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

Citation for published version (APA):
Buchin, K., Kostitsyna, I., Löffler, M., & Silveira, R. I. (2015). Region-based approximation of probability distributions (for visibility between imprecise points among obstacles). In U. Brandes, & D. Eppstein (Eds.), Proc. 17th Workshop on Algorithm Engineering and Experiments (ALENEX) (pp. 94-103). Society for Industrial and Applied Mathematics (SIAM). https://doi.org/10.1137/1.9781611973754.9

DOI:
10.1137/1.9781611973754.9

Document status and date:
Published: 01/01/2015

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:
- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 23. Oct. 2023
Region-based Approximation Algorithms for Visibility between Imprecise Locations

Kevin Buchin∗ Irina Kostitsyna∗ Maarten Löffler† Rodrigo I. Silveira‡

Abstract
In this paper we present new geometric algorithms for approximating the visibility between two imprecise locations amidst a set of obstacles, where the imprecise locations are modeled by continuous probability distributions. Our techniques are based on approximating distributions by a set of regions rather than on approximating by a discrete point sample. In this way we obtain guaranteed error bounds, and the results are more robust than similar results based on discrete point sets. We implemented our techniques and present an experimental evaluation. The experiments show that the actual error of our region-based approximation scheme converges quickly when increasing the complexity of the regions.

1 Introduction
Visibility has attracted much attention in geographic information science due to its many applications. The question of whether two objects can see each other amidst a group of obstacles arises naturally in plenty of situations. A basic operation in visibility analysis is computing the space visible from a certain location in a landscape, where the obstacles occluding vision originate from the terrain topography, and sometimes also from the vegetation and buildings in the surroundings [3, 8, 19, 23, 31, 32]. The portion of a terrain that is visible from an observation point is known as the viewsheet of the point. The equivalent concept for urban environments is usually called isovist [30]. For an overview on algorithms to compute viewsheets and other visibility structures, see the survey by De Floriani and Magillo [11]. Visibility and viewsheets also play a role in many other areas, such as animal behaviour studies. For instance, Aspbury and Gibson [1] use viewsheets to study grouse behavior in a lek, the area in which they perform their mating displays.

∗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. of Matemática & CIDMA, Universidade de Aveiro, and Dept. Matemática Aplicada II, Universitat Politècnica de Catalunya; rodrigo.silveira@ua.pt.

Camp et al. [28] use viewsheets of eagle nest locations to determine human disturbance of nesting activities.

The focus of the research presented in this paper is the study of visibility under location uncertainty.

In the GIS literature, it has long been recognized that data acquired with physical devices are subject to uncertainty (from measurement and sampling errors, among several others) [10, 22], and that this is particularly relevant for location data of moving objects [27, 29]. In this context, a widespread way to model an uncertain location—especially in the area of spatiotemporal databases—is by assuming that the exact location of an object is at distance at most d from its estimated location. This is equivalent to representing the object’s location with a disk of radius d, centered at the estimated location: the uncertainty region. In addition, it is often assumed that the exact object location follows a probability density function, defined as 0 for any point outside of the uncertainty region. Even though the choice of the probability density function depends on the application, most previous work uses uniform and Gaussian distributions (e.g., [6, 7, 27, 29]). In particular, Gaussian distributions have been shown to be appropriate to model the error in locations obtained from GPS [34] and to model the error in location between measurements [15]. To the best of our knowledge, previous work on uncertain visibility has focused on modeling the uncertainty present in the terrain model, usually by using some type of error propagation analysis based on Monte Carlo simulation (e.g., [9, 24]). While both visibility and uncertainty have been extensively studied in GIS, we are not aware of research combining visibility and location uncertainty.

Another field where both visibility and data uncertainty have been studied extensively is computational geometry. A lot of attention has been devoted to the design of efficient algorithms for computing visibility regions of points in environments like polygons [16] or terrains [18]. Multiple variants and extensions of visibility concepts have also been studied. For surveys of the vast literature on this topic, we refer to Ghosh [12] and De Floriani and Magillo [11].

