

## Distribution of Distances and Triangles in a Point Set and Algorithms for Computing the Largest Common Point Sets\*

T. Akutsu,<sup>1</sup> H. Tamaki,<sup>2</sup> and T. Tokuyama<sup>3</sup>

<sup>1</sup>Human Genome Center, Institute of Medical Science,  
University of Tokyo, Tokyo 108, Japan  
takutsu@ims.u-tokyo.ac.jp

<sup>2</sup>Department of Computer Science, School of Science and Technology,  
Meiji University, Tama-ku, Kawasaki,  
Kanagawa 214, Japan  
tamaki@cs.meiji.ac.jp

<sup>3</sup>IBM Tokyo Research Laboratory,  
Yamato, Kanagawa 242, Japan  
ttoku@trl.ibm.co.jp

**Abstract.** This paper considers the following problem, which we call the largest common point set problem (LCP): given two point sets  $P$  and  $Q$  in the Euclidean plane, find a subset of  $P$  with the maximum cardinality that is congruent to some subset of  $Q$ . We introduce a combinatorial-geometric quantity  $\lambda(P, Q)$ , which we call the inner product of the distance-multiplicity vectors of  $P$  and  $Q$ , show its relevance to the complexity of various algorithms for LCP, and give a nontrivial upper bound on  $\lambda(P, Q)$ . We generalize this notion to higher dimensions, give some upper bounds on the quantity, and apply them to algorithms for LCP in higher dimensions. Along the way, we prove a new upper bound on the number of congruent triangles in a point set in four-dimensional space.

### 1. Introduction

Matching of point sets in the plane or in space is an important problem that arises, in various forms, in computer vision [17] and computational biology [2], [30], [14]. In the simplest form of the problem, we are asked whether two finite sets of points  $P$  and  $Q$  are

---

\* The first author was partially supported by a Grant-in-Aid for Scientific Research on Priority Areas, "Genome Informatics," of the Ministry of Education, Science, and Culture of Japan.

congruent, that is, whether there is a transformation  $T$  consisting of translation, rotation, and possibly reflection such that  $T(P) = Q$  [5], [27], [4], [3]. A generalization of this question is to ask whether  $Q$  is congruent to some subset of  $P$  (*congruent copy detection* or CCD for short) [8], [15], [26]. We consider in this paper a yet further generalization: we ask for a set  $R$  of the maximum cardinality that is simultaneously congruent to some subset of  $P$  and to some subset of  $Q$ . We call such  $R$  the *largest common point set* between  $P$  and  $Q$ , and name this problem the *largest common point set problem*, or LCP. A version of LCP in which congruence is replaced by similarity (i.e., in which we allow scaling as well) was previously considered by Irani and Raghavan [18]. We set  $n = |P|$ ,  $m = |Q|$ , and assume without loss of generality that  $n \geq m$ .

To solve the LCP, we consider two standard schemes for point matching problems, namely, *voting* and *alignment*. We consider both deterministic and randomized algorithms based on these schemes. Since the randomization techniques we use are also rather standard, our focus is on combinatorial-geometric analyses of the performance of those algorithms. In particular, we study a quantity that we call the *inner product of distance-multiplicity vectors* of two point sets. This quantity is naturally related to the running time of the algorithms we consider for the LCP. We derive a new upper bound on this quantity and hence new upper bounds for the LCP.

Given a set  $P$  of  $n$  points in the plane and a positive real  $l$ , we define the *multiplicity of distance  $l$  in  $P$* , denoted by  $H_P[l]$ , to be the number of ordered pairs  $(\mathbf{p}, \mathbf{q})$  of  $P$  such that the distance between  $\mathbf{p}$  and  $\mathbf{q}$  is  $l$ . The *distance-multiplicity vector* of  $P$ , denoted by  $M(P)$ , is the vector  $(H_P[l_1], \dots, H_P[l_v])$  where  $l_1, \dots, l_v$  is the list of distinct distances with positive multiplicities, sorted in the nonincreasing order of their multiplicities in  $P$ . Erdős [11] asked for theoretical bounds on  $v = v(P)$ , the number of distinct distances in  $P$ , and  $H_P[l_1]$ , the largest multiplicity of a distance in  $P$ ; currently, the best-known upper bound on  $H_P[l_1]$  is  $O(n^{4/3})$  and the best-known lower bound on  $v(P)$  is  $\Omega(n^{4/5})$  [28], [29].

Given two point sets  $P$  and  $Q$ , we define  $\lambda(P, Q) = \sum_l H_P[l]H_Q[l]$ , where the summation is taken over all distances occurring in  $P$  or in  $Q$ . We call  $\lambda(P, Q)$  the “inner product” of the distance-multiplicity vectors  $M(P)$  and  $M(Q)$ , a natural term if we regard  $M(P)$  as a finite representation of a vector in  $(\mathbf{Z}^+)^{\mathbf{R}}$ . Since  $\lambda(P, Q)$  represents the number of matches between the distances in  $P$  and those in  $Q$ , it is not difficult to imagine that this quantity may play an essential role in the complexity analysis of algorithms based on distance comparisons—the more distance matches we have, the more potential congruences we have to consider. We will see that this is indeed the case for both voting- and alignment-based algorithms.

Let  $\lambda(n, m) = \max_{|P|=n, |Q|=m} \lambda(P, Q)$ . A trivial upper bound on  $\lambda(n, m)$  is  $O(n^2m^2)$ , while a result of Erdős [11] implies that  $\lambda(n, m) \geq \Omega(nm^2\sqrt{\log m})$ . The above known bound of  $O(n^{4/3})$  on the maximum multiplicity of a distance immediately implies that  $\lambda(n, m) = O(n^{4/3}m^2)$ . Our result is an upper bound of  $O(n^{1.43}m^{1.77})$ , which is stronger for all values of  $n \leq m^{2.3}$ . This bound is proved by extending the technique of Székely, who has recently developed and simplified proofs of the  $O(n^{4/3})$  bound on the maximum multiplicity and the  $\Omega(n^{4/5})$  bound on the number of distinct distances. We remark that further improving our upper bound, especially the special case of  $O(n^{3.2})$  on  $\lambda(n, n)$ , is arguably difficult, because it would mean improving the best-known lower bound  $\Omega(n^{4/5})$  on  $v(P)$ , a long standing open problem. To see this, first note that  $\lambda(P, P) \geq$

$\nu(P)(n(n-1)/\nu(P))^2 \simeq n^4/\nu(P)$  for any set  $P$  of  $n$  points. This is because, given  $\nu(P)$ ,  $\lambda(P, P)$  is minimized when the number of occurrences are equally distributed among the distinct distances. Therefore, if  $\lambda(n, n) = O(n^{3.2-\epsilon})$  for some positive constant  $\epsilon$ , then  $n^4/\nu(P) = O(n^{3.2-\epsilon})$  or  $\nu(P) = \Omega(n^{0.8+\epsilon})$  for any set  $P$  of  $n$  points.

Similarly to the importance of the distance multiplicities in the plane, the multiplicity of distances and triangles plays an important role in the analysis of LCP algorithms in three-dimensional space. Here, the multiplicity of a triangle  $ABC$  in a point set  $P$  is defined as the number of ordered triples  $(p, q, r)$  of  $P$  such that  $pqr$  is congruent to  $ABC$ . A set of three colinear points is considered as a degenerate triangle. Triangle-multiplicity vectors and their inner products are defined analogously to distance-multiplicity vectors. Given point sets  $P$  and  $Q$  in three-dimensional space, let  $\lambda^{(3,1)}(P, Q)$  (resp.  $\lambda^{(3,2)}(P, Q)$ ) denote the inner product of the distance-multiplicity (resp. triangle multiplicity) vectors of  $P$  and  $Q$ . Let  $\lambda^{(3,1)}(n, m)$  and  $\lambda^{(3,2)}(n, m)$  be defined as before, taking the maximum over  $P$  and  $Q$  with  $|P| = n$  and  $|Q| = m$ . Our upper bounds are  $\lambda^{(3,1)}(n, m) = n^{3/2}m^2(\log^* n)^{O(1)}$  and  $\lambda^{(3,2)}(n, m) = \min\{n^{1.8}m^3, n^{1.8875}m^{2.8} + n^{1.95}m^{2.68}\}(\log^* n)^{O(1)}$ .<sup>1</sup>

From the known bound on the multiplicity of a distance in three dimensions [9], we can show that the number of triangles in  $P$  congruent to a given triangle is  $O(n^{1.8}(\log^* n)^{O(1)})$ . Hence, the first bound in  $\lambda^{(3,2)}(n, m)$  is trivially implied; the second bound comes from the incidence bound of points and circles (with multiplicities) in the space.

These notions naturally carry over to higher dimensions: for a point set in  $d$ -dimensional space, we consider the multiplicity of an  $i$ -dimensional simplex in the point set, where  $1 \leq i \leq d - 1$ , and define  $\lambda^{(d,i)}(P, Q)$  to be the inner product of the multiplicity vectors of  $P$  and  $Q$ , as before.

We only have a naive (though not entirely trivial) bound of  $O(n^{d-1}m^d)$  on  $\lambda^{(d,d-1)}$  for  $d \geq 5$ . For  $d = 4$ , we have a slightly stronger bound of  $O(n^{2.87}m^4)$  on  $\lambda^{(4,3)}$ . This is proved by showing that, for any set  $P$  of  $n$  points in four-dimensional space, the number of triangles in  $P$  congruent to a given triangle is at most  $O(n^{2.83})$ .

From a practical point of view, the relevance of the running-time upper bounds of the LCP algorithms obtained from our combinatorial-geometric analyses might be questioned, because they depend heavily on the exactness of the input and computation. Because of the inevitable errors in the input and computation, several authors formulate the point matching problem as that of approximate matching [4], [8], [15]. For example, the CCD is formulated as asking for a transformation  $T$  that minimizes the directed Hausdorff distance from  $T(Q)$  to  $P$ , i.e.,  $\max_{q \in Q} \min_{p \in P} \text{dist}(T(q), p)$  [8]. Although this is a mathematically clean formulation and has been adopted by many authors, the time complexity of solving the CCD in this model is rather high (the best-known upper bound for the two-dimensional case is  $O(n^2m^3 \log^2 n)$  [8]). Goodrich et al. [15] propose giving up exact minimization and to look for a transformation that achieves a Hausdorff distance within some constant number of times the minimum possible, obtaining an algorithm with a reduced running time of  $O(n^2m \log^2 n)$ . It is not clear, however, if this approach can be extended to the general LCP to produce algorithms with a practical running time. In fact, it is not even clear how we should formulate the LCP in this model,

---

<sup>1</sup>  $\log^* n$  is the height of the lowest tower of the powers satisfying that  $2^{2^{2^{\dots}}} \geq n$ .

**Table 1.** Time complexities for LCP problems.

	Dimension			
	2	3	4	$\geq 5$
LCP	$n^{1.43}m^{1.77} + n^2$	$n^{1.89}m^{2.8} + n^3$	$n^{2.87}m^4 + n^4$	$n^{d-1}m^d + n^d$
LSP	$n^2m^2$ (trivial)	$n^{7/3}m^{8/3} + n^3$	$n^3m^4 + n^4$	$n^{d-1}m^d + n^d$
LCP (randomized)	$LCP/K$	$LCP/K^2$	NA	NA
$\varepsilon$ -Approximate	$LCP/K^2$	$LCP/K^3$	NA	NA
CCD	$\min\{n^{1.43}m^{0.77}, n^{4/3}m\}$	$n^{1.89}m^{0.8} + \min\{n^{2.5}, n^3m^{-2}\}$	$n^{2.87}m + n^{3.83}$	$n^{d-1}m + n^d$
SCD	$n^2m$ [26]	$n^{7/3}m^{2/3} + n^3m^{-2}$	$n^3m + n^4$	$n^{d-1}m + n^d$

since we have two quantities to optimize (the subset size and the Hausdorff distance between the subsets).

A more heuristic approach for approximate matching is to take an algorithm for the exact matching model and use its approximate version, where approximate equality is used instead of exact equality to test matches. This is the approach taken by Irani and Raghavan [18], and probably the one preferred in practice. The generalized Hough transformation method [6], [19] commonly used in computer vision may be interpreted as following such an approach. Our position is that it is meaningful to analyze idealized algorithms as long as their approximate versions are used in practice. The hope that such an analysis will tell us something about the performance of the approximate versions is only heuristic, as much as the hope that such approximate versions will produce answers of any well-defined correctness.

Table 1 summarizes the algorithmic results of this paper, omitting logarithmic factors and using  $K$  to denote the size of the largest common point set. Some minor terms are abbreviated from the complexity of the three-dimensional LCP. We have included some results on the largest similar point set problem (LSP) and the similar copy detection problem (SCD), where similarity rather than congruence is the condition for matches, which can be obtained in a similar approach. Note that our SCD results improve the  $O(n^d m)$  time bound of Rezende and Lee [26] in three or more dimensions (and when specialized to CCD, in two dimensions as well).

The rest of this paper is organized as follows. In Section 2 we review the algorithms based on voting and alignment and see how the inner-product of multiplicity vectors appears in their analysis. Combinatorial-geometric analyses of the inner products are given in subsequent sections.

