Probabilistic Analysis for Combinatorial Functions of Moving Points *

Julien Basch Harish Devarajan Piotr Indyk Li Zhang

Computer Science Department
Stanford University
Stanford, CA94305

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 the world of analysis of algorithms, 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 type of answer 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].

In the past years, a number of papers have considered a setting where points are allowed to move along low degree algebraic trajectories [OW84, Ata85, Meg86, GMR91, DG93, KTI95, BGH97]. Different questions have been asked in this context. In particular, Atallah, in his pioneering work that introduced Davenport-Schinzel sequences in computational geometry [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 data structures (called *kinetic*) 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

* Support was provided in part by National Science Foundation grant CCR–9623851 and by US Army MURI grant 5–23542–A.

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 akin 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, with its surprising experimental results, that motivated the present paper.

Given these considerations, it is natural to study theoretical bounds for the expected number of changes of combinatorial functions of moving points drawn from prescribed distributions, as well as expected time bounds for kinetic data structures that maintain these combinatorial functions. For instance, it is natural to ask what is the expected number of topological changes to the Delaunay triangulation (a question whose static equivalent is of little interest).

In this paper, we obtain 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 analyze the expected running time of one of the algorithms proposed in [BGH97] for the maintenance of the convex hull.

In the next section, we present the probabilistic setting. Section 3 is devoted to the convex hull, section 4 to the Voronoi diagram, and section 5 to the closest pair. In section 6, we list some related results, and conclude in the last section.

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). It is this model that we consider in the rest of this paper.

When a particle is drawn from the square distribution, the density $sq_t(x, y)$ of its position (x, y) at time t is equal to $f_t(x)f_t(y)$, where:

$$f_t(z) = \begin{cases} \frac{z}{t(1-t)} & \text{if } z \leq t \\ \frac{1}{1-t} & \text{if } t \leq z \leq 1-t \\ \frac{1-z}{t(1-t)} & \text{if } z \geq t \end{cases}$$

for $t \leq \frac{1}{2}$ (the case $t > \frac{1}{2}$ is symmetric). For future reference, we note that $t^2(1-t)^2 sq_t(x, y) = xy$ for $y \leq x \leq t \leq \frac{1}{2}$.

3 Convex hull

We briefly 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).

$$\begin{aligned}
\Pr[PQ \text{ on left side of convex hull}] &= \Pr[\forall i, R_i \text{ to the right of } PQ] \\
&= E[\Pr[\forall i, R_i \text{ to the right of } PQ|P,Q]] \\
&= E\left[\prod_i \Pr[R_i \text{ to the right of } PQ|P,Q]\right]
\end{aligned}$$

In words, 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 ?

Although these questions can be answered exactly, the computations for the second problem are quite involved. We focus here on simplified calculations that can be performed relatively easily, while still giving the right order of magnitude. At the heart of the simplification lies the intuitive idea that

$$\begin{aligned}
Pr[PQR \text{ collinear on } \ell \text{ at time } t] &= Pr[P(t),Q(t),R(t)\subset \ell] \\
&= Pr[P(t)\in\ell]Pr[Q(t)\in\ell]Pr[R(t)\in\ell]
\end{aligned}$$

The second equality holds, but only because both sides are zero. However, this hides a fatal flaw that would make the equation false even if these calculations were to be performed with the associated densities: the last line assumes that the three events are independent, which is not the case as ℓ and t are not fixed, but depend on all three points. Still, we show how this idea can be converted into a formal, and correct, argument at the loss of only a constant factor. The calculations are still heavy in detail, although considerably less so than the exact method. We present some details for the convex hull of particles drawn from a square distribution, and sketch how the same calculations can be done for the Delaunay triangulation.

3.1 Square distribution

In this section, we will find upper and lower bounds for the number of combinatorial changes to the convex hull of n particles drawn from the uniform square distribution.