In the computational geometry literature, various
models to deal with (spatial) data uncertainty have been suggested. Motivated by finite coordinate precision, Guibas et al. [13] have considered uncertainty regions for points. Similar concepts have been later studied in a variety of settings [2, 14, 20, 25, 26]. Most of the previous work on uncertainty in computational geometry does not consider uncertainty modeled by probability distributions.

Performing geometric algorithms on probability distributions directly is often computationally intractable. Instead, the distributions can be approximated. A popular approach is to describe a distribution by a point set. For instance, for tracking imprecise objects a particle filter uses a discrete set of locations to model uncertainty [33]. Löfler and Phillips [21] and Jørgenson et al. [17] discuss several geometric problems on points with probability distributions, and show how to solve them using discrete point sets that have guaranteed error bounds.

However, when combining visibility queries with point-based approximations of probability distributions, we may lose any theoretical guarantees on the computed probabilities. The choice of points may greatly influence the resulting probability, as illustrated in Figure 1. Instead, we propose to approximate distributions by regions.

**Contribution** In this paper we present a new method to approximate the probability that two imprecise points can see each other. We consider an abstract setting in which the imprecise points are modeled by probability distributions, and the visibility is constrained by a set of polygonal obstacles. Our techniques are based on using regions—rather than point samples—to approximate this probability, allowing to compute probabilities guaranteed to any desired accuracy. We have implemented and experimentally evaluated our method. We show experimentally that, in the settings considered, our techniques are more robust and reliable than those based on point sampling, and that the empirical error of our approximation reduces quickly with the complexity of the regions used in the approximation.

Finally, even though in this paper we use the term visibility, our results also apply to other contexts. For example, in sensor networks visibility problems are formulated as coverage problems. This type of monitoring can be performed by radars, antennas, routers and basically any device that is able to send or receive some sort of wireless signal. Each of these devices can be placed almost anywhere on a terrain, so coverage is the unit that measures the quality of the chosen device placement scheme. Numerous problems related to visibility arise in this area. We refer to the survey by Yick et al. [35] for a more detailed treatment of the subject.

### 2 Modeling

In this section we explain how we model the situation of two imprecise points amid a collection of obstacles. We want to approximate the probability distributions that govern the locations of the imprecise points by sets of uniform distributions. Here we only give an overview of the technical details; we refer the reader to the technical supplement of this paper [5] for more details and proofs.

Assume that an imprecise point \( p \) in the plane is given by a two-dimensional probability distribution \( \mu \). Let \( \mathcal{M} \) be a set of weighted regions in the plane, and let \( w(M) \) denote the uniform weight of a region \( M \in \mathcal{M} \). Let \( \mathcal{M}(p) = \{ M \in \mathcal{M} \mid p \in M \} \) be the subset of \( \mathcal{M} \) containing a point \( p \in \mathbb{R}^2 \). A set \( \mathcal{M} \) defines a function \( m(p) = \sum_{M \in \mathcal{M}(p)} w(M) \) that sums the weights of all regions containing \( p \). We say that \( \mathcal{M} \) \( \varepsilon \)-approximates \( \mu \) if the symmetric difference of the volumes under \( m \) and \( \mu \) is at most \( \varepsilon \); that is, if \( \int_{p \in \mathbb{R}^2} |m(p) - \mu(p)| \leq \varepsilon \). Figure 2(a) illustrates the concept.

Any probability distribution \( \mu \) can be approximated in this way, but the total complexity of \( \mathcal{M} \), i.e., the sum of the complexities of each of its regions, depends on various factors: the shape of \( \mu \), the shape of allowed regions in \( \mathcal{M} \), and the error parameter \( \varepsilon \). To focus the discussion, in this work we limit our attention to Gaussian distributions, since they are natural and, as mentioned in the introduction, have been shown to be appropriate for modeling the uncertainty in commonly-used types of location data [15, 34].

