Similar Simplices in a d-Dimensional Point Set*

Pankaj K. Agarwal
Department of Computer Science
Duke University

Roel Apfelbaum
School of Computer Science
Tel Aviv University

George Purdy
Department of Electrical and Computer Engineering
University of Cincinnati

Micha Sharir
School of Computer Science
Tel Aviv University
and
Courant Institute
New York University

ABSTRACT

We consider the problem of bounding the maximum possible number $f_{k,d}(n)$ of k-simplices that are spanned by a set of n points in $\mathbb{R}^d$ and are similar to a given simplex. We first show that $f_{2,3}(n) = O(n^{13/6})$, and then tackle the general case, and show that $f_{d-2,d}(n) = O(n^{d-8/5})$ and[1] $f_{d-1,d}(n) = O^*(n^{d-72/55})$, for any d. Our technique extends to derive bounds for other values of k and d, and we illustrate this by showing that $f_{2,5}(n) = O(n^{8/3})$.

Categories and Subject Descriptors

G.2.1 [**Mathematics of Computing**]: Discrete Mathematics—*Combinatorics*

General Terms

Theory

Keywords

Combinatorial geometry, Incidences, Point configurations, Repeated patterns, Points and Hyperplanes, Points and Hyperplanes

1. INTRODUCTION

Let P be a set of n points in $\mathbb{R}^d$, and let Δ be a prescribed k-dimensional simplex (k-simplex, for short), for some $2 \leq k \leq d-1$. Let $f(P, \Delta)$ denote the number of k-simplices spanned by P that are similar to Δ. Set

$$f_{k,d}(n) = \max f(P, \Delta),$$

where the maximum is taken over all sets P of n points in $\mathbb{R}^d$ and over all k-simplices Δ in $\mathbb{R}^d$. We wish to obtain sharp bounds on $f_{k,d}(n)$. It suffices to consider cases with $2 \leq k \leq d-1$, since, trivially, $f_{0,d}(n) = n$, $f_{1,d}(n) = \binom{n}{2}$, and $f_{d,d}(n) \leq 2f_{d-1,d}(n)$. (It is still conceivable that $f_{d,d}(n)$ is much smaller than $f_{d-1,d}(n)$; see below.)

The problem of obtaining sharp bounds on $f_{k,d}(n)$ is motivated by *exact pattern matching*: We are given a set P of n points in $\mathbb{R}^d$ and a "pattern set" Q of $m \leq n$ points (in most applications m is much smaller than n; let us assume $m \geq d+1$), and we wish to determine whether P contains a similar copy of Q, under some allowed class of transformations, or, alternatively, to enumerate all such copies, See [13] for a comprehensive review of this and related problems. A commonly used approach to this problem is to take a d-simplex Δ spanned by some points of Q, and find all congruent copies of Δ that are spanned by points of P. For each such copy Δ', take the similarity transformation(s) that map Δ to Δ', and check whether all the other points of Q map to points of P under that transformation. The efficiency of such an algorithm depends on the number of similar copies of Δ in P. Using this approach for congruences, de Rezende and Lee [22] developed an $O(mn^d)$-time algorithm to determine whether P contains a congruent copy of Q. For $d = 3$, Brass [10] developed an $O(mn^{7/4}\beta(n)\log n + n^{11/7+\varepsilon})$-time algorithm, which improves an earlier result by Boxer [9]. See also [8, 12] for related work. To recap, for applications of this kind, the main quantity of interest is $f_{d-1,d}(n)$.

In their recent monograph [13, pp. 265–266], Brass et al. review the known bounds on $f_{k,d}(n)$ and state various conjectures and open problems (see also [11]). In the plane, $f_{2,2}(n) = \Theta(n^2)$, where the lower bound can be obtained, e.g., from a section of the triangular grid (see [2] for an improved constant of proportionality). For more involved patterns, sufficient and necessary conditions for a tight quadratic bound are given in [18]; see also [15].

There are practically no known upper bounds in $d \geq 3$ dimensions, with the sole exception of a bound of $O(n^{2.2})$ on $f_{2,3}(n)$

*Work by Pankaj Agarwal and Micha Sharir was supported by a grant from the U.S.-Israel Binational Science Foundation. Work by Pankaj Agarwal was also supported by by NSF under grants CCR-00-86013, EIA-01-31905, and CCR-02-04118 and an ARO grant W911NF-04-1-0278. Work by Micha Sharir was also supported by NSF Grant CCF-05-14079, by Grant 155/05 from the Israel Science Fund, and by the Hermann Minkowski–MINERVA Center for Geometry at Tel Aviv University.

[1]The notation $O^*(\cdot)$ hides polylogarithmic factors.

SCG'07, June 6–8, 2007, Gyeongju, South Korea.

(and $f_{3,3}(n)$), given by Akutsu et al. [5]. Brass [11] conjectures, though, that $f_{3,3}(n) = o(n^2)$; the best known lower bound is $\Omega(n^{4/3})$.

The case of *congruent* simplices has also been studied; see Agarwal and Sharir [4] and references therein. Denote by $g_{k,d}(n)$ the maximum number of k-simplices that are spanned by a set of n points in $\mathbb{R}^d$ and are congruent to a given simplex Δ. Agarwal and Sharir have shown that $g_{2,3}(n) = O^*(n^{5/3})$ (and $\Omega(n^{4/3})$), $g_{2,4}(n) = O^*(n^2)$ (and $\Omega(n^2)$), $g_{2,5}(n) = \Theta(n^{7/3})$, and $g_{3,4}(n) = O^*(n^{9/4})$. A simple construction, attributed to Lenz [19], shows that $g_{k,d}(n) = \Omega(n^{d/2})$, for even values of d and for sufficiently large k. Erdős and Purdy [17] conjectured that this construction is asymptotically best possible, namely, that $g_{k,d}(n) = O(n^{d/2})$. Agarwal and Sharir also derive a recurrence for $g_{k,d}(n)$, for general values of k and d. The solution of this recurrence is $O(n^{\zeta(d,k)+\varepsilon})$, where $\zeta(d,k)$ is a rather complicated function of d and k. They show that $\zeta(d,k) \le d/2$ for $d \le 7$ and $k \le d-2$, and conjecture that $\zeta(d,k) \le d/2$ for all d and $k \le d-2$, in accordance with the Erdős-Purdy conjecture just mentioned.

