Hyperplane Separability and Convexity of Probabilistic Point Sets

Martin Fink[1], John Hershberger[2], Nirman Kumar[1], and Subhash Suri[1]

1 University of California, Santa Barbara, CA, USA
2 Mentor Graphics Corp., Wilsonville, OR, USA

Abstract

We describe an $O(n^d)$ time algorithm for computing the exact probability that two d-dimensional probabilistic point sets are linearly separable, for any fixed $d \geq 2$. A probabilistic point in d-space is the usual point, but with an associated (independent) probability of existence. We also show that the d-dimensional separability problem is equivalent to a $(d+1)$-dimensional convex hull *membership* problem, which asks for the probability that a query point lies inside the convex hull of n probabilistic points. Using this reduction, we improve the current best bound for the convex hull membership by a factor of n [6]. In addition, our algorithms can handle "input degeneracies" in which more than $k+1$ points may lie on a k-dimensional subspace, thus resolving an open problem in [6]. Finally, we prove lower bounds for the separability problem via a reduction from the k-SUM problem, which shows in particular that our $O(n^2)$ algorithms for 2-dimensional separability and 3-dimensional convex hull membership are nearly optimal.

1998 ACM Subject Classification I.3.5 Computational Geometry and Object Modeling, F.2.2 Nonnumerical Algorithms and Problems, G.3 Probability and Statistics

Keywords and phrases probabilistic separability, uncertain data, 3-SUM hardness, topological sweep, hyperplane separation, multi-dimensional data

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

Multi-dimensional point sets are a commonly used abstraction for modeling and analyzing data in many domains. The ability to leverage familiar geometric concepts, such as nearest neighbors, convex hulls, hyperplanes, or partitioning of the space, is both a powerful intuition-builder and an important analysis tool. As a result, the design of useful data structures and algorithms for representing, manipulating, and querying these kinds of data has been a major research topic not only in computational geometry and theoretical computer science but also many applied fields including databases, robotics, graphics and vision, data mining, and machine learning.

Many newly emerging forms of multi-dimensional data, however, are "stochastic": the input set is not fixed, but instead is a *probability distribution* over a finite population. A leading source of these forms of data is the area of machine learning, used to construct data-driven models of complex phenomena in application domains ranging from medical diagnosis and image analysis to financial forecasting, spam filtering, fraud detection, and recommendation systems. These machine-learned models often take the form of a probability distribution over some underlying population: for instance, the model may characterize users based on multiple observable attributes and attempt to predict the likelihood that a user will buy a new product, enjoy a movie, develop a disease, or respond to a new drug.

Conference title on which this volume is based on.
Editors: Sándor Fekete and Anna Lubiw; pp. 1–19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In these scenarios, the model can often be viewed as a multi-dimensional probabilistic point set in which each point (user) has an associated probability of being included in the sample. More formally, a *probabilistic point* is a tuple (p, π), consisting of a (geometric) point $p \in \mathbb{R}^d$ and its associated probability π, with $0 < \pi \leq 1$. (We assume that the point probabilities are independent, but otherwise put no restrictions on either the values of these probabilities or the positions of the points.) We are interested in computing geometric primitives over probabilistic data models of this kind. For instance, how likely is a particular point to be a vertex (extreme point) of the convex hull of the probabilistic input set? Or, how likely are two probabilistic data sets to be linearly separable, namely, lie on opposite sides of some hyperplane? The main computational difficulty here is that the answer seems to require consideration of an exponential number of subsets: by the independence of point probabilities, the sample space includes all possible subsets of the input. For instance, the probability that a point z lies on the convex hull is a weighted sum over exponentially many possible subsets for which z lies outside the subset's convex hull. These "counting type problems" are typically $\#P$-hard [31]. Indeed, many natural graph problems that are easily solved for deterministic graphs, such as connectivity, reachability, minimum spanning tree, etc., become intractable in probabilistic graphs [27], and in fact they remain intractable even for planar graphs [30] or geometric graphs induced by points in the plane [20]. Our work explores to what extent the underlying (low-dimensional) *geometry* can be leveraged to avoid this intractability.

Our contributions.

The *hyperplane separability* problem for probabilistic point sets is the following. Given two probabilistic points sets $\mathcal{A}$ and $\mathcal{B}$ in $\mathbb{R}^d$ with a total of n points, compute the probability that a random sample of $\mathcal{A}$ can be separated from a random sample $\mathcal{B}$ by a hyperplane. One can interpret this quantity as the *expectation* of $\mathcal{A}$ and $\mathcal{B}$'s linear separability. (Throughout the paper, we use hyperplane separability interchangeably with linear separability.) Because separability by any *fixed degree polynomial* is reducible to hyperplane separability, using well-known linearization techniques, our approach can be used to determine separability by non-linear functions such as balls or ellipsoids as well.

The *convex hull membership* problem asks for the probability that a query point p lies inside the convex hull of a random sample of a probabilistic point set $\mathcal{A}$. This is the complement of the probability that p is an extreme point (convex hull vertex) of $\mathcal{A}$. Finally, the *halfspace intersection* problem asks for the probability that a set of d-dimensional halfspaces, each appearing with an independent probability, has a non-empty common intersection. This is equivalent to the *expected feasibility* of a probabilistic linear program on d variables.

Throughout, we focus on problems in dimensions $d \geq 2$ because their 1-dimensional counterparts are easily solved in $O(n \log n)$ time. Our main results can be summarized as follows.

1. We present an $O(n^d)$ time and $O(n)$ space algorithm for computing the hyperplane separability of two d-dimensional probabilistic point sets with a total of n points. The same bound also holds for the *oriented* version of separability, in which the halfspace containing one of the sets, say $\mathcal{A}$, is prespecified.
2. We prove that the d-dimensional separability problem is at least as hard as the $(d+1)$-SUM problem [9, 16, 17, 18], which implies that our $O(n^2)$ bound for $d = 2$ is nearly tight. (The 3-SUM problem is conjectured to require $\Omega(n^{2-o(1)})$ time [22].) When the

dimension d is non-constant, we show that the problem is $\#P$-hard.

3. We show that the convex hull membership problem in d-space has a linear-time reduction to a hyperplane separability problem in dimension $(d-1)$, and therefore can be solved in time $O(n^{d-1})$, for $d \geq 3$, improving the previous best bound of Agarwal et al. [6] by a factor of n. Our lower bound for separability implies that this bound is nearly tight for $d = 3$.
4. We show that the non-empty intersection problem for n probabilistic halfspaces in d dimensions can be solved in time $O(n^d)$. Equivalently, we compute the exact probability that a random sample from a set of n probabilistic linear constraints with d variables has a feasible solution.
5. Finally, our algorithms can cope with input degeneracies. Thus, for the convex hull membership problem, our result simultaneously improves the previous best running time [6] as well as eliminates the assumption of general position.

Related work.

The topic of algorithms for probabilistic (uncertain) data is a subject of extensive and ongoing research in many areas of computer science including databases, data mining, machine learning, combinatorial optimization, theory, and computational geometry [7, 8, 19]. We will only briefly survey the results that are directly related to our work and deal with multi-dimensional point data. Within computational geometry and databases, a number of papers address nearest neighbor searching, indexing and skyline queries under the *locational uncertainty* model in which the position of each data point is given as a probability distribution [1, 2, 3, 4, 5, 10, 23, 26], as well as separability by a line in the plane [13].

The uncertainty model we consider, in which each point's position is known but its existence is probabilistic, has also been studied in a number of papers recently. The problem of computing the *expected* length of the Euclidean minimum spanning tree (MST) of n probabilistic points is considered in [20], and shown to be $\#P$-hard even in two dimensions. The closest pair problem and nearest neighbor searching for probabilistic points are considered in [21]. Suri, Verbeek, and Yıldız [29] consider the problem of computing the *most likely convex hull* of a probabilistic point set, give a polynomial-time algorithm for dimension $d = 2$, but show NP-hardness for $d \geq 3$. The complexity of the most likely Voronoi diagram of probabilistic points has been explored by Suri and Verbeek [28], and also by Li et al. [24]. In the work most closely related to ours, Agarwal et al. [6] consider a number of problems related to probabilistic convex hulls, including the convex hull membership probability. Their main result is an $O(n^d)$-time algorithm for computing the probability of convex hull membership, but it only works for points satisfying the following non-degeneracy condition: the projection of no $k+1$ points on a subspace spanned by any k coordinates may lie on a $(k-1)$-dimensional hyperplane, for any $2 \leq k \leq d$. Our new algorithm improves the running time by a factor of n as well as eliminates the need for non-degeneracy assumptions.