#### 2.1 Approximation with Disks

A natural way to approximate a Gaussian distribution by using a set of regions is by using concentric disks. Thus, given a Gaussian probability distribution \( \mu \) and a maximum allowed error \( \varepsilon \), we would like to compute a set \( \mathcal{M} \) of \( k \) disks that \( \varepsilon \)-approximate \( \mu \). We may assume \( \mu \) is centered at the origin, leaving only a parameter \( \sigma \) that governs the shape of \( \mu \), that is,

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

Figure 1: Two pairs of similar 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.
Figure 2: (a) A probability density function $\mu$ (yellow) can be approximated by a set of weighted regions $\mathcal{M}$, representing a function $m$ (purple). (b) A Gaussian distribution (green) is $\varepsilon$-approximated by a step function (red), which is represented by $k$ weighted disks.

or in polar coordinates,

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

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

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

**Lemma 2.1.** Let $\mu$ be a Gaussian distribution with standard deviation $\sigma$. Let $k$ be a given integer. Then the minimum-error approximation of $\mu$ by $k$ disks is given by

$$
\begin{align*}
  r_i &= \sqrt{2\sigma^2 \log \frac{k(i+1)}{k(i+1)(k-i+2)}} , \\
  w_i &= \frac{1}{2\pi\sigma^2} \frac{(k-i+1)(k-i+2)}{k(i+1)} ,
\end{align*}
$$

where $i \in (1, \ldots, k)$.

**Proof.** Let $V(\rho)$ be the volume under the probability distribution $\mu$ outside the disk of radius $\rho$:

$$
V(\rho) = \iint_{\mathcal{D}} \mu(x, y)\,dx\,dy = \int_0^\infty \int_0^{2\pi} \mu(r, \theta) r\,dr\,d\theta
$$

$$
= \int_0^\infty r e^{-\frac{r^2}{2\sigma^2}} dr = e^{-\frac{\rho^2}{2\sigma^2}} .
$$

Then the symmetric difference between the function $\mu$ and the cylindrical step function defined by $k$ disks of radii $r_1, r_2, \ldots, r_k$ and $k$ weights $w_1, w_2, \ldots, w_k$ is a function given by the following formula:

$$
\begin{align*}
  F &= w_1(\pi r_1^2 - \pi \rho_1^2) - (V(0) - V(r_1)) + V(r_1) - V(r_2) - w_2(\pi r_2^2 - \pi \rho_2^2) + (V(r_2) - V(r_3)) + \ldots + w_k(\pi r_k^2 - \pi \rho_k^2) - (V(r_k)) + V(r_k) ,
\end{align*}
$$

(2.2)

where $0 < r_1 < r_2 < \cdots < r_k$ are the radii of the disks approximating the distribution function $\mu$, $w_1, w_2, \ldots, w_k$ are the corresponding weights, and $0 = \rho_1 < \rho_2 < \cdots < \rho_k$ are the intersections of the step function with $\mu$ (refer to Figure 2(b)). After minimizing function $F$ over all $r_i$’s and $w_i$’s we attain Formulas 2.1. For the details refer to the technical supplement of this article [5].

Since the error allowed $\varepsilon$ is given, we can solve Equation 2.2 for $\varepsilon$. This leads to the following result:

**Theorem 2.1.** Let $\mu$ be a Gaussian distribution with standard deviation $\sigma$. Given $\varepsilon > 0$, we can $\varepsilon$-approximate $\mu$ by a set of $k$ disks.

$$
\begin{align*}
  k &= \left\lfloor \frac{1}{e^\varepsilon - 1} \right\rfloor = O(1/\varepsilon)
\end{align*}
$$

weighted disks.

**Proof.** When minimizing function $F$ given by Formula 2.2, we obtain the following dependency for $\rho_i$:
we can easily obtain a set of polygons at most twice and thus, the total complexity of the approximation.

Disks with guaranteed error as large as the minimum, by first computing a set of minimum total complexity is a challenging mathematical boundary of disks make geometric computations more

Function \( F \) gives the error of approximating the distribution function \( \mu \) by the set of disks:

\[
\varepsilon = F = \log \frac{k+1}{k},
\]

and thus,

\[
k = \left\lfloor \frac{1}{e^2 - 1} \right\rfloor = O \left( \frac{1}{\varepsilon} \right).\]

\[\square\]

Theorem 2.2. A Gaussian distribution with standard deviation \( \sigma \) can be \( \varepsilon \)-approximated by \( O(1/\varepsilon) \) polygons of complexity \( O(1/\sqrt{\varepsilon}) \) each.