Given $a, b > 0$ we denote ℓ_{ab} the line passing through the points $(a, 0)$ and $(0, b)$, by T_{ab} the triangle delimited by the two axes and ℓ_{ab}, and by $\Delta_{ab}(\epsilon_a, \epsilon_b)$ the quadrilateral $T_{a+\epsilon_a, b+\epsilon_b} - T_{ab}$. We denote by $\mathcal{L}_{tab}(X, \epsilon_t, \epsilon_a, \epsilon_b)$ the event that a particle X passes through $\Delta_{ab}(\epsilon_a, \epsilon_b)$ between time t and time $t + \epsilon_t$. And last, let $T(X, Y, Z)$ be the first time at which particles X, Y, Z are collinear, and $A(X, Y, Z), B(X, Y, Z)$ be the parameters (as described above) of the line (XYZ) at that time, which are defined almost surely.

The calculations, although simple, are lengthy. The lesson of the static convex hull, however, is that "everything happens at the corners". We will thus restrict the study to the combinatorial changes for which $A(X, Y, Z) \leq B(X, Y, Z) \leq T(X, Y, Z) \leq 1/2$, and such that the convex hull is entirely to the right of (XYZ). It can be checked that the other cases for A and B give only lower order terms, while events that happen when $t \geq 1/2$ can be analyzed by symmetry.

Lemma 3.1. *Let $a < b < t < 1/2$. Let X be a particle drawn from the uniform square distribution. Let $F(a, b, t)$ be the probability that X remains on the same side of ℓ_{ab}* **until** *time t. Then:*

$$\Pr[X(t) \in T_{ab}] = \Theta(1)\frac{a^2b^2}{t^2}$$

$$\frac{\partial}{\partial a}F(a, b, t) = \Theta\left(\frac{ab^2}{t^2}\right), \qquad \frac{\partial}{\partial b}F(a, b, t) = \Theta\left(\frac{a^2b}{t^2}\right), \qquad \frac{\partial}{\partial t}F(a, b, t) = \Theta\left(-\frac{a^2b^2}{t^3}\right)$$

Moreover, with the above definition of $\mathcal{L}_{tab}$,

$$\Pr[\mathcal{L}_{tab}(X, \epsilon_t, \epsilon_a, \epsilon_b)] = \frac{a^2b^2}{t^2(1-t)^2}O\left(\frac{\epsilon_a}{a} + \frac{\epsilon_b}{b} + \frac{\epsilon_t}{t}\right)$$

Proof. Omitted in this abstract. □

In the following lemma, we show that the density of the time and line parameters of a collinearity event involving X, Y, Z can be computed as if the fact that X, Y, Z belong to the line they define were independent events. This will only make us lose a constant factor.

Lemma 3.2. *Let X, Y, Z be three particles drawn independently from the uniform square distribution. Let $h(t, a, b)$ denote the joint density of $(T(X, Y, Z), A(X, Y, Z), B(X, Y, Z))$. Then, for $a \leq b \leq t \leq 1/2$,*

$$h(t, a, b) = \Theta\left(\frac{a^5b^5}{t^7}\right)$$

Proof. Fix a, b, t, and let $\mathcal{C}(\epsilon_t, \epsilon_a, \epsilon_b)$ be the event that, during the window of time between t and $t + \epsilon_t$, the three particles are collinear on a line whose parameters lie respectively in $[a,\, a + \epsilon_a]$ and $[b,\, b + \epsilon_b]$. With the above definition of $\mathcal{L}_{tab}$, we have:

$$\mathcal{C} \subseteq \mathcal{L}_{tab}(X, \epsilon_t, \epsilon_a, \epsilon_b) \cap \mathcal{L}_{tab}(Y, \epsilon_t, \epsilon_a, \epsilon_b) \cap \mathcal{L}_{tab}(Z, \epsilon_t, \epsilon_a, \epsilon_b)$$

and, this time, the three events on the right hand side are independent. Hence,

$$\begin{aligned}\Pr[\mathcal{C}] &\leq \Pr[\mathcal{L}(X,\epsilon_t,\epsilon_a,\epsilon_b)]\Pr[\mathcal{L}(Y,\epsilon_t,\epsilon_a,\epsilon_b)]\Pr[\mathcal{L}(Z,\epsilon_t,\epsilon_a,\epsilon_b)]\\ &= \frac{a^6b^6}{t^6}\left(\frac{\epsilon_a}{a}+\frac{\epsilon_b}{b}+\frac{\epsilon_t}{t}\right)^3\end{aligned}$$

Now, $\frac{\Pr[\mathcal{C}(\epsilon_t,\epsilon_a,\epsilon_b)]}{\epsilon_a\epsilon_b\epsilon_t} \xrightarrow{(\epsilon_t,\epsilon_a,\epsilon_t)\to 0} h(t,a,b)$ We can obtain an upper bound for h by taking the limit along an appropriate path to the origin of the upper bound to $\Pr[\mathcal{C}]$ found above (although it does not converge along all paths). In particular, setting $\epsilon = \frac{\epsilon_t}{t} = \frac{\epsilon_a}{a} = \frac{\epsilon_b}{b}$, and letting $\epsilon \to 0$, we obtain the desired upper bound.