Returning to the case of similar simplices, we note that the only known lower bounds for $f_{k,d}(n)$ are the same bounds for $g_{d,d}(n)$, namely $\Omega(n^{d/2})$ for $d \ge 4$ even, and $\Omega(n^{d/2-1/6})$, for $d \ge 3$ odd; see also [1, 2, 4].

Our results. We first obtain the bound $O(n^{13/6})$ for $f_{2,3}(n)$, improving upon the bound of Akutsu et al. [5]. Recall that Brass [11] conjectures that $f_{3,3}(n) = o(n^2)$ (see also [13, p. 265]), and that, in contrast, $f_{2,3}(n) = \Omega(n^2)$ (in fact, $f_{2,2}(n)$ is already $\Theta(n^2)$).

We then tackle the general case, obtaining the first nontrivial bounds for $f_{d-2,d}(n)$ and $f_{d-1,d}(n)$ (and thus also for $f_{d,d}(n)$); the trivial bounds are, respectively, $f_{d-2,d}(n) = O(n^{d-1})$ and $f_{d-1,d}(n) = O(n^d)$). Specifically, we show:

$$f_{d-2,d}(n) = O(n^{d-8/5}) \quad \text{and} \quad f_{d-1,d}(n) = O^*(n^{d-72/55}).$$

The above results imply that $f_{2,4}(n) = O(n^{12/5})$. Finally, we prove that $f_{2,5}(n) = O(n^{8/3})$. Note that this is the last interesting case for triangles because, by Lenz' construction, $f_{2,6}(n) = g_{2,6}(n) = \Omega(n^3)$.

Needless to say, none of these bounds is known (nor conjectured) to be tight. Our techniques are strongly based on bounds on the number of incidences between points and spheres (or circles). Considerable progress has been made on these problems in recent years; see [3, 6, 7, 20] and also [21] for a comprehensive survey of these and related results. We also use the recent bound of Elekes and Tóth [16] on the number of so-called *rich non-degenerate* hyperplanes. In this regard, our work can be regarded as an application of the recent developments in incidence problems, which raises several interesting basic open problems in this area, as discussed at the end of the paper.

We regard the present paper as an initial stab at the problem. As the reader will realize, there are ample opportunities for improvements of the bounds. In fact, we are presently exploring these possible improvements, and hope to include them in the full version of the paper.

2. SIMILAR TRIANGLES IN THREE DIMENSIONS

In this section we prove an improved bound on $f_{2,3}(n)$. Let Δ_0 be a fixed triangle, and let $\mathcal{S}(\Delta_0)$ denote the set of all triangles spanned by P and similar to Δ_0. As a warm-up exercise, we first derive a simple, albeit weaker, upper bound on $f_{2,3}(n)$, and then prove a tighter bound whose proof is considerably more involved.

2.1 A simpler and weaker bound

For each pair a, b of points of P, any triangle abc in $\mathcal{S}(\Delta_0)$, with $c \in P \setminus \{a, b\}$, has the property that c lies on a circle $\gamma_{a,b}$, which is orthogonal to ab and whose center lies at a fixed point on ab. Moreover, given a circle γ, there exist at most two (unordered) pairs a, b, such that $\gamma = \gamma_{a,b}$ (if there are two pairs, one is the reflection of the other through the center of γ). Hence, ignoring the multiplicity factor 2, $|\mathcal{S}(\Delta_0)|$ is at most the number of incidences between the points of P and the (at most) $\binom{n}{2}$ distinct circles $\gamma_{a,b}$.

As shown by Aronov et al. [6] (see also Agarwal et al. [3], and Marcus and Tardos [20]), the number of incidences between n points and c distinct circles in $\mathbb{R}^3$ (or, for that matter, in any dimension) is

$$O(n^{2/3}c^{2/3} + n^{6/11}c^{9/11}\log^{2/11}(n^3/c) + n + c). \qquad (1)$$

Substituting $c = O(n^2)$, the second term dominates, and we obtain

$$f_{2,3}(n) = O(n^{24/11}\log^{2/11} n) = O^*(n^{24/11}) = O(n^{2.182}).$$

We remark that a similar approach was taken by Akutsu et al. [5], except that they used a weaker bound on point-circle incidences (albeit the best known at that time). Finally, taking Δ_0 to be an equilateral triangle and P to be a section of a 2-dimensional triangular lattice, it is easy to verify that $f_{2,2}(n) = \Theta(n^2)$, which implies that $f_{2,d}(n) = \Omega(n^2)$ for any $2 \le d \le 5$ (as already noted, $f_{2,d}(n) = \Theta(n^3)$ for $d \ge 6$).

Remark: A consequence of (1) is the following useful variant, which will be exploited throughout the paper: The number of circles that contain at least t points of P is $O^*(n^3/t^{11/2}) + O(n^2/t^3 + n/t)$, and the number of incidences between the points of P and these circles is $O^*(n^3/t^{9/2}) + O(n^2/t^2 + n)$.