2 Separability of Probabilistic Point Sets

2.1 Preliminaries

A *probabilistic point* is a tuple (p, π), consisting of a (geometric) point $p \in \mathbb{R}^d$ and its associated probability π, with $0 < \pi \leq 1$. For notational convenience, we denote a set of probabilistic points as $\mathcal{P} = \{p_1, p_2, \ldots, p_n\}$ with an implicit understanding that $\pi(p_i)$ is the probability associated with p_i. We assume that the point probabilities are independent but otherwise place no restrictions on either the values of these probabilities or the positions of the points. We are interested in computing how often certain geometric properties occur for sets of probabilistic points. This requires reasoning about random samples in which each point p is drawn according to its probability $\pi(p)$. In particular, a fixed subset $A \subseteq \mathcal{P}$ occurs as a *random* sample with probability

$$\mathbf{Pr}\big[A\big] = \prod_{p \in A} \pi(p) \cdot \prod_{p \notin A} (1 - \pi(p)).$$

The central problem of our paper is to compute the probability that two probabilistic point sets $\mathcal{A}$ and $\mathcal{B}$ are linearly separable. We say that two sample sets $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$ are linearly separable if there exists a hyperplane H for which A and B lie in different (open) halfspaces of H. The *open* halfspace separation means that no point of $A \cup B$ lies on H, thus enforcing a strict separation. When there is no loss of generality, we assume that A lies *above* H, namely in the positive halfspace, and B lies *below* H. For ease of reference, we define an indicator function $\sigma(\mathcal{A}, \mathcal{B})$ for linear separability:

$$\sigma(A, B) = \begin{cases} 1 & \text{if } A, B \text{ are linearly separable} \\ 0 & \text{otherwise.} \end{cases}$$

We assume $\sigma(\emptyset, \emptyset) = 1$ to handle the trivial case. Given two probabilistic point sets $\mathcal{A}$ and $\mathcal{B}$, their *separation probability* is the joint sum over all samples:

$$\mathbf{Pr}\big[\sigma(\mathcal{A}, \mathcal{B})\big] = \sum_{A \subseteq \mathcal{A}, B \subseteq \mathcal{B}} \mathbf{Pr}\big[A\big] \cdot \mathbf{Pr}\big[B\big] \cdot \sigma(A, B)$$

This is also the *expectation* of the random variable $\sigma(A, B)$. Because each sample pair is deterministic, we can decide its linear separability in $O(n)$ time using fixed-dimensional linear programming algorithms of Megiddo or Clarkson [11, 25]. We can, therefore, *estimate* $\mathbf{Pr}\big[\sigma(\mathcal{A}, \mathcal{B})\big]$ in polynomial time by drawing many samples A, B and returning the fraction of separable samples, but we are interested in the complexity of computing this quantity *exactly*. We begin our discussion by describing a reduction to a special kind of separability.

2.2 Reduction to Anchored Separability

A natural idea is to compute the sum $\mathbf{Pr}\big[\sigma(\mathcal{A}, \mathcal{B})\big]$ by considering the $O(n^d)$ combinatorially distinct separating hyperplanes induced by the points of $\mathcal{A} \cup \mathcal{B}$. However, two point sets may be separable by many different hyperplanes, so we need to ensure that the probability is assigned to a unique *canonical* hyperplane.[1] Our main insight is the following: if we introduce an extra point z into the input, then the canonical hyperplane can be defined uniquely (and computed efficiently) with respect to z: in particular, we prove that the separating hyperplane at *maximum distance* from z is a canonical one. We call this artificially added point z the *anchor point.*

How does the true separation probability we want, namely $\mathbf{Pr}\big[\sigma(\mathcal{A}, \mathcal{B})\big]$, relate to this *anchored separability* that includes an artificially added point *anchor*? It turns out the former can be calculated from two instances of the latter and one *lower dimensional* instance of the former.

[1] Dualizing the points to hyperplanes can simplify the enumeration of separating planes for the summation but does not address the over-counting problem.

We initially assume that the input points are in general position, namely, no $k+1$ points of $\mathcal{A} \cup \mathcal{B}$ are affinely dependent, but revisit the degeneracy problem in Section 4. Without loss of generality, we also assume that all points have positive dth coordinate, and therefore lie above the hyperplane $x_d = 0$. Let the anchor point z be a point that lies *above* all the points of $\mathcal{A} \cup \mathcal{B}$ and is in general position with them. The probability of z is $\pi(z) = 1$, so it is always included in the sample.

If $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$ are two random samples and H a hyperplane separating them, then clearly z lies either (i) on the same side as A, (ii) on the same side as B, or (iii) on the hyperplane H. The cases (i) and (ii) are symmetric, and can be handled by including the anchor point once in A and once in B, but unfortunately they are *not disjoint:* A and B may admit separating hyperplanes with z lying on either side. Fortunately, the following lemma shows that this duplication of probability is precisely accounted for by case (iii). Finally, Lemma 3 shows that case (iii) itself is an instance of hyperplane separability in a lower dimension.

▶ **Lemma 1.** *Let z be the anchor point. Then there exist separating hyperplanes H_1, H_2 with z lying on the same side of H_1 as A but on the same side of H_2 as B if and only if there is another hyperplane H that passes through z and separates A from B.*

Proof. In the forward direction, if either H_1 or H_2 passes through z, we are done, so let us assume that neither contains z. Without loss of generality, assume that both hyperplanes contain A on their positive side, and B on their negative side. Thus, we have $A \subset H_1^+ \cap H_2^+$ and $B \subset H_1^- \cap H_2^-$. It follows that there are no points of $A \cup B$ in the region of the space $\Phi = \mathbb{R}^d \setminus \big((H_1^+ \cap H_2^+) \cup (H_1^- \cap H_2^-)\big)$. On the other hand, the anchor point z must lie in Φ because it lies on different sides of H_1 and H_2. See Figure 1 for illustration.

If H_1 and H_2 are parallel, then a hyperplane passing through z and parallel to H_1 is a separator, and we are done. On the other hand, if H_1 and H_2 intersect in a $(d-2)$-dimensional subspace, then we choose H as the hyperplane through z containing this subspace. This hyperplane lies in Φ, contains $H_1^+ \cap H_2^+$ and $H_1^- \cap H_2^-$ on opposite sides, and thus is a separating hyperplane for A and B.

To prove the reverse direction of the lemma statement, given a separating hyperplane H passing through z, we simply move H parallel to itself slightly, once toward A and once toward B. This completes the proof. ◀

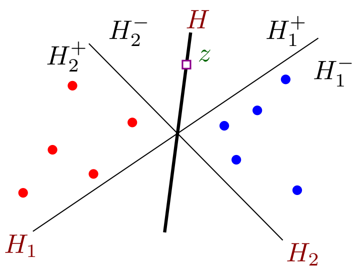


Figure 1 Illustration for the proof of Lemma 1.

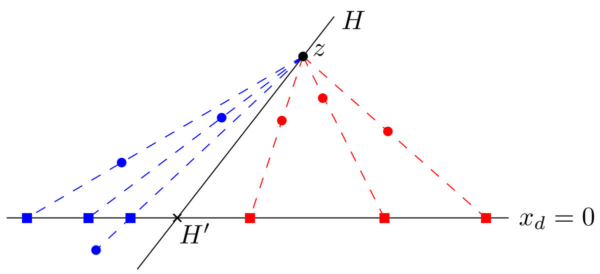


Figure 2 Illustration for the proof of Lemma 3.

Thus, event (iii) is precisely the intersection of events (i) and (ii). In the remainder of the paper, for notational convenience, we use $\mathcal{P} + z$ for the probabilistic point set $\mathcal{P} \cup \{(z, 1)\}$, where z is the anchor point with associated probability $\pi(z) = 1$. Let $\mathbf{Pr}\big[\sigma(z, \mathcal{A}, \mathcal{B})\big]$ denote the probability that sets $\mathcal{A}$ and $\mathcal{B}$ are linearly separable by a hyperplane passing through the anchor point z. Then, the preceding lemma gives the following result.

▶ **Lemma 2.** *Given two probabilistic point sets $\mathcal{A}$ and $\mathcal{B}$, we have the following equality:*

$$\mathbf{Pr}\big[\sigma(\mathcal{A},\mathcal{B})\big] \;=\; \mathbf{Pr}\big[\sigma(\mathcal{A}+z,\mathcal{B})\big] + \mathbf{Pr}\big[\sigma(\mathcal{A},\mathcal{B}+z)\big] - \mathbf{Pr}\big[\sigma(z,\mathcal{A},\mathcal{B})\big].$$

Computing the probabilities $\mathbf{Pr}\big[\sigma(\mathcal{A}+z,\mathcal{B})\big]$ and $\mathbf{Pr}\big[\sigma(\mathcal{A},\mathcal{B}+z)\big]$ requires solving two instances of *anchored separability*, once with z included in $\mathcal{A}$ and once in $\mathcal{B}$. This leaves the last term $\mathbf{Pr}\big[\sigma(z,\mathcal{A},\mathcal{B})\big]$, which as the following lemma shows can be reduced to an instance of separability in dimension $d-1$.