## 2. Algorithms

The goal of this section is to provide a motivation of our combinatorial-geometric analysis of the inner product of the multiplicity vectors by showing how it is related to the analysis of several algorithms for the LCP. We consider algorithms based on a voting scheme and algorithms based on an alignment scheme. In the description of the following algorithms, we consider congruence transformations that consist only of translations and rotations,

since reflections can be taken care of by running the algorithm once for  $P$  and  $Q$  and once for  $P$  and the mirror image of  $Q$ . We refer to such congruence transformations simply as transformations. We use  $K$  consistently to denote the size of the largest common point set between  $P$  and  $Q$ .

### 2.1. Voting Scheme

In this and the next subsection we assume that the point sets  $P$  and  $Q$  are in the plane. Let  $dist(x, y)$  be the distance between points  $x$  and  $y$ . Consider pairs  $p, p'$  and  $q, q'$  of points of  $P$  and  $Q$ , respectively. If  $dist(p, p') = dist(q, q')$ , these four points determine a unique transformation  $T$  such that  $T(p) = q$  and  $T(p') = q'$ . This transformation is denoted by  $T[p, p'; q, q']$ .

For each pair  $(p, q)$  of points  $p \in P$  and  $q \in Q$ , and each transformation  $T$ , we define  $mult_{p,q}(T)$  to be the number of pairs  $x \in P$  and  $y \in Q$  satisfying that  $T[p, x; q, y] = T$ . If  $|T(P) \cap Q| = k$ , it is easy to see that  $mult_{p,q}(T) = k - 1$  for any  $q \in T(P) \cap Q$  and  $p = T^{-1}q$ .

Hence,  $\max_{p,q}(\max_T mult_{p,q}(T)) = K - 1$  and the transformation achieving this maximum is the transformation that gives the largest common point set.

For each pair  $p \in P$  and  $q \in Q$ ,  $\max_T mult_{p,q}(T)$  can be computed by considering all matching pairs of edges  $(px, qy)$ ,  $x \in P$ ,  $y \in Q$ , and letting each pair cast a vote to the transformation  $T[p, x; q, y]$ . We call this process the *local voting* for pair  $(p, q)$ .

A similar voting idea is used in the generalized Hough transformation [6], where a vote is cast to a bucket in the transformation space rather than to an individual transformation as in our case. In other words, our scheme may be viewed as an extreme version of the generalized Hough transformation where the bucket size goes to zero.

In the deterministic version of the algorithm, we need to execute the local voting process for each pair  $(p, q)$ , with  $p \in P$  and  $q \in Q$ , and take the best result. The total number of votes cast in the algorithm execution is  $\lambda(P, Q)$ , because each matching pair  $(pp', qq')$  of edges contributes exactly one vote. Thus,  $\lambda(n, m)$  bounds the essential term in the running time.

**Theorem 2.1.** *In an efficient implementation, the deterministic voting algorithm computes the two-dimensional LCP in  $O((\lambda(n, m) + n^2) \log n)$  time and  $O(n)$  space.*

*Proof.* For each  $p \in P$ , we first sort the set  $\{(p, x) | x \in P\}$  of pairs in terms of  $dist(p, x)$  and store it in a binary search data structure  $D_p(P)$ . This requires  $O(n \log n)$  time for each  $p \in P$ . In the local voting for a fixed pair  $(p, q)$ , we search for  $dist(q, y)$ , for each  $y \in Q$ , in the data structure  $D_p(P)$  in  $O(\log n)$  time to obtain the matches  $px$  of the edge  $qy$ . This requires  $O(m \log n)$  time for each of  $nm$  local voting processes. The rest of the running time is at most  $O(\log n)$  per vote. Thus, the overall running time is  $O((\lambda(n, m) + nm^2 + n^2) \log n)$ , but we have  $\lambda(n, m) = \Omega(nm^2)$ , as noted earlier. See Appendix A for how to reduce the space complexity to  $O(n)$ .  $\square$

Observe that, in the above deterministic voting algorithm, the optimal transformation  $T$  receives the maximum vote of  $K - 1$  in the voting process for each  $(p_i, T(p_i))$ ,

$i = 1, \dots, K$ , where  $\{p_i\}$  is the largest common point set. This is a redundancy that is difficult to avoid deterministically. However, we can reduce the redundancy at the cost of a small failure probability by using a standard random sampling procedure: sample a random subset  $R$  of  $P$  and do local voting only for pairs in  $R \times Q$ . Here, once we fix a pair  $p \in R$  and  $q \in Q$ , the local voting for  $(p, q)$  is done exactly as before, scanning the entire sets  $P$  and  $Q$  (not just  $R$ ). This speeds up the algorithm by a factor of  $|P|/|R|$ : the terms  $n^2 \log n$  and  $nm^2 \log n$  of the deterministic time become  $|R|n \log n$  and  $|R|m^2 \log n$ , respectively, and the expected total number of votes cast is  $|R|n^{-1}\lambda(P, Q)$ . The algorithm makes an error only when  $R$  is disjoint from the largest common point set. This occurs only if all the  $|R|$  random draws are picked outside the largest common point set. Thus, the probability is  $(1 - K/n)^{|R|} = \{(1 - K/n)^{n/K}\}^{|R|K/n}$ , which is at most  $e^{-|R|K/n}$ , since  $(1 - K/n)^{n/K} < 1/e$ , where  $e$  is the base of the natural logarithm. We set a number  $c$  and apply the algorithm to  $|R|$  of size  $2^s c$  for  $s = 1, 2, \dots$  until we get an LCP which is larger than  $cn/|R|$ . When  $|R|$  becomes larger than  $cn/K$ , the algorithm outputs the correct LCP of the size  $K > cn/|R|$  with probability  $e^{-c}$ , and terminates. The total computation time is  $O(c(\lambda(n, m) + n^2)K^{-1} \log n)$ . Hence, we have the following theorem:

**Theorem 2.2.** *LCP can be solved in  $O(c(\lambda(n, m) + n^2)K^{-1} \log n)$  time with probability  $1 - e^{-c}$ .*

If we can be satisfied with an approximately optimal solution, which has a size of at least  $(1 - \varepsilon)K$ , we can further reduce the running time by adopting a different sampling strategy. We sample  $\tilde{P} \subset P$  with  $\alpha n$  points and  $\tilde{Q} \subset Q$  with  $\beta m$  points, where  $\alpha\beta = c \log n/K$  for some sufficiently large constant  $c$ , and compute the largest common set between  $\tilde{P}$  and  $\tilde{Q}$ .

To see why this sampling approach works, let  $T$  be any transformation. For each  $p \in P$  such that  $T(p) \in Q$ , the probability that both  $p$  and  $T(p)$  are sampled is  $\alpha\beta = c \log n/K$ . Therefore, if  $T$  gives a common point set of size  $K$  between  $P$  and  $Q$ , it gives a common point set of expected size  $c \log n$  between  $\tilde{P}$  and  $\tilde{Q}$  and, with high probability, of size at least  $(1 - \varepsilon/2)c \log n$ . From Chernoff's bound [23], the probability is at least  $1 - \exp(-c \log n \varepsilon^2/8) > 1 - n^{-c\varepsilon^2/8}$ .

On the other hand, a transformation  $T$  is called *bad* if it satisfies both  $|T(P) \cap Q| < (1 - \varepsilon)K$  (i.e., it is not a near-optimal solution for LCP) and  $|T(\tilde{P}) \cap \tilde{Q}| > (1 - \varepsilon/2)c \log n$ . Again from Chernoff's bound, the probability that a given transformation is bad is at most  $n^{-c\varepsilon^2/8}$ . Since there are at most  $\lambda(n, m) < n^2 m^2$  possible transformations mapping a pair of points of  $P$  to a pair of points of  $Q$ , the probability that there is any bad transformation is at most  $n^2 m^2 n^{-c\varepsilon^2/8}$ .

Thus, the solution of LCP for  $\tilde{P}$  and  $\tilde{Q}$  gives, with probability  $1 - (n^2 m^2 + 1)n^{-c\varepsilon^2/8}$ , a near-optimal transformation for  $P$  and  $Q$  as desired. If we take  $c$  larger than  $40\varepsilon^{-2}$ , the above probability is  $1 - o(n^{-1})$ . For the running time analysis, observe that the expected value of  $\lambda(\tilde{P}, \tilde{Q})$  is  $\alpha^2 \beta^2 \lambda(P, Q)$ , because the multiplicity of a fixed distance in  $P$  (resp.  $Q$ ) is reduced by a factor of  $\alpha^2$  (resp.  $\beta^2$ ) and because  $\tilde{P}$  and  $\tilde{Q}$  are sampled independently. (Note that this is a better bound than  $\lambda(\alpha n, \beta m)$ .)

We choose the pair  $(\alpha, \beta)$ , so as to reduce the preprocessing time of the deterministic algorithm applied to  $\tilde{P}$  and  $\tilde{Q}$  as much as possible: if  $n \log n/K$ , then we set

$(\alpha, \beta) = (c \log n / K, 1)$  and otherwise set  $(\alpha, \beta) = (\sqrt{cm \log n / nK}, \sqrt{cn \log n / mK})$  (so that  $\alpha n = \beta m = \sqrt{cnm \log n / K}$ ). In the former case, the preprocessing time is  $O((\alpha n)^2 \log n) = O(n^2 \log^3 n / K^2)$ . In the latter case, it is  $O((\alpha n)^2 \log n) = O(nm \log^2 n / K) = O(nm^2 \log^2 n / K^2)$ . Since  $\lambda(n, m) = \Omega(nm^2)$  as noted earlier, this preprocessing time in the latter case is subsumed by the time for the main part. Therefore, we have the following:

**Theorem 2.3.** *For any fixed constant  $\varepsilon > 0$ , we can find, with probability  $1 - n^{-1}$ , a common point set between  $P$  and  $Q$  of size at least  $(1 - \varepsilon)K$  in an expected  $O((\lambda(n, m) + n^2)K^{-2} \log^3 n)$  time.*

## 2.2. Alignment Scheme

In the *alignment* scheme, we consider each pair of edges of equal length  $pp'$  from  $P$  and  $qq'$  from  $Q$ , align  $pp'$  with  $qq'$ , and count the number of points of  $P$  that now coincide with a point of  $Q$ . In other words, we count the size of the intersection  $T(P) \cap Q$ , where  $T = T[\mathbf{p}, \mathbf{p}'; \mathbf{q}, \mathbf{q}']$ . In the deterministic version, we need to consider  $\lambda(P, Q)$  alignments and the counting for each alignment takes  $O(m \log n)$  time. Alignments can be enumerated in a manner similar to the voting-based algorithm. Thus, we have a running time bound of  $O((\lambda(n, m)m + n^2) \log n)$ . This is worse than that of the voting-based algorithm. However, when we apply random sampling (for exact optimization), the rate of reduction in the running time is larger in this case. Sample a subset  $R$  of  $P$  with size  $cn/K$ , with sufficiently large  $c$  so that  $R$  contains at least two points of the largest common point set with high probability. Then we need only consider each edge  $pp'$  in  $R$  for alignment. This is in contrast to the voting case, where we need to consider each edge  $pp'$  with  $p \in R$  and  $p' \in P$  for voting. Thus, the first term of the running time is reduced by a factor of  $c^2/K^2$  and, when  $K$  is close to  $m$ , is comparable with the voting scheme. Irani and Raghavan [18] apply random sampling to an alignment scheme for the LSP; this is reasonable, because  $K$  is linear in  $m$  in the applications they intend.

For the CCD, where our goal is to determine whether  $K = m$ , the alignment scheme has the advantage of giving a fast Las Vegas algorithm (i.e., an algorithm that does not risk an error), in contrast to the voting scheme, which gives a Monte Carlo randomized algorithm (i.e., an algorithm that has some probability of making an error). The idea is to sample from  $Q$  rather than from  $P$ . The reason we sample from  $P$  for the general LCP is because we want to reduce the  $n^2 \log n$  term in the deterministic time complexity, which accounts for the preprocessing of distances in  $P$ . For the CCD, this preprocessing can be avoided, and sampling from  $Q$  turns out to be advantageous.

**Theorem 2.4.** *The CCD can be solved in  $O(n^{4/3}m \log n + n^{4/3} \log^{8/3} n)$  deterministic time and in  $O(\min\{\lambda(n, m)m^{-1} \log n, n^{4/3}m \log n\} + n^{4/3} \log^{8/3} n)$  Las Vegas expected time.*

*Proof.* We randomly sample a pair  $\mathbf{q}, \mathbf{q}'$  of points from  $Q$ . Using the the methods of Agarwal et al. [1], we find all occurrences of the distance  $\text{dist}(\mathbf{q}, \mathbf{q}')$  in  $P$  in  $O(n^{4/3} \log^{8/3} n + k \log n)$  time, where  $k$  is the number of occurrences of the distance.

