Electronic Colloquium on Computational Complexity, Report No. 203 (2026)

Improved Algorithms for the Remote Point Problem

Ben Lee Volk*

Abstract

The Remote Point Problem (RPP) is an algorithmic problem that asks, given a linear subspace $L \subseteq \mathbb{F}^n$ of dimension k, to deterministically find a vector $v \in \mathbb{F}^n$ far in Hamming distance from L. This problem was introduced by Alon, Panigrahy and Yekhanin [APY09], motivated in part by the matrix rigidity approach for proving circuit lower bounds. An algorithm is said to achieve remoteness d if it finds a vector v whose Hamming distance from L is at least d.

We observe that over the rational numbers, the problem admits a deterministic polynomial-time algorithm that achieves optimal remoteness $n - k$.

Over finite fields, we obtain a (modest) improvement of a result of Alon, Panigrahy and Yekhanin [APY09], and give an algorithm that achieves remoteness $\Omega\left(\frac{n}{\max\{k, \log n\}} \log n\right)$.

1 Introduction

The *Remote Point Problem* (RPP) over a field $\mathbb{F}$ is the following natural algorithmic problem: given (the basis of) a linear subspace $L \subseteq \mathbb{F}^n$ of dimension k and a natural number d, find a vector $v \in \mathbb{F}^n$ which is d-far, in Hamming distance, from all points in L.

One motivation for studying this problem comes from the matrix rigidity approach for proving algebraic circuit lower bounds: Valiant [Val77] defined a matrix $M \in \mathbb{F}^{n \times n}$ to be (r, s)-rigid if it cannot be written as a sum $M = R + S$ of a matrix R of rank at most r and a matrix S with s-sparse rows.[1] In other words, for every r-dimensional subspace L, M contains a row that is $(s+1)$-far (in Hamming distance) from L. Thus, RPP can be seen as a "white-box" version of the problem: instead of constructing a set of vectors that work for all subspaces, the algorithm gets to "see" the subspace L and to construct a bespoke d-far vector that may depend on L.

Random matrices are rigid with very good parameters, and Valiant has shown that explicit constructions of rigid matrices with good enough parameters will imply new circuit lower bounds. See [Lok09, Ram20] for surveys on this topic. Similarly, in the RPP setting, a counting argument implies that as long as the dimension of L is not too large, most vectors are indeed pretty far from L. For concreteness, consider an $n/2$-dimensional space $L \subseteq \mathbb{F}_2^n$: L has only $2^{n/2}$ vectors, so a random vector $v \in \mathbb{F}_2^n$ has distance $\Omega(n)$ from L. The problem is to construct such a vector v *deterministically*. Hence, it is also a natural question in the theory of algebraic pseudorandomness.

RPP was introduced by Alon, Panigrahy and Yekhanin [APY09], who also gave a deterministic polynomial-time algorithm that finds an $\Omega(\log n)$-far vector from $n/2$-dimensional subspaces, or more generally, an $(n \log k/k)$-far vector from k-dimensional subspaces. Their algorithm is phrased over $\mathbb{F}_2$, but can be extended to any field.

Arvind and Srinivasan discovered more connections between RPP and circuit lower bounds [AS10a], and further generalized the algorithm of Alon, Panigrahy and Yekhanin [APY09] to finite groups, and obtained a parallel polynomial-time algorithm (with similar remoteness guarantees) for abelian groups [AS10b].

*Efi Arazi School of Computer Science, Reichman University, Israel. Email: benleevolk@gmail.com. The research leading to these results has received funding from the Israel Science Foundation (grant number 843/23).

[1] Sometimes it is more convenient to bound the total sparsity of S, rather than the sparsity per-row. However, the differences between the two definitions are quite inconsequential, and for the purpose of proving circuit lower bounds it is enough to consider matrices in which every row is sparse, which is the definition that is more convenient for us.

ISSN 1433-8092

In this paper we improve those algorithms in several settings. First, we consider RPP over the field of rational numbers $\mathbb{Q}$. This problem turns out to have an optimal and very simple algorithm that relies mostly on elementary linear algebra.