Consider any sample $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$. We centrally project all these points onto the hyperplane $x_d = 0$ from the anchor point z: that is, the image of a point $p \in \mathbb{R}^d$ is the point $p' \in \mathbb{R}^{d-1}$ at which the line connecting z to p intersects the hyperplane $x_d = 0$. Observe that all points of $\mathcal{A} \cup \mathcal{B}$ have a well-defined projection because z lies above all of them.

▶ **Lemma 3.** *Let $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$ be two sample sets, and let A', B' be their projections onto $x_d = 0$ with respect to z. Then A and B are separable by a hyperplane passing through z if and only if A' and B' are linearly separable in $x_d = 0$.*

Proof. First, suppose there is a hyperplane H passing through z that separates A and B. We may assume that H is not parallel to $x_d = 0$; otherwise, rotate the input slightly. The intersection of H with $x_d = 0$ is a hyperplane H' in the $(d-1)$-dimensional subspace $x_d = 0$. See Figure 2. Clearly, the projection of each point p lies on the same side of H as does p. Since H separates A from B, it follows that H' separates A' from B'.

Conversely, suppose H' separates A' from B' in $x_d = 0$. The hyperplane H spanned by H' and z clearly separates A' from B' in $\mathbb{R}^d$. Since each point p lies on the same side of H as its projection, the point sets A and B are also separated by H. ◀

3 Computing Anchored Separability

We now describe our main technical result: efficiently computing the separation probability of two probabilistic sets when one of the sets contains the anchor point z. Without loss of generality, we explain how to compute $\mathbf{Pr}\big[\sigma(\mathcal{A}+z,\mathcal{B})\big]$. We can restrict our search to the $O(n^d)$ "combinatorially distinct" hyperplanes induced by the set of points $\mathcal{A} \cup \mathcal{B}$. Indeed, any free hyperplane can be translated and rotated until it passes through d distinct points of the input, without changing the *closed* halfspace membership of any point. Conversely, any hyperplane that contains A and B on opposite *closed* halfspaces and passes through at most d points can be translated and rotated until the same separation is realized by open halfspaces. (Recall that the input set of points, including the anchor z, is assumed to be in general position. We discuss how to handle degeneracies in Section 4.)

Given a hyperplane H, we can easily compute the probability that $\mathcal{A}+z$ lies in H^+ and $\mathcal{B}$ lies in H^-. The separation probabilities for different hyperplanes, however, are not independent: a sample $A \subseteq \mathcal{A}, B \subseteq \mathcal{B}$ may be separated by many different hyperplanes, and the algorithm needs to "assign" each separable sample to a unique hyperplane. We will assign a *canonical* separator for every pair $(A+z, B)$ of separable samples and then sum the probabilities over all possible canonical separators. Geometrically, our canonical separator is the hyperplane that separates $A+z$ from B and lies at *maximum distance* from the anchor z. Before we formalize the definition of a canonical separator and prove its uniqueness (cf. Section 3.2), we need the following important concept of a shadow cone.

3.1 The Shadow Cone

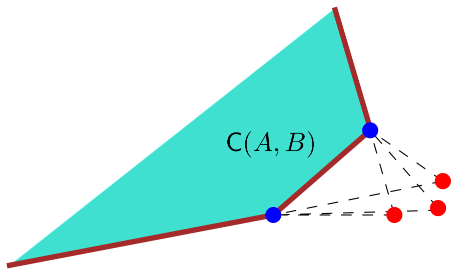


Figure 3: A shadow cone in two dimensions.

Given two points $u, v \in \mathbb{R}^d$, let $shadow(u, v) = \{\lambda v + (1-\lambda)u \mid \lambda \geq 1\}$ be the ray originating at v and directed along the line uv away from u. (If we place a light source at u, then this is the shadow cast by the point v.) Let $CH(P)$ denote the convex hull of a point set P. Given two sets of points A and B, with $A \cap B = \emptyset$, we define their *shadow cone* $\mathsf{C}(A, B)$ as the union of $shadow(u, v)$ for all $u \in CH(A)$ and $v \in CH(B)$. In other words, if we place light sources at all points of $CH(A)$, then $\mathsf{C}(A, B)$ is the shadow cast by the convex hull $CH(B)$. (The shadow cone $\mathsf{C}(A, B)$ includes both umbra and penumbra of the shadow.) In the trivial case of $A = \emptyset$, we define $\mathsf{C}(\emptyset, B)$ to be the same as $CH(B)$. Figure 3 gives an illustration in two dimensions. The proof of the following lemma appears in Appendix A.1.

▶ **Lemma 4.** *The shadow cone* $\mathsf{C}(A, B)$ *is a (possibly unbounded) convex polytope, and if* A *and* B *are nonempty,* $\mathsf{C}(A, B)$ *is the convex hull of the union of* $shadow(u, v)$*, for all* $u \in A, v \in B$.

Each face of $\mathsf{C}(A, B)$ is *defined* by a subset of (at most d) points in $A \cup B$, and the defining set always includes at least one point of B. When all the points defining the face are in B, the face must be a (bounded) face of $CH(B)$; otherwise, it is an unbounded face. (See Figure 3.) We will use the following simple but important fact: if u is a point of A and $p \in CH(B)$, then $shadow(u, p)$ is contained in $\mathsf{C}(A, B)$. We are now ready to state and prove the important connection between the shadow cone and hyperplane separability of two subsets $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$, with $A \cap B = \emptyset$.

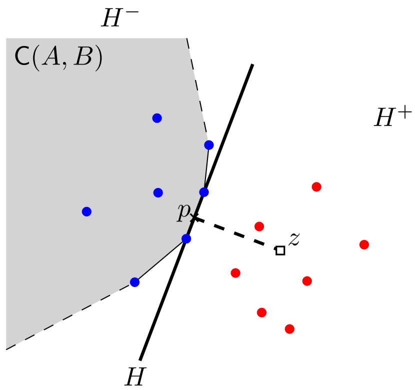
▶ **Lemma 5.** $A + z$ *and* B *can be separated by a hyperplane if and only if* $z \notin \mathsf{C}(A, B)$*.*

Proof. First, suppose there is a separating hyperplane H with $A + z \subset H^+$ and $B \subset H^-$. Then, we must also have $\mathsf{C}(A, B) \subset H^-$ because the shadow cone lies on the same side of H as B. Since $z \in H^+$ by assumption, this implies $z \notin \mathsf{C}(A, B)$.

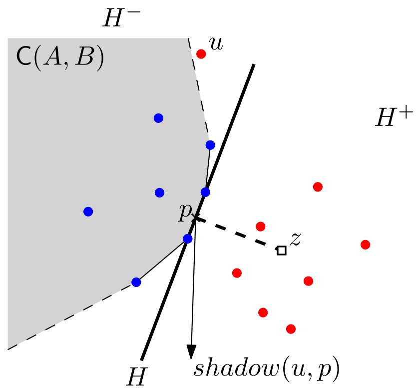
For the converse, we assume $z \notin \mathsf{C}(A, B)$ and exhibit a separating hyperplane. Let $p \in \mathsf{C}(A, B)$ be the point in the shadow cone with minimum distance to the anchor z. Then the hyperplane H passing through p and orthogonal to the vector $p - z$ necessarily has z and the shadow cone on opposite sides, which follows from the convexity of $\mathsf{C}(A, B)$: if $z \subset H^+$, then $\mathsf{C}(A, B)$ is in the closure of the halfspace H^-. See Figure 4a.

It still remains to show that we can achieve *open* half-space separability of the sets $A + z$ and B. This depends crucially on the assumption of general position—indeed, if degeneracies exist, $z \notin \mathsf{C}(A, B)$ is not sufficient to prove strict separability. First, observe that no point of A can be in the open halfspace H^-. If such a point $u \in A$ were to exist, then the ray $shadow(u, p)$ would be contained in $\mathsf{C}(A, B)$, and there is a point on this ray that is closer to z than p, contradicting the minimality of p. See Figure 4b.

Because H is a supporting hyperplane of the shadow cone, the intersection $F = H \cap \mathsf{C}(A, B)$ is a face of $\mathsf{C}(A, B)$. Let $I = A \cap H$ and $J = B \cap H$ be the subsets of the sample points defining F (at most d due to general position). Since F contains at least one point of B, we have $|J| \geq 1$ and, therefore, $|I| < d$. Because no point of J is contained in the affine span of I, we can perform an infinitesimal rotation of H around the subface determined by I in



(a) The separator H.



(b) $shadow(u,p)$ contains a point closer to z than p.

