Kernelization of the Subset General Position Problem in Geometry

Jean-Daniel Boissonnat[1], Kunal Dutta[2], Arijit Ghosh[3], and Sudeshna Kolay[4]

1 **INRIA Sophia Antipolis - Méditerranée, France**
2 **INRIA Sophia Antipolis - Méditerranée, France**
3 **Indian Statistical Institute, Kolkata, India**
4 **Eindhoven University of Technology, Eindhoven, The Netherlands**

Abstract

In this paper, we consider variants of the Geometric Subset General Position problem. In defining this problem, a geometric subsystem is specified, like a subsystem of lines, hyperplanes or spheres. The input of the problem is a set of n points in $\mathbb{R}^d$ and a positive integer k. The objective is to find a subset of at least k input points such that this subset is in general position with respect to the specified subsystem. For example, a set of points is in general position with respect to a subsystem of hyperplanes in $\mathbb{R}^d$ if no $d+1$ points lie on the same hyperplane. In this paper, we study the Hyperplane Subset General Position problem under two parameterizations. When parameterized by k then we exhibit a polynomial kernelization for the problem. When parameterized by $h = n - k$, or the dual parameter, then we exhibit polynomial kernels which are also tight, under standard complexity theoretic assumptions. We can also exhibit similar kernelization results for d-Polynomial Subset General Position, where a vector space of polynomials of degree at most d are specified as the underlying subsystem such that the size of the basis for this vector space is b. The objective is to find a set of at least k input points, or in the dual delete at most $h = n-k$ points, such that no $b+1$ points lie on the same polynomial. Notice that this is a generalization of many well-studied geometric variants of the Set Cover problem, such as Circle Subset General Position. We also study general projective variants of these problems. These problems are also related to other geometric problems like Subset Delaunay Triangulation problem.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Incidence Geometry, Kernel Lower bounds, Hyperplanes, Bounded degree polynomials

Digital Object Identifier 10.4230/LIPIcs.MFCS.2017.25

1 Introduction

In the *geometric subset general position* problem, the input is a *family* of *algebraic objects*, e.g. lines, circles, hyperplanes, zero set of quadratic functions, and a point set P in $\mathbb{R}^d$. The objective is to extract a large subset S of P such that the subset S is in *general position* with respect to the geometric objects. The definition of general position is different for different families of geometric objects. For the case of hyperplanes in $\mathbb{R}^d$, a set S, assume $|S| > d$, will be in general position with respect to the family of hyperplanes in $\mathbb{R}^d$ if no more than d points of S lie on a hyperplane. For the case of spheres in $\mathbb{R}^d$, a set S with $|S| > d+1$, will be in general position with respect to the family of spheres in $\mathbb{R}^d$ if no more than $d+1$ points of S lie on a sphere. In this paper, we will assume that d is a constant.

In computational geometry it is generally assumed that the point set is in general position, such as no more than d points lie on a hyperplane in the case of convex hull computation or no more than $d+1$ points lie on a sphere for Delaunay triangulation computation (see [6]). Also, algebraic techniques like *simulation of simplicity* have been introduced to handle degenerate cases in practice [10].

The problem of determining whether a given point set in $\mathbb{R}^d$ is in general position with respect to family of spheres and family of hyperplanes has been extensively studied in computational geometry. Edelsbrunner, O'Rourke and Seidel [11] gave an $O(n^d)$ (and $O(n^{d+1})$) space and time complexity algorithm to determine if a point set is in general position with respect to hyperplanes (resp. spheres) in $\mathbb{R}^d$. Edelsbrunner and Guibas [8, 9] later improved the space bound to $O(n)$. Erickson and Seidel [12, 13] showed in the worst case $\Omega(n^d)$ (and $\Omega(n^{d+1})$) sided queries are required to determine whether a set of n points in $\mathbb{R}^d$ is in general position with respect to hyperplanes (resp. spheres). We have mentioned a small sample of the papers on this topic and for a more complete picture of this area and the more general problem of *arrangement of hyperplanes* please refer to the survey by Agarwal and Sharir [1].

More recently, the problem of finding a maximum-cardinality subset of points in general position has been studied in parameterized complexity, approximation algorithm, and combinatorial geometry [14, 18, 4]. Payne et al. [18] and Cardinal [4] gave a non-trivial lower bound on the size of a largest size subset in general position from a point set with bounded *coplanarity* in $\mathbb{R}^2$ and $\mathbb{R}^d$. A point set in $\mathbb{R}^2$ (and $\mathbb{R}^d$) has bounded coplanarity if the number of points from the set that lie on a given plane (or hyperplane) is bounded. Cao [3] and Froese et al. [14] studied the geometric subset general position problem in $\mathbb{R}^2$ with respect to lines in $\mathbb{R}^2$ through the lens of parameterized complexity. Cao [3] also gave an $\mathcal{O}(\sqrt{\mathsf{opt}})$-factor approximation algorithm for the general position subset selection problem with respect to lines in $\mathbb{R}^2$.

In this paper we generalize the results of [14] by studying the kernelization aspect of the following *primal* problem:

> HYPERPLANE SUBSET GENERAL POSITION **Parameter:** k
> **Input:** An n point set P in $\mathbb{R}^d$, for a fixed constant d, and a positive integer k
> **Question:** Is there a subset $S \subseteq P$ of size at least k such that S is in general position with respect to hyperplanes in $\mathbb{R}^d$?

We will also study a more general version of the above problem for *bounded degree polynomial families* (see Section 2 and 3):

> d-POLYNOMIAL SUB. GENERAL POS. **Parameter:** k
> **Input:** A set P of n points in $\mathbb{R}^d$, for a fixed constant d, a bounded degree polynomial family $\mathcal{F}$ in $\mathbb{R}^d$ and a positive integer k
> **Question:** Is there a subset $S \subseteq P$ of size at least k such that S is in general position with respect to $\mathcal{F}$ in $\mathbb{R}^d$?