2.2 An improved bound

To simplify the presentation, let us assume[2] that Δ_0 is not isosceles, so its edges have distinct lengths. We denote by $\mathcal{S}(\Delta_0)$ the set of all triangles similar to Δ_0 and spanned by P. For each such triangle Δabc, we can order its vertices in a unique order, say a, b, c, so that a is incident to the two longest edges of the triangle, and b is the other endpoint of the longest edge. We call a the *main vertex* of the triangle. For any pair of points a, b, denote the sphere centered at a and containing b by $\sigma_{a,b}$. Let $\gamma = \gamma_{a,b}$ denote the circle that is the locus of all the points c' such that $\Delta abc' \sim \Delta_0$ (γ is the same circle constructed in the preceding proof). Clearly, γ is contained in a sphere centered at a (which is $\sigma_{a,c}$, for any $c \in \gamma$).

For $a \in P$, let Σ_a denote the set of all spheres $\sigma_{a,b}$, for $b \in P \setminus \{a\}$. Define a relation Π_a on $\Sigma_a \times \Sigma_a$, which contains all pairs $(\sigma_{a,b}, \sigma_{a,c})$ of spheres for which $\Delta abc \in \mathcal{S}(\Delta_0)$. We denote by Γ_a the set of all circles $\gamma_{a,b}$ (note that the pair (a, b) uniquely determines $\gamma_{a,b}$). Put $\Sigma = \bigcup_{a\in P} \Sigma_a$, $\Gamma = \bigcup_{a\in P} \Gamma_a$, and $\Pi = \bigcup_{a\in P} \Pi_a$. By construction, each sphere in Σ appears in at most two pairs of Π.

For an integer $i \le n$, let $\Pi_{\le i} \subseteq \Pi$ denote the set of pairs of spheres (σ, τ) such that either σ or τ contains at most i points of P, and let $\Pi_{>i} = \Pi \setminus \Pi_{\le i}$ denote those pairs in which each of the spheres contains more than i points. Define

$$\Sigma_{\le i} = \{\sigma \in \Sigma \mid (\sigma,\ \Pi(\sigma)) \in \Pi_{\le i}\}$$

and $\Sigma_{>i} = \Sigma \setminus \Sigma_{\le i}$.

The discussion above implies that each triangle $\Delta abc \in \mathcal{S}(\Delta_0)$ corresponds to an incidence between the point c and the circle $\gamma_{a,b}$.

[2]The proof can be carried out without this assumption, but the presentation is simplified.

Hence, as already noted,

$$\frac{1}{2}|\mathbb{S}(\Delta_0)| \le I(P,\Gamma) \le |\mathbb{S}(\Delta_0)|,$$

where $I(P,\Gamma)$ is the number of incidences between the points of P and the circles of Γ. Thus, as above, it suffices to bound $I(P,\Gamma)$.

Fix a threshold parameter t. We call a pair in $\Pi_{\le t}$ *light* and a pair in $\Pi_{>t}$ *heavy*. We classify the heavy pairs into non-degenerate and degenerate pairs, as follows: A heavy sphere σ (i.e., a sphere containing more than t points) is *degenerate* if there exists a circle $\gamma \subset \sigma$ (not necessarily from the family Γ) such that $|\gamma \cap P| \ge \beta|\sigma \cap P|$ for some sufficiently small constant $\beta > 0$; otherwise it is *non-degenerate*. A pair in $\Pi_{>t}$ is called *non-degenerate* if both the spheres in the pair are non-degenerate, and *degenerate* otherwise.

We bound separately the number of incidences between the points of P and the circles determined by each of these three types of pairs. Let I_L (resp., I_N, I_D) denote the number of incidences between P and the circles induced by light (resp., non-degenerate, degenerate) pairs. Then $I(P,\Gamma) = I_L + I_N + I_D$.

Handling light pairs. Let $a \in P$ and put

$$\Pi'_a = \Pi_a \cap \Pi_{\le t}.$$

For a sphere pair $(\sigma,\tau) \in \Pi'_a$, put

$$P_\sigma = P \cap \sigma, \qquad P_\tau = P \cap \tau, \quad \text{and} \quad \Gamma_\tau = \{\gamma_{a,b} \mid b \in P_\sigma\}.$$

Recall that either $|\Gamma_\tau| = |P_\sigma| \le t$, or $|P_\tau| \le t$. All the circles of Γ_τ lie on τ and have the same radius. Hence, as follows, e.g., from [23], the number of similar triangles associated with the pair (σ,τ) is bounded by

$$I(P_\tau,\Gamma_\tau) \;=\; O\big((|P_\sigma||P_\tau|)^{2/3} + |P_\sigma| + |P_\tau|\big).$$

Summing over all $(\sigma,\tau) \in \Pi'_a$, we get

$$\begin{aligned} I'_a &= \sum_{(\sigma,\tau)\in\Pi'_a} I(P_\tau,\Gamma_\tau) \\ &= \sum_{(\sigma,\tau)\in\Pi'_a} O((|P_\sigma||P_\tau|)^{2/3} + |P_\sigma| + |P_\tau|). \end{aligned}$$

The last two terms clearly sum up to $O(n)$. As for the first terms, one of $|P_\sigma|$, $|P_\tau|$ is at most t. We may assume, without loss of generality, that $|P_\tau| \le t$, and then obtain

$$\begin{aligned} \sum_{(\sigma,\tau)\in\Pi'_a} (|P_\sigma||P_\tau|)^{2/3} &\le t^{1/3}\sum |P_\sigma|^{2/3}|P_\tau|^{1/3} \\ &\le t^{1/3}\Big(\sum |P_\sigma|\Big)^{2/3}\Big(\sum |P_\tau|\Big)^{1/3} \\ &= O(nt^{1/3}). \end{aligned}$$

Altogether, summing over all points $a \in P$, the number of point-circle incidences with circles generated by light pairs is

$$I_L = \sum_{a\in P} I'_a = O(n^2 t^{1/3}). \tag{2}$$

Handling non-degenerate heavy pairs. Let $\Pi_N \subseteq \Pi_{>t}$ denote the set of non-degenerate heavy pairs. We bound the number of incidences between the points of P and the circles $\gamma_{a,b}$, for which there exists a pair of spheres $(\sigma,\tau) \in \Pi_N \cap \Pi_a$, such that $b \in \sigma$ and $\gamma_{a,b} \subset \tau$. Note that both spheres in each pair of Π_N contain more than t points of P. We first obtain an upper bound on $|\Pi_N|$.