Proof. First, we compute a set of \( k = \left\lfloor \frac{1}{e^2 - 1} \right\rfloor \) concentric disks by Formulas 2.1 that approximate the distribution function \( \mu \) with guaranteed error \( \varepsilon \). For each disk with radius \( r_i \), we find two radii \( r_i' \) and \( r_i'' \) from the following equations:

\[
\mu(r_i') = \frac{1}{2} (\mu(r_i) + \mu(r_{i+1})),
\]

\[
\mu(r_i'') = \frac{1}{2} (\mu(r_{i-1}) + \mu(r_i)).
\]

Then we choose \( 2k \) regular polygons that stay within annuli defined by pairs of radii \( \{r_i', r_i\} \) and \( \{r_i, r_i''\} \) with weights \( w_i \) and \( (w_i + w_{i+1})/2 \) correspondingly. These \( 2k \) polygons \( \varepsilon \)-approximate the probability distribution function \( \mu \). Refer to the technical supplement of this article [5] for the detailed proof of that fact.

The complexity of a regular polygon inscribed in an annulus depends only on the ratio of the radii. 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 in it. Similarly, 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 [5] for the proof that \( \lceil \pi / \arccos \frac{r'}{r} \rceil = O(\frac{1}{\sqrt{\varepsilon}}) \) and \( \lceil \pi / \arccos \frac{r''}{r} \rceil = O(\frac{1}{\sqrt{\varepsilon}}) \).

\[\square\]

2.3 The environment Now consider a set of obstacles \( \mathcal{R} \) in the plane. We assume that the obstacles are disjoint convex polygons with \( m \) vertices in total (non-convex obstacles can always be first decomposed into convex pieces). We approximate two imprecise points with probability distributions \( \mu_1 \) and \( \mu_2 \) with two sets of weighted regions \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \), each consisting of regular polygons. For every pair of polygons \( P_1 \in \mathcal{M}_1 \) and \( P_2 \in \mathcal{M}_2 \), 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 that two points \( p_1 = (x_1, y_1) \) and \( p_2 = (x_2, y_2) \) chosen uniformly at random from \( P_1 \) and \( P_2 \), respectively, can see each other can be computed by the formula:

\[
\text{prob} = \frac{\iiint v(x_1, y_1, x_2, y_2) \, dx_1 \, dy_1 \, dx_2 \, dy_2}{\iiint \, dx_1 \, dy_1 \, dx_2 \, dy_2},
\]

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

2.2 Approximation with Polygons The curved boundaries of disks make geometric computations more complicated. Therefore, next we consider approximating \( \mu \) 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 polygons at most twice as large as the minimum, by first computing a set of \( k \) disks with guaranteed error \( \varepsilon \) and then choosing \( 2k \) regular polygons that stay within the annuli defined by those disks. Figure 3 illustrates this idea; since the relative widths of the annuli change, polygons of different complexity are used for different annuli.

Knowing the widths of the annuli we can calculate the total complexity of the approximation.
To compute $\text{prob}$ we consider a dual space where a point with coordinates $(\alpha, \beta)$ corresponds to a line $y = \alpha x - \beta$ in the primal space (see [4] for a detailed exposition of geometric point-line duality). Our first goal is to decompose this dual space into a set of simple regions (called cells in the following) with the property that, for all lines in primal space represented in a cell, we can compute their contribution to $\text{prob}$ in constant time.

For this, we first construct a region $L^*$ in the dual space that corresponds to the set of all lines that intersect $P_1$ and $P_2$. This region can be partitioned into cells, each one corresponding to the set of all lines that cross the same four segments of $P_1$ and $P_2$ (refer to Figure 4(a)). The complexity of the resulting subdivision is summarized in the following lemma:

**Lemma 2.2.** Given two convex polygons $P_1$ and $P_2$ of total size $n$, the complexity of the partition $L^*$ in the dual space, which corresponds to the set of lines $L$ that stab $P_1$ and $P_2$, is $O(n^2)$.