Even for the worst pair  $q, q'$ , we have  $k = O(n^{4/3})$  from the bound on the repeated distances; moreover, the expected value of  $k$  is  $\lambda(P, Q)/m^2$ . For each such occurrence  $pp'$ , we align  $qq'$  with  $pp'$  and test whether the entire  $Q$  is matched into  $P$ , using  $O(m \log n)$  time. The total running time is  $O(km \log n + n^{4/3} \log^{8/3} n)$ , from which the claim follows when the bounds on  $k$  are plugged in.  $\square$

### 2.3. Higher Dimensions

Both the voting and alignment schemes can be easily generalized to higher dimensions. We discuss only the voting scheme here. The generalization of the alignment scheme (and especially its specialization to the CCD) can be done similarly. We start with the voting scheme in three dimensions.

Let the point sets  $P$  and  $Q$  be in three-dimensional space. Local voting is now done fixing a pair of edges of equal length  $pp'$  from  $P$  and  $qq'$  from  $Q$ . Each pair of points  $x \in P$  and  $y \in Q$  such that the triangles  $pp'x$  and  $qq'y$  are congruent casts a vote to the transformation that matches the triangles. When these triangles are degenerate, that is, when  $x$  lies on the line  $pp'$  and  $y$  on the line  $qq'$ , and hence does not determine a transformation, the vote becomes “public,” that is, it conceptually adds a count to every candidate in this particular local voting process. Public votes are counted separately and added to the maximum vote count in the end of the local voting. We do local voting for every matching pair  $(pp', qq')$  of edges and take the transformation with the maximum vote count. It is clear that the maximum vote count is  $K - 2$ .

The number of local voting processes to be performed is bounded by  $\lambda^{(3,1)}(n, m)$  and the number of total votes cast by  $\lambda^{(3,2)}(n, m)$ . To enumerate matching pairs of edges for which to perform local voting, we use a sorted list of distances in  $P$  as we did in the planar case, so that the time consumed for finding all matching pairs becomes  $O((n^2 + \lambda^{(3,1)}(n, m)) \log n)$ . To enumerate matching triangles that cast votes, we similarly prepare a sorted list of triangles in  $O(n^3 \log n)$  preprocessing time, so that the time consumed for finding all matching triangles becomes  $O(\lambda^{(3,1)}(n, m)m \log n + \lambda^{(3,2)}(n, m))$ . Thus, we have:

**Theorem 2.5.** *The LCP in three-dimensional space can be solved in time  $O((\lambda^{(3,2)}(n, m) + \lambda^{(3,1)}(n, m)m + n^3) \log n)$ .*

When we use our current bounds on  $\lambda^{(3,2)}(n, m)$  and  $\lambda^{(3,1)}(n, m)$ , which are roughly  $O(n^{1.89}m^{2.8})$  and  $O(n^{1.5}m^2)$ , respectively, it turns out that the second term is subsumed by the first. Random sampling approaches also work for three-dimensional space and give a speed-up of  $K^2$  for exact optimization and roughly  $K^3$  for approximate optimization. We omit the routine details.

**Corollary 2.1.** *The CCD in three-dimensional space can be solved in time  $O((\lambda^{(3,2)}(n, m)m^{-2} + \lambda^{(3,1)}(n, m)m^{-1}) \log n + \min(n^{2.5}, n^3m^{-2} \log n))$ .*

*Proof.* We set  $K = m$  to obtain  $O((\lambda^{(3,2)}(n, m)m^{-2} + \lambda^{(3,1)}(n, m)m^{-1} + n^3m^{-2}) \log n)$ . If  $m$  is smaller than  $n^{1/4}$ , we select a triangle  $\Delta$  and its edge  $e$  from  $Q$ . We find all possible



(at most  $O(n^{1.5})$ ) locations of  $e$  in  $P$  in  $O(n^2)$  time. Then we can find all possible locations of  $\Delta$  in  $P$  in  $O(n^{2.5})$  time to improve the  $n^3m^{-2} \log n$  term.  $\square$

Theorem 2.5 can be generalized to higher dimensions.

**Theorem 2.6.** *For any fixed  $d \geq 3$ , the LCP in the  $d$ -dimensional space can be solved in  $O((\lambda^{(d,d-1)}(n, m) + \lambda^{(d,d-2)}(n, m)m + n^d) \log n)$  time.*

Analysis of random sampling schemes for  $d \geq 4$  appears difficult if our goal is to achieve a speed-up of  $K^{d-1}$ , due to the possible degeneracy of the largest common point set.

### 3. Bounding $\lambda(n, m)$

We now turn to the combinatorial-geometric analysis of the inner product of multiplicity vectors. We start with the planar case.

**Theorem 3.1.**  $\lambda(n, m) = O(n^{1.43}m^{1.77})$ .

The following simple application of the Szemerédi–Trotter theorem [29] turns out to be useful.

**Lemma 3.1.** *Let  $P$  be a set of  $n$  points and let  $\mathcal{L}$  be a set of lines in the plane or higher-dimensional space such that each line in  $\mathcal{L}$  contains at least  $y$  points of  $P$ , where  $y \geq 2$ . Then the number of incidences between  $P$  and  $\mathcal{L}$  is  $O(n + n^2y^{-2})$ .*

*Proof.* We may assume that  $P$  and  $\mathcal{L}$  are configured in a plane: otherwise we may take an appropriate projection, such that at most one incidence in the space is mapped to each incidence in the plane, and bound the number of incidences in the plane. Let  $I$  denote the number of incidences between  $P$  and  $\mathcal{L}$ . The Szemerédi–Trotter theorem [29] (see also [20]) states that  $I \leq c(n + |\mathcal{L}| + n^{2/3}|\mathcal{L}|^{2/3})$ , where  $c$  is some constant. Since each line of  $\mathcal{L}$  contains at least two points of  $P$ ,  $|\mathcal{L}| \leq n^2$  and therefore  $I \leq 3cn^2$ . Thus, for  $y < 2c$ , we have  $I \leq 12c^3n^2y^{-2}$  so the bound of  $O(n + n^2y^{-2})$  certainly holds. We assume  $y \geq 2c$  in the following. From the assumption that each line of  $|\mathcal{L}|$  is incident to at least  $y$  points of  $P$ , we have  $y|\mathcal{L}| \leq I \leq c(n + |\mathcal{L}| + n^{2/3}|\mathcal{L}|^{2/3})$ . Since  $y \geq 2c$ , it follows that  $y|\mathcal{L}|/2 \leq c(n + n^{2/3}|\mathcal{L}|^{2/3})$  which implies that either  $|\mathcal{L}| \leq 4cn/y$  or  $|\mathcal{L}| \leq 4cn^{2/3}|\mathcal{L}|^{2/3}/y$ . In the first case, using the Szemerédi–Trotter theorem again, we have  $I = O(n + n^{4/3}y^{-2/3})$ . In the latter case, we have  $|\mathcal{L}| \leq (4c)^3n^2y^{-3}$  and hence  $I = O(n + n^2y^{-2})$ . Combining these cases we have  $I = O(n + n^{4/3}y^{-2/3} + n^2y^{-2})$ . The second term is subsumed by the first when  $y > n^{1/2}$  and by the third when  $y \leq n^{1/2}$ .  $\square$

We also need some lemmas due to Székely [28], [20] (Lemmas 3.2 and 3.3 below), which he used in proving the  $\Omega(n^{4/5})$  bound on the number of distinct distances in an arbitrary set of  $n$  points.

The *crossing number* of a drawing of a graph  $G$  on a plane is the number of intersections of drawn edges (see [24] and [25]). Let  $cr(G)$  denote the minimum of the crossing numbers over all drawings of  $G$ .

**Lemma 3.2.** *In the graph  $G$  with  $n$  vertices and  $t$  edges, if each edge has at most  $y$  edges parallel to it, then there are constants  $c_1$  and  $c_2$  satisfying  $cr(G) = c_1 t^3 / y n^2 - c_2 y^2 n$ .*

Recall the notation in the Introduction:  $v(P)$  is the number of distinct distances in  $P$ ,  $\mathcal{D}(P) = \{l_1, \dots, l_{v(P)}\}$  is the set of such distances listed in the nonincreasing order of multiplicity, and  $H_P[l]$  is the multiplicity of the distance  $l$  in  $P$ . Let  $1 \leq k \leq v(P)$ . We consider the subset  $\{l_1, \dots, l_k\}$  of  $\mathcal{D}(P)$ , consisting of the first  $k$  entries, and let  $f(k) = \sum_{i=1}^k H_P[l_i]$ . Seeking an upper bound on  $f(k)$ , we create  $k$  concentric circles of radii  $l_1, \dots, l_k$  around each point of  $P$ . Hence, we have  $nk$  circles in total: let  $\mathcal{C}(k)$  denote the set of these circles.

The circles in  $\mathcal{C}(k)$  have  $2f(k)$  incidences with the point set  $P$ . We delete circles containing at most two points of  $P$ , and let  $\mathcal{C}'(k)$  denote the set of remaining circles. Each circle of  $\mathcal{C}'(k)$  contains at least three points of  $P$ , and these points cut the circle into at least three arcs. If we let  $A$  denote the set of these arcs coming from all the circles of  $\mathcal{C}'(k)$ , then  $|A|$  is equal to the number of incidences between the circles of  $\mathcal{C}'(k)$  and the points of  $P$ , and hence is at least  $2f(k) - 2kn$ . Define a multigraph  $G(k)$  whose vertices are points of  $P$ , and whose edges are arcs of  $A$ . Since each pair of at most  $kn$  circles of  $\mathcal{C}'(k)$  intersect at most twice, the crossing number  $cr(G(k))$  of  $G(k)$  is at most  $k^2 n^2$ .

The following lemma is due (in a slightly different form) to Székely. The proof is a straightforward adaptation of his proof [20], [28].

**Lemma 3.3.** *For each  $y$ , the number of edges in  $G(k)$  having at least  $y$  edges parallel to them is at most  $O(n^2 k / y^2 + kn \log n)$ .*

*Proof.* Let  $M_i$  be the set of pairs  $\{\mathbf{u}, \mathbf{v}\}$  of vertices of  $G(k)$  such that there are at least  $2^i$  and fewer than  $2^{i+1}$  parallel edges connecting  $\mathbf{u}$  and  $\mathbf{v}$  in  $G(k)$ , and let  $E_i$  be the set of all edges each connecting a pair in  $M_i$ . Our goal is to find an upper bound for  $\sum_{y \leq 2^i \leq n} |E_i|$ .

Let  $L_{\mathbf{u}, \mathbf{v}}$  denote the perpendicular bisector of the segment  $\mathbf{u}\mathbf{v}$  and let  $\mathcal{L}_i = \{L_{\mathbf{u}, \mathbf{v}} \mid \{\mathbf{u}, \mathbf{v}\} \in M_i\}$ . Each edge connecting  $\mathbf{u}$  and  $\mathbf{v}$  determines one incidence of the line  $L_{\mathbf{u}, \mathbf{v}}$  with  $P$ , since the center of a circle through  $\mathbf{u}$  and  $\mathbf{v}$  must lie on  $L_{\mathbf{u}, \mathbf{v}}$ . On the other hand, we note that one such point-line incidence  $(\mathbf{p}, L)$ , where  $\mathbf{p} \in P$  and  $L \in \mathcal{L}_i$ , corresponds to at most  $2k$  edges of  $E_i$  because there are only  $k$  different radii. Therefore, the size of  $E_i$  is at most  $2kN_i$ , where  $N_i$  is the number of incidences of  $\mathcal{L}_i$  with  $P$ . By Lemma 3.1, we have  $N_i = O(n + n^2 2^{-2i} + n)$  and hence  $|E_i| = O(nk + n^2 k 2^{-2i})$ , from which the lemma follows when we sum over  $i$ . □

The following basic fact is useful.

**Lemma 3.4.** *Let  $\mathbf{X} = x_1, \dots, x_N$  be a nonincreasing sequence of positive real numbers. For sequences  $\mathbf{Y} = y_1, \dots, y_N$  and  $\mathbf{Z} = z_1, \dots, z_N$  of real numbers, if there exists a fixed positive constant  $c$  satisfying  $\sum_{i=1}^k z_i \geq c \sum_{i=1}^k y_i$  for  $k = 1, 2, \dots, N$ , the inner product  $(\mathbf{X}, \mathbf{Z})$  of  $\mathbf{X}$  and  $\mathbf{Z}$  is at least  $c(\mathbf{X}, \mathbf{Y})$ .*

**Proposition 3.1.**  $f(k) = O(\min\{n^{10/7}k^{5/7}, n^2\})$ .

*Proof.*  $O(n^2)$  is a trivial bound. The  $O(n^{10/7}k^{5/7})$  bound needs to be proved for  $k < n^{4/5}$ , where we can assume that  $f(k) \gg kn \log n$ .

