\varepsilon^2)$ time.

Keywords Visibility in polygonal domains · Probability distribution · Gaussian distribution

A preliminary version appeared at the 17th Workshop on Algorithm Engineering and Experiments (ALENEX) [6].

✉ Irina Kostitsyna
i.kostitsyna@tue.nl

Kevin Buchin
k.a.buchin@tue.nl

Maarten Löffler
m.loffler@uu.nl

Rodrigo I. Silveira
rodrigo.silveira@upc.edu

- ¹ Department of Mathematics and Computer Science, TU Eindhoven, Eindhoven, The Netherlands
- ² Department of Computing and Information Sciences, Utrecht University, Utrecht, The Netherlands
- ³ Departament de Matemàtiques, Universitat Politècnica de Catalunya, Barcelona, Spain

1 Introduction

Imprecision appears naturally in many applications. Imprecise points play an important role in databases [1,2,7–10,22], machine learning [4], and sensor networks [26], where a limited number of probes from a certain data set is gathered, each potentially representing the true location of a data point. Alternatively, imprecise points may be obtained from inaccurate measurements or may be the result of earlier inexact computations.

Data imprecision is an important obstacle to the application of geometric algorithms to real-world problems, since geometric algorithms typically assume exact input. In the computational geometry literature, various models to deal with data imprecision have been suggested. The most studied one is based on using regions to represent imprecise points. For instance, a set of imprecise points may be represented by a set of disks, where the (unknown) true location of a point is assumed to be somewhere within its corresponding disk. In such models the goal is usually to obtain worst-case bounds on the values of geometric measures [25].

In this paper we take a more general approach, we describe the location of each point by a probability distribution μ_i (for instance by a Gaussian distribution). This model has been rarely worked with directly because of the computational difficulties arising from its generality.

The standard technique to handle these difficulties is to approximate the distributions by point sets. For instance, for tracking uncertain objects a particle filter uses a discrete set of locations to model uncertainty [23]. Löffler and Phillips [17] and Jørgenson et al. [15] 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, they show in [17] how to compute for an xy -monotone function F (such as a cumulative probability density function) an ε -quantization: a 2-dimensional point set P such that for every point q in the plane the fraction of points in P that are Pareto-dominated¹ by q differs from $F(q)$ by at most ε .

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 discretize the distributions, the choice of points may greatly influence the resulting probability, as illustrated in Fig. 1. Therefore, approximating the distributions with point sets does not provide an approximation of the actual probability that the points see each other.

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. [12], motivated by finite coordinate precision, and later studied extensively in a variety of settings [3,13,16,18,19].

In this paper we approximate each density function by a weighted set of polygonal regions, and apply this to approximate the probability that two imprecise points can

¹ That is, having larger x - and y - coordinates.

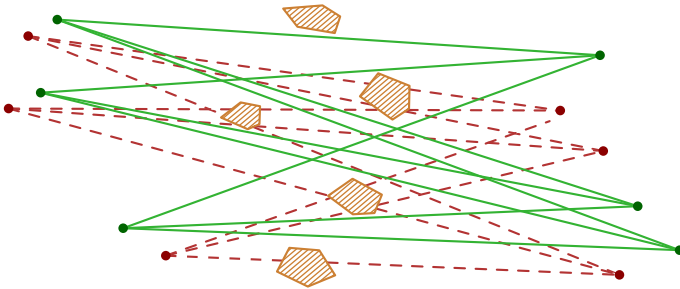


Fig. 1 Two pairs of point sets on opposite sides of a collection of obstacles. The green points can all see each other, whereas none of the red points can (Color figure online)

see each other. As part of our results, we introduce a novel technique to represent the placement space of pairs of points that can see each other amidst a set of disjoint obstacles. We believe this technique is interesting in its own right. For example, we show that it can be applied to compute the probability that two points inside a polygon see each other. Rote [20] suggested to use this probability as a measure of *degree of convexity* of a polygon and showed how to compute it in $O(n^9)$ time. Later it was observed that the running time of Rote’s algorithm can be reduced to $O(n^7)$ [21]. Applying our technique we can compute the degree of convexity in $O(n^2)$ time.

1.1 Results and Paper Outline

We are given two imprecise points p and q , modeled as probability density functions over \mathbb{R}^2 , and a set of disjoint polygonal obstacles \mathcal{O} , also in \mathbb{R}^2 .

Our goal is to compute the probability that p and q can see each other, or equivalently, that the segment connecting p and q does not cross any obstacle in \mathcal{O} .

The main result presented in this paper is that, if the probability density functions are Gaussian distributions, then this probability can be approximated efficiently. More precisely, our main result is the following.

Theorem 1 *Given two imprecise points p and q , modeled as Gaussian distributions μ_1 and μ_2 , and a set of obstacles of total complexity m , we can approximate the probability that p and q see each other with an error at most ε in $O(1/\varepsilon^3 + m^2/\varepsilon^2)$ time.*

Our approach is based on ε -approximating a Gaussian distribution using a set of $O(1/\varepsilon)$ convex weighted polygons, with $O(1/\sqrt{\varepsilon})$ complexity each. More precisely, the set of weighted polygons defines a step function that approximates the Gaussian probability density function in terms of the *volumetric error*: the volume of the symmetric difference between the step function and the probability density function.² The derivation of a set of weighted polygons of appropriate complexity that guarantees the approximation factor is the main goal of Sect. 2. In that section we also briefly discuss other error measures that could be considered instead of the volumetric error.

² Volumetric error can also be viewed as an L_1 distance between the two functions, and is twice the total variation distance used.

Once the probability density function of each imprecise point is approximated by a set of weighted polygons, we can focus on visibility between two of such polygons, one for each imprecise point. The key difference is that now we can use uniform probability for each polygon, simplifying the computations considerably. Therefore the problem becomes that of computing the probability that two points drawn uniformly at random from each polygon can see each other. At this stage we assume that the obstacles in \mathcal{O} are disjoint polygons, and that they are convex (this can always be achieved after pre-processing).

The main result in Sect. 3 is how to compute the probability that two points chosen uniformly at random inside two convex polygons (with obstacles in between) can see each other. We show how to do this in $O((n + m)^2)$ time, for n and m the total complexities of the two polygons and the set of obstacles, respectively. This result is technically involved, and requires computing an integral, which we defer to Sect. 4, and partially to the Appendix.

Note that the approximation factor in our main theorem only comes from approximating a Gaussian distribution with a set of weighted regions. For two points p and q , uniformly distributed in two polygons, the method presented in Sect. 3 allows to calculate the probability of them seeing each other exactly. We obtain a closed form expression for the probability making use of arithmetic operations and the logarithm function. Throughout the text we assume that the model of computation is the real RAM, standard in computational geometry, with the addition of the logarithm function.

Finally, once that the results in Sects. 2, 3 and 4 are in place, we show in Sect. 5 that it is straightforward to combine them to obtain our main result, Theorem 1.

2 Region-Based Approximation

Let \mathcal{P} be a set of weighted regions in the plane, and let $w(P) \geq 0$ denote the weight of a region $P \in \mathcal{P}$. Let $\mathcal{P}(p) = \{P \in \mathcal{P} \mid p \in P\}$ be the regions of \mathcal{P} containing point $p \in \mathbb{R}^2$. Set \mathcal{P} defines a step function

$$f_{\mathcal{P}}(p) = \sum_{P \in \mathcal{P}(p)} w(P),$$

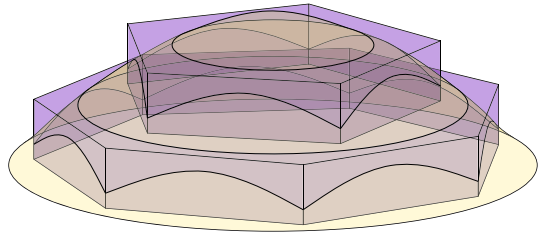
that sums the weights of all regions containing p . We will use the step function $f_{\mathcal{P}}$ to approximate a given density function, however, we will not require that $f_{\mathcal{P}}(p) \in [0, 1]$ or that $\int_{p \in \mathbb{R}^2} f_{\mathcal{P}}(p) = 1$.

To measure the quality of an approximation we will use the *volumetric error* between the approximation function and the distribution, i.e., the total volume of the symmetric difference between $f_{\mathcal{P}}$ and μ (refer to Fig. 2). More precisely, we say that

Definition 1 A step function $f_{\mathcal{P}}$ defined by a set of weighted regions \mathcal{P} ε -approximates a probability distribution with density function μ if

$$\int_{p \in \mathbb{R}^2} |\mu(p) - f_{\mathcal{P}}(p)| \leq \varepsilon.$$

Fig. 2 A probability density function μ (yellow) can be approximated by a set of weighted regions \mathcal{P} , representing a step function $f_{\mathcal{P}}$ (purple) (Color figure online)



Additive or Multiplicative? There are other, possibly more straightforward, approaches to approximating a given density function. Why did we opt for using the volumetric error instead of, for example, choosing a local *multiplicative* or *additive* approximation function? Next, we will formally introduce and discuss both approaches, and demonstrate why in our case a third approach was needed.

Consider a domain $D \subseteq \mathbb{R}^2$. We say that $f_{\mathcal{P}}$, for a given set of weighted regions \mathcal{P} , is a *local multiplicative δ -approximation* of μ on domain D if, for all points $p \in D$,

$$(1 - \delta)\mu(p) \leq f_{\mathcal{P}}(p) \leq (1 + \delta)\mu(p).$$

It is easy to verify that a local multiplicative approximation function is also a global approximation function, i.e., the total error between $f_{\mathcal{P}}$ and μ is bounded. However, observe, that in this case, no finite set \mathcal{P} can be a local multiplicative approximation of many natural distributions (like Gaussians, for instance). This is because no step function $f_{\mathcal{P}}$ with finite number of steps can be inscribed between functions $(1 - \delta)\mu(p)$ and $(1 + \delta)\mu(p)$, where $\mu(p)$ is a Gaussian distribution.

We say that $f_{\mathcal{P}}$, for a given set of weighted regions \mathcal{P} , is a *local additive δ -approximation* of μ on domain D if, for all points $p \in D$,

$$|\mu(p) - f_{\mathcal{P}}(p)| \leq \delta.$$

Finding set of such regions \mathcal{P} that locally additive δ -approximate a given distribution function is straightforward. However, bounding the absolute distance between $f_{\mathcal{P}}$ and μ at every point p in the plane implies no guarantee on the total error of the approximation.³ Figure 3 illustrates the difference between the locally additive and locally multiplicative approximation approaches.

As local approximations do not lead to a good global approximation, it is necessary to use a global method of bounding the total error. The total error between an approximation function and the distribution function is exactly the volume of the symmetric difference between the two. Thus, the total error of approximation of a distribution function μ by any function $f_{\mathcal{P}}$, that is an ε -approximation of μ by Definition 1, is bounded by ε . Any probability distribution μ can be approximated in this way, but the total complexity of \mathcal{P} , i.e., the sum of the complexities of each of its regions, depends on various factors: the shape of μ , the shape of the allowed regions in \mathcal{P} , and the error parameter ε . To focus the discussion, in this work we limit our attention to

³ An earlier version of this document [5] mistakenly claimed that local *additive* approximations imply global approximations.