**Proof.** Each vertex of $L^*$ corresponds to a line in primal space through two vertices of $P_1$ and $P_2$. Therefore, the total complexity of $L^*$ is $O(n^2)$. □

Next we need to extend the subdivision in order to take obstacles into account. To that end, we construct for each obstacle $h \in R$ a region $H^*$ in dual space, which corresponds to the set of all lines that intersect $h$. Since obstacles are convex, $H^*$ has an "hourglass" shape (see Figure 5). We now compute the subdivision of the dual plane resulting from overlaying the partition $L^*$ and the regions $H^*$. The cells of this overlay now provide the decomposition of the dual space that we need, allowing to distinguish the (primal space) lines that stab $P_1$ and $P_2$, but no obstacle.

**Lemma 2.3.** 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 that we can compute the visibility within each cell of the overlay of $L^*$ and $H^*$ in constant time.

**Proof.** The polygons $P_1$ and $P_2$, and the obstacles are bounded by a total of $O(m+n)$ line segments in primal space, thus the overlay has complexity $O((m+n)^2)$. If the visibility can be computed in constant time within each cell of the arrangement in the dual space, it takes $O((m+n)^2)$ total time to compute the probability of
visibility between a pair of points chosen uniformly at random from \( P_1 \times P_2 \).

It remains to show that we can compute the visibility within every cell of the overlay in constant time. In the following we show that this is indeed the case, that is, we provide the steps necessary to calculate how much the lines of a given cell contribute to \( \text{prob} \). For simplicity of presentation, we assume that \( P_1 \) and \( P_2 \) are separable by a vertical line, and that \( P_1 \) 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. In case that \( P_1 \) and \( P_2 \) cannot be separated by a vertical line, the convexity of \( P_1 \) and \( P_2 \), together with their disjointness, guarantee that we can find a non-vertical separating line, and rotate the scene to make the line vertical.

Consider line \( \ell \), given by the formula \( y = \alpha x - \beta \), which goes through points \( p_1 = (x_1, y_1) \in P_1 \) and \( p_2 = (x_2, y_2) \in P_2 \). In dual space, point \( \ell^* \), corresponding to line \( \ell \), has coordinates \((\alpha, \beta)\). Substitute variables \( y_1 \) and \( y_2 \) in Formula 2.3 with \( \alpha \) and \( \beta \): \((x_1, y_1, x_2, y_2) \leftarrow (x_1, \alpha, x_2, \beta)\), 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

\[
\text{prob} = \frac{\iiint v(x_1, \alpha, x_2, \beta)|J|dx_1 dx_2 d\alpha d\beta}{\iiint |J|dx_1 dx_2 d\alpha d\beta},
\]

where

\[
J = \det \begin{bmatrix}
\frac{d y_1}{d x_1} & \frac{d y_1}{d \alpha} \\
\frac{d y_2}{d x_1} & \frac{d y_2}{d \alpha}
\end{bmatrix} = \det \begin{bmatrix}
\frac{d y_1}{d x_2} & \frac{d y_1}{d \beta} \\
\frac{d y_2}{d x_2} & \frac{d y_2}{d \beta}
\end{bmatrix} = x_2 - x_1.
\]

The expression in Equation 2.4 can be solved analytically. We defer the tedious details to the technical supplement of this article [5].

**Theorem 2.3.** 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.

Combining Theorems 2.2 and 2.3, the main result of this section follows:

**Theorem 2.4.** Given two imprecise points, modeled as Gaussian distributions \( \mu_1 \) and \( \mu_2 \) with standard deviations \( \sigma_1 \) and \( \sigma_2 \), and obstacles of total complexity \( m \), we can approximate the probability that \( p \) and \( q \) see each other with an error at most \( \varepsilon \) in \( O(1/\varepsilon^2 + m/\varepsilon^2) \) time.

## 3 Experiments

We implemented our method in C++ using the Computational Geometry Algorithms Library (CGAL), and tested it on synthetically generated and real data.