We construct  $G(k)$ , set  $y = \sqrt{Cn^2k/f(k)}$  for some constant  $C$ , and delete all edges of multiplicity larger than  $y$ . From Lemma 3.3, at most  $f(k)$  edges are removed, if we set  $C$  to be sufficiently large. Then we have a graph  $G$  with  $\Omega(f(k))$  edges, whose edge multiplicity is at most  $y$ . From Lemma 3.2,  $k^2n^2 > cr(G) = \Omega(f(k)^3/yn^2) - O(y^2n)$ . Hence,  $k^2n^2 = \Omega(f(k)^{3.5}/n^3k^{0.5})$  (the  $O(y^2n)$  term is negligible). It follows that  $f(k)^{3.5} = O(k^{2.5}n^5)$  and hence  $f(k) = O(n^{10/7}k^{5/7})$ .  $\square$

We are now ready to prove Theorem 3.1. Let  $\mathcal{D}(P) = \{l_1, \dots, l_{\nu(P)}\}$  (resp.  $\mathcal{D}(Q) = \{l'_1, \dots, l'_{\nu(Q)}\}$ ) denote the set of distinct distances of  $P$  (resp.  $Q$ ) sorted in the nonincreasing order of their multiplicities. Note first that

$$\lambda(P, Q) = \sum_{l \in \mathcal{D}(P) \cap \mathcal{D}(Q)} H_P[l]H_Q[l] \leq \sum_{k=1}^{\min(\nu(P), \nu(Q))} H_P[l_k]H_Q[l'_k], \quad (1)$$

due to the monotonicity of the multiplicity sequences. The monotonicity also implies that  $H_P[l_k] \leq f(k)/k$  and hence, from the above bound on  $f(k)$ , that  $H_P[l_k] = O(n^{10/7}k^{-2/7})$  for  $k \leq O(n^{4/5})$ .  $H_P[l_k] : k = 1, 2, \dots, \nu(P)$  is a nonincreasing sequence.

Similarly, we have  $H_Q[l'_k] = O(m^{10/7}k^{-2/7})$  for  $k \leq O(m^{4/5})$ . We can replace  $H_Q[l'_k]$  by  $m^{10/7}k^{-2/7}$  for  $k \leq m^{4/5}$  and 0 for  $k > m^{4/5}$  in inequality (1) in order to give an asymptotic upper bound for  $\lambda(P, Q)$ . Indeed,  $\sum_{i=1}^t m^{10/7}k^{-2/7} \geq m^{10/7}k^{5/7} \geq c \sum_{i=1}^t H_Q[l'_k]$  for  $t = 1, 2, \dots, m^{4/5}$  and some constant  $c$ , and  $\sum_{i=1}^{m^{4/5}} m^{10/7}k^{-2/7} \geq m^2 > \sum_{i=1}^t H_Q[l'_k]$  for any  $t > m^{4/5}$ ; therefore, Lemma 3.4 works.

Hence,

$$\lambda(P, Q) \leq \sum_{k=1}^{m^{4/5}} O(k^{-4/7}n^{10/7}m^{10/7}) = O(n^{10/7}m^{62/35}).$$

Because  $n > m$ , we can write the bound as  $O(n^{1.43}m^{1.77})$ .

#### 4. The Three-Dimensional Case

We now turn to three-dimensional space. In the following,  $O'(f(n))$  is used as an abbreviation of  $(\log^* n)^{O(1)}f(n)$ , hiding polynomial factors of  $\log^* n$ . It is known that the number of occurrences of a given distance in a set of  $n$  points in three-dimensional space is  $O'(n^{3/2})$  [9]. This trivially implies that  $\lambda^{(3,1)}(n, m) = O'(n^{3/2}m^2)$ . In the following, we derive an upper bound on  $\lambda^{(3,2)}(n, m)$  by using a result on the circle-point incidence problem. We use a classical lemma in extremal graph theory:

**Lemma 4.1** (Bipartite Graph Lemma). *Let  $H = (U, V, E)$  be a bipartite graph, where  $E$  is the set of edges between vertex sets  $U$  and  $V$ . If  $H$  contains no subgraph isomorphic*

to the complete bipartite graph  $K_{s,t}$ , the following inequality holds:

$$|U| \binom{\lfloor |E|/|U| \rfloor}{t} \leq (s-1) \binom{|V|}{t}.$$

In particular, if  $s$  and  $t$  are constants,  $|E| = O(|U|^{1-1/t}|V| + |U|)$ .

*Proof.* This can be proved by counting the number of subgraphs isomorphic to  $K_{1,t}$  contained in  $H$ . See Lemma 2.1, p. 309 of [7] for the details of the proof. The proof for the special case where  $s$  and  $t$  are constants can also be found in [9] and [24].  $\square$

**Lemma 4.2.** *Given a set of  $N$  circles and  $n$  points in the  $d$ -dimensional space for  $d \geq 3$ , there are  $O(N + n + \min\{N^{2/3}n, Nn^{1/2}, N^{4/5}n^{3/5}\})$  circle-point incidences.*

*Proof.* We apply the scheme of Clarkson et al. [9]. First, we consider the bipartite graph  $H = (U, V, E)$ , where  $U$  corresponds to the set of circles,  $V$  corresponds to the set of points, and  $E$  is the edge set, where a vertex of  $U$  and a vertex of  $V$  are connected with an edge if the corresponding circle and point are incident. Since two circles intersect at no more than two points (in any-dimensional space), the graph does not contain  $K_{2,3}$ . From the Bipartite Graph Lemma applying to both  $H$  and the graph obtained by exchanging the roles of  $U$  and  $V$  in  $H$ , we have that  $H$  contains  $O(N + n + \min\{N^{2/3}n, Nn^{1/2}\})$  edges. We call this bound the *Canham-like* bound, abusing the terminology of [9].

Next, we project the figure to a two-dimensional plane, so that no pairs of projected points overlap. Since the projected image of a spatial circle is an ellipse, we have a set  $\mathcal{E}$  of ellipses. We construct a *cutting* from these ellipses as follows: We take  $r$  sample ellipses and construct the arrangement of ellipses. We subdivide each cell of the arrangement into “funnels” by vertical line segments each of which either is tangent to an ellipse or passes a vertex of the arrangement; here, a funnel is a region bounded by at most two ellipses and two vertical line segments. There are  $O(r^2)$  the funnels  $F_1, \dots, F_k$ .

Let  $n_i$  be the number of projected points located in  $F_i$ , except the vertices of the arrangement. A point on the boundary between two funnels is counted for each of the funnels.

A circle in the space has an incidence with a point only if its projected image intersects the funnel containing the projected image of the point. Let  $N_i$  be the number of projected ellipses of  $\mathcal{E}$  intersecting the interior of  $F_i$ . If we lift up these points and ellipses to the points and circles in three-dimensional space to the original position, there are  $O(N_i + N_i^{2/3}n_i)$  incidences between these points and circles because of the Canham-like bound.

The number of incidences at the points projected to the vertices of the arrangement is that between  $N$  ellipses and  $O(r^2)$  points, and hence  $O(N + r^2 + Nr) = O(Nr)$  (by definition,  $r < N$ ) from the Canham-like bound. Therefore,  $O(Nr + \sum_{i=1}^k (n_i N_i^{2/3} + N_i))$  incidences are counted in total.

Clarkson et al.’s “sampling lemma” (Lemma 5.3 in [9]) shows that there exists a cutting satisfying that  $\sum_{i=1}^k n_i N_i^{2/3} = O(n(N/r)^{2/3})$  and  $\sum_{i=1}^k N_i = O(rN)$ . We remark that the sampling lemma is stated for pseudolines and pseudocircles in [9]; however, it can be applied to ellipses in a straightforward way.

Thus, the total number of incidences is  $O(n(N/r)^{2/3} + rN)$ . We set  $r = n^{3/5}N^{-1/5}$  to obtain the  $O(N^{4/5}n^{3/5})$  bound.  $\square$

Erdős and Purdy [13] gave an  $O(n^{3-1/3})$  bound for the number of triangles with a given volume in the space, and asked a question about  $H_P[\Delta]$ , the number of pairwise congruent triangles.

**Proposition 4.1.** *The number of triangles in  $P$  congruent to a given triangle  $\Delta$  is  $O'(n^{1.8})$ . The number of triangles in  $P$  similar to  $\Delta$  is  $O(n^{2.2})$ .*

*Proof.* For the congruency, let  $\Delta = ABZ$ , and let  $h$  be the distance from  $Z$  to the line  $AB$ . For each pair  $\mathbf{p}, \mathbf{q}$  of points of  $P$  with  $dist(\mathbf{p}, \mathbf{q}) = dist(A, B)$ , the trajectory of the vertex  $Z$  of a triangle congruent to  $\Delta$  locating  $A$  and  $B$  at  $\mathbf{p}$  and  $\mathbf{q}$ , respectively, forms a circle of radius  $h$ . Thus, if  $N$  is the multiplicity of distance  $dist(A, B)$  in  $P$ , the multiplicity of  $\Delta$  is at most the number of incidences between  $N$  circles and  $n$  points. Since  $N = O'(n^{3/2})$  [9], it immediately follows from the above lemma that  $H_P[\Delta] = O'((n^{3/2})^{4/5}n^{3/5}) = O'(n^{1.8})$ .

For the similarity, there are  $O(n^2)$  circles and  $O(n)$  points. Hence, the incidence is  $O((n^2)^{4/5}n^{3/5}) = O(n^{2.2})$ .  $\square$

For each triangle  $\Delta$ , let  $H_P[\Delta]$  denote the multiplicity of  $\Delta$  in  $P$  (i.e., the number of ordered triples of  $P$  forming an occurrence of  $\Delta$ ) and recall that  $\lambda^{(3,2)}(P, Q) = \sum_{\Delta} H_P[\Delta]H_Q[\Delta]$ . Since  $H_P[\Delta] = O'(n^{1.8})$ , it can immediately be shown that  $\lambda^{(3,2)}(n, m) = O'(n^{1.8}m^3)$ . If  $n = m$ , the bound becomes  $O'(n^{4.8})$ . We can obtain another bound that is slightly better if  $m$  is large.

Let  $\mathcal{T}_l(P) = \{\Delta_1^l, \dots, \Delta_{k_l}^l\}$  denote the set of distinct triangles of  $\mathcal{T}_l$ , each of which contains an edge of length  $l$ , occurring in  $P$ , listed in the nonincreasing order of their multiplicities. We bound  $F_P(k; l) = \sum_{i=1}^k H_P(\Delta_i^l)$  for each  $l$ . Suppose that  $l$  occurs  $L_P(l) = n^\alpha$  times in  $P$ . Neglecting  $\log^* n$  factors, we can assume that  $\alpha \leq 1.5$  because of the upper bound on the repeated distances in the space.

**Lemma 4.3.**  $F_P(k; l) = O(kn^{7/5} + k^{4/5}n^{0.8+0.7\alpha})$ . Also,  $F_P(k; l) = O(n^{1+\alpha})$ .

*Proof.* The  $O(n^{1+\alpha})$  bound is almost trivial, since, for each location of  $l$ , there are at most  $n$  possible locations of the opposite vertex in  $P$ .

We show the other bound. We can assume  $k < n$ , since otherwise the  $O(n^{1+\alpha})$  bound is better. If we fix the location of the edge of length  $l$  of the triangle  $\Delta_i^l$ , the trajectory of the opposite vertex of the possible location of the triangle is a circle. Since the number of possible locations of the edge of length  $l$  is  $n^\alpha$ , and we consider  $k$  different triangles each containing an edge of length  $l$ , we have a family  $\mathcal{C}(k)$  of  $O(kn^\alpha)$  circles.  $F_P(k; l)$  is at most the number of incidences between the circles in  $\mathcal{C}(k)$  and  $n$  points. However, the family  $\mathcal{C}(k)$  should be considered as a multiset: If the distance  $l$  occurs  $s$  times on a line, it may happen that up to  $\min\{s, k\}$  triangles counted in  $F_P(k; l)$  define the same circle. Hence, we cannot use the incidence bound in a straightforward manner.

Let  $\mathcal{C}'(k)$  denote the set of circles obtained from the multiset  $\mathcal{C}(k)$  disregarding multiplicities. For each  $y > 0$ , we want to bound the number of circles of  $\mathcal{C}'(k)$  that have multiplicity at least  $y$  in  $\mathcal{C}(k)$ . Suppose a circle  $C \in \mathcal{C}'(k)$  has multiplicity at least  $y$  in  $\mathcal{C}(k)$ . Then the line  $L_C$  perpendicular to the plane containing  $C$  and going through the center of  $C$  contains at least  $y$  occurrences of the distance  $l$  between points of  $P$  and hence must contain at least  $y + 1$  points of  $P$ . We apply Lemma 3.1 to  $P$  and the lines  $L_C$  for all  $C \in \mathcal{C}'(k)$  with multiplicity at least  $y$  in  $\mathcal{C}(k)$ : the total number of incidences between  $P$  and those lines is  $O(n + n^2y^{-2})$ . Each occurrence of  $l$  contributes to at most  $k$  circles in  $\mathcal{C}(k)$ . Hence, the number of circles in  $\mathcal{C}'(k)$  whose multiplicities are at least  $y$  in  $\mathcal{C}(k)$  is  $O(kny^{-1} + kn^2y^{-3})$ .