LEMMA 2.1. $|\Pi_N| = O\left(\frac{n^4}{t^5} + \frac{n^3}{t^3}\right)$.

PROOF. Let Σ_N denote the set of all spheres that apper as the first component of a pair in Π_N. We apply to P and Σ_N the standard lifting transform to $\mathbb{R}^4$ [14], so that each point $p = (x,y,z) \in P$ is mapped to the point $p^+ = (x,y,z,x^2+y^2+z^2)$ in $\mathbb{R}^4$, and each sphere of Σ_N, centered at (x_0,y_0,z_0) and having radius r, is mapped to the hyperplane $w = 2x_0x + 2y_0y + 2z_0z + \big(r^2 - x_0^2 - y_0^2 - z_0^2\big)$. Let $P^+ = \{p^+ \mid p \in P\}$. The lifting preserves the incidence relation, and has the additional property that cocircular points of P (on a circle that lies on one of the lifted spheres σ) are lifted to coplanar points in $\mathbb{R}^4$ lying on a 2-flat in the hyperplane image of σ. Thus, a non-degenerate sphere is lifted to what Elekes and Tóth call a *β-degenerate* hyperplane [16] with respect to P^+. We can therefore apply the result of [16], which asserts that the number of β-degenerate hyperplanes that contain at least t points of P^+ (so-called *rich* hyperplanes), and hence, the number of spheres in Σ_N (and the number of pairs in Π_N), is bounded by $O(n^4/t^5 + n^3/t^3)$, provided that the constant β is chosen sufficiently small. This completes the proof of the lemma. □

We next partition Π_N into $O(\log n)$ classes, $\Pi_N^{(1)}, \Pi_N^{(2)}, \ldots$, where $\Pi_N^{(i)}$ consists of those pairs $(\sigma,\tau) \in \Pi_N$ such that

$$2^{i-1}t < \max\big\{|\sigma \cap P|, |\tau \cap P|\big\} \le 2^i t,$$

for $i = 1,\ldots,\log_2(n/t)$. We sum the incidence bounds within each class separately. Fix a class $\Pi_N^{(i)}$, and put $t_i = 2^i t$. For each pair $(\sigma,\tau) \in \Pi_N^{(i)}$, σ induces on τ a set of at most t_i congruent circles, and the number of points on τ is also at most t_i. Hence the number of incidences between these points and circles is $O(t_i^{4/3})$. By Lemma 2.1, the number of such pairs of spheres is $O(n^4/t_i^5 + n^3/t_i^3)$. Hence, summing over i, the overall number of incidences involving spheres in Σ_N is

$$I_N = O\left(\sum_i t_i^{4/3}\cdot\left(\frac{n^4}{t_i^5} + \frac{n^3}{t_i^3}\right)\right) = O\left(\frac{n^4}{t^{11/3}} + \frac{n^3}{t^{5/3}}\right). \tag{3}$$

Handling degenerate heavy pairs. Let $\Pi_D = \Pi_{>t} \setminus \Pi_N$ be the set of degenerate pairs. We apply the following pruning process on each pair $(\sigma,\tau) \in \Pi_D$. Suppose τ is a degenerate sphere containing more than t points of P (the case where σ is the degenerate sphere is handled in an essentially symmetric manner). Then there is a circle $\gamma_1 \subset \tau$ containing at least $\beta|P_\tau|$ points of $P_\tau = P \cap \tau$. If we remove the points of $P \cap \gamma_1$, then one of the following cases arises:

1. τ is incident to at most t of the remaining points of $P \setminus \gamma_1$.
2. τ becomes non-degenerate with respect to the remaining points.
3. τ is still degenerate and contains more than t points.

In the third case, we continue with the pruning process, until one of the first two events occurs, which will happen after at most $\log_{1/\beta} n$ iterations. At the end of the process, if τ contains t or fewer points, we include it, together with its remaining incidences and with the sphere σ, as one of the pairs of $\Pi_{\le t}$, meaning that we only consider the remaining points when bounding the number of point-circle incidences on τ. Otherwise, τ is non-degenerate with respect to the set of at least t remaining points. In other words, when considering only the surviving points on τ (or the surviving circles, if σ was the heavy sphere), at the end of the pruning process, (σ,τ) becomes a light or non-degenerate pair. Lemma 2.1 still holds for these latter pairs, because it relies on the bound of [16], and this bound still applies, as is easily verified.

We still have to count the incidences involving the removed points and/or circles on each of these spheres τ. Such a (yet uncounted) incidence $p \in \gamma$, on a degenerate sphere τ, is uncounted either because

(a) γ was one of the circles whose points were removed from P_τ, or

(b) γ was not such a circle, but p was removed because it lies on another circle $\gamma' \subset \tau$ that has been removed.