### 3.1 Setup

We generated several random scenarios, each consisting of two imprecise points modeled as Gaussian distributions and approximated by sets of layered regions, as explained in Section 2. These two Gaussian distributions are parameterized by \( \sigma \), their standard deviation, while the approximations are parameterized by \( \varepsilon \), the desired accuracy of the approximations.
natural, but clearly distinct, environments. The forest:
urban:
A set of 24 footprints of real buildings, representing a section of the old city center of Utrecht, spanning a 550m × 350m region. Figure 6(b) shows an example.

The two settings were chosen as to illustrate two natural, but clearly distinct, environments. The forest setting showcases a large number of obstacles of small size, whereas the urban setting contains just a few larger obstacles, which create a maze-like domain.

For the obstacles, we distinguish two settings.

forest: A set of 300 randomly generated obstacles, each representing the trunk of a tree, modeled as a regular pentagon, in a 100m × 100m domain. Figure 6(a) shows an example.

urban: A set of 24 footprints of real buildings, representing a section of the old city center of Utrecht, spanning a 550m × 350m region. Figure 6(b) shows an example.

We carried out a number of different experiments to, firstly, investigate how the region-based model changes when the parameters ε and σ vary, and secondly, to compare the region-based model to the point-based approach.

3.1.1 Goals of the experiments We begin our experimental evaluation with an analysis of the fundamental properties of our region-based approximation scheme. For this we first study the effect of changing the parameter of the approximation, specifically ε. We expect that the true value of the error of the visibility, calculated by our method, will be significantly smaller than our theoretical upper bounds. Thus, in applications where a provable worst-case error bound is not required but
an empirical error bound suffices, the value $\varepsilon$ of the provable error bound does not need to be very small and can be adjusted accordingly. We also evaluate the dependency of the visibility on the standard deviation $\sigma$ of the distributions.

With a suitable $\varepsilon$ in place, next we want to compare the point-based and region-based approximations of the probability distributions of the imprecise points. For this we study the convergence behavior of point-based methods. We also investigate the dependency of the visibility on $\sigma$. We want to know how our method works at different levels of precision of the coordinates of our entities.

In summary, our experiments address the following questions:

1. How do the results depend on various parameters? Our parameters are
   - $\varepsilon$: the approximation quality, and
   - $\sigma$: the standard deviation of the probability distributions.

2. How does the region-based approach compare to a simple point-based approach, where we approximate the probability by sampling a large number of pairs of points from the two distributions? Particularly, we wish to know
   - whether the region-based and point-based approaches give consistent results, that is, whether they yield similar probabilities in the limit case, and
   - how many samples in the point-based approach are needed before we get a reasonably stable approximation.

3.2 Results In the first experiment we evaluated the total size $n$ of the approximation regions depending on $\varepsilon$. Figure 7 shows the guaranteed error as a function of $\varepsilon$. The plot exhibits the inverse behavior of Theorem 2.2, as expected, and shows the constants as they occur in practice.

Secondly, knowing the dependency between $n$, the size of approximation, and $\varepsilon$, we plot the computed visibility as a function of $n$ (for the forest setting presented in Figure 6(a)). We can see in Figure 8 that the visibility seems to stabilise around $\approx 0.185$. This indicates that the true error, as expected, is much lower than the theoretically guaranteed error that our method provides. More specifically, we observe that the empirical error is already less than 0.03 for $n = 26$.

Figure 10 shows how the computed visibility varies with $\sigma$ in the same forest setting. We observe that the probability of visibility grows fast while $\sigma$ is less than 2.5, and then stabilizes with a slow steady grow, which can be explained by the fact that in the particular setting of the Figure 6(a) there are a number of trees blocking the visibility between the centers of the distributions. Furthermore, we again observe that the dependency of the result on $\varepsilon$ is very low, due to the fact that the empirical error is much lower than the provable error bound.