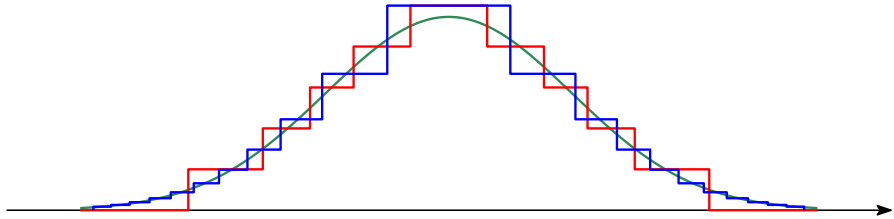


Fig. 3 Illustration of the difference between additive (red) and multiplicative (blue) approximations of the same (1-dimensional) Gaussian function (green). While similar at the center, at the tails the additive approximation drops to zero while the multiplicative one continues to add steps (Color figure online)

Gaussian distributions, since they are natural and have been shown to be appropriate for modeling the uncertainty in commonly-used types of location data, like location information obtained by a GPS (Global Positioning System) receiver [14,24].

In the following two sections we show how to find a set of polygons that ϵ -approximates a given 2-dimensional radially symmetric Gaussian distribution μ with covariance matrix $\Sigma = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$. First, we approximate μ with a set of concentric disks, and then, building up on this result, with a set of polygons. Note, that this method can also be applied to a general multivariate Gaussian distribution by first rescaling it to a radially symmetric one.

2.1 Approximating Gaussian Distributions with Disks

A natural way to approximate a Gaussian distribution with a set of regions is by using a set of concentric disks. Given a Gaussian distribution with probability density function μ , and a maximum allowed error ϵ , we would like to compute a set \mathcal{P} of k disks that ϵ -approximate μ . We may assume μ is centered at the origin, leaving only one parameter, the standard deviation σ , to govern the shape of μ ,

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

or in polar coordinates,

$$\mu(r, \theta) = \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}}.$$

The function μ does not depend on θ , therefore, in the following, we will omit it and write $\mu(r)$ for brevity.

We are looking for a set of radii r_1, \dots, r_k and corresponding weights w_1, \dots, w_k such that the set of disks centered at the origin with radii r_i and weights w_i ϵ -approximates μ . We use these disks to define a cylindrical step function $f_d(r)$. Figure 4 shows a 2-dimensional cross-section of the situation. Minimizing the volume between the step function and μ , we obtain the following lemma:

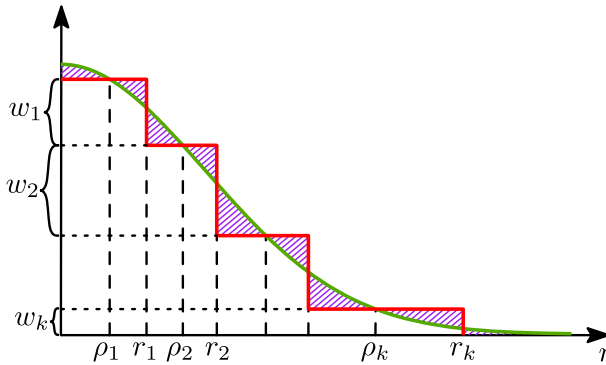


Fig. 4 Partial cross-section of a 2-dimensional Gaussian function (green) and the approximation function (red), illustrating the choice of radii (r_i) and weights (w_i) for the k regions in \mathcal{P} . We use ρ_i to indicate the i th radius where the approximation intersects with the Gaussian (Color figure online)

Lemma 1 *Given a Gaussian probability density function μ with standard deviation σ , and an integer k , a minimum-error approximation of μ by a cylindrical step function f_d , consisting of k disks with radii r_1, \dots, r_k centered at the origin with weights w_1, \dots, w_k respectively, is given by*

$$r_i = 2\sigma \sqrt{\log \frac{k+1}{k+1-i}}, \tag{1}$$

$$w_i = \frac{1}{\pi\sigma^2} \frac{(k+1-i)}{(k+1)^2}, \tag{2}$$

for all $1 \leq i \leq k$.

Proof Let $W_i = \sum_{j=i}^k w_j$. We introduce an additional set of parameters ρ_1, \dots, ρ_k , where ρ_i is such that $\mu(\rho_i) = W_i$, i.e., ρ_1, \dots, ρ_k are the radii of intersections of μ and the approximation function f_d (refer to Fig. 4). We will minimize the volume of the symmetric difference between the two functions over the $2k$ variables r_i and ρ_i , and derive the corresponding weights. Let $D(R)$ be the complement of the open disk of radius R centered at the origin, and let $V_D(R)$ be the volume under the probability density function μ on $D(R)$:

$$V_D(R) = \iint_{D(R)} \mu(x, y) \, dx \, dy = \int_0^{2\pi} \int_R^\infty \mu(r, \theta) r \, dr \, d\theta = \int_R^\infty \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \, dr = e^{-\frac{R^2}{2\sigma^2}}.$$

Then the volume of the symmetric difference between the functions μ and f_d can be found by the following formula:

$$V = (1 - V_D(\rho_1)) - W_1\pi\rho_1^2 + W_1\pi(r_1^2 - \rho_1^2) - (V_D(\rho_1) - V_D(r_1))$$

$$\begin{aligned}
 &+ (V_D(r_1) - V_D(\rho_2)) - W_2\pi \left(\rho_2^2 - r_1^2\right) \\
 &+ W_2\pi \left(r_2^2 - \rho_2^2\right) - (V_D(\rho_2) - V_D(r_2)) \\
 &+ \dots \\
 &+ (V_D(r_{k-1}) - V_D(\rho_k)) - W_k\pi \left(\rho_k^2 - r_{k-1}^2\right) \\
 &+ W_k\pi \left(r_k^2 - \rho_k^2\right) - (V_D(\rho_k) - V_D(r_k)) \\
 &+ V_D(r_k) \\
 = &1 + \pi \sum_{i=1}^{k-1} r_i^2 (\mu(\rho_i) + \mu(\rho_{i+1})) + \pi r_k^2 \mu(\rho_k) - 2\pi \sum_{i=1}^k \rho_i^2 \mu(\rho_i) \\
 &+ 2 \sum_{i=1}^k V_D(r_i) - 2 \sum_{i=1}^k V_D(\rho_i) \\
 = &1 + \frac{1}{2\sigma^2} \sum_{i=1}^{k-1} r_i^2 \left(e^{-\frac{\rho_i^2}{2\sigma^2}} + e^{-\frac{\rho_{i+1}^2}{2\sigma^2}} \right) + \frac{1}{2\sigma^2} r_k^2 e^{-\frac{\rho_k^2}{2\sigma^2}} - \frac{1}{\sigma^2} \sum_{i=1}^k \rho_i^2 e^{-\frac{\rho_i^2}{2\sigma^2}} \\
 &+ 2 \sum_{i=1}^k e^{-\frac{r_i^2}{2\sigma^2}} - 2 \sum_{i=1}^k e^{-\frac{\rho_i^2}{2\sigma^2}}. \tag{3}
 \end{aligned}$$

To minimize V , we compute the derivatives with respect to ρ_i and r_i :

$$\begin{aligned}
 \frac{dV}{d\rho_i} &= \begin{cases} \frac{\rho_i}{\sigma^4} e^{-\frac{\rho_i^2}{2\sigma^2}} \left(\rho_i^2 - \frac{1}{2} r_i^2 \right), & \text{for } i = 1 \\ \frac{\rho_i}{\sigma^4} e^{-\frac{\rho_i^2}{2\sigma^2}} \left(\rho_i^2 - \frac{1}{2} (r_{i-1}^2 + r_i^2) \right), & \text{for } 1 < i \leq k \end{cases} \\
 \frac{dV}{dr_i} &= \begin{cases} \frac{r_i}{\sigma^2} \left(e^{-\frac{\rho_i^2}{2\sigma^2}} + e^{-\frac{\rho_{i+1}^2}{2\sigma^2}} - 2e^{-\frac{r_i^2}{2\sigma^2}} \right), & \text{for } 1 \leq i < k \\ \frac{r_i}{\sigma^2} \left(e^{-\frac{\rho_i^2}{2\sigma^2}} - 2e^{-\frac{r_i^2}{2\sigma^2}} \right), & \text{for } i = k. \end{cases}
 \end{aligned}$$

Setting the derivatives to 0 results in the following identities:

$$2\rho_i^2 = \begin{cases} r_i^2, & \text{for } i = 1 \\ r_{i-1}^2 + r_i^2, & \text{for } 1 < i \leq k \end{cases} \tag{4}$$

$$2e^{-\frac{r_i^2}{2\sigma^2}} = \begin{cases} e^{-\frac{\rho_i^2}{2\sigma^2}} + e^{-\frac{\rho_{i+1}^2}{2\sigma^2}}, & \text{for } 1 \leq i < k \\ e^{-\frac{\rho_i^2}{2\sigma^2}}, & \text{for } i = k. \end{cases} \tag{5}$$

To find the closed-form expressions for r_i and ρ_i , we do the following. First, if we substitute Eq. (4) into Eq. (5) we get:

$$e^{-\frac{r_i^2}{4\sigma^2}} = \begin{cases} \frac{1}{2} \left(1 + e^{-\frac{r_{i+1}^2}{4\sigma^2}} \right), & \text{for } i = 1 \\ \frac{1}{2} \left(e^{-\frac{r_{i-1}^2}{4\sigma^2}} + e^{-\frac{r_{i+1}^2}{4\sigma^2}} \right), & \text{for } 1 < i < k \\ \frac{1}{2} e^{-\frac{r_{i-1}^2}{4\sigma^2}}, & \text{for } i = k. \end{cases}$$

Define a function $g[i] = e^{-\frac{r_i^2}{4\sigma^2}}$, and define $g[0] = 1$, then the expression above can be rewritten as:

$$g[i] = \frac{g[i - 1] + g[i + 1]}{2} \quad \text{for } 1 \leq i < k, \tag{6}$$

with the boundary cases $g[0] = 1$ and $g[k] = \frac{1}{2}g[k - 1]$. Notice that, because the domain of the function g is discrete, this relation holds only for linear functions, i.e.,

$$g[i] \equiv ai + b,$$

for some coefficients a and b . From the boundary cases we find that $a = -\frac{1}{k+1}$ and $b = 1$. Therefore,

$$g[i] = 1 - \frac{i}{k + 1}.$$

Finally, from this expression and Eq. (4) we can derive Eq. (1) and the expressions for ρ_i :

$$r_i = 2\sigma \sqrt{\log \frac{k + 1}{k + 1 - i}},$$

$$\rho_i = \sqrt{2\sigma^2 \log \frac{(k + 1)^2}{(k + 1 - i)(k + 2 - i)}}.$$

Substituting the expressions for ρ_i into $w_i = W_i - W_{i+1} = \mu(\rho_i) - \mu(\rho_{i+1})$, we attain Eq. (2), thus proving the lemma. □

Since the error ε is given, we can use the expressions derived in the proof of the previous lemma [in particular, Eq. (3)] to find the value of k such that the volume between the step function f_d with k disks and μ is at most ε . This leads to the following result.

Theorem 2 *Given a Gaussian distribution μ and $\varepsilon > 0$, we can ε -approximate μ with a cylindrical step function f_d that is defined by a set of k weighted disks, where*

$$k = \left\lceil \frac{1}{\varepsilon} - 1 \right\rceil = O(1/\varepsilon).$$

Proof Using Eqs. (4) and (5) and Lemma 1, the minimum symmetric difference between the functions μ and f_d found from Eq. (3) can be simplified to

$$V_{\min} = \frac{1}{k + 1}.$$

The volume V_{\min} gives the error of approximating the distribution density function μ by the set of disks:

$$\varepsilon = V_{\min} = \frac{1}{k + 1},$$

and thus,

$$k = \left\lceil \frac{1}{\varepsilon} - 1 \right\rceil = O\left(\frac{1}{\varepsilon}\right).$$

It follows that we can ε -approximate a Gaussian distribution by using $O(1/\varepsilon)$ disks. □

Notice that the result of the theorem does not depend on the shape of the Gaussian distribution, i.e., on σ . It only depends on the required error ε .