The lower bound can be obtained by considering three regions in $\Delta_{ab}(\epsilon_a,\epsilon_b)$ such that any line passing through these three regions has parameters in $[a,a+\epsilon_a]\times[b,b+\epsilon_b]$, putting one point in each region, and taking the limit. Here again, only a constant factor is lost compared to the exact density, which takes into account the correlations. We omit the details in this abstract. □

Lemma 3.3. *If n particles are drawn independently at random from the uniform distribution over the unit square, then the expected number of changes to their convex hull is $\Theta\left(\log^2 n\right)$.*

Proof. We will bound the probability that a given triplet is responsible for a combinatorial change of the convex hull, and sum over all triplets. As in the stationary case of the convex hull, the probability that three particles X, Y, Z participate a combinatorial event is

$$K = \int (1-G_{tab})^{n-3}\, h(t,a,b)\,dadbdt$$

where $G_{tab} = \Pr[W(t) \in T_{ab}]$ for any particle W. Here, we present only the calculations for $a \leq b \leq t \leq 1/2$, but the other cases can be done likewise. With the lower and upper bounds on h obtained from Lemma 3.2, we obtain the upper bound:

$$K = \Theta\left(\int \frac{a^5b^5}{t^7}\left(1-\frac{a^2b^2}{t^2}\right)^{n-3}[a \leq b \leq t \leq 1/2]\,dadbdt\right)$$

and it can be checked that this integral is $\Theta\left(\frac{\log^2 n}{n^3}\right)$. Summing over all $\binom{n}{3}$ triplets of particles gives the desired result. □

4 Voronoi diagram

The Voronoi diagram [Vor08], or its dual, the Delaunay triangulation, is probably the structure that has found the most varied applications in computational geometry and other fields. It is also a structure that is very easy to maintain in the kinetic sense, as it is *self-kinetizing*. Several authors have studied the Voronoi diagram of points moving along low degree algebraic curves [ISI89, GMR91, FL91]. In the case of points moving at constant speed, the worst case number of changes is known to be between $\Omega(n^2)$ and $O(n^3)$, and this gap is arguably the most significant open problem in dynamic computational geometry. In contrast, in our probabilistic setting, it is possible to find the exact expected number of changes.

We carry informally the calculations for n particles drawn independently from the uniform square distribution. An event in the Delaunay triangulation of moving points is the cocircularity of four points. We denote by D_{abr} the disk of radius r and center (a, b), and by C_{abr} its boundary. For a disk D_{abr} not too close to the boundary of the distribution (we leave this condition vague on purpose), and for $t > r$, we have $\Pr[X(t) \in D_{abr}] = \Theta(r^2)$. If $\mathcal{L}_{abrt}$ denotes the event that four particles become cocircular during $[t, t+\epsilon_t]$ on a circle whose parameters lie in $[r, r+\epsilon_r], [a, a+\epsilon_a]$, and $[b, b+\epsilon_n]$, we have $\Pr[\mathcal{L}_{abrt}] = \Theta(r(\epsilon_a + \epsilon_b + \epsilon_r + \epsilon_t))$. Hence the number of changes to the Delaunay is about:

$$\int r^4(1-r^2)^{n-4} dr dt da db = \Theta\left(\frac{1}{n^{5/2}}\right)$$

Summing over all 4-tuples of points gives the following:

Theorem 4.1. *The expected number of combinatorial changes to the Delaunay triangulation of n particles independently drawn at random from the uniform square distribution is $\Theta\left(n^{3/2}\right)$.*

The more precise calculations, with the treatment of the boundaries, are long, and we omit them in this abstract.