Therefore, the general position subset selection problems with respect to natural set families like the vector space of spheres, ellipses, etc. are special cases of the d-POLYNOMIAL SUBSET GENERAL POSITION problem. We also study the problems with respect to the dual parameter $h = n - k$. That is, the problem of HYPERPLANE SUBSET GENERAL POSITION or d-POLYNOMIAL SUBSET GENERAL POSITION still have the same input and aim for the same decision problem. However, the parameter for the problem becomes h.

Note that both these problems are NP-hard following from the results of [14] on SUBSET GENERAL POSITION in $\mathbb{R}^2$.

Our contribution

In this paper, we exhibit polynomial kernels for Hyperplane Subset General Position in $\mathbb{R}^d$. This is a generalization to higher dimensions of the results on kernelization obtained in [14], with more carefully designed reduction rules to take care of the higher dimension. We further generalize the result with the help of a variant of the Veronese mapping, to obtain polynomial kernels for d-Polynomial Subset General Position in $\mathbb{R}^t$, where the bounded degree polynomial family is a vector space of d-degree polynomials. Special cases of the d-Polynomial Subset General Position problem include variants where the polynomial family is that of spheres or quadratic surfaces. Also, Delaunay Subset Selection is a special case of this problem. We further study the general projective variants of these problems. These results are described in Section 3

We also give tight polynomial kernels for Hyperplane Subset General Position in $\mathbb{R}^d$ parameterized by h, the dual parameter, as described in Section 4. In Section 4, we obtain tight results for the number of elements in a polynomial kernel for Hyperplane Subset General Position in $\mathbb{R}^d$. These results are similar to those obtained in [15]. Finally, in Section 5, we are able to generalize this result for certain variants of d-Polynomial Subset General Position.

2 Preliminaries

Hypergraphs

A set of consecutive integers $\{1, 2, \dots n\}$ will be written as $[n]$ in short. A hypergraph G is a set system where $V(G)$ denotes the universe and $E(G)$ denotes the family of sets. We refer to the objects in the universe $V(G)$ by either vertices or elements, and each subset of $E(G)$ as a *hyperedge*. For a hyperedge $e \in E(G)$ the set of vertices belonging to e is denoted as V_e. A *d-uniform hypergraph* is a hypergraph where each hyperedge has exactly d vertices. Similarly, a *d-hypergraph* is a hypergraph where each hyperedge has at most d vertices. An *independent set* in a hypergraph G is a subset $I \subseteq V(G)$ such that there is no $e \in E(G)$ where all vertices in e belong to I. The d-Hypergraph Independent Set problem takes as input a d-hypergraph and a positive integer k and determines whether the input hypergraph has an independent set of size at least k.

The d-Hitting Set problem takes as input a d-hypergraph G and a positive integer k and determines whether there is a set $S \subseteq V(G)$ of size at most k such that for each $e \in E(G)$, $V_e \cap S \neq \emptyset$. Such a set S is called a d-hitting set.

General position in Geometry

An i-flat in $\mathbb{R}^d$ is the affine hull of $i+1$ affinely independent points. The dimension of a (possibly infinite) set of points P, denoted as $dim(P)$, is the minimum i such that the entire set P is contained in an i-flat of $\mathbb{R}^d$ [16]. We use the term hyperplanes interchangeably for $(d-1)$-flats. A set P of points in $\mathbb{R}^d$ is said to be in general position with respect to hyperplanes, if for each i-flat, $i \leq d-1$, in $\mathbb{R}^d$ there are at most $i+1$ points from P lying on the i-flat.

As described earlier, the Geometric Subset General Position problem, defined on a subsystem of geometric objects, takes as input a set of points P and a positive integer k and determines whether there is a subset of at least k points that are in general position with respect to the specified subsystem of geometric objects. For example, Hyperplane Subset General Position in $\mathbb{R}^d$ takes in a set of points in $\mathbb{R}^d$ and a positive integer k

and determines whether there is a subset of at least k points that are in general position with respect to hyperplanes in $\mathbb{R}^d$.

Similarly, we can define the notion of general position with respect to multivariate polynomials. Given a set $\{X_1, X_2, \ldots, X_t\}$ of variables, a real multivariate polynomial on these variables is of the form $P(X_1, \ldots, X_t) = \sum_{i_1,i_2,\ldots,i_t} a_{i_1 i_2 \ldots i_t} \prod_{j\in[t]} X_j^{i_j}$ where $[t] = \{1, \ldots, t\}$ and $a_{i_1 i_2 \ldots i_t} \in \mathbb{R}$. The set of all real multivariate polynomials in the variables $\{X_1, \ldots, X_t\}$ will be denoted by $\mathbb{R}[X_1, X_2, \ldots, X_t]$. The degree of such a polynomial $P(X_1, \ldots, X_t)$ is defined as $\deg(P) := \mathsf{max}\{i_1 + i_2 + \ldots + i_t \mid a_{i_1 i_2 \ldots i_t} \neq 0\}$. A polynomial is said to be a degree d polynomial is its degree is d.

In this paper, we are interested in the set/subsets of polynomials whose degree is bounded by d, for some $d \in \mathbb{N}$. In this context we define $\text{Poly}_d[X_1, \ldots, X_t] := \{f(X_1, \ldots, X_t) \in \mathbb{R}[X_1, X_2, \ldots, X_t] \mid \deg(f) \leq d\}$. Observe that $\text{Poly}_d[X_1, \ldots, X_t]$ is a vector space over $\mathbb{R}$ with the *monomials* $\left\{X_1^{i_1} \ldots X_t^{i_t} \mid 0 \leq \sum_{j=1}^t i_j \leq d\right\}$ as the basis. Notice that $\left|\left\{X_1^{i_1} \ldots X_t^{i_t} \mid 0 \leq \sum_{j=1}^t i_j \leq d\right\}\right| = \binom{d+t}{d}$.