2.2 Approximating Gaussian Distributions with Polygons

The curved boundaries of the disks of f_d make geometric computations rather complicated. Therefore, next we consider approximating μ by a set of polygons. Computing a set of polygons of minimum total complexity is a challenging mathematical problem that we leave to future investigation. However, we can easily obtain a set of at most twice as many polygons as the minimum, by first computing a set of k disks with guaranteed error ε , then defining $2k$ annuli (two for each disk), and finally choosing $2k$ regular polygons that stay within these annuli. Figure 5a illustrates this idea; since the relative widths of the annuli change, polygons of different complexity are used for different annuli. For each disk with radius r_i we define two radii r'_i and r''_i by the following formulas:

$$\begin{aligned} \mu(r'_i) &= \frac{1}{2}(\mu(\rho_i) + \mu(r_i)), \\ \mu(r''_i) &= \frac{1}{2}(\mu(r_i) + \mu(\rho_{i+1})). \end{aligned} \tag{7}$$

Knowing the widths of the annuli we can calculate the total complexity of the approximation.

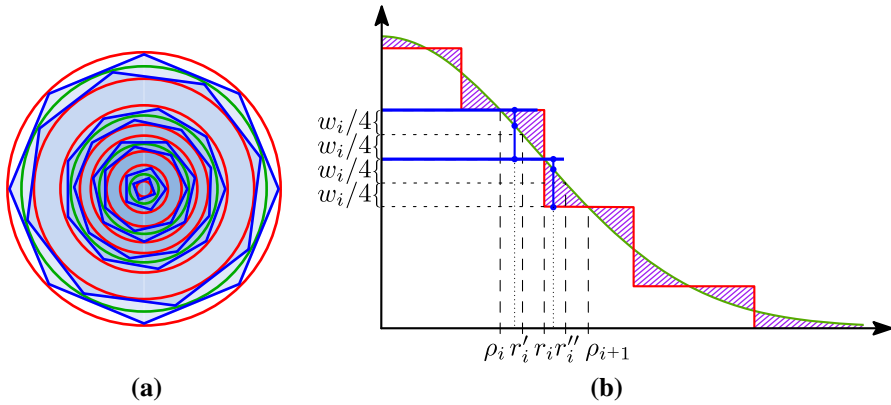


Fig. 5 **a** A Gaussian distribution, given by isolines at k levels (green), $2k$ annuli around each disk (red), and a set of polygons that can be used to obtain an approximation (blue). **b** We choose $2k$ regular polygons (blue) inscribed in annuli $\{r'_i, r_i\}$ and $\{r_i, r''_i\}$ with cumulative weights W_i and $(W_i + W_{i+1})/2$, respectively. For points within annuli $\{r'_i, r_i\}$ and $\{r_i, r''_i\}$ the error of approximation of the Gaussian function (green) by the weighted polygons is not greater than the error of approximation by k -level step function (red) (Color figure online)

Theorem 3 Given $\varepsilon > 0$, any Gaussian distribution can be ε -approximated by $O(1/\varepsilon)$ polygons of complexity $O(1/\sqrt{\varepsilon})$ each.

Proof First, we compute a set of $k = \left\lceil \frac{1}{e^\varepsilon - 1} \right\rceil$ concentric disks by Eqs. (1) and (2) that approximate the distribution function μ with guaranteed error ε . For each disk with radius r_i we find radii r'_i and r''_i from Eq. (7). Then we choose $2k$ regular polygons that stay within the annuli defined by pairs of radii $\{r'_i, r_i\}$ and $\{r_i, r''_i\}$ with weights $w_i/2$ each. These $2k$ polygons ε -approximate the probability distribution function μ . To prove this, we will show that this set of $2k$ weighted regular polygons approximates μ better than the cylindrical step function f_d with k disks. Consider all r such that $\rho_i \leq r \leq \rho_{i+1}$. The step function $f_d(r) = W_i$ for $r \leq r_i$, and $f_d(r) = W_{i+1}$ for $r > r_i$. The error of approximation of μ by f_d at point r , therefore, is $W_i - \mu(r)$ for $r \leq r_i$, and $\mu(r) - W_{i+1}$ for $r > r_i$. Now consider the approximation of μ with the polygons. For all points within two annuli $\{\rho_i, r'_i\}$ and $\{r'_i, \rho_{i+1}\}$, the error of approximation of μ by the weighted polygons is exactly the same as by the disks (for these points, the weight of the corresponding polygon is equal to the weight of the disks). For all points within two annuli $\{r'_i, r_i\}$ and $\{r_i, r''_i\}$, the error of approximation of μ by the weighted polygons is not greater than the error of approximation by the disks (refer to Fig. 5b). For these points, the cumulative weight (i.e., the value of the approximation) of the corresponding polygons equals the cumulative weight of the disks (W_i for annulus $\{r'_i, r_i\}$, or W_{i+1} for annulus $\{r_i, r''_i\}$), or is equal to $(W_i + W_{i+1})/2$. In the first case, again, the error of approximation of μ by the polygons in point r is the same as the error of approximating it by disks. In the second case, using Eq. (7), we conclude that the value of the approximation of μ by the polygons is closer to the true value of $\mu(r)$ than the one given by $f_d(r)$ (refer to Fig. 5b). Therefore, the error of approximating μ by $2k$ weighted regular polygons is less than ε .

It remains to show that the complexity of each polygon is $O(\frac{1}{\sqrt{\varepsilon}})$. The complexity of a regular polygon inscribed in an annulus depends only on the ratio of the radii. That is, given an annulus with an inner radius r_{in} , and an outer radius r_{out} , we can fit a regular $\lceil \pi / \arccos \frac{r_{in}}{r_{out}} \rceil$ -gon in it. Consider an annulus given by $\{r'_i, r_i\}$ (the calculations for the second type of annuli are alike). First, derive from Eq. (7) the formula for r'_i :

$$r'_i = \sqrt{2\sigma^2 \log \frac{2k(k+1)}{(k+1-i)(2k+3-2i)}}$$

then the number of vertices n'_i of the polygon inscribed in the annulus $\{r'_i, r_i\}$ is

$$n'_i = \left\lceil \frac{\pi}{\arccos \frac{r'_i}{r_i}} \right\rceil = \left\lceil \frac{\pi}{\arccos \sqrt{\frac{\log \frac{2k(k+1)}{(k+1-i)(2k+3-2i)}}{\log \frac{k(k+1)}{(k+1-i)^2}}}} \right\rceil.$$

Value n'_i reaches its maximum when r'_i/r_i is maximized. Consider

$$T(i) = \left(\frac{r'_i}{r_i}\right)^2 = \frac{\log \frac{2k(k+1)}{(2k+3-2i)(k+1-i)}}{\log \frac{k(k+1)}{(k+1-i)^2}}$$

as a continuous function of i , where i is defined on interval $[1, k]$, differentiate it and solve the following equation:

$$0 = \frac{dT}{di} = (4k + 6 - 4i) \log \frac{2k(k+1)}{(2k+3-2i)(k+1-i)} - (4k + 5 - 4i) \log \frac{k(k+1)}{(k+1-i)^2}.$$

After dividing both sides of the equation by $(4k + 6 - 4i) \log \frac{k(k+1)}{(k+1-i)^2}$ (notice that it is a non-zero value on the interval $[1, k]$) we get

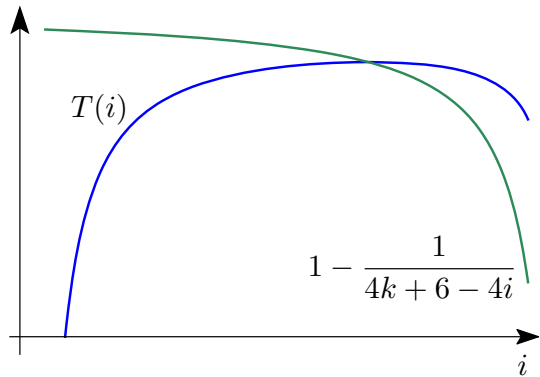
$$\frac{\log \frac{2k(k+1)}{(2k+3-2i)(k+1-i)}}{\log \frac{k(k+1)}{(k+1-i)^2}} = \frac{4k + 5 - 4i}{4k + 6 - 4i} = 1 - \frac{1}{4k + 6 - 4i}.$$

The left-hand side of this equation is $T(i)$ by definition. Therefore $T(i)$ passes through its critical point(s) where it intersects the function $(1 - \frac{1}{4k+6-4i})$ (refer to Fig. 6). By computing the second derivative of $T(i)$, and again using the fact that at the critical points,

$$\frac{\log \frac{2k(k+1)}{(2k+3-2i)(k+1-i)}}{\log \frac{k(k+1)}{(k+1-i)^2}} = \frac{4k + 5 - 4i}{4k + 6 - 4i},$$

we find that the second derivative is strictly negative. Testing the boundary cases $i = 1$ and $i = k$ we can conclude that for sufficiently large k ($k \geq 3$), the two functions will intersect on the interval $[1, k]$. Therefore, taking all the above considerations in

Fig. 6 Graphs of $T(i)$ and $1 - \frac{1}{4k+6-4i}$ intersect where $T(i)$ reaches its maximum



account, we conclude that there is a single critical point of $T(i)$ on the interval $[1, k]$, and it is a maximum. Then,

$$\max_{1 \leq i \leq k} T(i) \leq \max_{1 \leq i \leq k} \left(1 - \frac{1}{4k + 6 - 4i} \right) = 1 - \frac{1}{4k + 2},$$

and, using the Taylor series expansion,

$$n'_i \leq \left\lceil \frac{\pi}{\arccos \sqrt{1 - \frac{1}{4k+2}}} \right\rceil = 2\pi\sqrt{k} + O\left(\frac{1}{\sqrt{k}}\right) = O\left(\frac{1}{\sqrt{\varepsilon}}\right). \quad \square$$

Remark 1 A crude estimate on the volume under the $2k$ weighted polygons is $|V_{APX} - 1| = O(\varepsilon)$.

3 Visibility Between Two Regions

Consider a set of obstacles \mathcal{O} in the plane. Throughout the paper, we assume that the obstacles are disjoint simple polygons with m vertices in total.⁴ Let the obstacles be also convex. Otherwise we can partition the non-convex obstacles without increasing the total asymptotic complexity. Given two imprecise points with probability distributions μ_1 and μ_2 , we can approximate them with two sets \mathcal{P}_1 and \mathcal{P}_2 of weighted regions, each consisting of convex polygons. For every pair of polygons $P_1 \in \mathcal{P}_1$ and $P_2 \in \mathcal{P}_2$, 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. 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 following formula:

⁴ We could also start from non-disjoint polygons and decompose them into simple disjoint polygons, but this could possibly lead to higher complexity.

$$\text{Prob}\{p_1 \text{ sees } p_2\} = \frac{\int_{P_1} \int_{P_2} \int v(x_1, y_1, x_2, y_2) \, dx_2 \, dy_2 \, dx_1 \, dy_1}{\int_{P_1} \int_{P_2} \int dx_2 \, dy_2 \, dx_1 \, dy_1}, \tag{8}$$

where $v(x_1, y_1, x_2, y_2)$ is a visibility indicator function: it is 1 if the points see each other, and 0 otherwise. The denominator of the expression on the right-hand side is simply the product of the areas of the polygons P_1 and P_2 . Then, the probability that two points p_1 sampled from \mathcal{P}_1 and p_2 sampled from \mathcal{P}_2 see each other

$$\begin{aligned} & \text{Prob}\{p_1 \sim f_{\mathcal{P}_1} \text{ sees } p_2 \sim f_{\mathcal{P}_2}\} \\ &= \sum_{P_1 \in \mathcal{P}_1} \frac{w(P_1)\text{area}(P_1)}{V_{APX}(\mathcal{P}_1)} \sum_{P_2 \in \mathcal{P}_2} \frac{w(P_2)\text{area}(P_2)}{V_{APX}(\mathcal{P}_2)} \text{Prob}\{p_1 \in P_1 \text{ sees } p_2 \in P_2\} \\ &\approx \sum_{\substack{P_1 \in \mathcal{P}_1, \\ P_2 \in \mathcal{P}_2}} w(P_1)w(P_2) \int_{P_1} \int_{P_2} \int v(x_1, y_1, x_2, y_2) \, dx_2 \, dy_2 \, dx_1 \, dy_1, \end{aligned}$$

where $w(P_1)$ and $w(P_2)$ are the weights of polygons P_1 and P_2 , and $V_{APX}(\mathcal{P}_1)$ and $V_{APX}(\mathcal{P}_2)$ are the volumes under the weighted disks of \mathcal{P}_1 and \mathcal{P}_2 respectively. Recall from Remark 1 that these volumes are within $1 \pm O(\varepsilon)$. This expression also approximates the total probability that two points sampled from μ_1 and μ_2 see each other with an error at most ε .