5 Closest Pair

We now switch to the next archetypical problem of dynamic computational geometry, the closest pair among n moving points. Here, in order to vary the menu, we adopt a discrete approach, quite different in its vocabulary to the one used in the previous sections. We emphasize that the fundamental problems (handling the boundaries or the dependencies) remain the same in both settings.

Here, we split the time interval $[0, 1]$ into small subintervals of length Δt. By using the 'continuity' , i.e. the property that (for suitable Δt) the combinatorial properties of a set of particles cannot change significantly within one interval, we estimate the number of changes occurring during that interval. When multiplied by the number of intervals, this gives an estimate for the total number of changes.

For points P and Q, $|PQ|$ denotes the Euclidian distance between points P and Q. $S(P, r)$ denotes a ball of radius r centered at P.

Let $\Delta t = \frac{2}{n}$ and $T = \{i\Delta t | i = 0 \cdots \frac{n}{2} - 1\}$. Then

Lemma 5.1. *If n particles $X_i = (P_i, Q_i)$ for $i = 1 \cdots n$ are drawn independently at random from the uniform square distribution, then for any pair $i \neq j$ from $1, 2 \cdots n$, the probability that X_i, X_j is the closest pair at time 0, but not at time Δt, is $\Omega(\frac{1}{n^2})$.*

Proof. For simplicity, fix $i = 1$ and $j = 2$. Let $0 < \alpha < 1$ be some constant to be specified later. We define the events E_1, E_2, E_3, E_4, E_5 as follows:

- E_1 : $|X_1(0)X_2(0)| \leq \frac{\alpha}{n}$,
- E_2 : there exists $k, l \geq 3, k \neq l$ such that $|X_k(0)X_l(0)| \leq \frac{1}{n}$,
- E_3 : $|X_{i'}(0)X_{j'}(0)| > \frac{\alpha}{n}$, for every $(i', j') \neq (1, 2)$ or $(2, 1)$,
- E_4 : $|X_1(\Delta t)X_2(\Delta t)| > \frac{1}{n}$,

- E_5 : there exists $k, l \geq 3, k \neq l$ such that $|X_k(\Delta t) X_l(\Delta t)| < \frac{1}{n}$,
- $E = E_1 \wedge E_2 \wedge E_3 \wedge E_4 \wedge E_5$.

E implies that X_1, X_2 is the closest pair at time 0 but not at Δt. Hence it is sufficient to show that $\Pr[E] = \Omega(\frac{1}{n^2})$. To this end, we assume that we choose the parameters of $X_1 \cdots X_n$ by first choosing $P_1 \cdots P_n$ (in this order) and then $Q_1 \cdots Q_n$. Note that

$$\Pr[E] = \Pr[E_1] \cdot \Pr[E_2 \wedge E_3 | E_1] \cdot \Pr[E_4 | E_1 \wedge E_2 \wedge E_3] \cdot \Pr[E_5 | E_1 \wedge E_2 \wedge E_3] \wedge E_5$$

Let c, c_1, c_2 and c_3 be unspecified constants. Then the following facts hold:

Fact 5.2. $\Pr[E_1] \geq \frac{c}{n^2}$.

Fact 5.3. $\Pr[E_2 \wedge E_3 | E_1] > c_1 > 0$.

Consider now $\Pr[E_4 | E_1 \wedge E_2 \wedge E_3]$. First off, note that E_4 is independent of E_2. Now, the position of X_1 (or X_2) at time Δt depends on Q_1 (or Q_2), which are chosen independently of P_1, P_2, and hence of E_1 and E_3. Specifically, if the distances $|P_1 Q_1|$ and $|P_2 Q_2|$ are at least constant (so X_1 and X_2 move at a constant speed), and the points are moving in significantly different directions, then after Δt time the points are further away from each other by $\Omega(\frac{1}{n})$. Hence it can be easily verified that despite restrictions on P_1 and P_2 due to E_1 and E_3, the probability of choosing Q_1 and Q_2 such that E_4 holds is constant. Similar arguments hold also for E_5 - we have to use the fact that E_2 holds for some k, l and that the same pair of points satisfy E_5 with constant probability. □

Lemma 5.4. *For any pair $i \neq j = 1 \cdots n$ and any $t \in T$, the probability that (X_i, X_j) is the closest pair at time t, but not at time $t + \Delta t$, is $\Omega(\frac{1}{n^2})$.*