In d-Polynomial Subset General Position in $\mathbb{R}^t$, a subspace $\mathcal{F}$ of $\text{Poly}_d[X_1, \ldots, X_t]$ is given with a basis $\{f_1(X), f_2(X), \ldots, f_b(X), 1\}$, where $X = (X_1, \ldots, X_t)$ and $f_i(X) \in \text{Poly}_d[X_1, \ldots, X_t]$, and a set of n points from $\mathbb{R}^t$. The objective is to find a subset of points in general position with respect to the vector space of polynomials, i.e., no more than b points from the subset satisfy any equation of the form $f(X) := \sum_{i=1}^b \lambda_i f_i(X) + \lambda_{b+1} = 0$, where for each $j \in \{1, \ldots, b\}$, $\lambda_j \in \mathbb{R}$ and not all the λ_j's can be zero simultaneously. Here are some concrete examples of d-Polynomial Subset General Position.

► **Example 1.**
1. Hyperplanes in $\mathbb{R}^t$ are zero sets of linear combinations of polynomials $\{X_1, \ldots, X_t, 1\}$.
2. Union of spheres and hyperplanes in $\mathbb{R}^d$ are zero sets of linear combinations of polynomials $\left\{\sum_{i=1}^t X_i^2, X_1, \ldots, X_t, 1\right\}$.
3. Polynomial surfaces with degree bounded by d are zero sets of $\text{Poly}_d[X_1, \ldots, X_t]$.
4. Quadratic surfaces are zero set of polynomials in $\text{Poly}_2[X_1, \ldots, X_t]$.

Veronese mapping

In this paper, one of our strategies for generalizing our results is to convert d-Polynomial Subset General Position in $\mathbb{R}^t$ to Hyperplane Subset General Position in $\mathbb{R}^b$ by using a variant of Veronese mapping [17] from $\mathbb{R}^t \to \mathbb{R}^b$. The Veronese mapping of a vector space $\mathcal{F}$ of d-degree polynomials will be as the following: $\Phi_{\mathcal{F}} : \mathbb{R}^t \to \mathbb{R}^b$, where for a vector $X = (X_1, \ldots, X_t)$, $\Phi_{\mathcal{F}}(X) = (f_1(X), \ldots, f_b(X))$. Observe that if $p = (p_1, \ldots, p_t)$ satisfies the equation $f(X) := \sum_{i=1}^b \lambda_i f_i(X) + \lambda_{b+1} = 0$ then, for a vector of variables $Z = (Z_1, \ldots, Z_b)$, $\Phi_{\mathcal{F}}(p)$ will also satisfy the linear equation $\sum_{j=1}^b \lambda_j Z_j + \lambda_{b+1} = 0$. In other words, for any set of points P in $\mathbb{R}^t$ and the vector space $\mathcal{F}$, the incidences between P and $\mathcal{F}$ and incidences between $\Phi_{\mathcal{F}}(P)$ and $\text{Poly}_1[Z_1, \ldots, Z_b]$ (these are hyperplanes in $\mathbb{R}^b$) are preserved under the mapping $\Phi_{\mathcal{F}}$. Also, observe that there is a bijection between polynomials in $\mathcal{F}$ and hyperplanes in $\mathbb{R}^b$.

Parameterized Complexity

The instance of a parameterized problem/language is a pair containing the actual problem instance of size n and a positive integer called a parameter, usually represented as k. The problem is said to be in FPT if there exists an algorithm that solves the problem in $f(k)n^{\mathcal{O}(1)}$ time, where f is a computable function. The problem is said to *admit a $g(k)$-sized kernel*, if there exists an polynomial time algorithm that converts the actual instance to a reduced instance of size $g(k)$, while preserving the answer. When g is a polynomial function, then the problem is said to admit a polynomial kernel. A *reduction rule* is a polynomial time procedure that changes a given instance I_1 of a problem Π to another instance I_2 of the same problem Π. We say that the reduction rule is *safe* when I_1 is a YES instance of Π if and only if I_2 is a YES instance. Readers are requested to refer [5] for more details on Parameterized Complexity.

Lower bounds in Parameterized Algorithms

There are several methods of showing lower bounds in parameterized complexity under standard assumptions in complexity theory. One such technique is *polynomial parameter transformation*. For two parameterized problems Π, Π', a polynomial time algorithm $\mathcal{A}$ is called a polynomial parameter transformation (or ppt) from Π to Π' if, given an instance (x, k) of Π, $\mathcal{A}$ outputs in polynomial time an instance (x', k') of Π' such that $(x, k) \in \Pi$ if and only if $(x', k') \in \Pi'$ and $k' \leq k^{\mathcal{O}(1)}$. By a result of [2], if Π, Π' are two parameterized problems such that Π is NP-hard, $\Pi \ \in$ NP and there exists a polynomial parameter transformation from Π to Π', then, if Π does not admit a polynomial kernel neither does Π'.

We also require a lower bound technique given in [7]. This technique links kernelization to oracle protocols.

▶ **Definition 2.** [7] Given a language L, an oracle communication protocol for L is a two-player communication protocol. The first player gets an input x and can only execute computations taking time polynomial in $|x|$. The second player is computationally unbounded, but does not know x. At the end of the protocol, the first player has to decide correctly whether $x \in L$. The cost of the protocol is the number of bits of communication from the first player to the second player.

▶ **Proposition 3.** *[7] Let $d \geq 2$ be an integer, and ϵ be a positive real number. If* co-NP $\nsubseteq$ NP*/poly, then there is no protocol of cost $\mathcal{O}(n^{d-\epsilon})$ to decide whether a d-uniform hypergraph on n vertices has a d-hitting set of at most k vertices, even when the first player is co-nondeterministic.*

As noted in [7], this implies that for any $d \geq 2$ and any positive real number ϵ, if co-NP $\nsubseteq$ NP/poly, then there is no kernel of size $k^{d-\epsilon}$ for d-HITTING SET. In general, a lower bound for oracle communication protocols for a parameterized language L gives a lower bound for kernelization for L.

Kernels: size vs. number of elements

In literature, a lower bound on the kernel means the lower bound on the size of the kernel, but not necessarily on the number of input elements in the kernel. Kratsch et al. [15] were one of the first to study lower bounds in terms of the number of input elements in the kernel. They used the results of Dell and Melkebeek [7] along with results in two dimensional geometry to build a new technique to show lower bounds for the number of input elements in a kernel for a problem. In this paper, we have adhered to the general convention by saying that a kernel has a lower bound on its size if it has a lower bound on its representation in bits, while explicitly mentioning the cases where the kernel has a lower bound on the number of input elements.