Theorem 1.1. *There exists a deterministic polynomial-time algorithm that, given a basis of a k-dimensional subspace $L \subseteq \mathbb{Q}^n$, finds a vector v which is $(n-k)$-far from L.*

We remark that $(n-k)$ is the optimal remoteness parameter one could expect for a k-dimensional subspace. This follows for example from Lemma 2.1 that we prove in Section 2. While the magnitude of the coordinates of the vector v may be exponential in n, their bit complexity is merely polynomial in n (and in the bit complexity of the given basis for L). It is interesting, of course, to obtain a vector with smaller entries satisfying the remoteness property, and this might be helpful in making progress for RPP over finite fields.

We find it quite surprising that such a simple and optimal algorithm exists for RPP over $\mathbb{Q}$. While, to the best of our knowledge, RPP over $\mathbb{Q}$ was not explicitly considered in the literature before, constructing rigid matrices over $\mathbb{Q}$ is a well-known and equally interesting open problem as in the case of finite fields (we refer again to the surveys [Lok09, Ram20] for more background). It is therefore interesting that RPP seems to behave genuinely differently in this setting.

We also obtain a modest improvement to the algorithm of [APY09], when the dimension k is subpolynomial in n.

Theorem 1.2. *Let $\mathbb{F}$ be a finite field and $L \subseteq \mathbb{F}^n$ be a linear subspace of dimension $k \leq n/2$. There exists an algorithm that runs in time polynomial in n and finds a point that is $\Omega(\frac{n}{\max\{k, \log n\}} \cdot \log n)$ far from L.*

The assumption $k \leq n/2$ is not crucial. A similar theorem can be derived whenever $k \leq (1-\varepsilon)n$ for a constant $\varepsilon > 0$. This will only affect the constant hidden under the big Ω.

In [APY09], a similar theorem was proved with remoteness $\Omega(\frac{n}{k} \cdot \log k)$. When k is subpolynomial in n, Theorem 4.2 gives an improvement. For example, when $k = \log^C n$, the algorithm of [APY09] gives remoteness of $\frac{n}{k} \log k$ whereas we can get remoteness $\frac{n}{k} k^{1/C}$.

1.1 Technique

As mentioned above, the proof of Theorem 1.1 is remarkably and surprisingly simple. We first observe an equivalent formulation for remoteness from a subspace, by considering what is known in coding theory as the *syndrome* of the vector v, which is the vector $Hv \in \mathbb{Q}^m$ where $m = n - k$ and $H \in \mathbb{Q}^{m \times n}$ is a parity-check matrix for L, that is, $\ker H = L$. The vector v is d-close to L if and only if its syndrome s can be written as a linear combination of at most d columns of H. It follows that in order to solve RPP it is enough to find a syndrome vector s that cannot be written as a linear combination of $d-1$ columns of H (this is proved in Section 2). The description so far is in fact true over any field.

Over $\mathbb{Q}$, we find such s simply by defining $s = (1, K, K^2, \ldots, K^{m-1})$ where K is a certain (exponentially large) quantity that depends on the entries of H. We wish to show that s is not in the image of any $m \times (m-1)$ submatrix A of H. Elementary linear algebra guarantees the existence of a vector $a \in \mathbb{Q}^m$ with "small" entries such that $a^T A = 0$. In particular, the entries of a in absolute value are much smaller than K, which guarantees that $a^T s \neq 0$. This implies that s is not in the column span of A. The proof appears in Section 3.

Theorem 1.2 is proved using a completely different technique. The main ingredient is an algorithm of [APY09] that finds a vector $\Omega(\log n)$-far from subspaces of dimension $n/2$. To handle k-dimensional subspaces, [APY09] partition the n coordinates into $n/2k$ blocks of size $2k$ each. The projection of a k-dimensional subspace to each block (of size $2k$) has dimension k, which enables them to find a vector v_i that is $\Omega(\log k)$-far on the i-th block. Concatenating the $n/2k$ vectors $v_1, v_2, \ldots, v_{n/2k}$ gives a vector of length n that is $\Omega(n \log k/k)$-far.