Figure 4 Illustration for the proof of Lemma 5.

the direction of z, so that all points of J (and thus B) lie in the open halfspace H^-. We then can translate the hyperplane by an infinitesimal amount away from z to ensure that all points of I (and thus A) lie in the open halfspace H^+. We now have a hyperplane whose open halfspaces separate the sample sets, and the proof is complete. ◀

3.2 Canonical Separating Hyperplanes

By Lemma 5, sets $A+z$ and B can be separated by a hyperplane if and only if $z \notin \mathsf{C}(A,B)$. Since $\mathsf{C}(A,B)$ is a convex set, there is a *unique* nearest point $p = \mathrm{np}(z, \mathsf{C}(A,B))$ on the boundary of $\mathsf{C}(A,B)$ with minimum distance to z. We define our *canonical hyperplane* $H(z,A,B)$ as the one that passes through p and is orthogonal to the vector $p-z$. Indeed, the proof of Lemma 5 already argues that $H(z,A,B)$ is a separating hyperplane in the sense that with infinitesimal rotation and translation it achieves open half-space separation between $A+z$ and B.

Our main idea now is to turn the separation question around and instead of asking "which hyperplane separates a particular sample pair A,B," we ask "for which pairs of samples A,B is H a canonical separator?" The latter formulation allows us to compute the separation probability $\mathbf{Pr}\big[\sigma(\mathcal{A}+z,\mathcal{B})\big]$ by considering at most $O(n^d)$ possible hyperplanes. The following lemmas encapsulate our definition of canonical separators. (Proofs appear in Appendix A.2.)

▶ **Lemma 6.** *Let C be a d-dimensional convex polyhedron and z a point not contained in C. Then there is a unique point $p \in C$ that minimizes the distance to z, and a unique face of C whose relative interior contains p.*

▶ **Lemma 7.** *Let C be a d-dimensional convex polyhedron, z a point not contained in C, and p the point of C at minimum distance from z. If p lies in the relative interior of the face F of C, then the hyperplane H through p that is orthogonal to $p-z$ contains F. This hyperplane contains C in one of its closed halfspaces, and is the hyperplane farthest from z with this property.*

3.3 The Algorithm

Consider a random sample of input points $A \cup B$ such that A and B are linearly separable. By Lemma 5, we know that $z \notin \mathsf{C}(A, B)$. Since $\mathsf{C}(A, B)$ is a convex polyhedron, there is a unique face F with $p = \mathrm{np}(z, \mathsf{C}(A, B))$ in its relative interior, by Lemma 6. Finally, by Lemma 7, F lies in the canonical hyperplane $H(z, A, B)$, which is the hyperplane passing through p and orthogonal to $p - z$.

We now consider the defining set of F, which consists of two subsets $I \subseteq A$ and $J \subseteq B$, with $|I \cup J| \leq d$ and $|J| \geq 1$. It follows from the definition of the shadow cone that $F = \mathsf{C}(I, J)$. If F is finite, then it is the convex hull of its vertices, all of which belong to B, so $I = \emptyset$ and $F = CH(J) = \mathsf{C}(I, J)$. On the other hand, if F is unbounded, it is the convex hull of its finite vertices and a constant number of shadow rays. If F has dimension k, general position implies that its affine span includes exactly $k + 1$ vertices of $A \cup B$. F is the shadow hull of these $k + 1$ vertices, which constitute sets I and J.

Since F is the face of *smallest dimension* in $\mathsf{C}(A, B)$ containing p in its relative interior, we conclude that $I \cup J$ is the smallest subset of $A \cup B$ for which p lies in the relative interior of $\mathsf{C}(I, J)$. Equivalently, $I \cup J$ is the smallest subset of $A \cup B$ for which the canonical hyperplane $H(z, I, J)$ is the same as $H(z, A, B)$. Note that these sets are unique since the input is in general position.

This last property is the key to our algorithm: we simply enumerate all subsets $I \subseteq \mathcal{A}$ and $J \subseteq \mathcal{B}$, with $|I \cup J| \leq d$ and $|J| \geq 1$, and assign to the hyperplane $H(z, I, J)$ the separation probability of *all those samples* $A \cup B$ *that are separable and for which* $H(z, I, J)$ *is the canonical separator* $H(z, A, B)$. In particular, let us define the following function for the probability that the points defining the hyperplane $H(z, I, J)$ are in the sample and none of the remaining points of $\mathcal{A} \cup \mathcal{B}$ lies on its *incorrect side*.

$$\mathbf{Pr}\big[H(z, I, J)\big] = \prod_{u \in I \cup J} \pi(u) \cdot \prod_{u \in \mathcal{A} \cap H^-} (1 - \pi(u)) \cdot \prod_{u \in \mathcal{B} \cap H^+} (1 - \pi(u)).$$

The first term in the product is the joint probability that all points of $I \cup J$ are in the sample, while the second and third terms are the probabilities that none of the points of $\mathcal{A}$ (resp. $\mathcal{B}$) that lie in the negative halfspace (resp. positive halfspace) of $H(z, I, J)$ are chosen.

Finally, to decide whether $H(z, I, J)$ is the canonical hyperplane for a sample, we just need to check if the point closest to z in $\mathsf{C}(I, J)$ lies in the relative interior of $\mathsf{C}(I, J)$. The following algorithm **AnchoredSep** implements this construction.

Algorithm AnchoredSep:

Input: The point sets $\mathcal{A} + z$ and $\mathcal{B}$
Output: Their separation probability $\alpha = \mathbf{Pr}\big[\sigma(\mathcal{A} + z, \mathcal{B})\big]$
$\alpha = \prod_{u \in \mathcal{B}} (1 - \pi(u))$;
forall the $I \subseteq \mathcal{A}, J \subseteq \mathcal{B}$ *where* $|I \cup J| \leq d, J \neq \emptyset$ **do**
  let $p = \mathrm{np}(z, \mathsf{C}(I, J))$;
  if *p lies in the relative interior of* $\mathsf{C}(I, J)$ **then**
    $\alpha = \alpha + \mathbf{Pr}\big[H(z, I, J)\big]$;
  end
end
return α;

▶ **Theorem 8.** **AnchoredSep** *correctly computes the probability* $\mathbf{Pr}\big[\sigma(\mathcal{A}+z,\mathcal{B})\big]$.

Proof. The initial assignment $\alpha = \prod_{u\in\mathcal{B}}(1-\pi(u))$ accounts for the trivial case when none of the points of $\mathcal{B}$ is present. The separation probability of any other outcome is associated with the minimal defining set $I \cup J$, and computed exactly once within the **forall** loop. ◀

3.4 Implementation in $O(n^d)$ Time and $O(n)$ Space

A naïve implementation of Algorithm **AnchoredSep** runs in $O(n^{d+1})$ time and $O(n)$ space: there are $O(n^d)$ subset pairs $I \subseteq A, J \subseteq B$ with $|I \cup J| \leq d$, $J \neq \emptyset$, and evaluating $\mathbf{Pr}\big[H(z,I,J)\big]$ for each one individually takes $O(n)$ time. We show how to reduce the average evaluation time to $O(1)$ per subset pair, which reduces the overall running time to $O(n^d)$. The main result of our paper can be stated as follows.

▶ **Theorem 9.** *Let* $\mathcal{A},\mathcal{B} \subseteq \mathbb{R}^d$ *be two probabilistic sets of* n *points in general position, for* $d \geq 2$. *We can compute their probability of hyperplane separation* $\mathbf{Pr}\big[\sigma(\mathcal{A},\mathcal{B})\big]$ *in* $O(n^d)$ *worst-case time.*

Proof. We compute the separation probability for all subsets I, J with $|I \cup J| < d$ explicitly by the naïve algorithm. This takes $O(n)$ time for each of $O(n^{d-1})$ subset pairs, for a total time of $O(n^d)$. To handle subset pairs with $|I \cup J| = d$, we process instances in linear-size groups. Let $\mathcal{A} \cup \mathcal{B} = \mathcal{P} = \{p_1, \ldots, p_n\}$. We group d-element subsets of $\mathcal{P}$ according to their $(d-1)$-subsets with smallest indices, as follows. For each $(d-1)$-element subset $P \subseteq \mathcal{P}$, let p_k be the element with maximum index. For each $p \in P_{>k} = \{p_{k+1}, \ldots, p_n\}$, we compute $\mathbf{Pr}\big[H(z,I,J)\big]$ for $I \cup J = P \cup \{p\}$. As we show below, this can be done in $O(n)$ time for each $(d-1)$-subset of $\mathcal{P}$, for a total bound of $O(n^d)$.