For an incidence of type (a), we have $\gamma \in \Gamma_\tau$. There can be at most two spheres τ for which this situation arises (with the same γ). Thus, the number of uncounted incidences of type (a) is at most twice the number of incidences between the points of P and the removed circles. Using the fact that each of these circles contains at least βt points, the results of [3, 7, 20] imply, as noted above, that the number of these incidences is

$$O^*(n^3/t^{9/2}) + O(n^2/t^2 + n).$$

To bound the number of incidences of type (b), we observe that, for each pair $(\sigma, \tau) \in \Pi_D$, the number of removed circles on τ is $O(\log n)$, and each $\gamma \in \Gamma_\tau$ can lose at most two incidences for each removed circle. Thus, the number of such incidences on τ is $O(|\Gamma_\tau| \log n)$. Summing over all degenerate spheres, the overall number of type (b) incidences is $O(n^2 \log n)$.

A similar analysis applies to the case where the degenerate sphere in the pair (σ, τ) is σ rather than τ. Here each pruning step removes circles from Γ_τ, whose centers lie on a common circle. It is easily checked that, for any point $p \in \tau$, at most two such circles can pass through p, so each pruning step loses at most $2|P_\tau|$ incidences, for a total of $O(|P_\tau| \log n)$ incidences. Summing over all spheres τ, we obtain the same bound $O(n^2 \log n)$ as above. Hence,

$$\begin{aligned} I_D &= O^*\left(\frac{n^3}{t^{9/2}}\right) + O\left(\frac{n^2}{t^2} + n\right) + O(n^2 \log n) \\ &= O^*\left(\frac{n^3}{t^{9/2}}\right) + O(n^2 \log n). \qquad (4) \end{aligned}$$

Adding (2)–(4), we get

$$\begin{aligned} I(P,\Gamma) &\le I_L + I_N + I_D \\ &= O\left(n^2 t^{1/3} + \frac{n^4}{t^{11/3}} + \frac{n^3}{t^{5/3}} + n^2 \log n\right), \end{aligned}$$

for any t. By choosing $t = n^{1/2}$, we obtain the main result of this section:

THEOREM 2.2. *Let P be a set of n points in $\mathbb{R}^3$, and let Δ_0 be some fixed triangle. Then the number of triangles similar to Δ_0 spanned by P is $O(n^{13/6}) = O(n^{2.167})$.*

3. BOUNDS IN HIGHER DIMENSIONS

We next consider the general case of bounding the maximum number $f_{d-1,d}(n)$ of mutually similar $(d-1)$-simplices in a set of n points in $\mathbb{R}^d$. Since this bound depends on $f_{d-2,d}(n)$, we begin with bounding this quantity.

A bound on $f_{d-2,d}(n)$. Let Δ_0 be a fixed $(d-2)$-simplex, let Δ_1 be some fixed ($(d-3)$-dimensional) facet of Δ_0, and let z be the vertex of Δ_0 not incident to Δ_1. Let $\tau = (p_1, \ldots, p_{d-2})$ be a $(d-2)$-tuple of distinct points in P that span a simplex similar to Δ_1. The number of such tuples is at most $f_{d-3,d}(n)$, which we bound naively by $O(n^{d-2})$. (This is an overestimate, which we make to simplify the analysis. See Section 4 where a better bound is derived for $d = 5$.)

Any point $q \in P$ that forms with τ a $(d-2)$-simplex similar to Δ_0 must lie on a 2-sphere $\sigma = \sigma_\tau$ orthogonal to the $(d-3)$-flat spanned by the points of τ, and centered at a fixd point on this flat. Moreover, let σ be a given 2-sphere. Any tuple τ for which $\sigma = \sigma_\tau$ must be such that

(i) the $(d-3)$-simplex s spanned by τ is determined up to *congruence*;

(ii) s is contained in the $(d-3)$-flat orthogonal to σ and passing through its center o; and

(iii) o coincides with a fixed point rigidly attached to σ (the point corresponding to the foot of the perpendicular to Δ_1 from z in Δ_0).

In other words, given σ, its *multiplicity* (the number of times it arises as σ_τ for an appropriate tuple τ) can be estimated as follows. We have a set Q of $m \le n$ points in $\mathbb{R}^{d-3}$, and we want to bound the number of congruent full-dimensional simplices that are spanned by Q, where all these simplices are obtained from one another by rotation about (say) the origin o. An easy bound on this quantity is $O(n^{d-4})$. Indeed, if we fix $d-4$ of the points that span the simplex, then, together with o, they span a $(d-4)$-simplex s' (in $\mathbb{R}^{d-3}$), and each of the remaining two vertices can then be placed in at most two locations, depending on the orientation of the full $(d-3)$-simplex with respect to s', from which the bound follows.

To recap, we have a set Σ of $O(n^{d-2})$ 2-spheres, each occurring with multiplicity at most $O(n^{d-4})$ (note though that $O(n^{d-2})$ bounds the total number of spheres, counted *with multiplicity*), and we need to bound the number of incidences between Σ and P. We follow an approach similar to Section 2.2. We fix a parameter $t > 0$ and a sufficiently small constant $\beta > 0$ (that depends on d). As in Section 2.2, we call a 2-sphere *light* (resp., *heavy*) if it contains at most (resp., more than) t points of P. A heavy 2-sphere σ is *degenerate* if more than a β-fraction of the points of $P \cap \sigma$ are co-circular, and *non-degenerate* otherwise.

The number of incidences with the light 2-spheres is, trivially, at most $O(n^{d-2}t)$. We therefore focus on the heavy 2-spheres.