3 Kernel Upper Bounds for primal parameter

In this section, we consider the Hyperplane Subset General Position problem in $\mathbb{R}^d$ parameterized by the primal parameter k. We describe a polynomial kernelization for this problem. This method is similar to that described in [14]. However, there is an error in the analysis of kernel size in [14]. Our proof, when restricted to the case of Line Subset General Position problem gives the correct bound on the kernel. We will point out the place where there is an error in [14], while describing our proof. Moreover, using the well-known Veronese mapping, we can generalize this result to give polynomial kernels for d-Polynomial Subset General Position in $\mathbb{R}^d$ parameterized by k.

3.1 Hyperplane case

First, we consider an easy variant of the Hyperplane Subset General Position problem in $\mathbb{R}^d$, where the input point set P is such that for every subset S of P of size less than d, $\mathsf{dim}(S) = |S| - 1$. In this case, the i-flats are said to be *non-degenerate*. In this case, parameterization by k gives us a polynomial kernel by a generalization of the results obtained in [14]. For the sake of completeness, we describe the kernelization. We apply a reduction rule that bounds the coplanarity of hyperplanes in $\mathbb{R}^d$.

▶ **Reduction Rule 4.** *Given an instance (P, k) of* Hyperplane Subset General Position *in $\mathbb{R}^d$, if there is a hyperplane H with at least $(d-1)\binom{k-d}{d} + d$ points then we delete all the points in $H \cap P$ and set $k = k - d$.*

▶ **Lemma 5.** *Reduction Rule 4 is safe.*

We apply this Reduction Rule exhaustively. In the end, we know that each hyperplane can contain $\mathcal{O}(k^d)$ input points. Together with the bound on coplanarity, we will also use the following result by Cardinal et al. [4, Theorem 4.1] to get a kernel.

▶ **Theorem 6** (Cardinal et al. [4])**.** *Let P be a set of n points in $\mathbb{R}^d$ with at most ℓ cohyperplanar points, where $\ell = O(\sqrt{n})$. Then P contains a subset of size $\Omega\left((n/\log l)^{1/d}\right)$ of points in general position.*

▶ **Theorem 7.** Hyperplane Subset General Position *in $\mathbb{R}^d$, parameterized by k and with an input point set where all lower dimensional flats are non-degenerate, has a $\mathcal{O}(k^{2d})$ kernel.*

Proof. We know from Theorem 6 that for a point set of size n and *cohyperplanarity* ℓ, i.e., at most ℓ points from the point set can lie on a given hyperplane, such that $\ell \leq \sqrt{n}$, there is a point set in general position of size at least $C(\frac{n}{\log \ell})^{1/d}$ where $C = C(d)$ is a constant. Thus, when $\ell \leq \sqrt{n}$, if $C(\frac{n}{\log \ell})^{1/d} > k$, we correctly say Yes. Substituting ℓ by its upper bound of $\mathcal{O}(k^d)$, this equation is true when $n \geq \Omega(k^{d+1})$. When $n = \mathcal{O}(k^{d+1})$, then anyway we obtain a kernel of size $\mathcal{O}(k^{d+1})$. The remaining case is when $\ell > \sqrt{n}$. Then, substituting ℓ by its upper bound of $\mathcal{O}(k^d)$, we know that $n = \mathcal{O}(k^{2d})$. Thus, we obtain the required polynomial kernel. ◀

Now we consider the general problem. To design a kernel for the general problem, point subsets lying in lower dimensional flats also have to be kept in mind. The approach is to first reduce the number of points that can lie in a lower dimensional flat before we can employ a strategy similar to the kernelization of Theorem 7. First, we describe a reduction rule such that in the reduced instance, the coplanarity of each i-flat with $i \leq d$ is bounded by a function of k. This reduction rule is similar to Reduction Rule 4, except that it has to take care of point subsets lying in lower dimensional flats before it considers point subsets lying in hyperplanes of $\mathbb{R}^d$.

▶ **Reduction Rule 8.** *Let (P, k) be an instance of* Hyperplane Subset General Position *parameterized by k. Let i be the smallest integer between 2 and d, auch that there is an $(i-1)$-flat that contains at least $c(d) \cdot k^{id} + 1$ points. Then we delete all but $c(d)k^{id}$ points. Here $c(d) = 15(d-1)$ is a constant.*

▶ **Lemma 9.** *Reduction Rule 8 is safe.*

Proof sketch. We prove the correctness of the reduction rule by induction on i. In the base case, suppose $i = 2$. Suppose there is a line L containing at least $c(d)k^4 + 1$. Then by the reduction rule, we construct an instance (P', k) such that P is modified to P' by deleting all but $c(d)k^4$ points. We show that P has a k-sized set in general position if and only if P' has a k-sized set in general position. Since, $P' \subset P$, if P' has a k-sized set in general position, so does P.

In the forward direction, suppose P has a k-sized set S in general position. Let P_L be the set of points in $L \cap P$ and P'_L be the set of points in $L \cap P'$. By definition of general position, there are $\Sigma_{j \leq d}\binom{k}{j}$ flats of dimension at most d that can be formed by the points in S. Consider the intersection of these flats with the line L. Each intersection is of dimension at most 1. That is, if the intersection is not the line L itself, then it is a point in L. Thus there are at most $\Sigma_{j \leq d}\binom{k}{j}$ points of intersection. Since, the set S is in general position, at most two points from S lie on L. If there are no points or if all the points in $S \cap L$ belong to P' then S is also a k-sized set in general position for the instance (P', k). Otherwise, suppose there are at most $t \leq 2$ points from $P_L \setminus P'_L$. The number of intersection points is at most $\Sigma_{j \leq d}\binom{k}{j} < c(d)k^{2d}$. Thus there is a set $\hat{S} = \{p_t | t \leq 2\}$ on L that are not intersection points. Let $S' = S \setminus (P \cap L) \cup \hat{S}$. We show that S' is also a set in general position. Consider a flat defined by points from S'. If they do not contain points from $\hat{S}$, then they remain non-degenerate flats. Suppose a flat contains points from $\hat{S}$. Also, for the sake of contradiction, suppose the flat is degenerate. If the flat contains all the points in $\hat{S}$ then it contains the line L and therefore the points from $L \cap P$. Thus, in P the set S was not in general position, which is a contradiction. Now, suppose the flat F contains a single point, say p_1, from $\hat{S}$. Then the points from $F \cap S$ were in general position and therefore the flat was either the line L or $L \cap F$ was an intersection point. This contradicts the fact that p_1 belongs to F, as p_1 is chosen to be a point that is not an intersection point. Thus, $S' \subseteq P'$ is in general position and of size k. Note that this means that after exhaustive application of this rule all lines contain at most $\lambda_2 = c(d)k^{2d}$ points.