To compare our method with the point-based method we iteratively sampled the distributions in the urban setting (Figure 6(b)). Figure 9(a) shows the probability of visibility as a function of $(n + m)^2$, which is in the order of the complexity of the partition in the dual space. Figure 9(b) shows the estimated probability of the point-based method as a function of the number of samples. Since the probability of visibility in this particular setting is very small, it takes about 20,000 samples before the point-based method even detects that the probability is non-zero. In comparison, our method estimates a non-zero probability even for the smallest complexity ($\varepsilon = 0.5; n = 44$). Furthermore, we see that the probability converges very slowly for the point-based method. Additionally, the experiments provide empirical evidence of the fact that the point-based approach not only results in a relatively slow convergence rate for the error, but also that any guarantee on the error would need to come with a considerable probability of failure. In contrast, our region-based approach produces results with error guaranteed to be smaller than the input parameter $\varepsilon$. Again the empirical error of the region-based approach is significantly lower than the guaranteed error: for $\varepsilon = 0.05$ the approach computes a probability of 0.000032, which seems to only slightly underestimate the actual probability.
4 Conclusions

In this paper we presented geometric techniques that provide a sound estimation of the probability of visibility between imprecise points amidst a group of obstacles. These techniques are based on approximating these probability distributions by sets of polygons. This type of approximation allowed us to give guarantees on how closely the computed probabilities approximate the actual probabilities. Our experiments provide evidence that these techniques provide strong bounds, and in the settings tested give better results than point-based approximation techniques.

Given that the empirically obtained error is much smaller than the analytically obtained upper bound on the error, an interesting open problem is whether we can formulate conditions under which tighter theoretical guarantees on the error can be proven. Furthermore, with the region-based approximation proven successful in the setting of visibility, we plan to investigate other geometric problems on imprecise points for which the region-based approximation of probability distributions results in faster geometric approximation algorithms.

Acknowledgments

Kevin Buchin, Irina Kostitsyna, and Maarten Löffler are supported by the Netherlands Organisation for Scientific Research (NWO) under grant no. 612.001.106, 612.001.207, and 639.021.123, respectively. Rodrigo Silveira 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, Rodrigo Silveira was partially supported by projects MINECO MTM2012-30951/FEDER, Gen. Cat. DGR2014SGR46, and by ESF EUROCORES program EuroGIGA-ComPoSe IP04-MICINN project EUI-EURC-2011-4306.

References

[1] A. S. Aspbury and R. M. Gibson. Long-range visibility of greater sage grouse leks: a GIS-based analysis. Animal Behaviour, 67(6):1127 – 1132, 2004.
[2] D. Bandyopadhyay and J. Snoeyink. Almost-Delaunay simplices: Nearest neighbor relations for imprecise points. In SODA, 2004.
[3] P. Bartie, F. Reitsma, S. Kingham, and S. Mills. Advancing visibility modelling algorithms for urban environments. Computers, Environment and Urban Systems, 34(6):518 – 531, 2010. GeoVisualization and the Digital City Special issue of the International Cartographic Association Commission on GeoVisualization.
[4] M. J. Berg, O. Cheong, M. v. Kreveld, and M. Overmars. Computational Geometry: Algorithms and Applications. Springer-Verlag TELOS, Santa Clara, CA, USA, 3rd ed. edition, 2008.
[5] K. Buchin, I. Kostitsyna, M. Löffler, and R. I. Silveira. Region-based approximation of probability distributions (for visibility between imprecise points among obstacles), 2014. Preprint, http://arXiv.org/abs/1402.5681.
[6] J. Chen and R. Cheng. Efficient evaluation of imprecise location-dependent queries. In Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on, pages 586–595, April 2007.
[7] R. Cheng, D. Kalashnikov, and S. Prabhakar. Querying imprecise data in moving object environments. Knowledge and Data Engineering, IEEE Transactions on, 16(9):1112–1127, Sept 2004.
[8] D. J. Dean. Improving the accuracy of forest viewsheds using triangulated networks and the visual permeability method. Canadian Journal of Forest Research, 27:969–977, 1997.
[9] P. F. Fisher. Algorithm and implementation uncertainty in viewshed analysis. International Journal of Geographical Information Systems, 7(4):331–347, 1993.
[10] P. F. Fisher. Models of uncertainty in spatial data. In P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, editors, Geographical information systems: principles, techniques, management and applications. John Wiley & Sons, Inc., Chichester, 1999.
[11] L. D. Floriani and P. Magillo. Algorithms for visibility computation on terrains: a survey. Environ. Plann. B, 30(5):709–728, 2003.
[12] S. Ghosh. Visibility Algorithms in the Plane. Cambridge University Press, New York, NY, USA, 2007.
[13] L. J. Guibas, D. Salesin, and J. Stolfi. Epsilon geometry: building robust algorithms from imprecise computations. In SoCG, 1989.
[14] L. J. Guibas, D. Salesin, and J. Stolfi. Constructing strongly convex approximate hulls with inaccurate primitives. Algorithmica, 9:534–560, 1993.
[15] J. Horne, E. Garton, S. Krone, and J. Lewis. Analyzing animal movements using Brownian bridges. Ecology, 88(9):2354–2363, 2007.
[16] B. Joe and R. B. Simpson. Corrections to Lee’s visibility polygon algorithm. BIT, 27:458–473, 1987.
[17] A. Jørgensen, M. Löffler, and J. Phillips. Geometric computations on indecisive points. In Proc. 12th Algorithms and Data Structures Symposium, LNCS 6844, pages 536–547, 2011.
[18] M. J. Katz, M. H. Overmars, and M. Sharir. Efficient hidden surface removal for objects with small union size. Comput. Geom., 2:223–234, 1992.
[19] M. Llobera. Modeling visibility through vegetation. International Journal of Geographical Information Science, 21(7):799–810, 2007.
[20] M. Löffler. Data Imprecision in Computational Geometry. PhD thesis, Utrecht University, 2009.
[21] M. Löffler and J. Phillips. Shape fitting on point sets with probability distributions. In ESA, 2009.
[22] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind. Geographical information systems and science. John Wiley & Sons, 3rd edition, 2011.
[23] E. Morello and C. Ratti. A digital image of the city:
3d isovists in Lynch’s urban analysis. *Environment and Planning B: Planning and Design*, 36(5):837–853, 2009.