$$\begin{aligned} & \text{Prob}\{p_1 \sim \mu_1 \text{ sees } p_2 \sim \mu_2\} \\ &\approx \sum_{\substack{P_1 \in \mathcal{P}_1, \\ P_2 \in \mathcal{P}_2}} w(P_1)w(P_2) \int_{P_1} \int_{P_2} \int v(x_1, y_1, x_2, y_2) \, dx_2 \, dy_2 \, dx_1 \, dy_1. \end{aligned}$$

To compute this integral we will apply the point-line duality transformation (refer to Chapter 8 of [11]). Consider the dual space where a point with coordinates (α, β) corresponds to a line $y = \alpha x - \beta$ in the primal space. Let $\mathcal{L}(P_1, P_2)$ be the set of all lines that intersect both P_1 and P_2 , and let $\mathcal{L}^*(P_1, P_2)$ be a set of points in the dual space corresponding to $\mathcal{L}(P_1, P_2)$. We shall omit the arguments and write \mathcal{L} and \mathcal{L}^* when it is clear from the context which polygons are under consideration.

Let us rewrite the numerator of Eq. (8) in terms of integration over \mathcal{L}^* . This will correspond to integrating over all lines $\ell \in \mathcal{L}$. Consider some line $\ell \in \mathcal{L}$ such that the visibility along ℓ is unobstructed by obstacles. Then, intuitively, the contribution to the integral of this line ℓ is the product of lengths of the segments $\ell \cap P_1$ and $\ell \cap P_2$. However, if there are obstacles obstructing the visibility along ℓ , the contribution to the integral will depend on the relative positions of the polygons and the obstacles (refer to Fig. 7). First, we will show the change of variables that will lead to integration in the dual space, and then we will show how to account for obstacles and to compute the numerator of Eq. (8).

Change of variables Consider some line $\ell \in \mathcal{L}$ with equation $y = \alpha x - \beta$ that passes through two points $p_1 = (x_1, y_1) \in P_1$ and $p_2 = (x_2, y_2) \in P_2$. Then,

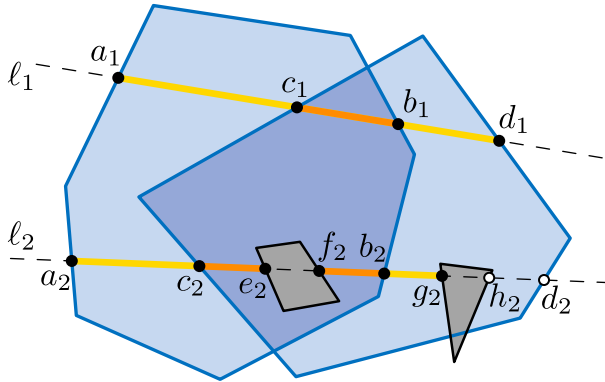


Fig. 7 The visibility along the line ℓ_1 is unobstructed by obstacles, the contribution to the integral is $|a_1b_1| \times |c_1d_1|$ (where $|s|$ denotes the length of segment s); due to two obstacles blocking the visibility along the line ℓ_2 , its contribution to the integral is $|a_2e_2| \times |c_2e_2| + |f_2b_2| \times |f_2g_2|$

$$y_1 = \alpha x_1 - \beta,$$

$$y_2 = \alpha x_2 - \beta.$$

In the dual space, point ℓ^* , corresponding to the line ℓ , has coordinates (α, β) . Using the following change of variables,

$$(x_1, y_1, x_2, y_2) \leftarrow (x_1, \alpha, x_2, \beta),$$

we obtain the numerator of Eq. (8) in the following form

$$\iint_{\mathcal{L}^*} \int_{X_1(\alpha, \beta)}^{X_2(\alpha, \beta)} \int_{X_3(\alpha, \beta)}^{X_4(\alpha, \beta)} v(x_1, \alpha, x_2, \beta) |J| dx_2 dx_1 d\alpha d\beta, \tag{9}$$

where $X_1(\alpha, \beta)$ and $X_2(\alpha, \beta)$ are the x -coordinates of the intersection points of the line $y = \alpha x - \beta$ with polygon P_1 , $X_3(\alpha, \beta)$ and $X_4(\alpha, \beta)$ are the x -coordinates of the intersection points of the line $y = \alpha x - \beta$ with polygon P_2 , and

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

In Sect. 4 we will show the details of how to compute this integral for a fixed combinatorial configuration, i.e., for α and β such that the functions $X_1(\alpha, \beta)$, $X_2(\alpha, \beta)$, $X_3(\alpha, \beta)$, and $X_4(\alpha, \beta)$ are given by the same expressions, and the function $v(x_1, \alpha, x_2, \beta)$ stays constant. For now, we assume this is possible and argue how to partition the domain of the integral into such fixed combinatorial configuration regions.

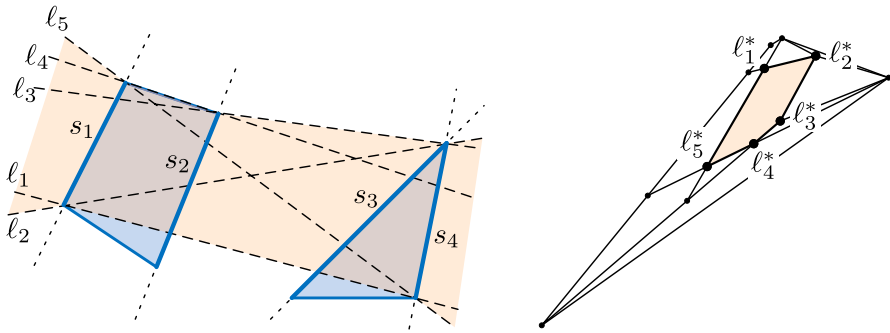


Fig. 8 Left: Two polygons P_1 and P_2 in the primal space. The shaded region represents the set of lines \mathcal{L} which intersect segments s_1 and s_2 of polygon P_1 , and segments s_3 and s_4 of polygon P_2 . Right: The partition of region \mathcal{L}^* in the dual space. The shaded cell corresponds to a set of lines \mathcal{L} in the primal space intersecting segments $s_1, s_2, s_3,$ and s_4

3.1 Computing Expression (9)

Partition region \mathcal{L}^* in the dual space into cells each corresponding to a set of lines in \mathcal{L} that cross the same four segments of P_1 and P_2 (refer to Fig. 8). Then within each cell of this partition the functions $X_1(\alpha, \beta), X_2(\alpha, \beta), X_3(\alpha, \beta),$ and $X_4(\alpha, \beta)$ are given by the same four expressions.

The following observation follows from the fact that each vertex of the partition of \mathcal{L}^* corresponds to a line in the primal space through two vertices of P_1 and P_2 .

Observation 4 *Given two convex polygons P_1 and P_2 of total size n , the complexity of the partition of $\mathcal{L}^*(P_1, P_2)$ into cells, each corresponding to a set of lines in $\mathcal{L}(P_1, P_2)$ that cross the same four segments of P_1 and P_2 , is $O(n^2)$.*

Next, we will introduce a more refined subdivision of the dual space. In each cell the visibility indicator function $v(x_1, \alpha, x_2, \beta)$ will stay constant (in addition to the functions $X_1(\alpha, \beta), X_2(\alpha, \beta), X_3(\alpha, \beta),$ and $X_4(\alpha, \beta)$ being described by the same expressions). We first assume that the polygons $P_1, P_2,$ and the obstacles \mathcal{O} are all disjoint. Later we will remove this assumption.

3.1.1 Disjoint Polygons

For each obstacle $H \in \mathcal{O}$ we construct a region H^* in the dual space, corresponding to the set of lines in the primal space that intersect H . Region H^* has an “hour-glass” shape (refer to Fig. 9). Let $\mathcal{S}^*(P_1, P_2, \mathcal{O})$ be a subdivision of the dual plane resulting from overlaying the partition of the region $\mathcal{L}^*(P_1, P_2)$ into cells and the regions H^* for all $H \in \mathcal{O}$. We shall simply write \mathcal{S}^* when it is clear from the context which polygons and obstacles are considered.

Since the objects involved have in total $O(m + n)$ vertices in the primal space, we observe:

Observation 5 *If the polygons P_1 and $P_2,$ and the obstacles \mathcal{O} are disjoint, subdivision $\mathcal{S}^*(P_1, P_2, \mathcal{O})$ of the dual space has complexity $O((m + n)^2)$, where n is the total*

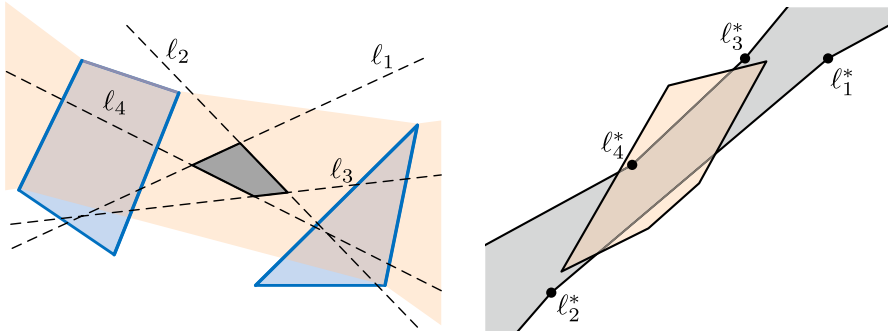


Fig. 9 Left: Primal space. Polygons P_1 and P_2 , and an obstacle between them. Right: Dual space. The “hourglass” shape H^* (shown gray) in the dual space that corresponds to a set of all lines in the primal space that intersect the obstacle

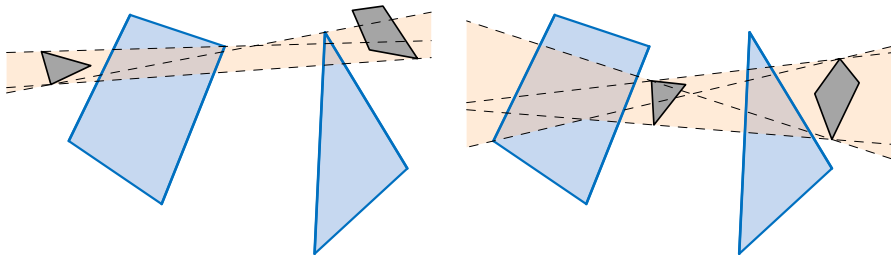
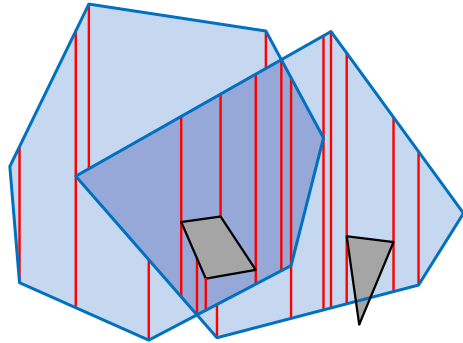


Fig. 10 A set of lines corresponding to one cell C of the subdivision \mathcal{S}^* is highlighted in orange. These lines intersect the same four segments of the polygons P_1 and P_2 , and the same subset of obstacles $\mathcal{O}_C \subset \mathcal{O}$. Left: Visibility function $v(C) = 1$. Right: Visibility function $v(C) = 0$ (Color figure online)

complexity of the polygons P_1 and P_2 , and m is the total complexity of the obstacles in \mathcal{O} .