The $d-1$ points in P define a $(d-2)$-dimensional subspace. The $n-d+1$ points in $\overline{P} = \mathcal{P} \setminus P$ can be rotationally ordered around this subspace. For the moment, let us assume this rotational order is known. If p_0 is an arbitrary element of $\overline{P}$, we compute $\mathbf{Pr}\big[H(z,I,J)\big]$ in $O(n)$ time for $I \cup J = P \cup \{p_0\}$. We then process the points of $\overline{P}$ in rotational order. Each point p contributes a multiplicative factor to $\mathbf{Pr}\big[H(z,I,J)\big]$: $\pi(p)$ if $p \in I \cup J$, $(1-\pi(p))$ if $p \in ((\mathcal{A} \cap H^-) \cup (\mathcal{B} \cap H^+))$, and 1 otherwise. When H rotates from one point $p \in \overline{P}$ to the next, the multiplicative factors for those two points change, and we can update $\mathbf{Pr}\big[H(z,I,J)\big]$ with two multiplications and two divisions. Whenever the conditions for acceptance of I, J are met—$J \neq \emptyset$, $H \cap P_{>k} \neq \emptyset$, $\mathrm{np}(z, \mathsf{C}(I,J))$ is in the relative interior of $\mathsf{C}(I,J)$—then we add $\mathbf{Pr}\big[H(z,I,J)\big]$ to the separation probability.

If we compute the rotational order of the points in $\overline{P}$ by sorting, we spend $O(n \log n)$ time per $(d-1)$-subset of $\mathcal{P}$, for a total running time of $O(n^d \log n)$. To do better, we use the ideas of *duality* [12] and *topological sweep* [14]. *Duality* is an order-preserving, invertible mapping between points and hyperplanes in $\mathbb{R}^d$. Each point $p \in \mathcal{P}$ dualizes to a hyperplane p^*, and the hyperplane H spanning d points $p_1, \ldots, p_d$ dualizes to the point H^* in dual space that is the intersection of the d hyperplanes $p_1^*, \ldots, p_d^*$. A subset $P \subseteq \mathcal{P}$ with $|P| = d-1$ dualizes to a line ℓ, and the rotational order of $\overline{P}$ (as defined above) around the $(d-1)$-dimensional subspace defined by P corresponds exactly to the order of intersections of the dual hyperplanes p^* (for $p \in \overline{P}$) with the dual line ℓ.

Ordering intersections along a line is still a sorting problem, but we can reduce the time by a logarithmic factor by considering arrangements of lines in two-dimensional planes. We consider all subsets $P \subseteq \mathcal{P}$ with $|P| = d-2$. Let p_k be the maximum-index point in a given P, and define $P_{>k} = \{p_{k+1}, \ldots, p_n\}$, as above. The intersection $\cap_{p \in P} p^*$ is a dual plane Q, and the intersection of Q with each p^*, for $p \in \overline{P}$, is a line. We use *topological* *sweep* [14] to visit the vertices of the arrangement of these $n-d+2$ lines in order along each line. We initialize $\mathbf{Pr}\big[H(z, I, J)\big]$ at the first vertex of each line, then update it in constant time per vertex during the sweep. At every vertex corresponding to two points in $P_{>k}$, if the acceptance criteria are met, we add the corresponding $\mathbf{Pr}\big[H(z, I, J)\big]$ to the separation probability. Topological sweep takes linear space and $O(n^2)$ time for each of the $O(n^{d-2})$ subsets $P \subseteq \mathcal{P}$ with $|P| = d-2$, so the total processing time is $O(n^d)$, and the total space is $O(n)$. This establishes the main result of our paper. ◀

4 Handling Input Degeneracies

Our algorithm so far has relied on the assumption that the input points or hyperplanes are in *general* (non-degenerate) position. That is, no $(k+2)$ points lie on a k-dimensional affine space, or no $k+1$ hyperplanes meet in a $(d-k)$-dimensional subspace. These assumptions, while convenient for theory, are rarely satisfied in practice. They are especially troublesome in our analysis because of the need to define unique *canonical* sets. Indeed, when the input is degenerate, the need to choose a single canonical subset is the reason why the convex hull membership algorithm of [6] does not work—there is no efficient way to isolate *witness faces*. In this section, we show how to handle inputs that are not in general position.

Let us consider computing the probability of anchored separability $\mathbf{Pr}\big[\sigma(\mathcal{A}+z, \mathcal{B})\big]$ when the input sets are in a degenerate position. Our characterization of separable instances using shadow cones, Lemma 5, fails in the presence of degeneracy. As a concrete example, consider the case when $B \subseteq \mathcal{B}$ consists of a single point that lies in the convex hull of points $A \subseteq \mathcal{A}$, and all points of $A \cup B$ lie on a hyperplane that does not contain z. Although z lies outside $\mathsf{C}(A, B)$, we clearly cannot separate $A+z$ from B.

To address the problem of degenerate inputs, we apply a symbolic perturbation to the points. Part of our solution is standard Simulation of Simplicity [15], but the more important part is problem-specific. We convert degenerate non-separable samples into non-degenerate samples that are still non-separable. We first choose the anchor z above all points in $\mathcal{P} = \mathcal{A} \cup \mathcal{B}$ and outside the affine span of every d-tuple of $\mathcal{P}$. This can be done in $O(n^d)$ time. For each point $a \in \mathcal{A}$, we define a perturbed point $a' = a + \epsilon \cdot (a - z)$, for an infinitesimal $\epsilon > 0$. This point lies on the line supporting $\overline{az}$, but slightly farther from z than a. Similarly, for each $b \in \mathcal{B}$, define $b' = b + \epsilon \cdot (z - b)$, a point contained in $\overline{bz}$, but slightly closer to z than b. Let $\mathcal{A}', \mathcal{B}'$ be the sets of perturbed points corresponding to $\mathcal{A}$ and $\mathcal{B}$.

▶ **Lemma 10.** *Let $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$ be two sample sets, and let A', B' be the corresponding perturbed sets. Then $A+z$ and B are strictly separable by a hyperplane if and only if $A'+z$ and B' are. Furthermore, if some hyperplane H with $z \notin H$ is a non-strict separator of $A'+z$ and B' for some ϵ, then H is a strict separator for any $\epsilon_0 < \epsilon$.*

Proof. First, suppose $A+z$ and B are strictly separable by a hyperplane H, with $A+z \subseteq H^+$. Let δ be the minimum distance between H and any point in $A \cup \{z\} \cup B$, and let Δ be the maximum distance between z and any point in $A \cup B$. The choice of any $\epsilon < \delta/\Delta$ ensures that $A'+z \subseteq H^+$ and $B' \subseteq H^-$. Conversely, if $A'+z \subseteq H^+$ and $B' \subseteq H^-$, then *a fortiori* $A+z \subseteq H^+$ and $B \subseteq H^-$ because each a is closer to z than a' and each b is farther from z than b' (and hence farther from H).

To prove the second part of the lemma, we simply note that if any point of $A' \cup B'$ lies on the separating hyperplane H, choosing any $\epsilon_0 < \epsilon$ moves the point off of H and into the desired halfspace. This completes the proof. ◀

We apply Simulation of Simplicity [15] to the point sets $\mathcal{A}'$ and $\mathcal{B}'$, with the symbolic perturbation of each point chosen to be of smaller order than the ϵ perturbation applied to produce $\mathcal{A}'$ and $\mathcal{B}'$. Simulation of Simplicity breaks any remaining degeneracies in the point set, so Lemma 5 holds and the algorithm of Section 3 works without modification. By Lemma 10, every separable point set in the symbolically perturbed data corresponds to a separable point set in the original data, and vice versa, so the Simulation of Simplicity computation correctly solves the original problem.

▶ **Theorem 11.** *Let $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^d$ be two probabilistic sets of n points, possibly in degenerate position, for $d \geq 2$. We can compute $\mathbf{Pr}\big[\sigma(\mathcal{A}, \mathcal{B})\big]$, their probability of hyperplane separation, in $O(n^d)$ worst-case time and $O(n)$ space.*

5 Lower Bounds

In this section we argue that the running time of any algorithm that computes the probability of hyperplane separability must have an exponential dependence on dimension d. For any fixed d, we show that the separability problem is at least as hard as the k-SUM problem for $k = d + 1$. The proof of this result is interesting in that it uses Simulation of Simplicity [15], a technique for removing geometric degeneracies, as a *means to detect degeneracies*. In addition, we show that the problem is $\#P$-hard when $d = \Omega(n)$.

The k-SUM problem is a generalization of 3-SUM, which is a classical hard problem in computational geometry [9, 16, 17, 18]. The current conjectures state that the 3-SUM problem requires time at least $\Omega(n^{2-o(1)})$ [17, 18, 22], and that the k-SUM problem, for $k > 3$, has a lower bound of $\Omega(n^{\lceil k/2 \rceil})$ under some models of computation [9, 16, 17, 18]. We use the following variant of the k-SUM problem:

> **Problem k-SUM:** Given k sets containing a total of n real numbers, grouped into a single set Q and $k - 1$ sets $R_1, R_2, \ldots, R_{k-1}$, determine whether there exist $k - 1$ elements $r_i \in R_i$, one per set R_i, and an element $q \in Q$ such that $\sum_{i=1}^{k-1} r_i = q$.