We observe that since the algorithm is allowed to be polynomial in n, we can use the basic algorithm of [APY09] with slightly different parameters on each block. Suppose for simplicity that $k \geq C \log n$ for some

large enough constant C. We use the algorithm of [APY09] to find a vector v_i that is $\Omega(\log n)$-far from the projection of L on the i-th block (instead of $\Omega(\log k)$): the distance is therefore larger than logarithmic in the ambient dimension $2k$. However, since we aim for an algorithm that is polynomial in n (rather than polynomial in k), this still fits within the polynomial time budget. The details appear in Section 4.

2 Syndrome-Based Characterization of Remote Points

We begin by showing the following elementary lemma, which relates the distance of v from a subspace L to the expressibility of the *syndrome* of v, the vector $Hv \in \mathbb{F}^{n-k}$ (where H is the parity-check matrix of L) as a linear combination of the columns of H. The results in this section hold over any field.

Lemma 2.1. *Let L be a linear subspace of dimension k, and let H be a parity-check matrix for L, i.e., an $(n-k) \times n$ matrix of full row rank such that $L = \ker H$. Let $v \in \mathbb{F}^n$ and $s = Hv \in \mathbb{F}^{n-k}$. Then v is d-close to L if and only if s can be written as a linear combination of at most d columns of H.*

Proof. In one direction, suppose v is d-close to L. Then $v = u + \alpha_1 e_{i_1} + \alpha_2 e_{i_2} + \cdots + \alpha_{d'} e_{i_{d'}}$ for $u \in L$, elementary basis vectors $e_{i_1}, e_{i_2}, \ldots, e_{i_{d'}}$, scalars $\alpha_1, \ldots, \alpha_{d'} \in \mathbb{F}$ and $d' \leq d$. Then $s = Hv = H(u + \alpha_1 e_{i_1} + \alpha_2 e_{i_2} + \cdots + \alpha_{d'} e_{i_{d'}}) = \sum_{j=1}^{d'} \alpha_j H_{i_j}$ where H_{i_j} is the i_j-th column of H (since $Hu = 0$ as $u \in L$).

Conversely, if $s = \sum_{j=1}^{d'} \alpha_j H_{i_j}$ for some $d' \leq d$ and $i_1, \ldots, i_{d'}$, then $v - \sum_{j=1}^{d'} \alpha_j e_{i_j}$ is in L: indeed, $H(v - \sum_{j=1}^{d'} \alpha_j e_{i_j}) = s - \sum_{j=1}^{d'} \alpha_j H_{i_j} = 0$. □

As a corollary, finding a vector v which is d-far from L reduces to finding a vector s that *cannot* be expressed as a linear combination of $d-1$ columns of H.

Corollary 2.2. *Suppose there exists a deterministic polynomial-time algorithm that, given a matrix $H \in \mathbb{F}^{(n-k)\times n}$ of full row-rank, finds a vector $s \in \mathbb{F}^{n-k}$ such that s cannot be written as a linear combination of $d-1$ columns of H. Then there exists an algorithm that finds a vector v which is d-far from $\ker H$.*

Proof. Let $L = \ker H$. By solving a system of linear equations we can find a vector v such that $Hv = s$. Note that a solution always exists since $\operatorname{rank}(H) = n - k$, so for every $s \in \mathbb{F}^{n-k}$, the linear system has a solution. By Lemma 2.1, v is d-far from L. □

3 The Remote Point Problem over the Rational Numbers

In this section, we observe a very simple polynomial-time algorithm with optimal parameters for the remote point problem over the field of rational numbers. The algorithm uses the reduction of Section 2 in order to find, given a basis of a k-dimensional subspace $L \subseteq \mathbb{Q}^n$, a vector v which is $(n-k)$-far from L.