Consider a set of lines L in the primal space that intersect the same four boundary segments of P_1 and P_2 , that intersect the same set of obstacles $\mathcal{O}_L \subseteq \mathcal{O}$, and that split the rest of the obstacles into the same two subsets. A set L^* in the dual space, corresponding to L , may consist of one or more cells of the subdivision \mathcal{S}^* . As the polygons and the obstacles are disjoint (by assumption), the value of the visibility indicator function $v(x_1, y_1, x_2, y_2)$ is constant for any pair of points $(x_1, y_1) \in P_1$ and $(x_2, y_2) \in P_2$ that lie on the lines in L : it is 1 if the set of obstacles \mathcal{O}_L is empty or if none of the obstacles in \mathcal{O}_L obstructs the visibility between the polygons (Fig. 10 left); and it is 0 if there is at least one obstacle in \mathcal{O}_L blocking the visibility between the polygons (Fig. 10 right). In this case we will write that $v(L^*) = 1$ or $v(L^*) = 0$. Thus, each cell of the subdivision \mathcal{S}^* is comprised of the points corresponding to the same fixed combinatorial configuration. Then, for each cell we can calculate the integral of Eq. (9). It can be written as a sum of integrals over all cells C of \mathcal{S}^* for which the value of the visibility indicator function v is 1:

Fig. 11 Trapezoidal decomposition of the configuration in the primal space



$$\sum_{\substack{C \subset \mathcal{S}^* \\ \nu(C)=1}} \iint_C \left(\int_{X_1(\alpha,\beta)}^{X_2(\alpha,\beta)} \int_{X_3(\alpha,\beta)}^{X_4(\alpha,\beta)} |x_2 - x_1| dx_2 dx_1 \right) d\alpha d\beta. \tag{10}$$

In Sect. 4 we will show how to compute this integral for each cell C in time linear in the complexity of C . The following lemma assumes this result.

Lemma 2 *If the polygons P_1 and P_2 , and the obstacles \mathcal{O} are disjoint, Expression (10) can be evaluated in $O((m + n)^2)$ time.*

Proof From Observation 5, the total complexity of the cells of \mathcal{S}^* is $O((m + n)^2)$. We explicitly compute the overlay of the partition of \mathcal{L}^* with regions \mathcal{H}^* [11]. While building the overlay, we can update the value of the visibility indicator function $\nu(C)$ in $O(1)$ time per event. Assuming that the integral can be computed in time linear in the size of each cell of \mathcal{S}^* , Expression (10) can be evaluated in $O((m + n)^2)$ time. \square

3.1.2 Intersecting Polygons

In the case when the polygons P_1, P_2 , and the obstacles are not disjoint (the obstacles are still disjoint among themselves), we partition regions $P_1 \setminus (P_2 \cup \mathcal{O})$, $P_2 \setminus (P_1 \cup \mathcal{O})$, and $(P_1 \cap P_2) \setminus \mathcal{O}$ into convex sub-regions (refer to Fig. 11). Note that any linear-size decomposition into convex sub-regions, for example a triangulation, could be used. However, here we settle on trapezoidal decomposition for simplicity of implementation, as the closed form solution of the probability integral will be much more concise for the pieces containing vertical segments in comparison to the general case (refer to Sect. 4 and Appendix A).

If the total complexity of the polygons P_1 and P_2 is n , and the total complexity of the obstacles in \mathcal{O} is m , then the total size of the convex decomposition is $O(m + n)$. For two convex sub-regions $T_1 \subset P_1$ and $T_2 \subset P_2$ we note the following: (a) T_1 and T_2 are either equal or disjoint, and (b) the size of the corresponding subdivision $\mathcal{S}^*(T_1, T_2, \mathcal{O})$ is $O((m + n)^2)$, as the sizes of T_1 and T_2 may be linear. If $T_1 = T_2$, then the corresponding contribution to the integral is $(\text{area}(T_1))^2$; and if T_1 and T_2 are disjoint, then we can evaluate the integral as in the previous case. Thus, for all pairs of

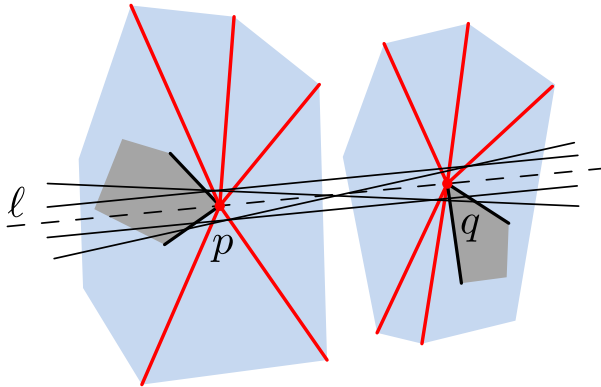


Fig. 12 Lines in ε -neighborhood of ℓ can be of a constant number of combinatorial types

convex sub-regions, we can evaluate Expression (10) in $O((m+n)^4)$ time. However, we are greatly overestimating the complexity of $\mathcal{S}^*(P_1, P_2, \mathcal{O})$ when counting in this manner, as we show next.

Lemma 3 *Given two convex polygons P_1 and P_2 of total complexity n , a set of obstacles \mathcal{O} of total complexity m , and a linear-size convex decomposition of $P_1 \setminus (P_2 \cup \mathcal{O})$, $P_2 \setminus (P_1 \cup \mathcal{O})$, and $(P_1 \cap P_2) \setminus \mathcal{O}$, the total size of all the subdivisions $\mathcal{S}^*(T_1, T_2, \mathcal{O})$, for all pairs of disjoint convex sub-regions $T_1 \subset P_1$ and $T_2 \subset P_2$ in the convex decomposition, is $O((m+n)^2)$.*

Proof Consider the following construction. Draw the subdivisions $\mathcal{S}^*(T_1, T_2, \mathcal{O})$ for all pairs of disjoint convex sub-regions T_1 and T_2 in layers, one layer for each pair. Glue the layers along the coinciding boundary segments of the subdivisions $\mathcal{S}^*(T_1, T_2, \mathcal{O})$. Let us call this construction \mathfrak{S} . Each vertex of \mathfrak{S} corresponds to a line in the primal space passing through two vertices of the polygons P_1 , P_2 , or the obstacles \mathcal{O} . Therefore, there are $O((m+n)^2)$ vertices in \mathfrak{S} .

Consider an ε -neighborhood of some vertex ℓ^* of \mathfrak{S} in the dual space. Let line ℓ in the primal space, corresponding to ℓ^* , pass through two vertices p and q of the convex decomposition. For small enough ε , the lines in the primal space, dual to the points in the ε -neighborhood, intersect the same set of segments of the convex subdivision, except maybe for the segments adjacent to p and q . The number of cells in \mathfrak{S} adjacent to the vertex ℓ^* is equal to the number of different *combinatorial types*⁵ of the lines dual to the points in the ε -neighborhood of ℓ^* . These combinatorial types are defined by the two sets of segments adjacent to p and q that can be intersected by a line from the ε -neighborhood (refer to Fig. 12). Notice that there can be only a constant number of such combinatorial types (four, if there is no boundary segment of the subdivision collinear to ℓ , and six at most, if there are boundary segments collinear to ℓ). Then, each vertex of \mathfrak{S} is adjacent to only a constant number of cells. Thus, the number of edges and faces of \mathfrak{S} is $O((m+n)^2)$ as well. Therefore, the total complexity of all the

⁵ Two lines have the same combinatorial type if they intersect the same set of segments in the same order.

subdivisions $\mathcal{S}^*(T_1, T_2, \mathcal{O})$ for all pairs of convex sub-regions $T_1 \subset P_1$ and $T_2 \subset P_2$ is $O((m + n)^2)$. □

Recall that we assume that, for each cell of \mathcal{S}^* , the integral in Expression (10) can be evaluated in time linear in the size of the cell. Then, as a corollary to the previous lemma, we obtain the following lemma.

Lemma 4 *Expression (10) can be evaluated in $O((m + n)^2)$ time.*

We sum up our results in this section with the following theorem.

Theorem 6 *Given two convex polygons P_1 and P_2 of total size n and a set of disjoint polygonal obstacles of total complexity 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.*

Observe that we do not necessarily need to require that the polygons P_1 and P_2 are convex for Theorem 6 to hold. It would be enough for the polygons to have a linear-size intersection, so that the convex decomposition of their overlay (together with the obstacles) would also be of linear size.

As an easy corollary to Theorem 6, we improve on a result by Rote [20] of computing the degree of convexity of a polygon, i.e., the probability that two points inside the polygon, chosen uniformly at random, can see each other, from $O(n^9)$ down to $O(n^2)$.

Corollary 1 *Let P be a polygon (possibly with holes) of total complexity n . We can compute the probability that two points chosen uniformly at random in P see each other in $O(n^2)$ time.*

4 Computing the Probability for a Fixed Combinatorial Configuration

This section contains technical details of computing the integral in Expression (10) for a fixed combinatorial configuration. The reader may safely skip it, and move to the next section that concludes the paper with a general summary of the results.

Next, we will show how to evaluate Expression (10) for a set L^* in the dual space corresponding to a set of lines L of some fixed combinatorial structure in the primal space. That is, all lines in L intersect the same four segments of P_1 and P_2 , intersect the same subset of obstacles, and split the rest of obstacles into the same two subsets. For all points (α, β) in L^* , functions $X_1(\alpha, \beta)$, $X_2(\alpha, \beta)$, $X_3(\alpha, \beta)$, and $X_4(\alpha, \beta)$ are given by the same four expressions, and $v(L^*) = 1$. Denote the term that the integration over L^* contributes to the Expression (10) as I_{L^*} :

$$I_{L^*} = \iint_{L^*} \left(\int_{X_1(\alpha, \beta)}^{X_2(\alpha, \beta)} \int_{X_3(\alpha, \beta)}^{X_4(\alpha, \beta)} |x_2 - x_1| dx_2 dx_1 \right) d\alpha d\beta.$$

Suppose the lines in L intersect segments s_1 and s_2 of the polygon P_1 , and s_3 and s_4 of the polygon P_2 (Fig. 8). Let segment s_i , for $1 \leq i \leq 4$, lie on the line with equation

$a_i x + b_i y + c_i = 0$. Then, the limits of integration in the general case can be expressed as:

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

Now, given a specific set of lines L with a fixed combinatorial structure, we can express the limits of integration as functions of α and β and evaluate the integral. Later we will also consider special cases, including when some of the terms and denominators can be zero.

Before we present the integration in detail, we will need to take a closer look at the shape of the set L^* . First, consider the case when L does not contain vertical lines. Let, as before, O_L be the set of obstacles that the lines of L intersect, let O_a be the set of obstacles lying above the lines of L , and let O_b be the set of obstacles lying below the lines of L . Then L^* is an intersection of (1) the four wedges corresponding to the four segments, (2) the hour-glass shapes corresponding to the obstacles of O_L , (3) the unbounded convex shapes above the hour-glass shapes corresponding to the obstacles of O_a , and (4) the unbounded convex shapes below the hour-glass shapes corresponding to the obstacles of O_b . As a result, the set L^* can consist of multiple cells of S^* , with each such cell being bounded and x -monotone.