Let  $\mathcal{C}'(k, y)$  denote the subset of  $\mathcal{C}'(k)$ , each of whose circles has a multiplicity between  $y$  and  $2y$  in  $\mathcal{C}(k)$ ; let  $\mathcal{C}(k, y)$  denote the corresponding subfamily of  $\mathcal{C}(k)$ . Since the multiplicity is at most  $2k$ , we have  $y \leq 2k$ . From the above analysis we have  $|\mathcal{C}'(k, y)| = O(kny^{-1} + kn^2y^{-3})$ . On the other hand, the number of circles in  $\mathcal{C}(k)$  is  $O(kn^\alpha)$  so we also have that  $|\mathcal{C}'(k, y)| = O(kn^\alpha y^{-1})$ . The terms  $kny^{-1}$  and  $kn^2y^{-3}$  balance when  $y = n^{1/2}$ ; the terms  $kn^\alpha y^{-1}$  and  $kn^2y^{-3}$  when  $y = n^{(2-\alpha)/2}$ . For  $y \leq n^{(2-\alpha)/2}$  (case 1), we use the second bound  $O(kn^\alpha y^{-1})$ . For  $n^{(2-\alpha)/2} \leq y \leq n^{1/2}$  (case 2), which is a nonempty case only if  $\alpha \geq 1$ , we use the first bound which is simplified to  $O(kn^2y^{-3})$ . Finally, for  $y \geq \max\{n^{1/2}, n^{(2-\alpha)/2}\}$  (case 3), we use the first bound which is simplified to  $O(kny^{-1})$ . We analyze the number of incidences between  $P$  and  $\mathcal{C}(k, y)$  according to these cases. This is done by bounding the number of incidences between  $P$  and  $\mathcal{C}'(k, y)$ , using Lemma 4.2, and multiplying the result by  $2y$ . If we denote  $|\mathcal{C}'(k, y)|$  by  $N$ , Lemma 4.2 gives the bound of  $O(N + n + N^{4/5}n^{3/5})$  on the number of incidences between  $P$  and  $\mathcal{C}'(k, y)$ . The term  $N^{4/5}n^{3/5}$  dominates  $N$  as long as  $N < n^3$ , which always holds in the following cases under our assumption that  $k < n$ , so we ignore the term  $N$  in this bound.

*Case 1:*  $y \leq n^{(2-\alpha)/2}$  and we have  $N = O(kn^\alpha y^{-1})$ . Thus, the number of incidences between  $P$  and  $\mathcal{C}(k, y)$  is  $O(y(N^{4/5}n^{3/5} + n)) = O(k^{4/5}y^{1/5}n^{0.6+0.8\alpha} + ny)$ .

*Case 2:*  $n^{(2-\alpha)/2} \leq y \leq n^{1/2}$  and we have  $N = O(kn^2y^{-3})$ . The number of incidences between  $P$  and  $\mathcal{C}(k, y)$  is  $O(y(N^{4/5}n^{3/5} + n)) = O(k^{4/5}y^{-7/5}n^{11/5} + ny)$ .

*Case 3:*  $y \geq \max\{n^{1/2}, n^{(2-\alpha)/2}\}$  and we have  $N = O(kny^{-1})$ . The number of incidences between  $P$  and  $\mathcal{C}(k, y)$  is  $O(y(N^{4/5}n^{3/5} + n)) = O(k^{4/5}y^{1/5}n^{7/5} + ny)$ .

Our bound on  $F_P(k; l)$  is obtained by summing up these bounds for  $y = 2^0, 2^1, 2^2, \dots, 2k$ . The common term  $ny$  sums to  $O(kn)$  since it forms a geometric series with the maximum term  $2kn$ . The term  $k^{4/5}y^{1/5}n^{0.6+0.8\alpha}$  in case 1 also forms a geometric series with the maximum term  $k^{4/5}(n^{(2-\alpha)/2})^{1/5}n^{0.8\alpha+0.6} = k^{4/5}n^{0.8+0.7\alpha}$  and thus sums up to  $O(k^{4/5}n^{0.8+0.7\alpha})$ . Similarly, the term  $k^{4/5}y^{-7/5}n^{11/5}$  in case 2 takes its maximum at  $y = n^{(2-\alpha)/2}$  and thus sums up to  $O(k^{4/5}(n^{(2-\alpha)/2})^{-7/5}n^{11/5}) = O(k^{4/5}n^{0.8+0.7\alpha})$ . Finally the term  $k^{4/5}y^{1/5}n^{7/5}$  in case 3 takes its maximum at  $y = 2k$  and thus sums up to  $O(kn^{7/5})$ . Adding all these four terms, we obtain the bound in the first part of the lemma. □

**Corollary 4.1.**  $H_P(\Delta_k^l)$  is  $O(n^{7/5} + k^{-1/5}n^{0.8+0.7\alpha})$ . Also,  $H_P(\Delta_k^l) = O(k^{-1}n^{1+\alpha})$ .

**Theorem 4.1.**  $\lambda^{(3,2)}(P, Q) = O'(\min\{m^3n^{1.8}, m^{2.68}n^{1.95} + m^{2.8}n^{1.8875}\})$ .

*Proof.* For convenience sake, we omit  $O'()$  for the complexities and all equalities and inequalities in this proof are given in an asymptotic sense ignoring factors of polynomials of  $\log^* n$ . The  $m^3n^{1.8}$  bound has already been shown. Hence it suffices to give the other bound.

Let  $\sum_{\Delta \in \mathcal{T}_l(P) \cap \mathcal{T}_l(Q)} H_P(\Delta)H_Q(\Delta)$  be the inner product of triangle sequences  $\mathcal{T}_l(P)$  and  $\mathcal{T}_l(Q)$ . This inner product is bounded by  $S(P, Q, l) = \sum_i H_P(\Delta_i^l)H_Q(\nabla_i^l)$ , where Let  $\{\nabla_1^l, \dots, \nabla_{k_l}^l\}$  is the set of distinct triangles of  $\mathcal{T}_l(Q)$ , listed in the nonincreasing order of their multiplicities.

We want to estimate

$$S(P, Q) = \sum_l S(P, Q; l) = \sum_l \sum_i H_P(\Delta_i^l)H_Q(\nabla_i^l) \tag{2}$$

to bound  $\lambda^{(3,2)}(P, Q)$ .

Consider a length  $l$  such that  $2^s < L_P(l) \leq 2^{s+1}$  and  $2^t < L_Q(l) \leq 2^{t+1}$ . For convenience sake, we denote  $2^s = n^\alpha$  and  $2^t = m^\beta$ . Of course, both  $\alpha$  and  $\beta$  are at most 1.5. Let  $G(s, t)$  and  $g(s, t)$  be the set and number of different lengths  $l$  satisfying the above condition, respectively. Naturally,  $S(P, Q) = \sum_{s,t} \sum_{l \in G(s,t)} S(P, Q, l)$ , where  $s$  and  $t$  run up to  $1.5 \log m$  and  $1.5 \log n$ , respectively. We also denote the union  $\bigcup_{1 \leq t \leq \lfloor 1.5 \log m \rfloor} G(s, t)$  by  $G(s, *)$ , and let  $g(s, *)$  be its cardinality.

We examine expression (2). The terms where  $H_P(\Delta_i^l) \leq n^{7/5}$  in (2) contributes at most  $m^3n^{7/5} < m^{2.8}n^{1.8875}$ , and hence can be ignored. Next, we consider the terms where  $H_Q(\nabla_i^l) \leq m^{7/5}$  in the expression. For  $z \leq m^{7/5}$ , we consider the triangles in  $\mathcal{T}_l(Q)$  which occur in  $Q$  at most  $2z$  but at least  $z$  times. Let  $\varphi(l; z)$  be the number of different such triangles. From the  $k^{0.8}n^{0.8+0.7\alpha}$  bound in Lemma 4.3, contribution of such triangles to  $\sum_{l \in G(s,*)} S(P, Q, l)$  is bounded by  $\sum_{l \in G(s,*)} z\varphi(l; z)^{0.8}n^{0.8+0.7\alpha}$ . Since  $\sum_l \varphi(l; z) < m^3/z$ , this is maximized when  $\varphi(l; z) = (m^3/z)/g(s, *)$  for all  $l \in G(s, *)$ . The maximum value is  $g(s, *)^{0.2}z^{0.2}m^{2.4}n^{0.8+0.7\alpha}$ . Since  $g(s, *) \leq n^{2-\alpha}$  and  $z \leq m^{7/5}$ , the value is  $m^{2.68}n^{1.2+0.5\alpha}$ . Hence the contribution to  $S(P, Q)$  is  $\sum_{s=1}^{1.5 \log n} m^{2.68}n^{1.2}2^{0.5s}$ , which is  $m^{2.68}n^{1.2+0.75} = m^{2.68}n^{1.95}$ .

Therefore, from now on, we remove all indices  $i$  where either  $H_Q(\nabla_i^l) \leq m^{7/5}$  or  $H_P(\Delta_i^l) \leq n^{7/5}$  from expression (2), and hence we can assume that  $H_P(\Delta_k^l) \leq k^{-1/5}n^{0.8+0.7\alpha}$  and  $H_Q(\nabla_k^l) \leq k^{-1/5}m^{0.8+0.7\beta}$  for any  $k$ . We want to estimate  $g(s, t)S(P, Q, l)$  for the distance  $l = l(s, t) \in G(s, t)$  maximizing  $S(P, Q, l)$ . Comparing the two bounds in Corollary 4.1,  $k^{-1}m^{1+\beta} < k^{-1/5}m^{0.8+0.7\beta}$  if  $k > m^{0.25+0.375\beta}$ . Let  $y = \min\{n^{0.25+0.375\alpha}, m^{0.25+0.375\beta}\}$ .

If  $y = m^{0.25+0.375\beta}$ ,

$$S(P, Q, l) \leq \sum_{k=1}^y k^{-2/5}m^{0.8+0.7\beta}n^{0.8+0.7\alpha} + \sum_{k=y}^{m^{1+\beta}} k^{-6/5}n^{0.8+0.7\alpha}m^{1+\beta}.$$

Hence,  $S(P, Q, l) \leq y^{0.6}n^{0.8+0.7\alpha}m^{0.8+0.7\beta} = n^{0.8+0.7\alpha}m^{0.95+0.925\beta}$ .

The same bound can be similarly obtained if  $y = n^{0.25+0.375\alpha}$ . Since  $g(s, t) \leq \min\{n^{2-\alpha}, m^{2-\beta}\} = \min\{n^2 2^{-s}, m^2 2^{-t}\}$ ,  $g(s, t) \leq n^{x(2-\alpha)} m^{(1-x)(2-\beta)}$  for any  $0 \leq x \leq 1$ . Thus, setting  $x = 0.075$ ,  $g(s, t)S(P, Q, l(s, t)) \leq n^{0.8+0.15+0.625\alpha} m^{0.95+1.85} = n^{0.95+0.625\alpha} m^{2.8} = n^{0.95} 2^{0.625s} m^{2.8}$ , and equality (in an asymptotic sense) holds only if  $n^2 2^{-s} = m^2 2^{-t}$ . This attains the maximum value  $n^{1.8875} m^{2.8}$  when  $s$  takes the maximum value  $\lfloor 1.5 \log n \rfloor$ .

More precisely, if  $n^2 2^{-s} > m^2 2^{-t}$ ,

$$g(s, t)S(P, Q, l(s, t)) \leq n^{0.95} 2^{0.625s} m^{2.8} (n^2 2^{-s} / m^2 2^{-t})^{-0.025},$$

and otherwise  $g(s, t)S(P, Q, l(s, t)) \leq n^{0.95} 2^{0.625s} m^{2.8} (n^2 2^{-s} / m^2 2^{-t})^{0.975}$ . Hence,  $g(s, t)S(P, Q, l(s, t))$  behaves as a geometric sequence with respect to  $s$  and  $t$  for each of above cases, and the peak is when  $s$  is  $\lfloor 1.5 \log n \rfloor$  and  $n^2/2^s = m^2/2^t$  holds.

Hence,  $\sum_{s=1}^{\lfloor 1.5 \log n \rfloor} \sum_{t=1}^{\lfloor 1.5 \log m \rfloor} \sum_{l \in G(s, t)} S(P, Q, l) = n^{1.8875} m^{2.8}$ . Combining the obtained two upper bounds, we have the theorem.  $\square$

**Corollary 4.2.**  $\lambda^{(3,2)}(n, n) = O'(n^{4.6875})$ .

**Corollary 4.3.** A set of  $n$  points of space defines at least  $n^{1.3125} / (\log^* n)^{O(1)}$  distinct triangles under congruency.

*Proof.* This comes from the fact that  $(n^{3-1.3125})^2 n^{1.3125} = n^{4.6875}$ .  $\square$

**Theorem 4.2.** The LSP can be solved in  $O((\min\{n^{7/3} m^{8/3} \log^2 n, n^{2.2} m^3\} + n^3) \log n)$  time.

*Proof.* We compute all  $O(n^3)$  (resp.  $O(m^3)$ ) triangles in  $P$  (resp.  $Q$ ), and classify them under similarity in  $O(n^3 \log n)$  time (resp.  $O(m^3 \log m)$  time). For each triangle  $\Delta$ , we search similar triangles in  $P$  and  $Q$  by using the above classification, and give one vote to the corresponding similar transformation for each pair of triangles, one in  $P$  and the other in  $Q$ . Each vote can be done in  $O(\log n)$  time, and the transformation gathering the largest number of votes gives the solution of LSP. The number of votes is bounded by  $\sum_{\Delta} H_P^{\text{sim}}(\Delta) H_Q^{\text{sim}}(\Delta)$ , where  $H_P^{\text{sim}}(\Delta)$  is the number of similar triangles to  $\Delta$  in  $P$ . It is  $O(n^{2.2} m^3)$  by Proposition 4.1.