Incidences on non-degenerate heavy 2-spheres. Let μ be the number of *distinct* non-degenerate 2-spheres in Σ. We can bound μ by lifting the 2-spheres to $\mathbb{R}^{d+1}$, as in the proof of Theorem 2.2. We obtain a collection of μ 3-flats, each containing at least t points of the lifted set P^+, and none of them contains a 2-flat with more than a β-fraction of its points (for the same reason as in the preceding proof). Project the lifted collection of points and 3-flats onto some generic 4-space E, and apply the Elekes-Tóth bound [16], to conclude that the number μ of the projected 3-flats (now hyperplanes) is $\mu = O(n^4/t^5 + n^3/t^3)$. Moreover, the number of incidences between the original 2-spheres and the points of P is at most $O(n^4/t^4 + n^3/t^2)$. Multiplying this bound by the maximum multiplicity of a 2-sphere, which is $O(n^{d-4})$, we obtain the bound

$$O(n^d/t^4 + n^{d-1}/t^2).$$

Incidences on degenerate heavy 2-spheres. We can replace each degenerate heavy 2-sphere $\sigma \in \Sigma$ by a circle γ that contains more than a β-fraction of the points on σ, so it contains at least βt points. It then suffices to bound the number of incidences between these circles, *counted with multiplicity*, and the points of P, as $I(P, \{\gamma\}) \ge \beta I(P, \{\sigma\})$.

Let us first fix such a circle γ, and bound the multiplicity of γ, namely the number of $(d-2)$-tuples τ such that σ_τ contains γ. Any such tuple τ, together with the center o of γ, spans the $(d-2)$-dimensional flat h orthogonal to γ and passing through o.

The number of such tuples is at most $O(n^{d-3})$. Indeed, if we fix a sub-tuple τ' of $d-3$ points of τ, then the size of the simplex is determined (because the distance from any of the fixed points to any point on γ is now fixed). Let o' denote the center of σ_τ. Then o' has to lie on a fixed $(d-3)$-sphere in h centered at o. Moreover, since o' is rigidly attached to the simplex, it follows that it must lie on a circle orthogonal to the $(d-4)$-flat spanned by τ', and this circle intersects the sphere in at most two points,[3] each of which corresponds to a unique τ.

Let C_t (resp., $C_{\geq t}$) denote the number of circles that are spanned by P and contain exactly (resp., at least) t points of P. As remarked above, it follows from the analysis of [6] that

$$C_{\geq \beta t} = O^*\left(\frac{n^3}{t^{11/2}} + \frac{n^2}{t^3} + \frac{n}{t}\right),$$

and that the number of incidences with these circles is

$$O^*(n^3/t^{9/2} + n^2/t^2 + n).$$

The preceding argument implies that the number of $(d-2)$-simplices that correspond to heavy degenerate 2-spheres is at most

$$O(n^{d-3}) \cdot \left[O^*\left(\frac{n^3}{t^{9/2}}\right) + O\left(\frac{n^2}{t^2} + n\right)\right] =$$

$$O^*\left(\frac{n^d}{t^{9/2}}\right) + O\left(\frac{n^{d-1}}{t^2} + n^{d-2}\right).$$

Hence the overall number of simplices is $O(n^{d-2}t + n^d/t^4 + n^{d-1}/t^2)$. Choosing $t = n^{2/5}$, we obtain $I(P, \Sigma) = O(n^{d-8/5})$, which also bounds the number of simplices similar to Δ_0 spanned by the points of P. Hence, we obtain the following.

THEOREM 3.1. $f_{d-2,d}(n) = O(n^{d-8/5})$.

A bound on $f_{d-1,d}(n)$. We can now obtain the main result of this section (and of the paper). Let Δ_0 be a fixed $(d-1)$-simplex, let Δ_1 be some fixed ($(d-2)$-dimensional) facet of Δ_0, and let z be the vertex of Δ_0 not incident to Δ_1.

Let $\tau = (p_1, \ldots, p_{d-1})$ be a $(d-1)$-tuple of distinct points in P that span a simplex similar to Δ_1. The number of such tuples is $\mu \leq f_{d-2,d}(n) = O(n^{d-8/5})$.

Any point $q \in P$ that forms with τ a $(d-1)$-simplex similar to Δ_0 must lie of a circle $\gamma = \gamma_\tau$ orthogonal to the $(d-2)$-flat spanned by the points of τ, and centered at a fixed point on that flat. Moreover, let γ be a given circle. In analogy with the preceding analysis, any tuple τ for which $\gamma = \gamma_\tau$ must be such that

(i) the $(d-2)$-simplex s spanned by τ is determined up to *congruence*;

(ii) s is contained in the $(d-2)$-flat orthogonal to γ and passing through its center o; and

(iii) o is rigidly attached to s (it is the point corresponding to the foot of the perpendicular to Δ_1 from z in Δ_0), and all possible simplices s are obtained from one another by rotation about o.

In other words, given γ, to estimate its *multiplicity*, we need to bound the number of congruent full-dimensional simplices that are spanned by a set Q of $m \leq n$ points in $\mathbb{R}^{d-2}$, where the simplex is unique up to rotation about (say) the origin o. Arguing as above, this number is at most $O(m^{d-3})$, because any tuple of $d-3$ points of Q, together with o, determines a $(d-3)$-simplex (in $\mathbb{R}^{d-2}$), which can be completed into a full-dimensional one in only two ways.

We thus proceed as before, fixing a threshold parameter t, and distinguishing between *light* circles (those with at most t points), and the remaining *heavy* circles. The number of incidences with light circles is at most μt, and the number of incidences with heavy circles is

$$O(n^{d-3}) \cdot O^*\left(\frac{n^3}{t^{9/2}} + \frac{n^2}{t^2} + n\right) = O^*\left(\frac{n^d}{t^{9/2}} + \frac{n^{d-1}}{t^2} + n^{d-2}\right).$$

By choosing $t = (n^d/\mu)^{2/11}$, we obtaining the main result:

THEOREM 3.2. $f_{d-1,d}(n) = O^*(n^{d-72/55})$.