If L contains vertical lines, the above-below relation between the lines and the obstacles O_a and O_b depends on the lines' slopes. The non-vertical lines of L can be split into two sets L_1 and L_2 such that the obstacles of O_a lie above the lines of L_1 and below the lines of L_2 , and the obstacles of O_b lie below the lines of L_1 and above the lines of L_2 . Thus, to obtain the set L^* we need to take the union of two sets defined above for L_1 and L_2 . The set L^* may consist then of multiple cells of S^* . Each such cell is x -monotone, but is not necessarily bounded. However, for each unbounded cell C^+ with α going to $+\infty$, there is a corresponding cell C^- with α going to $-\infty$, and the four infinite rays on the boundaries of C^+ and C^- lie on the same pair of lines.

When computing the integral we will consider each bounded cell C separately, and each pair of unbounded cells C^+ and C^- together. As all cells are x -monotone, we can partition them into vertical strips and get a closed form integral for each such strip.

4.1 Bounded C

First consider the simpler case of a bounded cell C . Recall that by construction the polygons P_1 and P_2 in the primal space are disjoint. We can assume without loss of generality that $X_1 < X_2 \leq X_3 < X_4$, and therefore $|x_2 - x_1| = x_2 - x_1$. Then,

$$I_C = \iint_C \left(\int_{X_1(\alpha,\beta)}^{X_2(\alpha,\beta)} \int_{X_3(\alpha,\beta)}^{X_4(\alpha,\beta)} (x_2 - x_1) dx_2 dx_1 \right) d\alpha d\beta.$$

After solving the inner two integrals we get:

$$\begin{aligned}
 I_C &= \iint_C \frac{(X_2 - X_1)(X_4 - X_3)(X_3 + X_4 - X_1 - X_2)}{2} \, d\alpha \, d\beta \\
 &= \frac{1}{2} \iint_C (-X_1^2 X_3 + X_2^2 X_3 + X_1 X_3^2 - X_2 X_3^2 \\
 &\quad + X_1^2 X_4 - X_2^2 X_4 - X_1 X_4^2 + X_2 X_4^2) \, d\alpha \, d\beta.
 \end{aligned}$$

For some i and j ,

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

Define I_{ij} in the following way:

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

where C is split into vertical strips C_v , with each C_v bounded from left and right by 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 $\beta = A_2 \alpha + B_2$ (Fig. 13). Then,

$$I_C = \frac{1}{2} (I_{13} - I_{31} - I_{23} + I_{32} - I_{14} + I_{41} + I_{24} - I_{42}). \tag{11}$$

Denote with $F_{C_v,ij}(\alpha)$ the indefinite integral corresponding to the vertical strip C_v with the constant of integration equal to 0:

$$F_{C_v,ij}(\alpha) = \int \left(\int_{A_1 \alpha + B_1}^{A_2 \alpha + B_2} \frac{(b_i \beta - c_i)(b_j \beta - c_j)^2}{(b_i \alpha + a_i)(b_j \alpha + a_j)^2} \, d\beta \right) \, d\alpha. \tag{12}$$

Then the terms of the integral can be evaluated by the following formula:

$$I_{ij} = \sum_{C_v \subset C} (F_{C_v,ij}(\alpha_2) - F_{C_v,ij}(\alpha_1)).$$

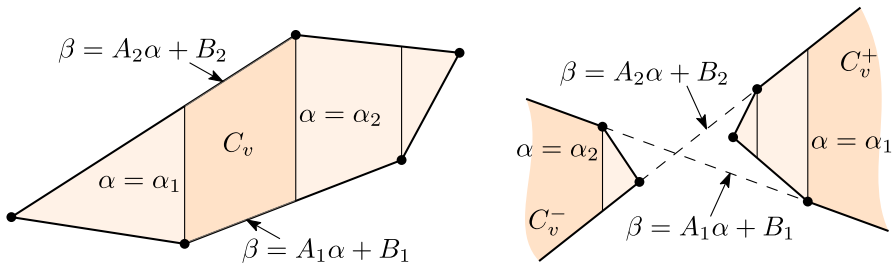


Fig. 13 Left: A bounded cell C , divided into vertical strips. The darker strip is bounded by vertical lines $\alpha = \alpha_1$ and $\alpha = \alpha_2$ from the sides, and by the lines $\beta = A_1\alpha + B_1$ and $\beta = A_2\alpha + B_2$ from below and above respectively. Right: A pair of unbounded cells C_v^+ and C_v^- , divided into vertical strips. The darker strips C_v^+ and C_v^- are bounded by the lines $\beta = A_1\alpha + B_1$ and $\beta = A_2\alpha + B_2$, C_v^+ is bounded by $\alpha = \alpha_1$ from the left, and C_v^- is bounded by $\alpha = \alpha_2$ from the right

The complete case analysis for the closed form of $F_{C_v,ij}$ after the integration is presented in Appendix A.

4.2 Unbounded C

In the case when $C = C^+ \cup C^-$ is unbounded, we can assume without loss of generality that $X_1 < X_2 \leq X_3 < X_4$ and $|x_2 - x_1| = x_2 - x_1$ in the right component of C , and $X_1 > X_2 \geq X_3 > X_4$ and $|x_2 - x_1| = -(x_2 - x_1)$ in the left component of C . Therefore, the integral I_C can be expressed in the following way:

$$I_C = \iint_{C^+} \left(\int_{X_1(\alpha,\beta)}^{X_2(\alpha,\beta)} \int_{X_3(\alpha,\beta)}^{X_4(\alpha,\beta)} (x_2 - x_1) dx_2 dx_1 \right) d\alpha d\beta - \iint_{C^-} \left(\int_{X_2(\alpha,\beta)}^{X_1(\alpha,\beta)} \int_{X_4(\alpha,\beta)}^{X_3(\alpha,\beta)} (x_2 - x_1) dx_2 dx_1 \right) d\alpha d\beta,$$

and one term of the integral for given i and j as:

$$I_{ij} = \iint_{C^+} X_i X_j^2 d\alpha d\beta - \iint_{C^-} X_i X_j^2 d\alpha d\beta = \iint_{C^+} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} d\alpha d\beta - \iint_{C^-} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} d\alpha d\beta.$$

The evaluation is equivalent to the bounded cell case except for the two unbounded vertical strips C_v^+ and C_v^- . In this case we will have to evaluate the integral I_C as it tends to $+\infty$ and $-\infty$, and show how to join the two terms of the unbounded strips into one term that depends only on two bounded values of α_1 and α_2 .

$$\begin{aligned}
 I_{ij}(C_v^+) + I_{ij}(C_v^-) &= \int_{\alpha_1}^{+\infty} \left(\int_{A_1\alpha+B_1}^{A_2\alpha+B_2} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} d\beta \right) d\alpha \\
 &\quad - \int_{-\infty}^{\alpha_2} \left(\int_{A_2\alpha+B_2}^{A_1\alpha+B_1} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} d\beta \right) d\alpha \\
 &= \int_{\alpha_1}^{+\infty} \left(\int_{A_1\alpha+B_1}^{A_2\alpha+B_2} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} d\beta \right) d\alpha \\
 &\quad + \int_{-\infty}^{\alpha_2} \left(\int_{A_1\alpha+B_1}^{A_2\alpha+B_2} \frac{(b_i\beta - c_i)(b_j\beta - c_j)^2}{(b_i\alpha + a_i)(b_j\alpha + a_j)^2} d\beta \right) d\alpha \\
 &= \lim_{\alpha \rightarrow \infty} \left(F_{C_v^+,ij}(\alpha) - F_{C_v^+,ij}(-\alpha) \right) - F_{C_v^+,ij}(\alpha_1) + F_{C_v^+,ij}(\alpha_2).
 \end{aligned}$$

Hence,

$$\begin{aligned}
 I_{ij} &= \sum_{C_v \subset C\{C_v^+ \cup C_v^-\}} (F_{C_v,ij}(\alpha_2) - F_{C_v,ij}(\alpha_1)) \\
 &\quad + F_{C_v^+,ij}(\alpha_2) - F_{C_v^+,ij}(\alpha_1) + \lim_{\alpha \rightarrow \infty} (F_{C_v^+,ij}(\alpha) - F_{C_v^+,ij}(-\alpha)).
 \end{aligned}$$

In Appendix B we will show that, after plugging in the formula for I_{ij} into Eq. (11), the indefinite integrals $F_{C_v^+,ij}(\cdot)$ cancel out when their argument tends to $-\infty$ and $+\infty$. Intuitively, this could be explained by the fact that the slopes A_1 and A_2 are equal for the strips C_v^- and C_v^+ . Recall that the complete case analysis for the closed form of $F_{C_v,ij}$ after the integration is presented in Appendix A.

Therefore, using the closed form expressions for the indefinite integral $F_{C_v^+,ij}$ we can evaluate the value of the integral I_C in time proportional to the complexity of the cell, and thus calculate the total probability that two points sampled from two weighted sets \mathcal{P}_1 and \mathcal{P}_2 see each other in $O((m + n)^2)$ time.

5 Main Result

Combining Theorems 3 and 6, our main result follows:

Theorem 1 *Given two imprecise points p and q , modeled as Gaussian distributions μ_1 and μ_2 , and a set of obstacles of total complexity m , we can approximate the probability that p and q see each other with an error at most ε in $O(1/\varepsilon^3 + m^2/\varepsilon^2)$ time.*

Proof We apply Theorem 3 to ε -approximate μ_1 and μ_2 with two sets of weighted regions \mathcal{M}_1 and \mathcal{M}_2 , each consisting of $O(1/\varepsilon)$ convex polygons of complexity

$O(1/\sqrt{\varepsilon})$. For every pair of polygons $P_1 \subset \mathcal{M}_1$ and $P_2 \subset \mathcal{M}_2$, we use Theorem 6 to compute the probability that a point chosen uniformly at random from P_1 can see a point chosen uniformly at random from P_2 . We need to solve a total of $O(1/\varepsilon^2)$ such problems. For each of them, we have $n = O(1/\sqrt{\varepsilon})$, so applying Theorem 6 leads to an $O((1/\sqrt{\varepsilon} + m)^2)$ running time. Overall, we obtain a total running time of $O((1/\varepsilon^2)(1/\sqrt{\varepsilon} + m)^2) = O(1/\varepsilon^3 + m^2(1/\varepsilon^2))$. \square

6 Conclusions

Motivated by approximating the probability that two imprecise points given by probability distributions can see each other, we presented a technique to approximate each probability density function by a weighted set of polygons. This constitutes a novel approach for dealing with probability density functions in computational geometry. We also showed how to apply this technique to solve our original visibility problem efficiently, improving a recent result along the way.

Regarding future work, it would be interesting to study how our technique can be improved in at least two aspects. Firstly, by finding a minimum complexity approximation of μ by weighted polygons, instead of resorting to a 2-approximation, as we currently do. Secondly, we used approximation to go from a continuous distribution to a set of polygons, but then used exact computations for the visibility between such polygons. Another interesting question is if one could speed up the computation considerably and/or use simpler data structures by using approximation also in this second step. Other related topics worth studying would be imprecise points that move, and the application of this general imprecision model to other problems on points beyond visibility, e.g., geodesic shortest paths.

