Probabilistic Analysis for Combinatorial Functions of Moving Points*

Li Zhang† Harish Devarajan Julien Basch‡ Piotr Indyk§

Computer Science Department
Stanford University
Stanford, CA94305
{lizhang,harish,jbasch,indyk}@cs.stanford.edu

Abstract

We initiate a probabilistic study of configuration functions of moving points. In our probabilistic model, a *particle* is given an initial position and a velocity drawn independently at random from the same distribution $\mathcal{D}$. We show that if n particles are drawn independently at random from the uniform distribution on the square, their convex hull undergoes $\Theta(\log^2 n)$ combinatorial changes in expectation, their Voronoi diagram undergoes $\Theta(n^{3/2})$ combinatorial changes, and their closest pair undergoes $\Theta(n)$ combinatorial changes.

1 Introduction

Given a set of n points, what is the description complexity of their convex hull? In our world, this question is understood with an implicit "in the worst case", and the answer is $n^{\lfloor d/2 \rfloor}$ where d is the dimension of the underlying space. This is not entirely satisfactory, as this description complexity can vary tremendously depending on the positions of the points. Another approach is to look at the expected description complexity when the points are drawn from a given distribution. This type of analysis, initiated by Rényi and Sulanke [RS63] and pursued by others [Efr65, Ray70, Dwy90], gets its value from the fact that this expectation is in general much smaller than in the worst case, and, more importantly, in that it often allows one to design algorithms that have expected running times against which worst case aware algorithms cannot compete [Sei86, BCL90, Dwy91]. For instance, the convex hull of n points drawn independently uniformly at random from a d-dimensional hypercube has expected complexity $O(\log^{d-1} n)$, and can be computed in expected linear time.

In parallel, in the past decade, a number of papers have considered a setting where points are allowed to move along low degree algebraic trajectories [OW84, Ata85, GMR91, KTI95, BGH97]. Different questions have been asked in this context. In particular, Atallah [Ata85], studied the number of times the combinatorial description of the convex hull or closest pair can change, in the worst case ("dynamic computational geometry"). More recently, Basch, Guibas, and Hershberger [BGH97] have designed *kinetic* data structures to maintain these attributes in an *online* setting, measuring the quality of a kinetic data structure by the ratio of the worst case number of changes to the configuration of interest, to the worst case number of changes to the data structure itself, for low degree algebraic motions. This measure is not ideal and would gain to be replaced by one similar to the competitive ratio, but there is no result in this direction yet. In the meanwhile, an experimental study has been undertaken to assess the quality of these data structures in practice [BGSZ], showing that the worst case analysis can hide vastly different results in terms of expectation when the point positions and speeds are drawn at random from some distributions. It is this study that motivated the present paper.

Indeed, in the light of previous works on static problems, it is natural to study theoretical bounds for the expected number of changes of combinatorial functions of moving points, when these points are drawn from prescribed distributions, as well as expected time bounds for kinetic data structures that maintain these combinatorial functions.

We report here tight bounds on the expected number of changes to the convex hull, the Delaunay triangulation, and the closest pair of points in the plane for a distribution uniform on a unit square. We also give related (not tight) results, and analyze the expected run-

*The authors' names appear in random order.

†Supported in part by National Science Foundation grant CCR-9623851 and by US Army MURI grant DAAH04-96-1-0007.

‡Supported in part by the Okawa Foundation.

§Supported in part by ARO MURI Grant DAAH04-96-1-0007 and by NSF Award CCR-9357849, with matching funds from IBM, Mitsubishi, Schlumberger Foundation, Shell Foundation, and Xerox Corporation.

ning time of one of the algorithms proposed in [BGH97] for the maintenance of the convex hull.

2 Probabilistic Setting

As a point in the plane cannot move without losing its identity, we call the objects of our study *particles.* A particle X is given an initial position X_0 at time $t = 0$ and a constant velocity V_X, so that it has a position $X(t) = X_0 + tV_X$ at any time $t \geq 0$. Given a distribution $\mathcal{D}$ in $\mathcal{R}^d$, we say that a particle is drawn from $\mathcal{D}$ if its initial position and velocity are each drawn independently at random from $\mathcal{D}$. In [BGSZ], a number of particles were independently drawn at random this way, and the number of events of several kinetic data structures, for several underlying distributions, were studied experimentally.