Moreover, abusing the notations for the congruent case, let  $F_P(k)$  be the upper bound for the number of triangles in  $P$  similar to one of the elements of a set  $X$  of given  $k$  triangles. We will show below that  $F_P(k) = O(\min\{n^3, k^{2/3} n^{7/3} \log n\})$ . Without loss of generality, we assume that  $F_P(k) - F_P(k-1) \geq F_P(k+1) - F_P(k)$ , since we can consider the  $k$  triangles maximizing  $F_P(k)$ .

Thus,  $\sum_{\Delta} H_P^{\text{sim}}(\Delta) H_Q^{\text{sim}}(\Delta)$  is bounded by  $\sum_k (F_P(k) - F_P(k-1))(F_Q(k) - F_Q(k-1))$ . Here, we can assume that  $F_Q(k)$  so that  $F_Q(m) = m^3$  and  $F_Q(k) - F_Q(k-1) = 0$  for  $k > m$  because of Lemma 3.4. Thus,  $\sum_k (F_P(k) - F_P(k-1))(F_Q(k) - F_Q(k-1))$  is asymptotically bounded by  $\sum_{k=1}^m (k^{-1/3} m^{7/3} \log m)(k^{-1/3} n^{7/3} \log n)$ , which is  $O(n^{7/3} m^{8/3} \log n \log m)$ . Hence, LSP can be solved in  $O((\min\{n^{7/3} m^{8/3} \log^2 n, n^{2.2} m^3\} + n^3) \log n)$  time.



Now, it suffices to prove  $F_P(k) = O(k^{2/3}n^{7/3} \log n)$  for  $k \leq n$ . There are  $n^2$  possible edges, and for each distance there are  $k$  circles, each corresponding to a triangle in  $X$ . A circle  $C$  can appear for  $\min\{k, y(y+1)/2\}$  different triangles if the line perpendicular to  $C$  through the center contains  $y$  points of  $P$ .

We consider the incidence between a multiset of  $O(kn^2)$  circles and a set of  $O(n)$  points. The number of lines containing between  $y$  and  $2y$  points of  $P$  is  $\max\{ny^{-1}, n^2y^{-3}\}$ . If  $y^2 < k \leq n$ , the number of triangles created by edges on these lines is  $O(y^2(kn^2y^{-3})^{2/3}n)$ , which is  $O(k^{2/3}n^{7/3})$ . If  $y^2 > k$ , and  $y \leq n^{0.5}$ , it is  $O(k(kn^2y^2)^{2/3}n)$ , which is  $O(k^{2/3}n^{7/3})$ . If  $y^2 > k$  and  $y > n^{0.5}$ , for each line  $L$  containing  $y$  points, each point  $v$  of  $P$  defines at most one triangle similar to a given triangle  $\Delta$  which has its longest edge on  $L$  and the opposite vertex at  $v$ . Hence, the line defines at most  $n$  similar triangles to each of  $k$  given triangles. Thus, there are  $kn$  triangles associated with  $L$ , and there are  $ny^{-1}$  such lines. Hence, we have an  $O(ky^{-1}n^2)$  bound, which is  $O(k^{1/2}n^2) = o(k^{2/3}n^{7/3})$ . Since we have cut the interval of  $y$  into  $\log n$  subintervals, the total complexity is  $O(k^{2/3}n^{7/3} \log n)$ .  $\square$

## 5. The Four-Dimensional Case

### 5.1. Number of Congruent Triangles

In this subsection we prove the following theorem:

**Theorem 5.1.** *In the given set  $P$  of  $n$  points in four-dimensional space, the multiplicity of any given triangle is  $O(n^{65/23}) = O(n^{2.83})$ .*

In the following, all geometric objects are in four-dimensional space unless otherwise stated. We consider triangles congruent to a given triangle  $\Delta = ABZ$ , whose longest edge  $AB$  has a length of 1, and whose height  $ZX$  is  $h$ . We call  $AB$  the *base edge* of  $\Delta$ . We need the following basic fact:

**Proposition 5.1.** *The intersection of four mutually distinct three-dimensional unit spheres is either a circle or a set of no more than two points.*

We call two circles  $C_1$  and  $C_2$  a *distance-circle pair* if the distance  $\text{dist}(\mathbf{u}, \mathbf{v}) = 1$  for any  $\mathbf{u} \in C_1$  and  $\mathbf{v} \in C_2$ . Note that for any circle  $C_1$  of radius  $r < 1$ , there is a unique circle  $C_2$  such that  $C_1$  and  $C_2$  form a distance-circle pair. To show the existence, choose a coordinate system so that  $C_1 = \{(x, y, 0, 0) \mid x^2 + y^2 = r^2\}$  and set  $C_2 = \{(0, 0, z, w) \mid z^2 + w^2 = 1 - r^2\}$ . For uniqueness, take an arbitrary four points from  $C_1$  and observe that  $C_2$  must lie in the intersection of the four three-dimensional unit spheres centered at these four points. Therefore, we denote the circle forming the distance pair with a circle  $C$  by  $D(C)$ . The existence of a distance-circle pair kills the hope of getting a nontrivial asymptotic bound on the multiplicity of a distance in four dimensions, because one may place  $n/2$  points on each circle of a distance-circle pair to get  $n^2/4$  occurrences of the unit distance.

Given a point set  $P$  of  $n$  points, we say that a circle  $C$  is *heavy* if it contains more than  $n^\alpha$  points of  $P$ , where  $\alpha = \frac{7}{23}$ . Let  $N$  denote the number of heavy circles.

**Lemma 5.1.**  $N = O(n^{3-5\alpha})$ .

*Proof.* By Lemma 4.2 on the number of circle-point incidences, we have  $Nn^\alpha = O(N + n + N^{4/5}n^{3/5})$ . Hence,  $N = O(n^{3-5\alpha})$ , since  $\alpha < 0.5$ . □

We construct a bipartite graph  $G = (U, V, E)$ , so that each of  $U$  and  $V$  corresponds to the point set  $P$ , and  $(u, v) \in E$  if and only if  $dist(u, v) = 1$ . An edge  $(u, v)$  is called *special* if  $u$  (or  $v$ ) is in a heavy circle  $C$ ,  $v \in D(C)$  (or  $u \in D(C)$ ), and  $D(C)$  has at least four points of  $P$ . Let  $E_0$  be the set of special edges, and  $E_1 = E - E_0$ .

**Lemma 5.2.**  $|E_1| = O(n^{(7+\alpha)/4})$ .

*Proof.* First we show that the bipartite graph  $(U, V, E_1)$  has no  $K_{n^\alpha, 4}$  as a subgraph (with the  $n^\alpha$  points on  $U$ 's side and the four points on  $V$ 's side). Suppose otherwise, namely, that  $(U, V, E_1)$  has a subgraph isomorphic to  $K_{n^\alpha, 4}$ . Then the  $n^\alpha$  points of  $U$  in this subgraph must lie in the intersection of four unit three-dimensional spheres centered at the four points  $v_1, v_2, v_3, v_4$  of  $V$  in this subgraph. By Proposition 5.1, these  $n^\alpha$  points must lie on a circle  $C$ , making the circle heavy. Moreover, the four points  $v_1, v_2, v_3, v_4$  must lie in the intersection of  $n^\alpha$  spheres, and hence must lie on  $D(C)$ . This contradicts the definition of  $E_1$ . Hence,  $(U, V, E_1)$  has no  $K_{n^\alpha, 4}$  as a subgraph, and, from the Bipartite Graph Lemma, we have

$$|U| \binom{|E_1|/|U|}{4} < n^\alpha \binom{|V|}{4}.$$

Since  $|U| = |V| = n$ , we have  $|E_1| = O(n^{(7+\alpha)/4})$ . □

**Corollary 5.1.** *The number of triangles congruent to  $\Delta$  whose base edge is in  $E_1$  is  $O(n^{(11+\alpha)/4})$ .*

*Proof.* Each edge defines at most  $n$  triangles in  $P$ ; Hence the corollary holds. □

We next consider triangles whose base edge is located at edges in  $E_0$ . We fix a distance-circle pair  $(C, D(C))$  in which  $C$  is heavy, and consider the embedding of  $\Delta$  so that its base edge is mapped to an edge between  $C_1$  and  $D(C)$ . We assume each edge is directed from  $C$  to  $D(C)$ , so that the copy of the vertex  $A$  is in  $C$ , and the copy of  $B$  is in  $D(C)$ . Recall that  $X$  is the point on the line  $AB$  such that the line segment  $ZX$  defines the height  $h$  of  $\Delta$ , and let  $X(e)$  denote the image of  $X$  when  $AB$  is mapped to a directed edge  $e$  of unit length. Then the locus of the opposite vertex  $Z$  of  $\Delta$  when we transform  $AB$  into  $e$  forms a two-dimensional sphere  $S^{(2)}(e)$  with radius  $h$ , centered at  $X(e)$ . Moreover,  $e$  is the normal vector of the three-dimensional space containing the two-dimensional sphere. We consider the family of spheres  $\mathcal{F}(C, D(C)) = \{S^{(2)}(\mathbf{x}\mathbf{y}) \mid \mathbf{x} \in C, \mathbf{y} \in D(C)\}$ .

**Lemma 5.3.** *Consider the family of two-dimensional spheres with radius  $h$  containing a fixed circle  $C$ . The centers of these spheres are located on a circle  $C'$ . Moreover, the hyperplane containing each sphere has its normal vector that is tangent to  $C'$  at the center of the sphere.*

*Proof.* Let  $\mathbf{p}$  be the center and  $r$  the radius of  $C$ . Let  $L$  be the two-dimensional plane determined by  $C$  and let  $L'$  be the two-dimensional plane containing  $\mathbf{p}$  and orthogonal to  $L$ . Let  $S$  be any sphere belonging to the family in the lemma, and let  $\mathbf{q}$  be the center of  $S$ . Then  $\mathbf{q}$  is on  $L'$  and at a distance  $\sqrt{h^2 - r^2}$  from  $\mathbf{p}$ ; that is to say, it is on a circle  $C'$  of radius  $\sqrt{h^2 - r^2}$  centered at  $\mathbf{p}$  and contained in  $L'$ . The three-dimensional space containing  $S$  is spanned by the line  $\mathbf{pq}$  and  $L$ ; and hence is perpendicular to the tangent of  $C'$  at  $\mathbf{q}$ . □

**Lemma 5.4.** *For any circle  $C_0$ , there are at most four spheres in  $\mathcal{F}(C, D(C))$  that contain  $C_0$ .*

*Proof.* Let  $F(C_0)$  denote the set of spheres of  $\mathcal{F}(C, D(C))$  that contain  $C_0$  and assume that  $F(C_0)$  is nonempty. By the above lemma, there is some circle  $C'$  such that each sphere of  $F(C_0)$  has its center on  $C'$  and has its normal tangent to  $C'$ . Let  $e = \mathbf{xy}$ ,  $\mathbf{x} \in C$ ,  $\mathbf{y} \in D(C)$ , be an edge such that  $S^{(2)}(e) \in F(C_0)$ . Then, from the observation preceding the above lemma, the normal of the space containing  $S^{(2)}(e)$  is  $e$  and the center of  $S^{(2)}(e)$  is  $X(e)$ . Therefore,  $e$  must be tangent to  $C'$  and hence contained in the two-dimensional plane  $L$  containing  $C'$ .  $L$  cannot contain  $C$ , because if it does then it does not intersect  $D(C)$  and hence does not contain any edge between  $C$  and  $D(C)$ . Similarly,  $L$  does not contain  $D(C)$ . Therefore,  $L$  intersects each of  $C$  and  $D(C)$  at no more than two points and hence there are at most four edges between  $C$  and  $D(C)$  contained in  $L$ . Therefore,  $S^{(2)}(e) \in F(C_0)$  for at most four edges  $e$ . □

We are now ready to bound the number of triangles in  $P$  congruent to  $\Delta$  that have their base edge in  $E_0$ . We consider the bipartite graph  $(E_0, P, T)$ , which has vertex set  $E_0$  and  $P$ , and connect  $e \in E_0$  and  $\mathbf{p} \in P$  by an edge of  $T$  if and only if  $e$  and  $\mathbf{p}$  form a triangle congruent to  $\Delta$ . Each edge  $e \in E_0$  defines a sphere corresponding to the locus of the opposite vertex of  $\Delta$ , and hence  $\mathbf{p}$  must lie on this sphere. Each  $e \in E_0$  has one end on a heavy circle  $C$  and the other end on the circle  $D(C)$ .