Acknowledgements K.B. and M.L. are supported by the Netherlands Organisation for Scientific Research (NWO) under Grant Nos. 612.001.207 and 639.021.123, respectively. I.K. was supported in part by the Netherlands Organisation for Scientific Research (NWO) under Grant 639.023.208 and by F.R.S.-FNRS. R.S. is partially supported by Projects MTM2015-63791-R (MINECO/FEDER) and Gen. Cat. DGR2014SGR46, and by MINECO through the Ramón y Cajal program.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

A Evaluation of $F_{C_v,ij}$

The closed form of $F_{C_v,ij}$ after the integration depends on the relative position of the segments s_i and s_j . There are the following cases to consider:

1. In the general case, when $b_i \neq 0, b_j \neq 0$ (segments s_i and s_j are not vertical), and $a_i/b_i \neq a_j/b_j$ (segments s_i and s_j are not parallel), after the integration we obtain that

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

$$\begin{aligned}
 & -\alpha (a_j b_i - a_i b_j) \left(6a_j b_j (2a_j b_i + a_i b_j) (A_2^4 - A_1^4) \right. \\
 & - 24a_j b_i b_j^2 (A_2^3 B_2 - A_1^3 B_1) + 8a_j b_j (b_j c_i + 2b_i c_j) (A_2^3 - A_1^3) \Big) \\
 & + 6a_j^4 b_i^2 (A_2^4 - A_1^4) + 6b_i^2 b_j^4 (B_2^4 - B_1^4) + 36a_j^2 b_i^2 b_j^2 (A_2^2 B_2^2 - A_1^2 B_1^2) \\
 & - 24a_j^3 b_i^2 b_j (A_2^3 B_2 - A_1^3 B_1) - 24a_j b_i^2 b_j^3 (A_2 B_2^3 - A_1 B_1^3) \\
 & + 8 (b_j c_i + 2b_i c_j) (a_j^3 b_i (A_2^3 - A_1^3) - b_i b_j^3 (B_2^3 - B_1^3) \\
 & - 3a_j^2 b_i b_j (A_2^2 B_2 - A_1^2 B_1) + 3a_j b_i b_j^2 (A_2 B_2^2 - A_1 B_1^2) \Big) \\
 & + 12 (2b_j c_i + b_i c_j) (a_j^2 b_i c_j (A_2^2 - A_1^2) + b_i b_j^2 c_j (B_2^2 - B_1^2) \\
 & - 2a_j b_i b_j c_j (A_2 B_2 - A_1 B_1) \\
 & \left. + 24a_j b_i b_j c_i c_j (A_2 - A_1) - 24b_i b_j^2 c_i c_j (B_2 - B_1) \right].
 \end{aligned}$$

1.a. If still $b_i \neq 0, b_j \neq 0$ and $a_i/b_i \neq a_j/b_j$, but the segment s_i is pointing towards the segment s_j (the supporting line of s_i intersects s_j), and one of the corners of the integration strip C_v corresponds to this line going through s_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' is the x -coordinate of the corner of the strip C_v in the dual space. We obtain these equalities from considering the intersection point of the lines $\beta = A_1 \alpha + B_1$ and $\beta = A_2 \alpha + B_2$ in the dual space, and the corresponding line $a_i x + b_i y + c_i = 0$ in the primary space, to which s_i belongs. Then, the argument of the log function of the first term of the general formula becomes 0 at $\alpha = \alpha'$, and we need to treat this as a special case. Plugging in these equalities into the Eq. (12) before the integration, and evaluating the integrals, we obtain the following closed form for the $F_{C_v,ij}(\alpha)$:

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

$$\begin{aligned}
 &+ \left(6a_j^2b_i^2 - 9a_i^2b_j^2 \right) \left(A_2^4 - A_1^4 \right) + 16a_jb_i \left(b_ic_j - b_jc_i \right) \left(A_2^3 - A_1^3 \right) \\
 &+ 12 \left(b_jc_i - b_ic_j \right)^2 \left(A_2^2 - A_1^2 \right) \Big].
 \end{aligned}$$

1.b. In the symmetrical case, when $b_i \neq 0, b_j \neq 0$ and $a_i/b_i \neq a_j/b_j$, but the segment s_j is pointing towards segment s_i (the supporting line of s_j intersects s_i), and one of the corners of the integration strip corresponds to the supporting line of s_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 strip. In that case,

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

2. In case when $b_i \neq 0, b_j \neq 0$, but the segments s_i and s_j are parallel, the following equality holds:

$$\frac{a_i}{b_i} = \frac{a_j}{b_j}.$$

Then after integrating Eq. (12) we get:

$$\begin{aligned}
 F_{C_v,ij}(\alpha) &= \frac{\log(|a_i + \alpha b_i|)}{2b_i^2b_j^2} \Big[3a_i^2b_j^2 \left(A_2^4 - A_1^4 \right) + 3b_i^2b_j^2 \left(A_2^2B_2^2 - A_1^2B_1^2 \right) \\
 &- 6a_i b_i b_j^2 \left(A_2^3B_2 - A_1^3B_1 \right) + 2b_j \left(b_j c_i + 2b_i c_j \right) \left(a_i \left(A_2^3 - A_1^3 \right) \right. \\
 &\left. - b_i \left(A_2^2B_2 - A_1^2B_1 \right) \right) + \left(2b_j c_i + b_i c_j \right) b_i c_j \left(A_2^2 - A_1^2 \right) \Big] \\
 &+ \frac{1}{24b_i^2b_j^2 \left(a_i + \alpha b_i \right)^2} \Big[3\alpha^4 \left(A_1^4 - A_2^4 \right) b_i^4 b_j^2 \\
 &+ \alpha^3 \left(12a_i b_i^3 b_j^2 \left(A_2^4 - A_1^4 \right) - 24b_i^4 b_j^2 \left(A_2^3 B_2 - A_1^3 B_1 \right) \right. \\
 &+ 8b_i^3 b_j \left(b_j c_i + 2b_i c_j \right) \left(A_2^3 - A_1^3 \right) \Big) \\
 &+ \alpha^2 \left(33a_i^2 b_i^2 b_j^2 \left(A_2^4 - A_1^4 \right) - 48a_i b_i^3 b_j^2 \left(A_2^3 B_2 - A_1^3 B_1 \right) \right. \\
 &\left. + 16a_i b_i^2 b_j \left(b_j c_i + 2b_i c_j \right) \left(A_2^3 - A_1^3 \right) \right) \Big]
 \end{aligned}$$

$$\begin{aligned}
 & -\alpha \left(6a_i^3 b_i b_j^2 (A_2^4 - A_1^4) + 72a_i b_i^3 b_j^2 (A_2^2 B_2^2 - A_1^2 B_1^2) \right. \\
 & - 48a_i^2 b_i^2 b_j^2 (A_2^3 B_2 - A_1^3 B_1) - 24b_i^4 b_j^2 (A_2 B_2^3 - A_1 B_1^3) \\
 & + 8b_i b_j (b_j c_i + 2b_i c_j) \left(2a_i^2 (A_2^3 - A_1^3) \right. \\
 & \left. - 6a_i b_i (A_2^2 B_2 - A_1^2 B_1) + 3b_i^2 (A_2 B_2^2 - A_1 B_1^2) \right) \\
 & + 24b_i^2 c_j (2b_j c_i + b_i c_j) \left(a_i (A_2^2 - A_1^2) - b_i (A_2 B_2 - A_1 B_1) \right) \\
 & + 24b_i^3 c_i c_j^2 (A_2 - A_1) \left. \right) \\
 & - 21a_i^4 b_j^2 (A_2^4 - A_1^4) + 3b_i^4 b_j^2 (B_2^4 - B_1^4) - 54a_i^2 b_i^2 b_j^2 (A_2^2 B_2^2 - A_1^2 B_1^2) \\
 & + 60a_i^3 b_i b_j^2 (A_2^3 B_2 - A_1^3 B_1) + 12a_i b_i^3 b_j^2 (A_2 B_2^3 - A_1 B_1^3) \\
 & - 4b_j (b_j c_i + 2b_i c_j) \left(5a_i^3 (A_2^3 - A_1^3) - 9a_i^2 b_i (A_2^2 B_2 - A_1^2 B_1) \right. \\
 & + 3a_i b_i^2 (A_2 B_2^2 - A_1 B_1^2) + b_i^3 (B_2^3 - B_1^3) \left. \right) \\
 & - 6b_i c_j (2b_j c_i + b_i c_j) \left(3a_i^2 (A_2^2 - A_1^2) - 2a_i b_i (A_2 B_2 - A_1 B_1) \right. \\
 & \left. - b_i^2 (B_2^2 - B_1^2) \right) \\
 & \left. - 12a_i b_i^2 c_i c_j^2 (A_2 - A_1) - 12b_i^3 c_i c_j^2 (B_2 - B_1) \right].
 \end{aligned}$$

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

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

3.a. If $b_i = 0, b_j \neq 0$, and segment s_i is pointing towards s_j then $A_1 = A_2 = -c_i/a_i$. Note that the value of α corresponding to the line passing through s_i is at ∞ , thus

there is no extra consideration to be made of evaluating the integral at the corner of the strip in this case.

$$\begin{aligned}
 F_{C_v,ij}(\alpha) = & \frac{\log(|a_j + \alpha b_j|)}{a_i^3 b_j} \left[c_i^2 (B_2 - B_1) (a_i b_j (B_1 + B_2) - 2a_i c_j + 2a_j c_i) \right] \\
 & + \frac{c_i (B_2 - B_1)}{3a_i b_j (a_j + \alpha b_j)} \left[-3\alpha^2 b_j^2 c_i^2 - 3\alpha a_j b_j c_i^2 \right. \\
 & + a_i^2 b_j^2 (B_1^2 + B_1 B_2 + B_2^2) + 3a_i b_j (a_j c_i - a_i c_j) (B_1 + B_2) \\
 & \left. + 3(a_i c_j - a_j c_i)^2 \right].
 \end{aligned}$$

3.b. If $b_i = 0$, and segment s_j is pointing towards segment s_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 strip. Then,

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

4. If $b_i \neq 0$, and $b_j = 0$ (the segment s_j is vertical), then

$$\begin{aligned}
 F_{C_v,ij}(\alpha) = & \frac{\log(|a_i + \alpha b_i|)}{2a_j^2 b_i^2} c_j^2 \left[(A_2^2 - A_1^2) a_i^2 + (B_2^2 - B_1^2) b_i^2 - 2(A_2 B_2 - A_1 B_1) a_i b_i \right. \\
 & \left. + 2(A_2 - A_1) a_i c_i - 2(B_2 - B_1) b_i c_i \right] \\
 & + \frac{c_j^2 \alpha}{4a_j^2 b_i} \left[\alpha (A_2^2 - A_1^2) b_i - 2a_i (A_2^2 - A_1^2) \right. \\
 & \left. + 4b_i (A_2 B_2 - A_1 B_1) - 4c_i (A_2 - A_1) \right].
 \end{aligned}$$

4.a. If $b_i \neq 0$, $b_j = 0$, and segment s_i is pointing towards s_j then

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

We get:

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

4.b. If $b_i \neq 0, b_j = 0$, and segment s_j is pointing towards s_i then $A_1 = A_2 = -c_j/a_j$, and

$$F_{C_v,ij}(\alpha) = \frac{\log(|a_i + b_i \alpha|)(B_2 - B_1)c_j^2}{2a_j^3 b_i} [(B_1 + B_2)a_j b_i + 2a_i c_j - 2a_j c_i] - \alpha \frac{c_j^3}{a_j^3} (B_2 - B_1).$$

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

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

B Evaluation of $F_{C_v,ij}$ at $+\infty$ and $-\infty$

In this appendix we show that the indefinite integrals $F_{C_v,ij}(\cdot)$ cancel out when their argument tends to $-\infty$ and $+\infty$.