Using an appropriate space-time transformation, it is easy to see that the above model is combinatorially equivalent to a model where a particle is given an initial position X_0 and a final position X_1 independently drawn at random from $\mathcal{D}$. In this setting, the position of X at time $t \in [0,1]$ is given by $X(t) = (1-t)X_0 + tX_1$ (X moves at a constant speed from X_0 to X_1). The proofs are easier to carry out in this second model.

3 Results

In this communication, we report some results in the plane, when points are drawn from the uniform distribution on the unit square.

Theorem 1 *If n particles are drawn independently at random from the uniform distribution on the unit square, then the expected number of changes to the following combinatorial functions is:*

Closest pair	$\Theta(n)$
Delaunay triangulation	$\Theta\left(n^{3/2}\right)$
Convex hull	$\Theta\left(\log^2 n\right)$

Some comparisons will help put these results in perspective. In the case of the convex hull, its expected size in the static setting $\Theta(\log n)$, while the worst case number of changes in the dynamic setting is $\Theta\left(n^2\right)$. On the other hand, the closest pair has a static description complexity of $O(1)$, and a worst case number of changes of $\Theta\left(n^2\right)$, but the expected number of changes seems surprisingly high. Finally, the static Delaunay triangulation always has linear size, while no tight bounds are known for the number of changes in the worst case when points move (there is a trivial lower bound of $\Omega(n^2)$ and an upper bound of $O(n^3)$). It is pleasant to see that the expected value was rather easy to obtain.

The probabilistic analysis can be applied to all configuration functions that are typically looked at in the setting of moving points, and to the kinetic data structures that are used to maintain them.

For instance, as the minimum spanning tree is a subgraph of the Delaunay triangulation, the result on the Delaunay triangulation coupled with a standard batching argument imply that, in our model, the MST in the plane changes $O(n^{5/2})$ times on average. Although this is probably not a tight bound, it is to be compared with the best known worst case upper bound of $O\left(n^3 2^{\alpha(n)}\right)$ [KTI95].

The furthest pair, on the other hand, is always constituted of two points that are on the convex hull. From this and Devroye's moment inequalities [Dev83], we deduce that the number of changes to the furthest pair is polylogarithmic in expectation, for the square distribution.

In [BGH97], a kinetic data structure is described to maintain the convex hull of a set of moving points. This data structure divides the point set arbitrarily into a blue and a red half, recursively computes the value and red convex hulls, and maintains a set of certificates between edge-vertex pairs and edge-edge pairs to certify the convex hull of the whole set. Our convex hull result directly imply:

Theorem 2 *If n particles are drawn independently at random from the uniform distribution on the unit square, then the kinetic data structure of [BGH97] processes a linear number of events in expectation to maintain the convex hull.*

Hence, whatever the worst case number of changes to the Delaunay triangulation turns out to be, this type of result suggest strongly that it is not a good idea to use it to maintain the convex hull. Our probabilistic framework provides a way that is different from and complementary to [BGH97] for analysing and comparing kinetic data structures.

We now review several kinetic data structures for the maintenance of the closest pair. The Delaunay triangulation can be used to maintain the closest pair, and will do it in roughly $n^{3/2}$ time in our probabilistic model. How do other methods compare? A method is to maintain the L_1 Delaunay triangulation [CD85], as one of its edges is the closest pair, but our method shows that it undergoes roughly the same number of changes as the L_2 Delaunay. Another method was proposed in [BGH97] and modified in [BGSZ], whose average running time is experimentally also roughly $\Theta\left(n^{3/2}\right)$. At last, one may consider a more straightforward algorithm, which cuts the square (say) into n cells and tracks every particle as it goes from cell to cell [KSG]: although extremely bad in the worst case, it is easy to see that the average case of this algorithm in our probabilistic model is also precisely $\Theta\left(n^{3/2}\right)$. There is probably a good reason for that.

4 Some proof ideas

The basic idea behind the proofs of the three results of theorem 1 is as follows. Firstly, we recall the approach of Rényi and Sulanke [RS63] for computing the expected size of the static convex hull. The idea is to consider a given pair of random points (P, Q), and to compute the probability that this pair forms an edge of the convex hull, i.e. that all other points $R_1, \cdots, R_n$, whose positions are chosen independently, are on the same side (say to the right) of the line passing through (P, Q). The positions of the other points with respect to the line (PQ), although not independent (as they all depend on P, Q), are independent conditionally on P, Q. The recipe, then, for obtaining the required probability, is to compute the probability density of the line defined by two random points, and integrate $(1 - G(\ell))^n$ with respect to this density, where $G(\ell)$ is the probability that a point lies to the left of line ℓ.