The arguments of the other values of i are similar but with more case analysis. The full proof can be found in the full version of this paper. ◀

This Reduction Rule is exhaustively applied and in the end the reduced instance is such that for any hyperplane in $\mathbb{R}^d$, the coplanarity is $\mathcal{O}(k^{d^2})$. Now, we can exhibit a polynomial kernel for Hyperplane Subset General Position in $\mathbb{R}^d$ parameterized by k. The proof is same as that of Theorem 7, while taking into account that the collinearity bound in this case is $\mathcal{O}(k^{d^2})$.

▶ **Theorem 10.** Hyperplane Subset General Position *in $\mathbb{R}^d$ parameterized by k has a kernel of size $\mathcal{O}(k^{2d^2})$.*

Using the techniques in the proof of Theorem 10 we can also solve the *projective version* of the general position problem. For a given point set P in $\mathbb{R}^d$, $S \subseteq P \setminus \{0\}$ is said to be in *projective general position* if no more than $d-1$ points from S lie on a hyperplane in $\mathcal{R}^d$ that passes through the origin. The parameterized version of the problem is the following:

> Projective Hyperplane Subset General Position **Parameter:** k
> **Input:** An n point set P in $\mathbb{R}^d$, for a fixed constant d, and a positive integer k
> **Question:** Is there a subset $S \subseteq P$ of size at least k such that S is in projective general position?

Notice that this problem in $\mathbb{R}^2$ is polynomial time solvable, as the problem is equivalent to asking whether the projection of the points on a unit sphere, centered at the origin, equals to at least k points where no two lie on the same line through the origin.

We will apply the following reduction rule to reduce the coplanarity of the hyperplanes passing through the origin.

▶ **Reduction Rule 11.** *Let (P,k) be an instance of* Projective Hyperplane Subset General Position *parameterized by k. Let i be the smallest integer between 2 and $d-1$, such that there is an $(i-1)$-flat passing through the origin that contains at least $c'(d) \cdot k^{id} + 1$ points. Then we delete all but $c'(d)k^{i(d-1)}$ points. Here, as in the case with Reduction Rule 8, $c'(d)$ is a large constant depending linearly on d.*

The correctness of this Reduction Rule is same as the inductive proof of Reduction Rule 8. Applying the above reduction rule exhaustively, as was the case with Hyperplane Subset General Position problem, we will get that any hyperplane passing through the origin has $O(k^{(d-1)^2})$ input points on them. To get a polynomial kernel for the Projective Hyperplane Subset General Position problem we will need the following analogue to Theorem 6 with bounded coplanarity.

▶ **Lemma 12** (Projective version of Theorem 6). *Let P be a set of n points in $\mathbb{R}^d$ such that for any hyperplane H passing through the origin, we have $|H \cap P| \leq \ell$, where $\ell = O(n^{1/3})$. Then P contains a subset of size $\Omega\left(\left(\frac{n^{2/3}}{\log \ell}\right)^{1/(d-1)}\right)$ of points in projective general position.*

Proof. The proof of this lemma will use Theorem 6. Without loss of generality we will assume that the hyperplane $X_1 = 1$, intersects all the lines passing through the origin and one point of P. Let L denotes the set of lines passing through the origin and one point of P. Observe that since for any hyperplane H passing through the origin, $|H \cap P| \leq \ell$, therefore $|L| \geq \frac{n}{\ell}$. Let P' be the set points we get by intersecting lines in L with the hyperplane $X_1 = 1$. Again observe that $|P'| \geq \frac{n}{\ell}$, and for any $(d-2)$-hyperplane H' contained in the hyperplane $X_1 = 1$, we have $|H' \cap P'| \leq \ell$, otherwise we can get a hyperplane passing through origin containing more than ℓ points from P. Using the fact that $\ell = O\left(\sqrt{n/\ell}\right)$ and Theorem 6, we get that P' contains a subset P'_1 of size $\Omega\left(\left(\frac{n^{2/3}}{\log \ell}\right)^{1/(d-1)}\right)$ with no more than ℓ points from P'_1 lying on any $(d-2)$-dimensional subflat of the hyperplane $X_1 = 1$. For each point $p' \in P'_1$, include in the set P_1 a point p from the set P such that p and p' are on the same line passing through the origin. By construction the set P_1 is in projective general position and $|P_1| = \Omega\left(\left(\frac{n^{2/3}}{\log \ell}\right)^{1/(d-1)}\right)$. ◀

Now, using the fact that any hyperplane passing through the origin has $O(k^{(d-1)^2})$ input points on them after application of Reduction Rule 11 and Lemma 12, we can prove the following result in the same way we proved Theorem 10.