Consider the case 1 from Appendix A. The rest of the cases are similar. For ease of reference, we will split the terms summed up in the expression of $F_{C_v,ij}$ into five parts F_1, \dots, F_5 . Let F_1 be the first term of the expression, i.e., the one which is multiplied by $\log(|a_i + \alpha b_i|)$. Similarly, let F_2 be the second term, which is multiplied by $\log(|a_j + \alpha b_j|)$. The remaining terms F_3, F_4, F_5 will be defined later, such that $F(\alpha) = F_1(\alpha) + F_2(\alpha) + F_3(\alpha) + F_4(\alpha) + F_5(\alpha)$. We will argue that $\lim_{\alpha \rightarrow \infty} (F_i(\alpha) - F_i(-\alpha)) = 0$, for each $i = 1, \dots, 5$.

B.1 Case of F_1 and F_2

Denote the coefficient of $\log(|a_i + \alpha b_i|)$ in F_1 as Z , and consider the limit of this term for $F_{C_v^+,ij}(\alpha)$, when α tends to infinity.

$$\lim_{\alpha \rightarrow \infty} (F_1(\alpha) - F_1(-\alpha)) = \lim_{\alpha \rightarrow \infty} (Z \log |a_i + \alpha b_i| - Z \log |a_i - \alpha b_i|),$$

Then,

$$\begin{aligned} &\lim_{\alpha \rightarrow \infty} (F_i(\alpha) - F_i(-\alpha)) \\ &= \lim_{\alpha \rightarrow \infty} \left(\left(Z \log \left| \frac{a_i}{\alpha} + b_i \right| + Z \log |\alpha| \right) - \left(Z \log \left| \frac{a_i}{\alpha} - b_i \right| + Z \log |\alpha| \right) \right) \end{aligned}$$

$$= Z \log |b_i| - Z \log |-b_i| + \lim_{\alpha \rightarrow \infty} (Z \log |\alpha| - Z \log |\alpha|) = 0.$$

Similarly,

$$\lim_{\alpha \rightarrow \infty} (F_2(\alpha) - F_2(-\alpha)) = 0.$$

B.2 Case of F_3, F_4 and F_5

The remaining terms of $F_{C_v,ij}$ will not cancel out as easily as in the case of F_1 and F_2 , therefore we will split them into three groups F_3, F_4 and F_5 . We will need to evaluate these terms for all the $F_{C_v,ij}$ for all i and j . Denote F_{C_v} to be

$$F_{C_v} = F_{C_v,13} - F_{C_v,31} - F_{C_v,23} + F_{C_v,32} - F_{C_v,14} + F_{C_v,41} + F_{C_v,24} - F_{C_v,42},$$

analogously to Eq. (11).

Consider the terms with the variable expression of the form $\alpha^3/(a_j + \alpha b_j)$. After we collect all such terms for all i and j they will cancel out. Indeed, let

$$f_3(i, j) = \frac{\alpha^3}{a_j + \alpha b_j} \cdot \frac{3(A_2^4 - A_1^4) b_i b_j^3 (a_j b_i - a_i b_j)}{24 b_i b_j^2 (a_j b_i - a_i b_j)}.$$

Note that $f_3(i, j)$ is a function of α ; however, we will omit α in this and subsequent definitions for ease of presentation. We can now define F_3 as

$$F_3 = f_3(1, 3) - f_3(3, 1) - f_3(2, 3) + f_3(3, 2) - f_3(1, 4) + f_3(4, 1) + f_3(2, 4) - f_3(4, 2).$$

After simplifying the expression above, we obtain $F_3(\alpha) = 0$, and therefore

$$\lim_{\alpha \rightarrow \infty} (F_3(\alpha) - F_3(-\alpha)) = 0.$$

Now, let

$$f_4(i, j) = \frac{\alpha^2}{a_j + \alpha b_j} \cdot \frac{a_i b_j - a_j b_i}{24 b_i b_j^2 (a_j b_i - a_i b_j)} \left(3(A_2^4 - A_1^4) b_j^2 (3 a_j b_i + 2 a_i b_j) - 24(A_2^3 B_2 - A_1^3 B_1) b_i b_j^3 + 8(A_2^3 - A_1^3) b_j^2 (b_j c_i + 2 b_i c_j) \right),$$

and let

$$F_4 = f_4(1, 3) - f_4(3, 1) - f_4(2, 3) + f_4(3, 2) - f_4(1, 4) + f_4(4, 1) + f_4(2, 4) - f_4(4, 2).$$

Then, after calculating the limits of F_4 we get:

$$\lim_{\alpha \rightarrow +\infty} F_4(\alpha) = \frac{(A_1^3 - A_2^3)((a_3b_4 - a_4b_3)(b_1c_2 - b_2c_1) - (a_2b_1 - a_1b_2)(b_4c_3 - b_3c_4))}{3b_1b_2b_3b_4}$$

$$\lim_{\alpha \rightarrow -\infty} F_4(\alpha) = \frac{(A_1^3 - A_2^3)((a_3b_4 - a_4b_3)(b_1c_2 - b_2c_1) - (a_2b_1 - a_1b_2)(b_4c_3 - b_3c_4))}{3b_1b_2b_3b_4}.$$

Hence

$$\lim_{\alpha \rightarrow \infty} (F_4(\alpha) - F_4(-\alpha)) = 0.$$

Finally, let

$$f_5(i, j) = -\frac{\alpha}{a_j + \alpha b_j} \cdot \frac{(a_j b_i - a_i b_j)}{24b_i b_j^2 (a_j b_i - a_i b_j)} \left(-24a_j (A_2^3 B_2 - A_1^3 B_1) b_i b_j^2 \right. \\ \left. + 8 (A_2^3 - A_1^3) a_j b_j (b_j c_i + 2b_i c_j) + 6 (A_2^4 - A_1^4) a_j b_j (2a_j b_i + a_i b_j) \right),$$

and let

$$F_5 = f_5(1, 3) - f_5(3, 1) - f_5(2, 3) + f_5(3, 2) - f_5(1, 4) \\ + f_5(4, 1) + f_5(2, 4) - f_5(4, 2).$$

After calculating the limits of F_5 we get:

$$\lim_{\alpha \rightarrow +\infty} F_5(\alpha) = \frac{(A_1^3 - A_2^3)((a_4b_3 - a_3b_4)(b_1c_2 - b_2c_1) + (a_2b_1 - a_1b_2)(b_4c_3 - b_3c_4))}{3b_1b_2b_3b_4},$$

$$\lim_{\alpha \rightarrow -\infty} F_5(\alpha) = \frac{(A_1^3 - A_2^3)((a_4b_3 - a_3b_4)(b_1c_2 - b_2c_1) + (a_2b_1 - a_1b_2)(b_4c_3 - b_3c_4))}{3b_1b_2b_3b_4},$$

and

$$\lim_{\alpha \rightarrow \infty} (F_5(\alpha) - F_5(-\alpha)) = 0.$$

The last term of $F_{C_v,ij}$, with $1/(a_j + \alpha b_j)$ will tend to 0 when α tends to plus or minus infinity. Thus, collecting all the terms together, we get that

$$\lim_{\alpha \rightarrow \infty} (F_i(\alpha) - F_i(-\alpha)) = 0.$$

References

1. Agarwal, P.K., Cheng, S.-W., Tao, Y., Yi, K.: Indexing uncertain data. In: 28th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS), pp. 137–146 (2009)

2. Agrawal, P., Benjelloun, O., Sarma, A.D., Hayworth, C., Nabar, S., Sugihara, T., Widom, J.: Trio: a system for data, uncertainty, and lineage. In: 32nd International Conference on Very Large Data Bases (VLDB), pp. 1151–1154 (2006)
3. Bandyopadhyay, D., Snoeyink, J.: Almost-Delaunay simplices: nearest neighbor relations for imprecise points. In: 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 410–419 (2004)
4. Bi, J., Zhang, T.: Support vector classification with input data uncertainty. In: Advances in Neural Information Processing Systems (NIPS) (2004)
5. Buchin, K., Kostitsyna, I., Löffler, M., Silveira, R.I.: Region-based approximation of probability distributions (for visibility between imprecise points among obstacles). In: 30th European Workshop on Computational Geometry (EuroCG) (2014)
6. Buchin, K., Kostitsyna, I., Löffler, M., Silveira, R.I.: Region-based approximation algorithms for visibility between imprecise locations. In: 17th Workshop on Algorithm Engineering and Experiments (ALENEX), pp. 94–103 (2015)
7. Cormode, G., Garofalakis, M.: Histograms and wavelets of probabilistic data. In: IEEE 25th International Conference on Data Engineering (ICDE), pp. 293–304 (2009)
8. Cormode, G., Li, F., Yi, K.: Semantics of ranking queries for probabilistic data and expected ranks. In: IEEE 25th International Conference on Data Engineering (ICDE), pp. 305–316 (2009)
9. Cormode, G., McGregor, A.: Approximation algorithms for clustering uncertain data. In: 27th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS), pp. 191–200 (2008)
10. Dalvi, N., Suciu, D.: Efficient query evaluation on probabilistic databases. *Int. J. Very Large Data Bases* **16**, 523–544 (2007)
11. de Berg, M., Cheong, O., van Kreveld, M., Overmars, M.: *Computational Geometry: Algorithms and Applications*, 3rd edn. Springer, Berlin (2008)
12. Guibas, L.J., Salesin, D., Stolfi, J.: Epsilon geometry: building robust algorithms from imprecise computations. In: 5th Annual Symposium on Computational Geometry (SoCG), pp. 208–217 (1989)
13. Guibas, L.J., Salesin, D., Stolfi, J.: Constructing strongly convex approximate hulls with inaccurate primitives. *Algorithmica* **9**, 534–560 (1993)
14. Horne, J., Garton, E., Krone, S., Lewis, J.: Analyzing animal movements using Brownian bridges. *Ecology* **88**(9), 2354–2363 (2007)
15. Jørgensen, A., Löffler, A., Phillips, J.: Geometric computations on indecisive points. In: 12th International Symposium on Algorithms and Data Structures (WADS), pp. 536–547 (2011)
16. Löffler, M.: Data imprecision in computational geometry. Ph.D. thesis, Utrecht University (2009)
17. Löffler, M., Phillips, J.: Shape fitting on point sets with probability distributions. In: 17th Annual European Symposium on Algorithms (ESA), pp. 313–324 (2009)
18. Nagai, T., Tokura, N.: Tight error bounds of geometric problems on convex objects with imprecise coordinates. In: Japanese Conference on Discrete and Computational Geometry, pp. 252–263 (2000)
19. Ostrovsky-Berman, Y., Jostkovic, L.: Uncertainty envelopes. In: 21st European Workshop on Computational Geometry (EuroCG), pp. 175–178 (2005)
20. Rote, G.: The degree of convexity. In: 29th European Workshop on Computational Geometry (EuroCG), pp. 69–72 (2013)
21. Rote, G.: Private communication (2017). <http://page.mi.fu-berlin.de/rote/Papers/abstract/The+degree+of+convexity.html>
22. Tao, Y., Cheng, R., Xiao, X., Ngai, W.K., Kao, B., Prabhakar, S.: Indexing multi-dimensional uncertain data with arbitrary probability density functions. In: 31st International Conference on Very Large Data Bases (VLDB) (2005)
23. van der Merwe, R., Doucet, A., de Freitas, N., Wan, E.: The unscented particle filter. In: Advances in Neural Information Processing Systems (NIPS), vol. 8, pp. 351–357 (2000)
24. Van Diggelen, F.: GNSS accuracy: lies, damn lies and statistics. *GPS World* **18**(1), 26–32 (2007)
25. van Kreveld, M., Löffler, M.: Largest bounding box, smallest diameter, and related problems on imprecise points. *Comput. Geom. Theory Appl.* **43**, 419–433 (2010)
26. Zou, Y., Chakrabarty, K.: Uncertainty-aware and coverage-oriented deployment of sensor networks. *J. Parallel Distrib. Comput.* **64**, 788–798 (2004)