▶ **Theorem 12.** *The d-dimensional hyperplane separability problem is at least as hard as $(d + 1)$-SUM.*

Proof. Let P be a regular $(d - 1)$-simplex, embedded in the hyperplane $x_d = 0$ in $\mathbb{R}^d$. Let $p_1, p_2, \ldots, p_d$ be the vertices of P, and let c be its barycenter. Given an instance $(Q, R_1, R_2, \ldots, R_d)$ of $(d+1)$-SUM, we define $d+1$ sets of d-dimensional points, one for each of the input sets, as follows. The sets $\mathcal{B}_i = \{p_i + (0, \ldots, 0, r) \mid r \in R_i\}$ correspond to the input sets R_i, for $i = 1, 2, \ldots, d$; let $\mathcal{B} = \cup_i \mathcal{B}_i$. The set $\mathcal{A} = \{c + (0, \ldots, 0, q/d) \mid q \in Q\}$ corresponds to the input set Q. Finally, add one extra point z to $\mathcal{A}$ that is higher than all other points (to serve as anchor) and lies on the same line as all points of $\mathcal{A}$. All points in $\mathcal{A} \cup \mathcal{B}$ lie on $d+1$ parallel lines perpendicular to the hyperplane $x_d = 0$. By construction, the $(d + 1)$-SUM instance has a TRUE value if and only if there exists a hyperplane H defined by d vertices, one from each set $\mathcal{B}_i$, and a vertex $a \in \mathcal{A}$, where $a \neq z$, that lies in H.

We solve the separability problem twice, for two symbolically perturbed versions of $\mathcal{A}$ and $\mathcal{B}$. In particular, let $\mathbf{v}$ be the unit vector in direction x_d. For a given real parameter $\epsilon > 0$, denote by $\mathcal{A}^{+\epsilon}$ the set $\{a + \epsilon\mathbf{v} \mid a \in \mathcal{A}\}$; this is the result of slightly shifting the entire set $\mathcal{A}$ in direction $\mathbf{v}$. Define sets $\mathcal{A}^{-\epsilon}$, $\mathcal{B}^{+\epsilon}$, and $\mathcal{B}^{-\epsilon}$ analogously.

We assign every point in $\mathcal{A} \cup \mathcal{B}$ a probability of $1/2$, except z, which is assigned probability 1. We then compute the probability that $\mathcal{A}^{+\epsilon}$ is separable from $\mathcal{B}^{-\epsilon}$ by a non-vertical hyperplane H. We use Simulation of Simplicity [15] to compute the result for an infinitesimal perturbation value ϵ. An algorithm with running time $T(n)$ on ordinary points will run in time $O(T(n))$ on the symbolically perturbed points. Similarly, we compute the probability of separability for $\mathcal{A}^{-\epsilon}$ and $\mathcal{B}^{+\epsilon}$.

If there exists a hyperplane defined by d points of $\mathcal{B}$ that contains a point of $\mathcal{A}$, then the probability values returned by the two computations will differ—the $(d+1)$-tuple is strictly separable in $(\mathcal{A}^{+\epsilon}, \mathcal{B}^{-\epsilon})$ and strictly not separable in $(\mathcal{A}^{-\epsilon}, \mathcal{B}^{+\epsilon})$ (because z must lie above the hyperplane). If no such hyperplane exists, then the probability values will be equal, because the only sets $A \subseteq \mathcal{A}$, $B \subseteq \mathcal{B}$ whose separation probabilities are affected by the perturbation are those containing such a hyperplane.

By computing a separation probability twice, we solve an instance of $(d+1)$-SUM: the $(d+1)$-SUM instance is TRUE if and only if the two probabilities are not equal. Thus d-dimensional probabilistic separability is at least as hard as $(d+1)$-SUM. ◀

The reduction from $(d+1)$-SUM to our problem is evidence that the algorithm of Section 3 is nearly optimal in two dimensions, and that an algorithm with running time $n^{o(d)}$ is unlikely for $d > 2$. Finally, we prove that the problem is $\#P$-hard if d can be as large as $\Omega(n)$.

▶ **Lemma 13.** *Computing $\mathbf{Pr}\left[\sigma(\mathcal{A}, \mathcal{B})\right]$ is $\#P$-hard if the dimension d is not a constant.*

Proof. We reduce the $\#P$-hard problem of counting independent sets in a graph [31] to the separability problem. Consider an undirected graph $G = (V, E)$ on the vertex set $\{1, 2, \ldots, n\}$. For each i, we construct an n-dimensional point $a_i = (0, \ldots, 1, \ldots, 0)$, namely, the unit vector along the ith axis. The collection of points $\{a_1, \ldots, a_i, \ldots, a_n\}$, each with associated probability $\pi_i = 1/2$, is our point set $\mathcal{A}$. Next, for each edge $e = (i, j) \in E$, we construct a point b_{ij} at the midpoint of the line segment connecting a_i and a_j. The set of points b_{ij}, each with associated probability 1, is the set $\mathcal{B}$. It is easy to see that there is a one-to-one correspondence between separable subsets of $\mathcal{A} \cup \mathcal{B}$ and the independent sets of G. Each separable sample occurs precisely with probability $(1/2)^n$, and therefore we can count the number of independent sets using the separation probability $\mathbf{Pr}\left[\sigma(\mathcal{A}, \mathcal{B})\right]$. ◀

6 Convexity, Halfspace Emptiness and Related Problems

Given the fundamental role of hyperplanes in geometry, it is not surprising that many other problems can be reduced to hyperplane separability of points, possibly in a transformed space. In the following, we discuss a few sample problems that can be solved by reducing them to hyperplane separability of point sets. Full details appear in Appendix B.

Convex Hull Membership.

Given a probabilistic set of points $\mathcal{P}$, the convex hull membership probability of a query point z is the probability that z lies in the convex hull of $\mathcal{P}$. We write this as

$$\mathbf{Pr}\left[z \in CH(\mathcal{P})\right] \;=\; \sum_{P \subseteq \mathcal{P},\; z \in CH(P)} \mathbf{Pr}\left[P\right].$$

We show that convex hull membership in $\mathbb{R}^d$ is equivalent to point set separability in $\mathbb{R}^{d-1}$. This gives the following bound, which improves by a factor of n the bound in [6].

▶ **Theorem 14.** *Given a probabilistic set of n points $\mathcal{P}$ in general position in $\mathbb{R}^d$, for any fixed $d \geq 3$, and a query point z in general position with $\mathcal{P}$, we can compute the convex hull membership probability $\mathbf{Pr}\left[z \in CH(\mathcal{P})\right]$ in time $O(n^{d-1})$.*

Our lower bounds imply that the time complexity is nearly optimal for $d = 3$, and that the convex hull membership problem is also $\#P$-hard when $d = \Omega(n)$.

Halfspace Emptiness and Linear Programming.

Suppose we are given a set of n probabilistic halfspaces in $\mathbb{R}^d$, defined by a set of hyperplanes $\mathcal{H}$, where each hyperplane H is associated with an independent probability $\pi(H)$. What is the probability that a random sample of these halfspaces has non-empty common intersection? By using an order-preserving duality between points and hyperplanes, we can map this problem to an instance of point set separability, obtaining the following result:

▶ **Theorem 15.** *Given a set of n probabilistic halfspaces in $\mathbb{R}^d$, we can compute the probability that their common intersection is non-empty in time $O(n^d)$.*

7 Concluding Remarks

We considered the problem of hyperplane separability for probabilistic point sets. Our main result is that given two sets of n probabilistic points in $\mathbb{R}^d$, we can compute in $O(n^d)$ time the exact probability that their random samples are linearly separable. The same technique and result lead to similar bounds for several other problems, including the probability that a query point lies inside the convex hull of n probabilistic points, or the probability that n probabilistic halfspaces have non-empty intersection. One of the interesting connections we establish is the equivalence between d-dimensional hyperplane separability and $(d+1)$-dimensional convex hull containment. Another useful feature of our approach is its ability to handle degeneracies in input.

We also proved that the d-dimensional separability problem is at least as hard as the $(d+1)$-SUM problem [9, 16, 17, 18], which implies that our $O(n^2)$ algorithms for 2-dimensional separability or 3-dimensional convex hull membership are nearly optimal.

A number of open problems are suggested by our work. Our lower bounds suggest that an exponential dependence on d is probably unavoidable for the exact computation of separability, but better bounds may be possible for $d > 2$. In particular, can the hyperplane separability problem be solved with running time $\tilde{O}(n^{\lceil (d+1)/2 \rceil})$, instead of $O(n^d)$? Another important direction is to explore algorithms that can compute the probability within a small *multiplicative* factor.