▶ **Theorem 13.** *Projective Hyperplane Subset General Position in $\mathbb{R}^d$ parameterized by k has a kernel of size $\mathcal{O}(k^{3(d-1)^2})$.*

3.2 Bounded degree polynomials

The following lemma will show the direct connection between the d-Polynomial Subset General Position problem and Hyperplane Subset General Position problem:

▶ **Lemma 14.** *Let P be a set of points in $\mathbb{R}^t$, and $\mathcal{F}$ be a subspace of $\text{Poly}_d[X_1, \ldots, X_t]$ with a basis $\{f_1(X), \ldots, f_b(X), 1\}$, where $X = (X_1, \ldots, X_t)$.*

1. *If P is a set of ℓ points in general position with respect to the polynomial family $\mathcal{F}$ (defined earlier in the section) then $\Phi_{\mathcal{F}}(P)$ ($\Phi_{\mathcal{F}}$ is defined earlier in Section 2) is a set of ℓ points in general position with respect to hyperplanes in $\mathbb{R}^b$.*
2. *Let $S = \{q_1, \ldots, q_\ell\} \subseteq \Phi_{\mathcal{F}}(P)$ be a set of ℓ points in general position with respect to hyperplanes in $\mathbb{R}^b$. Then the set $S' = \{p_1, \ldots, p_\ell\}$, where $p_i \in \Phi_{\mathcal{F}}^{-1}(q_i) \cap P$, will be a set of ℓ points in general position with respect to $\mathcal{F}$.*

Proof.

1. First, observe that it is enough to show that the map $\Phi_{\mathcal{F}}$ is injective on P. In general, the map $\Phi_{\mathcal{F}}$ need not be an injective mapping on an arbitrary set of n points in $\mathbb{R}^t$. However, we show that Φ is injective when restricted to P if P is in general position with respect to $\mathcal{F}$. To reach a contradiction, let $\Phi_{\mathcal{F}}(p_1) = \Phi_{\mathcal{F}}(p_2)$ where $p_1, p_2 (\neq p_1) \in P$. Let $S \subseteq P$ be of size $b+1$ and $p_1, p_2 \in S$. Observe that the set $\Phi_{\mathcal{F}}(S)$ will have less than $b+1$ points and this will imply that there exists a hyperplane $\sum_{i=1}^{b} \lambda_i Z_i + \lambda_{b+1} = 0$ on which the set $\Phi_{\mathcal{F}}(S)$ will lie. But this implies that the polynomial $\sum_{i=1}^{b} \lambda_i f_i(X) + \lambda_{b+1} = 0$ will be satisfied by all the points in S. This is a contradiction to the fact that the point set P is in general position.
2. This result follows directly from the construction of the mapping $\Phi_{\mathcal{F}}$. ◀

With the above result and Theorem 10 we get the following theorem.

▶ **Theorem 15.** *d-Polynomial Subset General Position in $\mathbb{R}^t$ parameterized by k, has a polynomial kernel of size $\mathcal{O}(k^{2b^2})$. Here b is the size of the basis generating the underlying vector space of polynomials.*

As in the case with Theorem 10 we can also get a projective version of Theorem 15. Let $\mathcal{F}$ be a subspace of $\text{Poly}_d[X_1, \ldots, X_t]$ with basis $\{f_1(X), \ldots, f_b(X)\}$ where $X = (X_1, \ldots, X_t)$ and none of the polynomial functions $f_i(X)$ are constants. Then we can define projective analog of the general position problem for polynomial families like $\mathcal{F}$. For a given point set P in $\mathbb{R}^t$, a subset S of P will be in general position with respect to $\mathcal{F}$ if no more than $b-1$ points from S lie on any $f(X) \in \mathcal{F}$.

Using the same techniques as in the proof of Theorem 10 we will get the following result:

▶ **Corollary 16.** *Projective Polynomial Subset General Position in $\mathbb{R}^t$ parameterized by k has a kernel of size $\mathcal{O}(k^{3(b-1)^2})$.*

Upper bounds for non-vector space families

Observe that we may be interested in getting general position point set with respect to families of polynomials that are not vector spaces of $\text{Poly}_d[X_1, \ldots, X_t]$. For example, consider the case of hyperplanes in R^t of the following type $H := \sum_{i=1}^{t-1} \lambda_i X_i + X_t + \beta$, where λ_i's and β

are in $\mathbb{R}^+$. One might be interested in getting a general position set in $\mathbb{R}^t$ with respect to these hyperplanes.

Note that our upper bound on the kernel size in the primal parameter extends to these families as well.

▶ **Corollary 17.** *Let $\mathcal{F}$ be a subfamily of $\mathrm{Poly}_d[X_1, \ldots, X_t]$ such that there exists polynomial functions $f_1(X), \ldots, f_b(X)$ in $\mathrm{Poly}_d[X_1, \ldots, X_t]$ and for any $f(X) \in \mathcal{F}$, $f(X) = \sum_{i=1}^{b} \lambda_i f_i(X)$ where the λ_i's are in $\mathbb{R}$. Subset general position problem with respect to $\mathcal{F}$ parameterized by primal parameter k has a kernel of size $\mathcal{O}(k^{3(b-1)^2})$.*

4 Tight kernels for hyperplanes in dual parameter

In this Section, we show that Hyperplane Subset General Position in $\mathbb{R}^d$, parameterized by the dual parameter h, cannot have a kernel of size $h^{d-\epsilon}$ if co-NP $\nsubseteq$ NP/poly. We show this result by the standard technique of polynomial parameter transformation. For a fixed d, we reduce the d-Hitting Set problem on d-uniform graphs to the problem of Hyperplane Subset General Position in $\mathbb{R}^d$. By Proposition 3, this gives us a lower bound for Hyperplane Subset General Position in $\mathbb{R}^d$.

For the main result, we construct for each positive integer n and each d, a set of n points in $\mathbb{R}^d$ with some special properties.

▶ **Lemma 18.** *For every d-uniform hypergraph G in n vertices and m hyperedges, there is a transformation to a set $P \uplus B = \{p_1, p_2, \ldots, p_n\} \uplus \{b_1, b_2, \ldots, b_m\}$ of $n + m$ points in $\mathbb{R}^d$ that have the following properties:*

1. *The points $\{p_1, p_2, \ldots, p_n\}$ are in general position.*
2. *Each vertex $v_i \in V(G)$ is mapped to the point $p_i \in P$.*
3. *Each hyperedge $e_j \in E(G)$ is mapped to the point $b_j \in B$.*
4. *For a hyperedge $e_j \in E(G)$, if there are d points $\{p_{i_1}, p_{i_2}, \ldots, p_{i_d}\} \in P$ such that $b_j, p_{i_1}, p_{i_2}, \ldots, p_{i_d}$ are cohyperplanar, then it must be the case that $e_j = \{v_{i_1}, v_{i_2}, \ldots, v_{i_d}\}$.*
5. *For any set of $i \leq d$ points of B and $d + 1 - i$ points of P cannot be cohyperplanar.*
6. *The points in B are in general position. In other words, no $d + 1$ points in B are cohyperplanar.*

Proof Idea. The main idea behind the proof is that the sets P and B satisfying the conditions of Lemma 18 can be generated in a greedy manner from considering large grids. We will first construct the point set P of n points in a greedy manner such that P comes from a large grid and is in general position. After P is constructed, the set B is again greedily constructed, this time using a lower dimensional grid lying in a particular hyperplane. ◀

This transformation from a graph to a point set leads to the following observation.