[24] K. Nackaerts, G. Govers, and J. V. Orshoven. Accuracy assessment of probabilistic visibilities. *International Journal of Geographical Information Science*, 13(7):709–721, 1999.

[25] 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., LNCS 2008, pages 252–263, 2000.

[26] Y. Ostrovsky-Berman and L. Joskowicz. Uncertainty envelopes. In *Abstracts 21st European Workshop on Comput. Geom.*, pages 175–178, 2005.

[27] D. Pfoser and C. S. Jensen. Capturing the uncertainty of moving-object representations. In R. H. Güting, D. Papadias, and F. H. Lochovsky, editors, *SSD*, volume 1651 of Lecture Notes in Computer Science, pages 111–132. Springer, 1999.

[28] D. T. S. Richard J. Camp and R. L. Knight. Viewsheds: A complementary management approach to buffer zones. *Wildlife Society Bulletin*, 25(3):612–615, 1997.

[29] P. Sistla, O. Wolfson, S. Chamberlain, and S. Dao. Querying the uncertain position of moving objects. In *In Temporal Databases: Research and Practice*, pages 310–337. Springer, 1998.

[30] C. Tandy. The isovist method of landscape survey. In *Symposium Methods of landscape analysis*, pages 9–10, 1967.

[31] M. Tomko, F. Trautwein, and R. S. Purves. Identification of practically visible spatial objects in natural environments. In M. Sester, L. Bernard, and V. Paelke, editors, *AGILE Conf.*, Lecture Notes in Geoinformation and Cartography, pages 1–23. Springer, 2009.

[32] A. Turner, M. Doxa, D. O’Sullivan, and A. Penn. From isovists to visibility graphs: a methodology for the analysis of architectural space. *Environment and Planning B: Planning and Design*, 28:103–121, 2001.

[33] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan. The unscented particle filter. In *Advances in Neural Information Processing Systems*, volume 8, pages 351–357, 2000.

[34] F. Van Diggelen. GNSS Accuracy: Lies, Damn Lies and Statistics. *GPS World*, pages 26–32, 2007.

[35] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. *Comput. Netw.*, 52(12):2292–2330, 2008.