Finding Points in General Position*

VINCENT FROESE[†] IYAD KANJ[‡] ANDRÉ NICHTERLEIN* ROLF NIEDERMEIER*

Abstract

We study computational aspects of the GENERAL POSITION SUBSET SELECTION problem defined as follows: Given a set of points in the plane, find a maximum-cardinality subset of points in general position. We prove that GENERAL POSITION SUBSET SELECTION is NP-hard, APX-hard, and give several fixed-parameter tractability results.

Problem definition and motivation. For a set $P = \{p_1, \dots, p_n\} \subseteq \mathbb{Q}^2$ of n points in the plane, a subset $S \subseteq P$ is in *general position* if no three points in S are *collinear* (that is, lie on the same line). Whereas a frequent assumption for point set problems in computational geometry is that the given point set is in general position, surprisingly, the problem of choosing a maximum-cardinality subset of points in general position, from a given set of points, has received little attention from the computational complexity perspective, although not from the combinatorial geometry perspective. In particular, up to our knowledge, the classical complexity of the aforementioned problem is unresolved. Formally, the decision version of the problem is:

GENERAL POSITION SUBSET SELECTION
Input: $P \subseteq \mathbb{Q}^2$ and $k \in \mathbb{N}$.
Question: Is there a subset $S \subseteq P$ in general position of cardinality at least k?

The special case of GENERAL POSITION SUBSET SELECTION, referred to as the no-three-in-line problem, which asks to place a maximum number of points in general position on an $n \times n$-grid, received considerable attention in discrete geometry. Since at most two points can be placed on any grid-line, the maximum number of points in general position that can be placed on an $n \times n$ grid is at most $2n$. Indeed, only for small n it is known that $2n$ points can always be placed on the $n \times n$ grid. Erdős [7] observed that, for sufficiently large n, one can place $(1 - \epsilon)n$ points in general position on the $n \times n$ grid, for any $\epsilon > 0$. This lower bound was improved by Hall et al. [4] to $(\frac{3}{2} - \epsilon)n$. It was conjectured by Guy and Kelly [3] that, for sufficiently large n, one can place more than $\frac{\pi}{\sqrt{3}}n$ many points in general position on the $n \times n$ grid. This conjecture remains unresolved, hinting at the challenging combinatorial nature of the no-three-in-line problem, and hence of the GENERAL POSITION SUBSET SELECTION problem as well.

Observing that the computational complexity of the closely-related POINT LINE COVER problem (given a point set in the plane, find a minimum-cardinality set of lines, the size of which is called the *line cover number*, that cover all points) was intensively studied, we try to fill the gap by providing both computational hardness and fixed-parameter tractability results for the much less studied GENERAL POSITION SUBSET SELECTION. In doing so, we particularly consider the parameters solution size k (size of the sought subset in general position) and its dual $h := n - k$, and investigate their impact on the computational complexity of GENERAL POSITION SUBSET SELECTION.

---

*This work is currently under submission to a journal. A full version is available at: http://arxiv.org/pdf/1508.01097v1.pdf

[†]Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany. Email: `<firstname>.<lastname>@tu-berlin.de`.

[‡]School of Computing, DePaul University, Chicago, USA. Email: `ikanj@cs.depaul.edu`. Supported by the DFG project DAPA (NI 369/12) during a Mercator fellowship when staying at TU Berlin.

Related Work. Payne and Wood [6] provide improved lower bounds on the size of a point set in general position, a question originally studied by [2]. In his master's thesis, Cao [1] gives a problem kernel of $O(k^4)$ points for General Position Subset Selection (there called Non-Collinear Packing problem) and a simple greedy $O(\sqrt{\text{opt}})$-factor approximation algorithm for the maximization version. He also presents an Integer Linear Program formulation for General Position Subset Selection, and shows that it is in fact the dual of an Integer Linear Program formulation for the Point Line Cover problem. As to results for the much better studied Point Line Cover, we refer to [5].

Our Contributions. We show the NP-hardness and the APX-hardness of General Position Subset Selection. Our main algorithmic results concern the power of polynomial-time data reduction for General Position Subset Selection: we give an $O(k^3)$-point kernel and an $O(h^2)$-point kernel, and show that the latter kernel is asymptotically optimal under a reasonable complexity-theoretic assumption. Table 1 summarizes our results. The NP-hardness and the APX-hardness results are obtained using reductions from variants of the Independent Set problem. The kernelization results with respect to the parameter solution size are obtained using: (1) an extremal result by Payne and Wood [6] that gives a lower bound on the the cardinality of a subset in general position in a point-set of bounded collinearity, and (2) data reduction rules. The upper bound on the kernel size with respect to the dual parameter is obtained via a reduction to the 3-Hitting Set problem, and the lower bound on the kernel size is obtained via a reduction from the Vertex Cover problem combined with techniques that are based on oracle communication protocols.

Table 1: Overview of the results we obtain for General Position Subset Selection, where n is the number of input points, k is the parameter size of the sought subset in general position, $h = n - k$ is the dual parameter, and ℓ is the line cover number.

	Results
Hardness	NP-hard
	APX-hard
	no $2^{o(n)} \cdot n^{O(1)}$-time algorithm[a]
	no $O(h^{2-\epsilon})$-point kernel[b]
Tractability	$(15k^3)$-point kernel (in $O(n^2 \log n)$ time)
	$O(n^2 \log n + 41^k \cdot k^{2k})$-time solvable
	$(120\ell^3)$-point kernel (in $O(n^2 \log n)$ time)
	$O(n^2 \log n + 41^{2\ell} \cdot \ell^{4\ell})$-time solvable
	$O(h^2)$-point kernel (in $O(n^3)$ time)
	$O(2.08^h + n^3)$-time solvable

[a] Unless the Exponential Time Hypothesis fails.
[b] Unless coNP $\subseteq$ NP/poly.

References

[1] Cheng Cao. Study on Two Optimization Problems: Line Cover and Maximum Genus Embedding. Master's thesis, Texas A&M University, May 2012. 2

[2] Paul Erdős. On some metric and combinatorial geometric problems. *Discrete Mathematics*, 60:147–153, 1986. 2

[3] Richard K. Guy and Patrick A. Kelly. The no-three-in-line problem. *Canadian Mathematical Bulletin*, 11:527–531, 1968. 1

[4] R.R Hall, T.H Jackson, A Sudbery, and K Wild. Some advances in the no-three-in-line problem. *Journal of Combinatorial Theory, Series A*, 18(3):336 – 341, 1975. 1

[5] Stefan Kratsch, Geevarghese Philip, and Saurabh Ray. Point line cover: The easy kernel is essentially tight. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '14)*, pages 1596–1606, 2014. Available on arXiv:1307.2521. 2

[6] Michael S. Payne and David R. Wood. On the general position subset selection problem. *SIAM Journal on Discrete Mathematics*, 27(4):1727–1733, 2013. 2

[7] Klaus F Roth. On a problem of Heilbronn. *Journal of the London Mathematical Society*, 1(3):198–204, 1951. 1