Proof. We sketch how the proof of Lemma 5.1 may be extended to the case when $t > 0$. In such a case the density of distribution for both coordinates of $P_i(t)$ is defined as in section 2. Clearly, $f_t(z) \leq 2$ for any $t, z \in [0,1]$, and $f_t(z) \geq 1$ for all $t \in [0,1]$ and $\frac{1}{4} \leq z \leq \frac{3}{4}$. Hence, the distribution is 'close' to uniform and the bounds for E_1 and E_2 can be proved in the same way as before. The bound for E_3 needs to be slightly modified, as we have to ensure that the measure of the region denoted U (instead of just its area) is large with high probability, under the new density. This is however easy to obtain, as from the Chernoff bound a constant fraction of points P_i for $i \geq 5$ belong to $[\frac{1}{4}, \frac{3}{4}]$, and then the area of $|U|$ is close to the probability measure of U.

In order to bound E_4 and E_5, we add the constraints that $P_1(t), P_2(t), P_k(t), P_l(t) \in [\frac{1}{4}, \frac{3}{4}]$ (which changes the respective probabilities only by a constant). Under this restriction, the distribution of $Q_1(t), Q_2(t), Q_k(t), Q_l(t)$ can be easily seen to be close to the uniform distribution some subset of $[0,1]^2$ of a constant area (with respect to the usual Lebesgue measure). Then the probability bounds can be computed as before. □

Now we proceed with the upper bound. For any pairwise unequal indices i, j, k, l from $\{1 \cdots n\}$ and $t \in T$ we define events:

- $E_6(i,j,k,l,t)$: there exists a moment $t' \in [t, t + \Delta t]$ when the closest pair changes from (X_i, X_j) to (X_k, X_l),
- $E_7(i,j,k,t)$: there exists a moment $t' \in [t, t + \Delta t]$ when the closest pair changes from (X_i, X_j) to (X_i, X_k).

Clearly $O(\sum_{i,j,k,l,t}[E_6(i,j,k,l,t) \text{ is true}] + \sum_{i,j,k,t}[E_7(i,j,k,t) \text{ is true}])$ is an upper bound for the total number of changes. It is hence sufficient to prove Lemma 5.5.

Lemma 5.5. *for any different $i, j, k, l = 1 \cdots n$ and $t \in T$,*

$$\begin{aligned} \Pr[E_6(i,j,k,l,t)] &= \frac{O(1)}{n^4} \\ \Pr[E_7(i,j,k,t)] &= \frac{O(1)}{n^4}. \end{aligned}$$

Proof. We show the first upper bound. The second one can be handled in the same way. Fix $i, j, k, l = 1, 2, 3, 4$. We can write

$$\Pr[E_6(1,2,3,4,t)] = \sum_s \Pr[E_6(1,2,3,4,t) \wedge |X_1(t)X_2(t)| \in [s\Delta t, (s+1)\Delta t]]$$

As during any period of time Δt any point can traverse a distance at most $\sqrt{2}\Delta t$, we can bound the above sum by

$$\begin{aligned} \sum_s \Pr[\quad & |X_1(t)X_2(t)| \in [s\Delta t, (s+1)\Delta t] \\ \wedge \quad & |X_3(t)X_4(t)| \leq (s+1+4\sqrt{2})\Delta t \\ \wedge \quad & |X_{i'}(t)X_{j'}| \geq (s-4\sqrt{2})\Delta t \text{ for all different } i', j' \geq 5] \end{aligned}$$

Note that for each s the triple of events described above is independent. As before, consider first the case $t = 0$. We will bound each element of the latter sum as follows. Assume WLOG that we choose first X_1, X_2, then X_3, X_4 and then $X_i, i \geq 5$. Then :

- the probability that $|X_1(t)X_2(t)| \in [s\Delta t, (s+1)\Delta t]$ is $O(s\Delta t^2) = \frac{O(s)}{n^2}$,
- the probability that $|X_3(t)X_4(t)| \leq (s+1+4\sqrt{2})\Delta t$ is $O(s^2\Delta t^2) = O(\frac{s^2}{n^2})$,
- the probability that $|X_{i'}(t)X_{j'}| \geq (s-4\sqrt{2})\Delta t$ for all different $i', j' \geq 5$ can be bounded (as in the proof of Lemma 5.1) by $(1 - cns^2\Delta t^2)^{n/2}$ for some constant c, which in turn is less than $e^{-cs^2/2}$.