References

1 A. Abdullah, S. Daruki, and J. M. Phillips. Range counting coresets for uncertain data. In *Proc. 29th SCG*, pages 223–232. ACM, 2013.
2 P. Afshani, P. K. Agarwal, L. Arge, K. G. Larsen, and J. M. Phillips. (Approximate) uncertain skylines. *Theory of Comput. Syst.*, 52:342–366, 2013.
3 P. K. Agarwal, B. Aronov, S. Har-Peled, J. M. Phillips, K. Yi, and W. Zhang. Nearest neighbor searching under uncertainty II. In *Proc. 32nd ACM PODS*, pages 115–126, 2013.
4 P. K. Agarwal, S. W. Cheng, and K. Yi. Range searching on uncertain data. *ACM Trans. on Algorithms*, 8(4):43:1–43:17, 2012.
5 P. K. Agarwal, A. Efrat, S. Sankararaman, and W. Zhang. Nearest-neighbor searching under uncertainty. In *Proc. 31st ACM PODS*, pages 225–236. ACM, 2012.
6 P. K. Agarwal, S. Har-Peled, S. Suri, H. Yıldız, and W. Zhang. Convex hulls under uncertainty. In *Proc. 22nd ESA*, pages 37–48, 2014.
7 C. C. Aggarwal. *Managing and Mining Uncertain Data*. Springer, 2009.
8 C. C. Aggarwal and P. S. Yu. A survey of uncertain data algorithms and applications. *IEEE TKDE.*, 21(5):609–623, 2009.
9 N. Ailon and B. Chazelle. Lower bounds for linear degeneracy testing. *J. ACM*, 52(2):157–171, 2005.

10 C. Böhm, F. Fiedler, A. Oswald, C. Plant, and B. Wackersreuther. Probabilistic skyline queries. In *Proc. CIKM*, pages 651–660, 2009.
11 K. L. Clarkson. Las Vegas algorithms for linear and integer programming when the dimension is small. *J. ACM*, 42(2):488–499, 1995.
12 M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars. *Computational Geometry: Algorithms and Applications*. Springer, 3rd edition, 2008.
13 M. de Berg, A. D. Mehrabi, and F. Sheikhi. Separability of imprecise points. In *Proc. 14th SWAT*, pages 146–157. Springer, 2014.
14 H. Edelsbrunner and L. J. Guibas. Topologically sweeping an arrangement. *J. Comput. Syst. Sci.*, 38(1):165–194, 1989.
15 H. Edelsbrunner and E. P. Mücke. Simulation of simplicity: A technique to cope with degenerate cases in geometric algorithms. *ACM Trans. on Graphics*, 9(1):66–104, 1990.
16 J. Erickson. Lower bounds for linear satisfiability problems. *Chicago J. Theoret. Comp. Sci.*, 1999(8), August 1999.
17 A. Gajentaan and M. H. Overmars. On a class of $O(n^2)$ problems in computational geometry. *CGTA*, 5(3):165–185, 1995.
18 A. Gronlund and S. Pettie. Threesomes, degenerates, and love triangles. In *Proc. 55th FOCS*, pages 621–630, 2014.
19 A. Jørgensen, M. Löffler, and J. M. Phillips. Geometric computations on indecisive and uncertain points. *CoRR*, abs/1205.0273, 2012.
20 P. Kamousi, T. M. Chan, and S. Suri. Stochastic minimum spanning trees in Euclidean spaces. In *Proc. 27th SCG*, pages 65–74, 2011.
21 P. Kamousi, T. M. Chan, and S. Suri. Closest pair and the post office problem for stochastic points. *CGTA*, 47(2):214–223, 2014.
22 T. Kopelowitz, S. Pettie, and E. Porat. Higher lower bounds from the 3SUM conjecture. In *SODA*, 2016.
23 H. P. Kriegel, P. Kunath, and M. Renz. Probabilistic nearest-neighbor query on uncertain objects. In *Advances in Databases: Concepts, Systems and Applications*, volume 4443, pages 337–348. Springer, 2007.
24 Y. Li, J. Xue, A. Agrawal, and R. Janardan. On the arrangement of stochastic lines in $\mathbb{R}^2$. Unpublished manuscript (personal communication), 2015.
25 N. Megiddo. Linear programming in linear time when the dimension is fixed. *J. ACM*, 31(1):114–127, 1984.
26 C. F. Olson. Probabilistic indexing for object recognition. *IEEE PAMI*, 17(5):518–522, 1995.
27 J. S. Provan and M. O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. Comput.*, 12(4):777–788, 1983.
28 S. Suri and K. Verbeek. On the most likely Voronoi diagram and nearest neighbor searching. In *Proc. 25th ISAAC*, pages 338–350, 2014.
29 S. Suri, K. Verbeek, and H. Yıldız. On the most likely convex hull of uncertain points. In *Proc. 21st ESA*, pages 791–802, 2013.
30 S. P. Vadhan. The complexity of counting in sparse, regular, and planar graphs. *SIAM J. Comput.*, 31:398–427, 1997.
31 L. G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comput.*, 8(3):410–421, 1979.

Appendix

A Omitted Proofs and Details

A.1 Proof of Lemma 4

If $A = \emptyset$, then $\mathsf{C}(A, B) = CH(B)$, which is a convex polytope. If $A \neq \emptyset$, then $\mathsf{C}(A, B)$ certainly includes all of $CH(B)$, since each point $b \in CH(B)$ belongs to $shadow(a, b)$ for every point $a \in CH(A)$. For any pair (a, b) with $a \in CH(A)$, $b \in CH(B)$, $shadow(a, b)$ intersects a face of $CH(B)$, and $shadow(b, a)$ intersects a face of $CH(A)$. If $\{a_1, \ldots, a_k\} \subseteq A$ and $\{b_1, \ldots, b_l\} \subseteq B$ are the vertices of these faces, then it is not hard to see that the part of $shadow(a, b)$ outside $CH(B)$ is contained in the convex hull of $\cup_{1\leq i\leq k, 1\leq j\leq l} shadow(a_i, b_j)$. Thus $shadow(a, b) \subseteq CH(U)$, for $U = \cup_{u\in A, v\in B} shadow(u, v)$. Each ray $shadow(u, v)$ is the convex hull of one finite point and one point at infinity. The convex hull of U is therefore a convex hull of a finite number of points, and hence is a convex polytope.

A.2 Proofs for Section 3.2

We repeat the lemma statements from Section 3.2:

▶ **Lemma 6.** *Let C be a d-dimensional convex polyhedron and z a point not contained in C. Then there is a unique point $p \in C$ that minimizes the distance to z, and a unique face of C whose relative interior contains p.*

Proof. The existence of p is a well-known fact in convex geometry. For the second part, observe that if p lies in the closure of multiple faces, then we simply pick the face of the smallest dimension containing p in its interior. By convention, a 0-dimensional face, namely, a vertex, is defined to be its own relative interior. This completes the proof. ◀

▶ **Lemma 7.** *Let C be a d-dimensional convex polyhedron, z a point not contained in C, and p the point of C at minimum distance from z. If p lies in the relative interior of the face F of C, then the hyperplane H through p that is orthogonal to $p - z$ contains F. This hyperplane contains C in one of its closed halfspaces, and is the hyperplane farthest from z with this property.*

Proof. Since p lies in the relative interior of F, if the face is not entirely contained in H, then an ϵ-neighborhood of F around p contains a point closer to z than p, which contradicts the choice of p.

The distance from z to H is $|p - z|$, since H is orthogonal to $p - z$. Because $p = \mathrm{np}(z, C)$, no halfspace containing C can be farther from z than $|p - z|$. In fact, H is unique: any separating hyperplane H' that does not pass through p is closer to z than H along the segment $\overline{zp}$, and any separating hyperplane $H' \neq H$ passing through p is closer to z than H in the neighborhood of p. ◀

B Convexity, Halfspace Emptiness and Related Problems

Considering the fundamental role of hyperplanes in geometry, it is not surprising that many other problems can be reduced to hyperplane separability of points, possibly in a transformed space. In the following, we discuss a few sample problems whose complexity can be expressed in terms of $T(n, d)$, the time needed to compute the hyperplane separability of n probabilistic points in d-space.

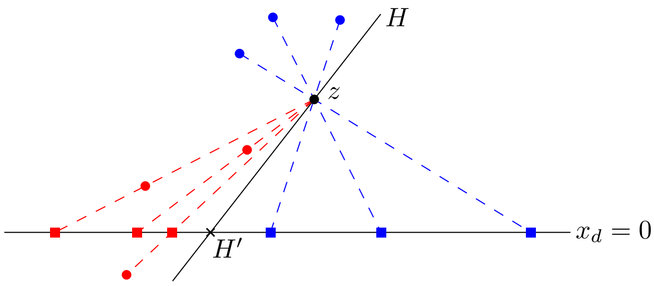


Figure 5 The central projection. Mapping 2-dimensional convex hull membership problem to 1-dimensional separability.