Consider any  $4N + 1$  elements of  $E_0$  and its corresponding spheres. These spheres must be in  $\bigcup \mathcal{F}(C, D(C))$ , where union is taken over all heavy circles  $C$ . Since there are only  $N$  heavy circles, there is a distance-circle pair  $(C, D(C))$  such that at least five of those spheres are in  $\mathcal{F}(C, D(C))$ . By Lemma 5.4, the intersection of these five spheres cannot contain a circle and hence is a set of at most two points. Thus, our bipartite graph does not contain  $K_{4N+1,3}$  as a subgraph. Since  $|E_0| < n^2$  and  $|P| = n$ , from the Bipartite Graph Lemma, we have

$$n^2 \binom{\lfloor |T|/n^2 \rfloor}{3} \leq 4N \binom{n}{3},$$

and hence  $|T| = O(N^{1/3}n^{7/3})$ . Since  $N = O(n^{3-5\alpha})$ ,  $|T| = O(n^{10/3-5\alpha/3})$ .

Putting this together with Corollary 5.1, the number of triangles congruent to  $\Delta$  in  $P$  is  $O(n^{(11+\alpha)/4} + n^{10/3-5\alpha/3})$ . We set  $\alpha = \frac{7}{23}$  to obtain the bound of  $O(n^{65/23})$ .

### 5.2. LCP in Four-Dimensional Space

We consider three-dimensional simplices in four-dimensional space.

**Theorem 5.2.** *The maximum number of three-dimensional simplices of  $n$  points in four-dimensional space congruent to a given three-dimensional simplex  $\Gamma$  is  $O(n^{2.87})$ .*

*Proof.* If we fix a location of a triangle face  $\Delta$  of  $\Gamma$ , the locus of the opposite vertex  $v$  of  $\Gamma$  forms a circle. The number of triangles congruent to  $\Delta$  in  $P$  is  $O(n^\gamma)$  for  $\gamma = \frac{65}{23}$ . Thus, we have a set  $C(\Gamma)$  of  $n^\gamma$  circles corresponding to the possible locations of the three-dimensional simplices.

We would like to use the circle-point incidence to bound the number of three-dimensional simplices. Unfortunately, different locations of  $\Delta$  may create the same circle  $C$ , and therefore we cannot use Lemma 4.2 directly.

Suppose that  $s$  locations of  $\Delta$  can create  $C$ . These triangles must lie on the plane that contains the center  $\mathbf{o}$  of  $C$  and is perpendicular to the plane containing  $C$ . Moreover, the set of vertices of the triangles must lie on (at most) three circles  $Circ_1(C)$ ,  $Circ_2(C)$ , and  $Circ_3(C)$  centered at  $\mathbf{o}$ .  $Circ_1(C)$  contains  $s$  vertices, and  $Circ_1(C)$  cannot coincide with  $Circ_1(C')$  if  $C \neq C'$ : Consider the case in which  $Circ_1(C)$  has its center at the origin and is contained in the  $x$ - $y$  plane. If the height of  $\Gamma$  with respect to the face  $\Delta$  is  $h$ ,  $C$  must be the unique circle defined by  $x = y = 0, z^2 + t^2 = h^2$ .

We consider the subset  $\mathcal{C}(s)$  of  $C(\Gamma)$  in which each member has multiplicities equal to or greater than  $s$  but less than  $2s$ . Then, for each  $C \in \mathcal{C}(s)$ ,  $Circ_1(C)$  contains at least  $s$  points of  $P$ . From Lemma 4.2, there are at most  $O'(ns^{-1} + n^3s^{-5})$  circles, each containing more than  $s$  points. Hence,  $\mathcal{C}(s)$  contains  $O'(\min\{ns^{-1} + n^3s^{-5}, n^\gamma s^{-1}\})$  circles.

If  $s > n^{0.5}$ , the incidence between  $P$  and  $\mathcal{C}(s)$  is  $O(s(ns^{-1})^{0.8}n^{0.6}) = O(s^{0.2}n^{1.4})$ , which is small even if  $s = n$ . If  $s > n^{1/23}, n^3s^{-5} < n^\gamma s^{-1}$ , and the incidence between  $P$  and  $\mathcal{C}(s)$  is  $O(s(n^3s^{-5})^{0.8}n^{0.6})$ , which is  $O(n^3s^{-3})$ . Hence, there are  $O(n^{66/23})$  incidences created by circles whose multiplicities are at least  $n^{1/23}$ .

If  $s < n^{1/23}$ , the incidence between  $P$  and  $\mathcal{C}(s)$  is the incidence between  $n$  points and  $O(n^\gamma s^{-1})$  circles each of which has multiplicity  $s$ ; this is  $O(s^{0.2}n^{0.8\gamma+0.6})$ . Hence, there are  $O(n^{66/23})$  incidences created by circles whose multiplicities are less than  $n^{1/23}$ . Hence, the number of 3-dimensional simplices is  $O(n^{66/23}) = O(n^{2.87})$ .  $\square$

**Proposition 5.2.** *The maximum number of three-dimensional simplices of  $n$  points in four-dimensional space similar to a given three-dimensional simplex  $\Gamma$  is  $O(n^3)$ .*

*Proof.* If we fix the location of a similar copy of the triangle face  $\Delta$  of  $\Gamma$ , the locus of the opposite vertex  $v$  of the similar copy of  $\Gamma$  forms a circle. The number of triangles congruent to  $\Delta$  in  $P$  is  $O(n^3)$ . Thus, we have a set  $C(\Gamma)$  of  $n^3$  circles corre-

sponding to the possible locations of the similar copies of three-dimensional simplices. The rest is almost the same as the proof of Theorem 5.2. Abusing the notations,  $\mathcal{C}(s)$  contains  $O(ns^{-1} + n^3s^{-5})$  circles, and the incidence between  $P$  and  $\mathcal{C}(s)$  is  $O(s(ns^{-1} + n^3s^{-5})^{0.8}n^{0.6}) = O(s^{0.2}n^{1.4} + n^3s^{-3})$ . By summing this over  $s = 1, 2, \dots, n$ , we have the  $O(n^3)$  bound. □

**Theorem 5.3.** *The LCP in four-dimensional space can be solved in  $O((n^{2.87}m^4 + n^4) \log n)$  time. The LSP in four-dimensional space can be solved in  $O((n^3m^4 + n^4) \log n)$  time. The CCD can be solved in  $O(n^{3.83} + n^{2.87}m \log n)$  time.*

*Proof.* From the bounds of triangles and three-dimensional simplices,  $\lambda^{(4,3)}(n, m) = n^{2.87}m^3$  and  $\lambda^{(4,2)}(n, m) = n^{2.83}m^2$ . Hence, from Theorem 2.6, we have the bound for the LCP.

Analysis for the LSP is analogous. For the CCD, we choose a three-dimensional simplex  $\Gamma$  and its face  $\Delta$  from  $Q$ , and find all possible  $O(n^{2.83})$  locations of  $\Delta$  in  $P$  within  $O(n^3)$  time. Next, we find all possible  $O(n^{2.87})$  locations of  $\Gamma$  in  $O(n^{3.83})$  time from all three-dimensional simplices containing copies of  $\Delta$  in  $P$ . Then we check within  $O(m \log n)$  time whether each copy of  $\Gamma$  is associated with a congruent embedding of  $Q$  in  $P$ . □

## 6. The Higher-Dimensional Case

**Lemma 6.1.** *In  $d$ -dimensional space,  $d \geq 5$ , the number of  $(d - 1)$ -dimensional simplices in  $P$  similar to a given  $(d - 1)$ -dimensional simplex is  $O(n^{d-1})$ .*

*Proof.* For  $d = 5$ , we need special care, as we show in Appendix B. For  $d \geq 6$ , given a  $(d - 1)$ -dimensional simplex  $\Gamma$ , suppose that we fix  $(d - 1)$  points of  $P$  and the  $(d - 2)$ -dimensional simplex  $\Lambda$  spanned by these points, which is a facet of a copy (under similarity) of  $\Gamma$ . Then the locus of the remaining vertex  $u$  of a copy of  $\Gamma$  forms a circle.

More than one location of  $\Lambda$  can have the same circle  $C$  as the locus of  $u$ . Suppose there are  $K$  locations of  $\Lambda$  creating  $C$ . Let  $v_1, \dots, v_{d-1}$  be the set of vertices of  $\Lambda$ . The copies of  $v_i$  in the  $K$  locations of  $\Lambda$  must lie on a  $(d - 3)$ -dimensional sphere  $C_i$  which is centered at the center  $o$  of  $C$  and contained in the  $(d - 2)$ -dimensional plane perpendicular to  $C$ .

We project the points to  $C_1$  using the central projection with the center  $o$ . Let  $F$  be a face of  $\Lambda$ . Then copies of  $F$  in the  $K$  locations are projected to the  $(d - 3)$ -dimensional simplices which are congruent to each other.

Thus,  $K$  must be smaller than the number of congruent  $(d - 3)$ -dimensional simplices in the  $(d - 3)$ -dimensional sphere. This is bounded by the number of congruent  $(d - 4)$ -dimensional simplices, since there are at most two  $(d - 3)$ -dimensional congruent simplices sharing a  $(d - 4)$ -dimensional simplex in the sphere. The same  $O(n^{d-4})$  bound

as in the case of  $(d - 3)$ -dimensional space can be applied to the sphere, which holds if  $d - 3 = 3, 4$ , as we have seen, and inductively for  $d - 3 \geq 4$ .

Thus, we have a set of circles with multiplicities, so that the maximum multiplicity is  $O(n^{d-4})$  and the total sum of the multiplicities is  $O(n^{d-1})$ . The worst scenario for us is that we have a set of  $O(n^3)$  circles, each of which has multiplicity  $O(n^{d-4})$ . From Lemma 4.2, the incidence of a set of  $O(n^3)$  circles and  $n$  points is  $O(n^3)$ . Hence, the total number of incidences is  $O(n^{d-1})$ .  $\square$

**Theorem 6.1.** *The LSP and LCP in  $d$ -dimensional space can be solved in  $O((n^{d-1}m^d + n^d) \log n)$  time for  $d \geq 5$ . The SCD and CCD can be solved in  $O'(n^d + n^{d-1}m \log n)$  time for  $d \geq 5$ .*

*Proof.* It trivially follows from Lemma 6.1 that  $\lambda^{(d,d-1)}(n, m) = O(n^{d-1}m^d)$ . This and the trivial  $O(n^{d-1}m^{d-1})$  bound on  $\lambda^{(d,d-2)}(n, m)$  make the first two terms comparable in the running-time upper bound of Theorem 2.6. Therefore, the bound for LCP holds. Since Lemma 6.1 holds for similarity, we can give the same argument for LSP. For the SCD (and CCD), we pick a  $(d - 1)$ -dimensional simplex from  $Q$ , find all  $(O(n^{d-1}))$  copies of it in  $P$  within  $O(n^d)$  time, and check whether each associated transformation gives an embedding of  $Q$  in  $O(m \log n)$  time.  $\square$

Bounding the number of pairwise congruent (similar)  $k$ -dimensional simplices in a set of  $n$  points in  $d$ -dimensional space is a classical problem posed by Moser [21], Erdős and Purdy [13], and Elekes and Erdős [10]. A better solution to this problem for  $k = d - 1$  and  $k = d - 2$  will improve the bounds on  $\lambda^{(d,d-1)}(n, m)$  and  $\lambda^{(d,d-2)}(n, m)$ , and hence the bound on the analysis of the algorithms for the LCP (or LSP).

## Acknowledgments

We wish to thank Jiří Matoušek for communicating Székely's results [28], and also sending us a copy of his unpublished lecture notes [20]. We would also like to thank Janos Pach and Pankaj Agarwal for introducing us to several useful references.

## Appendix A. Linear Space Implementation

The following argument can be extended to the voting algorithms in any dimensions. However, for simplicity, we give the description for the planar case.

We first note that we process the local voting processes at the  $\mathbf{p}_i, \mathbf{q}_j$  in lexicographic order with respect to  $i$  and  $j$ . Hence, we keep the sorted list of  $\{\text{dist}(\mathbf{p}, \mathbf{u}) : \mathbf{u} \in P\}$  only during local voting processes at the  $\mathbf{p}, \mathbf{q}_j$  ( $j = 1, 2, \dots, m$ ) run.

The space complexity of the voting algorithm is naively  $O(nm)$ , because the size of the voting table may become  $O(nm)$ . We can improve it to  $O(n)$  by avoiding explicit construction of the voting table, as follows: In the local voting process, without loss of generality, we assume that  $\mathbf{p} = \mathbf{q} = \mathbf{O}$ . The transformation  $T$  corresponds to the rotation

angle  $\tau(T)$ , which equals  $\theta(\mathbf{u}) - \theta(\mathbf{v})$  if  $T$  maps  $\mathbf{u}$  to  $\mathbf{v}$ , where  $\theta(\mathbf{u})$  is the argument of the vector  $\mathbf{u}$ .

Let  $A'_l(P)$  be the family of lines consisting of  $y = x + \theta(\mathbf{u})$  for all  $\mathbf{u} \in P$  satisfying  $|\mathbf{u}| = l$ . Let  $A'_l(Q)$  be the family of lines consisting of  $y = \theta(\mathbf{v})$  for all  $\mathbf{v} \in Q$  satisfying  $|\mathbf{v}| = l$ . The total number (over all  $l$ ) of lines is at most  $n + m$ , and thus  $O(n)$ . For a real value  $x$ , let  $g_l(x)$  be the number of segment intersections of  $A_l = A'_l(P) \cup A'_l(Q)$  that have  $x$  as their  $x$ -coordinate value. Then the number of votes collected by a transformation  $T$  is  $\sum_l g_l(\tau(T))$ .