▶ **Observation 19.** *Let G be a d-uniform hypergraph and $P \uplus B$ be the set of points in $\mathbb{R}^d$ corresponding to G. For any maximal set S of points in general position, there is a set of size at least $|S|$ that contains all the points in B.*

This helps us to design a reduction from d-Hypergraph Independent Set to Hyperplane Subset General Position in $\mathbb{R}^d$.

▶ **Lemma 20.** *There is a many-one reduction from d-*Hypergraph Independent Set *on d-uniform hypergraphs to* Hyperplane Subset General Position *in $\mathbb{R}^d$.*

Finally, we are ready to prove the main result.

▶ **Theorem 21.** Hyperplane Subset General Position *in* $\mathbb{R}^d$ *parameterized by the dual parameter cannot have a kernel of size* $\mathcal{O}(h^{d-\epsilon})$ *if* co-NP $\not\subseteq$ NP/*poly.*

Proof. We give a reduction from d-Hitting Set on d-uniform hypergraphs. Given an instance (G, h) of d-Hitting Set on d-uniform hypergraphs, we consider the equivalent d-Hypergraph Independent Set instance $(G, |V(G)| - h)$. By Lemma 20, we construct an instance $(P \uplus B, |V(G)| + |E(G)| - h)$. Note that the transformation is such that $|P| = |V(G)|$ and $|B| = |E(G)|$. Thus, G has a d-hitting set of size k if and only if G has an independent set of size $|V(G)| - h$, which by Lemma 20 happens if and only if $P \uplus B$ has a point subset of size $k' = |P \uplus B| - h$ that is in general position with respect to hyperplanes in $\mathbb{R}^d$. This means that the dual parameter $|P \uplus B| - k'$ is equal to h, which is the d-hitting set size in G. This implies the lower bound on the kernel size of Hyperplane Subset General Position in $\mathbb{R}^d$ parameterized by the dual parameter. ◀

We obtain tight polynomial kernels from the following Proposition, derived from a folklore result.

▶ **Proposition 22.** Hyperplane Subset General Position *in* $\mathbb{R}^d$ *parameterized by the dual parameter* h *has a kernel of size* h^d.

Proof. We state the folklore reduction from Hyperplane Subset General Position in $\mathbb{R}^d$ to $(d+1)$-Hitting Set. Given an instance (P, h) of Hyperplane Subset General Position in $\mathbb{R}^d$, we construct the following instance (G, h) of $(d+1)$-Hitting Set. Corresponding to each point in P we create a vertex in $V(G)$. For any $d+1$ point in P that are coplanar in $\mathbb{R}^d$, we create a hyperedge with the corresponding vertices. Consider the vertices in a hyperedge of G. At least one of the corresponding points has to be deleted in order to construct a subset of P that is in general position with respect to hyperplanes in $\mathbb{R}^d$. Thus, the set of points deleted correspond to a hitting set of G. Therefore, the reduction is correct.

$(d+1)$-Hitting Set has a kernel where the universe size if $\mathcal{O}(h^d)$ [19]. This gives us an $\mathcal{O}(h^d)$ kernel for Hyperplane Subset General Position in $\mathbb{R}^d$ parameterized by the dual parameter. ◀

Lower bounds on elements in a kernel for hyperplanes in dual parameter

In this section, we show that by the method suggested by Dell and Melkebeek [7], we can show a lower bound on the number of points in a polynomial kernel for Hyperplane Subset General Position in $\mathbb{R}^d$, for each fixed positive integer d. This result is a direct extension of the results obtained in [15] and [14].

▶ **Theorem 23.** Hyperplane Subset General Position *in* $\mathbb{R}^d$*, parameterized by* h*, cannot have a kernel with* $\mathcal{O}(h^{d-\epsilon})$ *points if* co-NP $\not\subseteq$ NP/*poly.*

5 Bounded degree polynomials and the dual parameter

In this section we discuss about the generalization of Theorems 21 and 23. Note that, for any given point set P with n points, both theorems can be proved for finding a point set of size $n - h$, h being the dual parameter, if the construction of Lemma 18 can be replicated for a particular family $\mathcal{F}$. In particular, consider a family $\mathcal{F}$ of polynomials of degree at most d with basis $\{f_1(X), \ldots, f_b(X), 1\}$, where $X = \{X_1, \ldots, X_t\}$. Suppose for each b-uniform hypergraph G with n vertices and m edges it is possible to make a transformation as follows:

1. The points $\{p_1, p_2, \ldots, p_n\}$ are in general position with respect to $\mathcal{F}$ and have bounded representation.
2. Each vertex $v_i \in V(G)$ is mapped to the point $p_i \in P$.
3. Each hyperedge $e_j \in E(G)$ is mapped to the point $b_j \in B$.
4. For a hyperedge $e_j \in E(G)$, if there are d points $\{p_{i_1}, p_{i_2}, \ldots, p_{i_b}\} \in P$ such that $b_j, p_{i_1}, p_{i_2}, \ldots, p_{i_b}$ lie on a polynomial from $\mathcal{F}$, then it must be the case that $e_j = \{p_{i_1}, p_{i_2}, \ldots, p_{i_b}\}$.
5. For any set of $i \leq b$ points of B and $b+1-i$ points of P cannot be on any polynomial of $\mathcal{F}$.
6. The points in B are in general position. In other words, no $b+1$ points in B can be on any polynomial of $\mathcal{F}$.