B.1 Convex Hull Membership

Given a probabilistic set of points $\mathcal{P}$, the convex hull membership probability of a query point z is the probability that z lies in the convex hull of $\mathcal{P}$. We write this as

$$\mathbf{Pr}\big[z \in CH(\mathcal{P})\big] \; = \sum_{P \subseteq \mathcal{P},\; z \in CH(P)} \mathbf{Pr}\big[P\big].$$

Without loss of generality, assume that the query point is $z = (0, 0, \ldots, 0, 1)$. We further assume that none of the points of $\mathcal{P}$ has dth coordinate equal to 1, which is easily achieved by a rotation of the space. As a result, none of the lines pz, for $p \in \mathcal{P}$, is parallel to the hyperplane $x_d = 0$.

Given a point $p \in \mathcal{P}$, we define its *central projection* as the point p' at which the line pz meets the plane $x_d = 0$; see Figure 5. Let set $\mathcal{A}$ (resp. $\mathcal{B}$) be the central projections of all those points in $\mathcal{P}$ with $x_d > 1$ (resp. with $x_d < 1$), where each point inherits the associated probability of its corresponding point in $\mathcal{P}$. The sets $\mathcal{A}$ and $\mathcal{B}$ are sets of $(d-1)$-dimensional probabilistic points, with $|\mathcal{A}| + |\mathcal{B}| = n$.

▶ **Lemma 16.** $\mathbf{Pr}\big[z \in CH(\mathcal{P})\big] \; = \; 1 - \mathbf{Pr}\big[\sigma(\mathcal{A}, \mathcal{B})\big]$.

Proof. Consider a random sample $P \subseteq \mathcal{P}$. This sample is associated with a unique pair of sets $A \subseteq \mathcal{A}, B \subseteq \mathcal{B}$. Specifically, A is the projection of those points in P that are higher than z (in the dth dimension) and B is the projection of points in P lower than z. We argue that $z \notin CH(P)$ iff A and B are linearly separable in the plane $x_d = 0$.

If $z \notin CH(P)$, then there exists a hyperplane H passing through z that contains all points of P in an open halfspace. If H is horizontal, namely, parallel to $x_d = 0$, then clearly either A or B is empty, and separability holds trivially. Otherwise, consider the intersection of H with $x_d = 0$, which is a $(d-1)$-dimensional hyperplane H'. It is easy to see that all points of P with $x_d > 1$ map to one side of H', creating A, while those with $x_d < 1$ map to the other side, creating B; compare Figure 5. Therefore, A and B are linearly separable by H'.

Conversely, if A and B are linearly separable by H', then the d-dimensional hyperplane H spanned by H' and z certifies that $z \notin CH(P)$. This completes the proof. ◀

If the input set $\mathcal{P} \cup \{z\}$ is in general position in $\mathbb{R}^d$, then so are the sets $\mathcal{A}$ and $\mathcal{B}$ in $\mathbb{R}^{d-1}$. We show the following equivalence. Hence, as a consequence of the equivalence of the probabilities, we can compute the convex hull membership probability in $O(n^{d-1})$ time, improving by a factor of n the bound achieved in [6].

▶ **Theorem 17.** *Given a probabilistic set of n points $\mathcal{P}$ in general position in $\mathbb{R}^d$, and a query point z in general position with $\mathcal{P}$, we can compute the convex hull membership probability* $\mathbf{Pr}\big[z \in CH(\mathcal{P})\big]$ *in time $O(n^{d-1})$.*

Note that the reduction in Lemma 16 does not rely on general position. Hence, by using the improved algorithm for linear separability of degenerate inputs presented in Section 4, the result also holds without the requirement that the input be in general position.

The central projection employed in the preceding proof can be used to prove the converse result as well: d-dimensional hyperplane separability is equivalent to convex hull membership in dimension $d+1$. In particular, given an instance of hyperplane separability $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^d$, we create an instance in $\mathbb{R}^{d+1}$ as follows: add a new (anchor) point $z = (0, 0, \ldots, 0, 1)$; perform a central *lifting* that maps each point $a \in \mathcal{A}$ to a point a' on the line az with the $(d+1)$st coordinate greater than 1 (above z), while each point $b \in \mathcal{A}$ maps to a point b' on the line bz with the $(d+1)$st coordinate less than 1 (below z). The anchor point z is outside the convex hull of these lifted points if and only if the original points are linearly separable.

We next show that the probability of non-empty intersection for n probabilistic halfspaces in d-space can also be solved in $O(n^d)$ time, but first we introduce a variation of separability, called *oriented hyperplane separability*, which is both useful in its own right and also needed to solve the halfspace emptiness problem.

B.2 Oriented Hyperplane Separability

Given two probabilistic sets of points $\mathcal{A}$ and $\mathcal{B}$ in $\mathbb{R}^d$, we want the probability that they are separable by a hyperplane that contains the points of $\mathcal{A}$ *above* it. In other words, what is the probability that random samples $A \subseteq \mathcal{A}, B \subseteq \mathcal{B}$ admit a non-vertical hyperplane H with $A \subset H^+$, where H^+ is the halfspace containing the point $(0, 0, \ldots, 0, \infty)$?

Oriented separability is easily reduced to hyperplane separability by adding an extra point $z = (0, 0, \ldots, 0, M)$ to $\mathcal{A}$, with probability 1, where M is sufficiently large to ensure that z lies above any non-vertical hyperplane determined by a d-tuple among $\mathcal{A} \cup \mathcal{B}$. Thus, the probability of oriented separability can also be computed in time $O(n^d)$.

▶ **Theorem 18.** *Given two probabilistic point sets in d-space with a total of n points, we can compute the probability of their oriented separability in time $O(n^d)$.*

It turns out that the *converse* is also true: if oriented separability can be solved in time $T(n, d)$, then hyperplane separability can also be solved in time $O(T(n, d))$, using the following observation: if $A \subseteq \mathcal{A}$ and $B \subseteq \mathcal{B}$ admit two separating hyperplanes H_1, H_2 such that $A \subseteq H_1^+$ but $B \subseteq H_2^+$, then there also exists a *vertical* hyperplane separating A and B. Thus, we can compute the (unoriented) hyperplane separability of $\mathcal{A}$ and $\mathcal{B}$ by solving two instances of d-dimensional oriented separability and one instance of $(d-1)$-dimensional unoriented hyperplane separability.

B.3 Halfspace Emptiness and Linear Programming

Suppose we are given a set of n probabilistic halfspaces in $\mathbb{R}^d$, defined by a set of hyperplanes $\mathcal{H}$, where each hyperplane H is associated with an independent probability $\pi(H)$. What is the probability that a random sample of these halfspaces has non-empty common intersection? (In the language of linear programming, what is the probability that a set of n probabilistic linear inequalities is feasible?)

Without loss of generality, we assume that none of the defining hyperplanes is vertical; otherwise we rotate the space slightly. Let $\mathcal{H}_\mathcal{A} \subseteq \mathcal{H}$ be the hyperplanes whose halfspaces face downward, namely, include the point $(0, 0, \ldots, -\infty)$, and let $\mathcal{H}_\mathcal{B} \subseteq \mathcal{H}$ be the remaining (upward facing) hyperplanes. We make use of the standard point-hyperplane duality to map these hyperplanes into points. In this transformation, a point $\mathbf{a} = (a_1, a_2, \ldots, a_d)$ maps to the hyperplane $\mathbf{a}^*$ given by the equation $x_d = a_1x_1 + a_2x_2 + \ldots + a_{d-1}x_{d-1} - a_d$. Conversely, a hyperplane H given by the equation $x_d = a_1x_1 + a_2x_2 + \ldots + a_{d-1}x_{d-1} - a_d$ maps to the point $H^* = (a_1, a_2, \ldots, a_d)$. The transform is order-preserving in the following sense: if point $\mathbf{a}$ lies *above* a hyperplane H, then the *dual* hyperplane $\mathbf{a}^*$ lies above the dual point H^*.

Let $\mathcal{A}$ be the set of points that are duals of the hyperplanes of $\mathcal{H}_{\mathcal{A}}$ and $\mathcal{B}$ the set of points that are duals of the hyperplanes of $\mathcal{H}_{\mathcal{B}}$. Now, the set of halfspaces defined by a random sample from $\mathcal{H}$ has a non-empty intersection if and only if there exists a point p that has (1) all the sampled hyperplanes of $\mathcal{H}_{\mathcal{A}}$ *above* it, and (2) all the sampled hyperplanes of $\mathcal{H}_{\mathcal{B}}$ *below* it. In dual space, this is equivalent to the existence of a hyperplane H that contains all the sampled points of $\mathcal{A}$ *above* and all the sampled points of $\mathcal{B}$ *below* it. (Indeed the dual of the point p is such a hyperplane.) However, this is exactly the problem of oriented separability for the point sets $\mathcal{A}, \mathcal{B}$, as described above. It can be solved in time $T(n, d)$, giving us the following result.

▶ **Theorem 19.** *Given a set of n probabilistic halfspaces in $\mathbb{R}^d$, we can compute the probability that their common intersection is non-empty in time $O(n^d)$.*