We process the segment intersection computation by using Bentley–Ottman’s sweep algorithm, with a slight modification that the algorithm shares a priority queue controlling the sweeping processes for all  $A(l)$ , so that it can find the next  $x$ -coordinate value of the intersection (more precisely, the lexicographically smallest intersection) among all arrangements.

Using  $O(n)$  space, the algorithm maintains the  $x$ -coordinate value (on the left of the sweep line) of the largest number of intersection points (summed over all  $l$ ). We remark that, in a real implementation, we need not consider segments explicitly.

### Appendix B. The Five-Dimensional Case for Lemma 6.1

We are given a 4-simplex  $\Gamma$ , its facet  $\Lambda$ , which is a 3-simplex, and a triangle  $\Delta$ , which is a facet of  $\Lambda$ . The vertex opposite to  $\Delta$  is  $\mathbf{v}$ , and the vertices of  $\Delta$  are  $\mathbf{w}(1)$ ,  $\mathbf{w}(2)$ ,  $\mathbf{w}(3)$ , and  $\mathbf{w}(4)$ . We assume that the vertical distance from  $\mathbf{v}$  to  $\Lambda$  is 1, and that the projection of  $\mathbf{v}$  to  $\Lambda$  is  $\mathbf{v}'$ . The distance from  $\mathbf{v}'$  to  $\mathbf{w}(s)$  is  $d(s)$  for  $s = 1, 2, 3, 4$ .

The set  $P$  of  $n$  points in  $\mathcal{R}^5$  creates less than  $n^4$  3-simplices. Thus, the number of possible locations of  $\Delta$  using points of  $P$  as vertices is less than  $n^4$ . For a fixed location  $\Lambda_0$  of  $\Lambda$ , the trajectory of the possible locations of  $\Delta$  is a unit circle  $C$  centered by  $\mathbf{v}'_0$  and perpendicular to the hyperplane (of codimension two)  $H$  containing  $\Lambda_0$ ; here,  $\mathbf{v}'_0$  is the copy of  $\mathbf{v}'$  in  $\Lambda_0$ .

Different locations of  $\Delta$  may define the same circle  $C$ . Suppose that each member of  $t$  copies  $\Delta_i$  ( $i = 1, \dots, t$ ) of  $\Delta$  defines the same circle  $C$ . Let  $\mathbf{w}(s)_i$  be the copy of  $\mathbf{w}(s)$  in  $\Delta_i$ . Then  $\mathbf{w}(s)_i$  must be contained in  $H$ , and, moreover, is in the sphere  $\mathbf{S}(C; s)$  with diameter  $d(s)$  centered at  $\mathbf{v}'_1$ .

Without loss of generality, we only consider the circles  $C$  such that the number of points of  $P$  located on  $\mathbf{S}(C; 1)$  is at least a quarter of that on  $\bigcup_{s=1,2,3,4} \mathbf{S}(C; s)$ , since otherwise we choose another index. Note that  $C$  is uniquely defined from  $\mathbf{S}(C; 1)$  so that it is the circle centered at the center of  $\mathbf{S}(C; 1)$  and perpendicular to the hyperplane  $H$ .

We project the points of  $\bigcup_{j=1,2,3,4} \mathbf{S}(C; j)$  onto  $\mathbf{S}(C; 1)$ . They must define at least  $t$  copies of the same triangle (say,  $\Delta'$ ), which is the projective image of the copies of  $\Delta$  onto  $\mathbf{S}(C; 1)$ . On the sphere  $\mathbf{S}(C; 1)$ , at most two possible locations of  $\Delta'$  are permitted if its shortest edge is fixed. From the distance bound on the plane, which holds for the point set on the sphere,  $t = O(m^{4/3})$  if  $\mathbf{S}(C; 1)$  contains  $m$  points of  $P$ . (Note: the only difference between the case of sphere and plane is that more than two diameter circles on the sphere intersect at two points (north and south poles.) This can be resolved by decomposing the sphere into three parts: the equator and upper and lower hemispheres).

For a fixed  $t$ , we consider a set  $C_1, C_2, \dots, C_k$  of different circles, each of which

is defined by at least  $t$  but at most  $2t$  different locations of  $\Lambda$ , and there are  $m_i$  points  $\mathbf{S}(C_i; 1)$ . Since  $t = O(m_i^{4/3})$ ,  $m_i = \Omega(t^{3/4})$ .

We can give an upper bound of  $k$  in terms of  $t$  and  $n$ .  $\mathbf{S}(C_i; 1)$  has  $\Omega(m_i)$  points. We say  $C_i$  is a special circle if all the points in  $\mathbf{S}(C_i; 1)$  are located on a circle.

These points define  $\Omega(m_i^4)$  three-dimensional simplices in the space; moreover, if  $C_i$  is not special,  $\Omega(m_i^2) = \Omega(t^{8/3})$ , three-dimensional simplices among them are nondegenerate. For the time being, we omit special circles. Since there are at most  $n^4$  three-dimensional simplices in  $P$ , and it cannot happen that a nondegenerate three-dimensional simplex lies on the intersection of  $\mathbf{S}(C_i; 1)$  and  $\mathbf{S}(C_j; 1)$  for  $i \neq j$ ,  $kt^{8/3} = O(n^4)$ . Hence,  $k = O(n^4/t^{3/2})$ .

The number of incidences between the circles (counting the multiplicities) and  $P$  is  $O(tk^{4/5}n^{3/5} + tk + tn)$ . Thus, it is  $O(n^{19/5}t^{-0.2} + n^4t^{-0.5} + tn)$ . By summing this term over  $t = 2^j$  for  $j = 1, 2, \dots, \lfloor \log n^{4/3} \rfloor$ , we obtain an  $O(n^4)$  bound.

Finally, we consider the special circles. If  $C$  is special, all points on  $\mathbf{S}(C; 1)$  lie on a circle  $Y$ . Consider the projected images of points on  $\mathbf{S}(C; 2)$  onto  $\mathbf{S}(C; 1)$ . If the original point on  $\mathbf{S}(C; 2)$  contributes to a copy of  $\Lambda$  defining  $C$ , its projected point must lie on a circle centered at the points on  $Y$  with diameter  $\text{dist}(w(1)w'(2))$ . However, the number of incidences between  $O(m)$  equidiameter circles centered at points on  $Y$  and  $O(m)$  points on  $\mathbf{S}(C; 1)$  is  $O(m)$ , since three such circles cannot intersect at a point (except for two points corresponding to north and south poles).

Hence, they can create at most  $O(m)$  edges of the copies of the triangle  $\Delta'$ , and hence  $t = O(m)$ . If we set  $t = 2^j$ , and suppose there are  $k(j)$  special circles defined by at least  $t$  but at most  $2t$  different locations of  $\Lambda$ .

The sum of the multiplicities of the circles gives the formula  $\sum_{j=0}^{\lfloor \log n \rfloor} 2^j k(j) = O(n^4)$ . The upper bound of the associated incidences is  $\sum_{j=0}^{\lfloor \log n \rfloor} 2^j k(j)^{4/5} n^{3/5} + 2^j k(j) + 2^j n$ , which is  $O(n^4) + O(\sum_{j=0}^{\lfloor \log n \rfloor} 2^{0.2j} n^{19/5}) = O(n^4)$ .

## References

1. P. Agarwal, B. Aronov, M. Sharir, and S. Suri, Selecting Distances in the Plane, *Algorithmica*, vol. 9, pp. 495–514, 1993.
2. T. Akutsu, Substructure Search and Alignment Algorithms for Three-Dimensional Protein Structures, *Proc. SIGAL-41*, pp. 1–8, IPSJ, 1994.
3. T. Akutsu, On Determining the Congruency of Point Sets in Higher Dimensions, *Proc. ISAAC '94* (LNCS, vol. 834), pp. 38–55, Springer-Verlan, Berlin, 1994.
4. H. Alt, K. Melhorn, H. Wagnener, and E. Welzl, Congruence, Similarity, and Symmetries of Geometric Objects, *Discrete Comput. Geom.*, vol. 3, pp. 237–256, 1988.
5. M. Atkinson, An Optimal Algorithm for Geometric Congruence, *J. Algorithms*, vol. 8, pp. 159–172, 1987.
6. D. H. Ballard, Generalizing the Hough Transformation to Detect Arbitrary Shapes, *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.
7. B. Bollobás, *Extremal Graph Theory*, Academic Press, New York, 1978.
8. P. Chew, M. Goodrich, D. Huttenlocher, K. Kedem, J. Kleinberg, and D. Kravets, Geometric Pattern Matching under Euclidean Motion, *Proc. 5th Canadian Conf. Comput. Geom.*, pp. 151–156, 1993.
9. K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl, Combinatorial Complexity Bounds for Arrangements of Curves and Spheres, *Discrete Comput. Geom.*, vol. 5, pp. 99–160, 1990.
10. G. Elekes and P. Erdős, Similar Configurations and Pseudo Grids, in *Intuitive Geometry*, Colloq. Math. Soc. Janos Bolyai, vol. 63, pp. 85–104, North-Holland, Amsterdam, 1994.



11. P. Erdős, On the Set of Distances of  $N$  Points, *Amer. Math. Monthly*, vol. 53, pp. 248–250, 1946.
12. P. Erdős, On the Set of Distances of  $N$  Points in Euclidean Space, *Magyar Tud. Akad. Mat. Kut. Int. Közl.*, vol. 5, pp. 165–169, 1960. Also in *Paul Erdős: The Art of Counting (Selected Writing)* (J. Spencer, ed.) pp. 676–679, MIT Press, Cambridge, MA, 1973.
13. P. Erdős and G. Purdy, Some Extremal Problems in Geometry, *J. Combin. Theory*, vol. 10, no. 3, pp. 246–252, 1971.
14. P. Finn, L. E. Kavradi, J. C. Latombe, R. Motwani, C. Shelton, S. Ventakatasubramanian, and A. Yao, RAPID: Randomized Pharmacophore Identification for Drug Design, *Proc. 13th ACM Symp. on Computational Geometry*, pp. 324–333.
15. M. T. Goodrich, J. S. B. Mitchell, and M. W. Orletsky, Practical Methods for Approximate Geometric Pattern Matching under Rigid Motions, *Proc. 10th ACM Symp. on Computational Geometry*, pp. 103–112, 1994.
16. P. J. Heffernan, Generalized Approximate Algorithms for Point Sets Congruence, *Proc. Workshop on Algorithms and Data Structures*, pp. 373–384, 1993.
17. J. E. Hopcroft and D. P. Huttenlocher, *Geometric Invariance on Computer Vision*, Chapter 18, pp. 354–374, MIT Press, Cambridge, MA, 1992.
18. S. Irani and P. Raghavan, Combinatorial and Experimental Results for Randomized Point Matching Algorithms, *Proc. 12th ACM Symp. on Computational Geometry*, pp. 68–77, 1996.
19. N. Kiryai, Y. Eldar, and A. M. Bruckstein, A Probabilistic Hough Transform, *Pattern Recognition*, vol. 24, no. 4, pp. 303–316, 1991.
20. J. Matoušek, Combinatorial and Algorithmic Geometry, Unpublished lecture notes.
21. L. Moser, On the Different Distances Determined by  $N$  Points, *Amer. Math. Monthly*, vol. 59, pp. 85–91, 1952.
22. L. Moser and J. Pach, Recent Developments in Combinatorial Geometry, in *New Trends in Discrete and Computational Geometry*, (J. Pach, ed.), Algorithms and Combinatorics, vol. 10, Springer-Verlag, Berlin, 1993.
23. R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, Cambridge, 1995.
24. J. Pach and P. Aggarwal, *Combinatorial Geometry*, Series in Discrete Mathematics and Optimization, Wiley, New York, 1995.
25. J. Pach and G. Toth, Graphs Drawn with Few Crossings per Edge, Preprint, 1996.
26. P. J. de Rezende and D. T. Lee, Point Set Pattern Matching in  $d$ -Dimensions, *Algorithmica*, vol. 3, pp. 387–404, 1995.
27. K. Sugihara, An  $n \log n$  Algorithm for Determining the Congruity of Polyhedra, *J. Comput. System. Sci.*, vol. 29, pp. 36–47, 1984.
28. L. Székely, Crossing Numbers and Hard Erdős Problems in Discrete Geometry, *Combinatorics, Probability and Computing*, vol. 6, pp. 353–358, 1997.
29. E. Szemerédi and W.T. Trotter, Extremal Problems in Discrete Geometry, *Combinatorica*, vol. 3, pp. 381–392, 1983.
30. G. Vriend and C. Sander, Detection of Common Three-Dimensional Substructures in Proteins, *PROTEINS: Structure, Function, and Genetics*, vol. 11, pp. 52–58, 1991.

Received July 17, 1997, and in revised form March 6, 1998.