Hence

$$\Pr[E_6(1,2,3,4,0)] \leq \sum_s c_1 \cdot \frac{1}{n^4} e^{-cs^2/2} = \frac{O(1)}{n^4}.$$

The case $t > 0$ can be handled as before. □

From Lemma 5.4 and Lemma 5.5, we obtain the main

Theorem 5.6. *The expected number of changes to the closest pair for n particles drawn independently at random from the uniform square distribution is $\Theta(n)$.*

6 Related results

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.

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.

Our computations can be extended in arbitrary dimensions. For instance, in the setting we considered, in d dimensions, the Voronoi diagram changes $\Theta(n^{1+1/d})$ times in expectation, while the closest pair changes $\Theta(n^{2/d})$ times.

The results we obtained for the expected number of changes to the convex hull of particles drawn from the uniform square distribution directly imply an expected linear bound for the convex hull kinetic data structure described in [BGH97]. 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. If N_b and N_r are the (random) total number of edges that ever appear on the blue and red convex hull, the number of events involved at the top level is $O(N_b N_r)$. As these two quantities are independent, the expectation is $E(N_b N_r) = E(N_b)E(N_r) = \Theta(\log^4 n)$ in the case of the square distribution. The expected running time $T(n)$ of the whole data structure thus obeys the recurrence $T(n) = O(\log^4 n) + 2T(n/2)$, which solves to $T(n) = O(n)$. It is easy to see that this is also a lower bound.

The above result sheds an interesting light on the best way to compute the total number of combinatorial changes to the convex hull of moving points in the off-line setting. Atallah [Ata85] gave a $O(n^2)$ worst case off-line method, but it is easily seen that this worst case is also the best case. Thus, even if the motion of points is fully known in advance, it might be a good idea to use the on-line algorithm of [BGH97]. The results above also make it clear that it is not a good idea to use the Delaunay triangulation to maintain the convex hull.

7 Conclusion

In this paper we initiated the study of average case behavior of some combinatorial properties defined on points moving on random trajectories. We obtained tight bounds for convex hull, closest pair and Delaunay triangulation, and for the kinetic data structure of [BGH97] to maintain the convex hull.

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, e.g. points moving along Bezier curves with anchor points selected according to some distribution. The same questions can also be asked for other combinatorial functions, like the diameter, the width, or the minimum spanning tree, and for the recent kinetic algorithms that have been proposed for them [AGHV, BGZ]

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. The first author would like to acknowledge the Okawa Foundation for its generous support.

References

[AGHV] P.K. Agarwal, L. J. Guibas, J. Hershberger, and E. Veach. The extent of moving points. Submitted to CG97.

[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 *Symposium on Discrete algorithms*, page xxx, 1997.

[BGSZ] J. Basch, L. J. Guibas, C. Silverstein, and L. Zhang. A practical evaluation of kinetic data structures. Submitted to CG 97.

[BGZ] J. Basch, L. J. Guibas, and L. Zhang. Proximity problems for moving points. Submitted to CG97.

[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.

[DG93] O. Devillers and M. Golin. Dog bites postman: Point location in the moving Voronoi diagram and related problems. In *Proc. 1st Annu. European Sympos. Algorithms (ESA '93)*, volume 726 of *Lecture Notes in Computer Science*, pages 133–144. Springer-Verlag, 1993.

[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.

[FL91] J.-J. Fu and R. C. T. Lee. Voronoi diagrams of moving points in the plane. *Internat. J. Comput. Geom. Appl.*, 1(1):23–32, 1991.

[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.

[ISI89] K. Imai, S. Sumino, and H. Imai. Minimax geometric fitting of two corresponding sets of points. In *Proc. 5th Annu. ACM Sympos. Comput. Geom.*, pages 266–275, 1989.

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

[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.

[Meg86] N. Megiddo. Dynamic location problems. *Oper. Res.*, 6:313–319, 1986.

[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.

[Vor08] G. M. Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième Mémoire: Recherches sur les parallélloèdres primitifs. *J. Reine Angew. Math.*, 134:198–287, 1908.