Remark: We have actually shown that

$$f_{d-1,d}(n) = O^*\left(n^{2d/11}\right) \cdot (f_{d-2,d}(n))^{9/11},$$

which suggests an inductive approach, where each $f_{k,d}(n)$ is bounded in terms of $f_{k-1,d}(n)$, by such an "exponential averaging" expression. We have abandoned this induction already for $k = d-2$ (although an expression of this sort is implicit in the preceding analysis of this case).

4. SIMILAR TRIANGLES IN FIVE DIMENSIONS

Let P be a set of n points in $\mathbb{R}^5$, and let Δ_0 be a triangle. For each pair of points $a, b \in P$, let $\sigma_{a,b}$ denote the 3-sphere orthogonal to ab and containing all the points c for which $abc \sim \Delta_0$. Let Σ be the set of resulting 3-spheres. As above, it is easily checked that no 3-sphere can arise in this way more than twice. Ignoring this constant multiplicity, we face the problem of bounding the number of incidences between a set Σ of $O(n^2)$ 3-spheres and P.

As in earlier sections, we fix a parameter t and a sufficiently small constant $\beta > 0$. We define a 3-sphere to be light, heavy, non-degenerate, or degenerate as in the earlier sections. The number of incidences involving light 3-spheres is $O(n^2 t)$, so we concentrate on the heavy 3-spheres.

Consider first the *non-degenerate* heavy 3-spheres of Σ. In this case, by lifting the points and 3-spheres into $\mathbb{R}^6$ (so the 3-spheres become 4-flats), and then projecting them onto some generic 5-space, we can apply, as before, the Elekes-Tóth bound [16], to conclude that the number of such 3-spheres is $O(n^5/t^6 + n^4/t^4)$, and that the number of incidences between these 3-spheres and the points of P is $O(n^5/t^5 + n^4/t^3)$.

Consider next the heavy 3-spheres in Σ that are degenerate, so each of them contains a 2-sphere that contains more than a β-fraction of the points on the 3-sphere. We replace the 3-spheres by the respective 2-spheres, bound the number of incidences with these 2-spheres, counted with the appropriate multiplicity, and lose only a constant factor (about $1/\beta$) using this bound. Consider first the case where the 2-spheres themselves are non-degenerate, in the sense that none of them contains a circle that contains more than a β-fraction of the points on the 2-sphere. Since these 2-spheres are (βt)-heavy, the number of distinct such 2-spheres is, as above, $O(n^4/t^5 + n^3/t^3)$, and the number of incidences with them is $O(n^4/t^4 + n^3/t^2)$.

Here however the 2-spheres may appear with multiplicity, but we claim that the maximum multiplicity of a 2-sphere is $O(n)$. Indeed, given a 2-sphere σ' with center o', if we fix one point a in

[3]It is easily checked that one can always choose τ' so that the circle is not *contained* in the 2-sphere; if this were impossible, τ would not have spanned a $(d-3)$-space.

the defining pair (a,b) of a 3-sphere $\sigma_{a,b}$ containing σ', the size of the triangle is determined, and the center o of $\sigma_{a,b}$ must then lie at a fixed distance from a, within the 2-plane π orthogonal to σ' and passing through its center o'. Also, o lies on another fixed circle in π centered at o'. These two circles intersect at most twice (assuming $a \neq o'$, which can always be guaranteed), so at most two points b can form with a a pair (a,b) for which $\sigma_{a,b} \supset \sigma'$. Hence, the number of triangles similar to Δ_0 that fall into this subcase is $O(n^5/t^4 + n^4/t^2)$.

Finally, consider the subcase where the 2-spheres themselves are degenerate, so we replace each of them by a respective $(\beta^2 t)$-heavy circle, and bound the number of incidences between these circles and the points of P. The multiplicity of a circle is, trivially, at most $O(n^2)$. Hence, arguing as above, the number of triangles that arise in this subcase is $O^*(n^5/t^{9/2} + n^4/t^2)$. Hence, the overall number of triangles is $O(n^2 t + n^5/t^4 + n^4/t^2)$. Choosing $t = n^{2/3}$, we thus obtain

THEOREM 4.1. $f_{2,5}(n) = O(n^{8/3})$.

As already noted, $d = 5$ is the last interesting case for triangles, since, already for the congruent case, $g_{2,6}(n) = \Theta(n^3)$ [4].

5. DISCUSSION

Examining the proof of the general bound in Section 3, we note that there are three sources for potential improvements. First, the proof starts with the naive estimate $f_{d-3,d}(n) = O(n^{d-2})$; one should be able to get a better, nontrivial bound. Indeed, the previous section shows that this is the case for $d = 5$. Two other possibilities for improvements are in the estimation of the multiplicities of the circles and 2-spheres (and 3-spheres for $f_{2,5}(n)$) that arise in the analysis. We look at the flat h orthogonal to the circle, 2-sphere, or 3-sphere, and make the worst case assumption that $|h \cap P| = n$. With a more careful analysis (e.g., using the Elekes-Tóth bound), we expect to be able to improve this considerably. Also, when counting multiplicities, we are probably not exploiting all the restriction on the possible positions of the congruent subsimplex. We are currently exploring these possible improvements, and expect to include them in the full version.

Another open problem is to improve the bound $O(n^{13/6})$ for triangles in $\mathbb{R}^3$. A potential source for such an improvement is the fact that, when we lift P into $\mathbb{R}^4$, the resulting set P^+ lies on the 3-dimensional paraboloid, and the hope is that the Elekes-Tóth bound could be improved for such point sets, or, more generally, for point sets in convex position.