Let us call such a transformation a *good* transformation. Then with respect to such a family $\mathcal{F}$, equivalent tight kernelizations for the dual parameter can be given. When $\mathcal{F}$ is the family of spheres, then such a construction is possible.

▶ **Corollary 24.** *Given the family of spheres in $\mathbb{R}^d$, for each $(d+1)$-uniform hypergraph with n vertices and m points there is a good transformation to a set $P \uplus B$ of $n+m$ points.*

Proof. The construction is similar as that of Lemma 18, as we again can construct the sets greedily. The point set P can be extracted from a large enough grid as in the construction given in Lemma 18. The construction of the points in B is also done inductively. Suppose the points of a subset $B' \subset B$ have already been placed on rational points such that all the necessary conditions are satisfied. Let $b_e \in B \setminus B'$. Consider the sphere S_e defined by the d-sized point set P_e corresponding to the vertices of $e \in E(G)$. Consider the family $\mathcal{F}$ of spheres formed by (i) a set of any d points in P other than the set P_e, (ii) any d points from B', and (iii) a set S of d points with at least one point from P and at least one point from B'. The intersection of this family of spheres with S_e is a family $\mathcal{F}'$ of lower dimensional spaces. Since the points in P have bounded representation, so do the intersection spaces. It is possible to determine in polynomial time the arrangement of the lower dimensional intersections on S_e [1]. From this arrangement, we select a point with bounded representation that does not belong to any of the lower dimensional flats in $\mathcal{F}'$ and set the point to b_e. The set $B' \cup b_e$ again satisfies all the necessary conditions. We continue till all the points in B have been determined. Thus, we construct the required set $P \uplus B$. ◀

6 Open Problems

One of the major open questions in this area is regarding techniques for showing kernel lower bounds with respect to the primal parameter. Currently, no non-trivial lower bound is known for even Hyperplane Subset General Position in $\mathbb{R}^2$.

In Section 5, we gave tight lower bounds with respect to the dual parameter under some restricted vector spaces of d-degree polynomials. It will be interesting to understand the problem better for general vector spaces of d-degree polynomials. In fact, it might be useful for both algorithmic as well as combinatorial studies to understand the Geometric Subset General Position in both the primal and dual parameter for families of d-degree polynomials that are not vector spaces or a subset of a vector space. We are interested in studying families that fall outside these frameworks.

General position with respect to spheres is also connected to Delaunay Subset Selection problem where we are given a point set $P \subset \mathbb{R}^d$ as input and the problem is to extract a maximum size subset S of P such that the Delaunay complex $\mathrm{Del}(S)$ is a triangulation of the convex hull conv(S) of S. Although the upper bounds for kernelization given in this paper hold for this problem, lower bound questions remain open for both primal and dual parameter.

References

1 Pankaj K Agarwal and Micha Sharir. Arrangements and their applications. *Handbook of Computational Geometry*, pages 49–119, 2000.
2 Hans L. Bodlaender, Stéphan Thomassé, and Anders Yeo. Kernel bounds for disjoint cycles and disjoint paths. *Theor. Comput. Sci.*, 412(35):4570–4578, 2011.
3 C. Cao. Study on Two Optimization Problems: Line Cover and Maximum Genus Embedding. Master's thesis, Texas A&M University, May 2012.
4 Jean Cardinal, Csaba D. Tóth, and David R. Wood. General position subsets and independent hyperplanes in d-space. *Journal of Geometry*, pages 1–11, 2016. doi:10.1007/s00022-016-0323-5.
5 Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015.
6 Mark De Berg, Marc Van Kreveld, Mark Overmars, and Otfried Cheong Schwarzkopf. Computational geometry. In *Computational geometry*, pages 1–17. Springer, 2000.
7 Holger Dell and Dieter Van Melkebeek. Satisfiability allows no nontrivial sparsification unless the polynomial-time hierarchy collapses. In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, pages 251–260. ACM, 2010.
8 Herbert Edelsbrunner and Leonidas J. Guibas. Topologically Sweeping an Arrangement. *J. Comput. Syst. Sci.*, 38(1):165–194, 1989.
9 Herbert Edelsbrunner and Leonidas J. Guibas. Corrigendum: Topologically Sweeping an Arrangement. *J. Comput. Syst. Sci.*, 42(2):249–251, 1991.
10 Herbert Edelsbrunner and Ernst Peter Mücke. Simulation of Simplicity: A Technique to Cope with Degenerate Cases in Geometric Algorithms. *ACM Transactions on Graphics*, 9(1):66–104, 1990.
11 Herbert Edelsbrunner, Joseph O'Rourke, and Raimund Seidel. Constructing Arrangements of Lines and Hyperplanes with Applications. *SIAM J. Comput.*, 15(2):341–363, 1986.
12 Jeff Erickson and Raimund Seidel. Better Lower Bounds on Detecting Affine and Spherical Degeneracies. *Discrete & Computational Geometry*, 13:41–57, 1995.
13 Jeff Erickson and Raimund Seidel. Erratum to Better Lower Bounds on Detecting Affine and Spherical Degeneracies. *Discrete & Computational Geometry*, 18(2):239–240, 1997.
14 Vincent Froese, Iyad Kanj, André Nichterlein, and Rolf Niedermeier. Finding Points in General Position. *CCCG*, pages 7–14, 2016.
15 Stefan Kratsch, Geevarghese Philip, and Saurabh Ray. Point Line Cover: The Easy Kernel is Essentially Tight. *ACM Trans. Algorithms*, 12(3):40, 2016.
16 Stefan Langerman and Pat Morin. Covering things with things. *Discrete & Computational Geometry*, 33(4):717–729, 2005.
17 Jiří Matoušek. *Lectures on Discrete Geometry*, volume 108. Springer New York, 2002.
18 Michael S Payne and David R Wood. On the general position subset selection problem. *SIAM Journal on Discrete Mathematics*, 27(4):1727–1733, 2013.
19 René van Bevern. Towards optimal and expressive kernelization for d-hitting set. *Algorithmica*, 70(1):129–147, 2014.