When points start to move, we can use basically the same approach, but we now need three points to characterize the combinatorial event that the convex hull changes. Hence, the question becomes: Given three particles P, Q, R, what is the probability that, when they become collinear, they are on the convex hull ? Conditioning on the position of these three points at that time, we are left with the following two problems:

1. what is the probability that, at a given time t, a particle is to the left of a line ℓ ?

2. what is the probability density of the time and line on which three particles are collinear ?

For the closest pair and the Delaunay triangulation, these questions are replaced by similar questions that involve four points instead of three. The proof of Theorem 1 requires a detailed case analysis for different regions of the unit square. To our knowledge, this case analysis is unavoidable.

5 Conclusion

Further work is called for to investigate the probabilistic behavior of the number of changes of geometric structures on moving points. In particular, the results in this paper could be generalized to arbitrary dimensions, to more general distributions, and to algebraic motion within an appropriate probabilistic model. We also hope that some problems that are extremely difficult to analyse in the worst case will be amenable to our analysis to obtain tight answers. The minimum spanning tree is the most interesting of these problems (in part due to its connection with the k-level)

Acknowledgments

We would like to express our gratitude to Alexei Novikov and Suresh Venkatasubramanian for several very useful discussions, and Leo Guibas for his constant encouragements.

References

[Ata85] M. J. Atallah. Some dynamic computational geometry problems. *Comput. Math. Appl.*, 11:1171–1181, 1985.

[BCL90] J. L. Bentley, K. L. Clarkson, and D. B. Levine. Fast linear expected-time algorithms for computing maxima and convex hulls. In *Proc. 1st ACM-SIAM Sympos. Discrete Algorithms*, pages 179–187, 1990.

[BGH97] J. Basch, L. J. Guibas, and J. Hershberger. Data structures for mobile data. In *8th Symposium on Discrete algorithms*, pages 747–756, 1997.

[BGSZ] J. Basch, L. J. Guibas, C. Silverstein, and L. Zhang. A practical evaluation of kinetic data structures. This volume.

[CD85] L. P. Chew and R. L. Drysdale, III. Voronoi diagrams based on convex distance functions. In *Proc. 1st Annu. ACM Sympos. Comput. Geom.*, pages 235–244, 1985.

[Dev83] L. P. Devroye. Moment inequalities for random variables in computational geometry. *Computing*, 30:111–119, 1983.

[Dwy90] R. A. Dwyer. Kinder, gentler average-case analysis for convex hulls and maximal vectors. *SIGACT News*, 21(2):64–71, 1990.

[Dwy91] R. A. Dwyer. Higher-dimensional Voronoi diagrams in linear expected time. *Discrete Comput. Geom.*, 6:343–367, 1991.

[Efr65] B. Efron. The convex hull of a random set of points. *Biometrika*, 52(3):331–343, 1965.

[GMR91] L. Guibas, J. S. B. Mitchell, and T. Roos. Voronoi diagrams of moving points in the plane. In *Proc. 17th Internat. Workshop Graph-Theoret. Concepts Comput. Sci.*, volume 570 of *Lecture Notes in Computer Science*, pages 113–125. Springer-Verlag, 1991.

[KSG] D.-J. Kim, S.-Y. Shin, and L. J. Guibas. Fast collision detection among multiple moving spheres. This volume.

[KTI95] N. Katoh, T. Tokuyama, and K. Iwano. On minimum and maximum spanning trees of linearly moving points. *Discrete Comput. Geom.*, 13:161–176, 1995.

[OW84] T. Ottmann and D. Wood. Dynamical sets of points. *Comput. Vision Graph. Image Process.*, 27:157–166, 1984.

[Ray70] H. Raynaud. Sur l'enveloppe convexe des nuages de points aléatoires. *J. Appl. Prob.*, 7:35–48, 1970.

[RS63] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gerwähten Punkten I. *Z. Wahrsch. Verw. Gebiete*, 2:75–84, 1963.

[Sei86] R. Seidel. Constructing higher-dimensional convex hulls at logarithmic cost per face. In *Proc. 18th Annu. ACM Sympos. Theory Comput.*, pages 404–413, 1986.