Another observation is that we can relate $f_{k,d}(n)$ to $g_{k-2,d}(n)$ (the maximum number of $(k-2)$-simplices *congruent* to a given simplex), as follows. Let Δ_0 be a given k-simplex. Fix a pair a, b of points in P. If we use a, b as two (fixed) vertices of a k-simplex similar to Δ_0, then the size of that simplex is fixed, so the number of such simplices is at most $g_{k-2,d}(n)$, implying that

$$f_{k,d}(n) = O(n^2 g_{k-2,d}(n)). \tag{5}$$

(In fact, the bound is probably smaller, because all the possible $(k-2)$-simplices that go with a fixed edge ab are "anchored" about ab, so their number should be smaller.) Recall the Erdős-Purdy conjecture that $g_{k-2,d}(n) = O(n^{d/2})$ (for even d). If the conjecture were true then we would have $f_{k,d}(n) = O(n^{2+d/2})$, which, for large values of d, is significantly smaller than the general bounds derived in this paper.

As a matter of fact, we recall that Agarwal and Sharir [4] have shown that $g_{k,d}(n) = O(n^{d/2+\varepsilon})$, for any $\varepsilon > 0$, when $d \le 7$ and $k \le d-2$. This yields an improved bound $f_{6,7}(n) = O(n^{11/2+\varepsilon})$, for any $\varepsilon > 0$. Interestingly, this is so far the only case where the bound in (5) can be shown to be better than the previous bounds.

Another comment to observe is that the proof technique is essentially a careful analysis of incidences between points and spheres of various dimensions. While the case of circles has already been studied fairly intensively, the case of higher dimensional spheres has not received much attention. The bounds that we obtain via the Elekes-Tóth bound seem to be weak. For example, using this technique for estimating the number of k-rich circles would yield a bound of $O(n^3/k^4 + n^2/k^2)$, whereas the bound using (1) is $O(n^3/k^{11/2} + n^2/k^3)$. One would hope that similar improvements could be obtained for incidences with higher-dimensional spheres too. We propose to study this problem in more general settings, and regard the present paper as an initial step in this direction.

6. REFERENCES

[1] B. M. Ábrego, G. Elekes, and S. Frenándes-Merchant, Structural results for plane sets with many similar subsets, *Combinatorica* 24 (2004), 541–554.

[2] B. M. Ábrego and S. Frenándes-Merchant, On the maximum number of equilateral triangles, I, *Discrete Comput. Geom.* 23 (2000), 129–136.

[3] P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir and S. Smorodinsky, Lenses in arrangements of pseudocircles and their applications, *J. ACM* 51 (2004), 139–186.

[4] P. K. Agarwal and M. Sharir, On the number of congruent simplices in a point set, *Discrete Comput. Geom.* 28 (2002), 123–150.

[5] T. Akutsu, H. Tamaki and T. Tokuyama, Distribution of distances and triangles in a point set and algorithms for computing the largest common point set, *Discrete Comput. Geom.* 20 (1998), 307–331.

[6] B. Aronov, V. Koltun and M. Sharir, Incidences between points and circles in three and higher dimensions, *Discrete Comput. Geom.* 33 (2005), 185–206.

[7] B. Aronov and M. Sharir, Cutting circles into pseudo-segments and improved bounds for incidences, *Discrete Comput. Geom.* 28 (2002), 475–490.

[8] M. D. Atkinson, An optimal algorithm for geometric congruence, *J. Algorithms* 8 (1987), 159–172.

[9] L. Boxer, Point set pattern matching in 3-D, *Pattern Recog. Letts.* 17 (1996), 1293–1297.

[10] P. Brass, Exact point pattern matching and the number of congruent triangles in a three-dimensional point set, *Proc. 8th Annu. European Sympos. Algos.*, 2000, 112–119.

[11] P. Brass, Combinatorial geometry problems in pattern recognition, *Discrete Comput. Geom.* 28 (2002), 495–510.

[12] P. Brass and C. Knauer, Testing the congruence of d-dimensional point sets, *Proc. 16th Annu. Sympos. Comput. Geom.*, 2000, 310–314.

[13] P. Brass, W. Moser and J. Pach, *Research Problems in Discrete Geometry*, Springer Verlag, New York, 2005.

[14] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, Springer Verlag, Heidelberg, 1987.

[15] G. Elekes and P. Erdős, Similar configurations and pseudo-grids, in *Intuitive Geometry* (K. Böröczky and G. Fejes Tóth, eds.), *Colloq. Math. Soc. János Bolyai*, vol. 63, Elsevier, 1994, 85–104.

[16] G. Elekes and C. D. Tóth, Incidences of not too degenerate hyperplanes, *Proc. 21st Annu. Sympos. Comput. Geom.*, 2005, 16–21.

[17] P. Erdős and G. Purdy, Some extremal problems in geometry IV, *Proc. 7th South-Eastern Conf. Combinatorics, Graph Theory, and Comput.*, 1976, 307–322.

[18] M. Laczkovich and I. Ruzsa, The number of homothetic subsets, in: *The Mathematics of Paul Erdős*, vol. 2, (R. Graham and J. Nešetřil, eds.), Springer Verlag, 1997, 294–302.

[19] H. Lenz, Zur Zerlegung von Punktmengen in solche kleineren Durchmessers, *Arch. Math.* 6 (1955), 413–416.

[20] A. Marcus and G. Tardos, Intersection reverse sequences and geometric applications, *J. Combinat. Theory* Ser. A, 113 (2006), 675–691.

[21] J. Pach and M. Sharir, Geometric incidences, in *Towards a Theory of Geometric Graphs* (J. Pach, ed.), *Contemporary Mathematics*, Vol. 342, Amer. Math. Soc., Providence, RI, 2004, 185–223.

[22] P. J. de Rezende and D.-T. Lee, Point set pattern matching in d-dimensions, *Algorithmica* 13 (1995), 387–404.

[23] L.A. Székely, Crossing numbers and hard Erdős problems in discrete geometry, *Combinatorics, Probability and Computing* 6 (1997), 353–358.