Let $H \in \mathbb{Q}^{m\times n}$ be the parity-check matrix of L, where $m = n - k$. By clearing denominators, we may assume H has integer entries. Further, clearing denominators increases the bit complexity of each entry of H by at most a polynomial factor. Now, let $M = \max_{i,j} \{|H_{i,j}|\}$ be the maximum magnitude of an element in H, and define $K' := m!M^m$ and $K = 2K' + 1$. Finally, let $s \in \mathbb{Q}^m$ be defined as

$$s := (1, K, K^2, \ldots, K^{m-1}). \tag{1}$$

Observe once again that the bit complexity of K is polynomial in the bit complexities of the elements of H (and hence polynomial in the input length), and thus the same is true for the entries of the vector s.

Claim 3.1. *For any $m' \leq m - 1$, s cannot be written as a linear combination of m' columns of H.*

Proof. Consider any set of $m' \leq m - 1$ columns of H. Clearly, we may assume the columns are linearly independent, and by adding more columns if necessary, we may assume the set has size exactly $m - 1$. Let A denote the resulting $m \times (m-1)$ submatrix of H.

We now show, by a Cramer's rule-type argument, that there exists a non-zero vector $a \in \mathbb{Q}^m$ with "small" entries such that $a^T A = 0$ but $a^T s \neq 0$, implying that s is not in the image of the columns of A.

For $i \in [m]$, let A_i denote the $(m-1) \times (m-1)$ submatrix of A obtained by deleting the i-th row, and let $a_i = (-1)^i \det(A_i)$. Note that $|a_i| \leq m! M^m = K'$. Define $a = (a_1, \ldots, a_m)$, and observe that a is not the zero vector, since $\operatorname{rank}(A) = m - 1$ so at least one of the A_i's has full rank and $\det(A_i) \neq 0$.

We claim that $a^T A = 0$: indeed, if c_j denotes the j-th column of A, then the fact that $a^T c_j = 0$ can be seen by considering the $m \times m$ matrix $A^{(j)}$ obtained by appending c_j as the m-th column of A, and expanding the determinant of $A^{(j)}$ (which equals 0, since the column c_j appears in it twice) along the last column.

Finally, we show that $a^T s \neq 0$. Let j_0 be the largest index for which a_{j_0} is non-zero. Then $|a_{j_0} s_{j_0}| \geq K^{j_0 - 1}$, whereas

$$\left| \sum_{i=1}^{j_0 - 1} a_i s_i \right| \leq \sum_{i=1}^{j_0 - 1} |a_i s_i| \leq K'(1 + K + K^2 + \cdots + K^{j_0 - 2}) < K^{j_0 - 1}$$

(since $K' < K$, and by the uniqueness of expansion in base K). It follows that

$$a^T s = \sum_{i=1}^{j_0} a_i s_i \neq 0. \qquad \square$$

Corollary 3.2. *There exists a deterministic polynomial-time algorithm that, given a basis of a k-dimensional subspace $L \subseteq \mathbb{Q}^n$, finds a vector v which is $(n-k)$-far from L.*

Proof. The algorithm first constructs $s \in \mathbb{Q}^{n-k}$ as in (1) (as commented above, this can be done in polynomial time). By Claim 3.1, s cannot be expressed as a linear combination of less than $n - k$ columns of H. The statement now follows from Corollary 2.2. $\square$

4 The Remote Point Problem over Finite Fields

In this section we obtain an improvement to the algorithm of Alon, Panigrahy and Yekhanin [APY09]. As a part of their algorithm, they prove the following theorem:

Theorem 4.1 ([APY09])**.** *Let $L \subseteq \mathbb{F}^n$ be a linear subspace, and let B_d denote the Hamming ball of radius d. There exists a collection $\mathcal{S}$ of $\binom{t}{d}$ subspaces, each of dimension at most $\dim L + \frac{d}{t} n$, such that $L + B_d \subseteq \bigcup_{V \in \mathcal{S}} V$.*

If $\dim L \leq n/2$, $t = 4d$ and $d = O(\log n)$, $\mathbb{F}^n$ contains a vector not in $\bigcup_{V \in \mathcal{S}} V$ which is found using a potential-function-based algorithm in [APY09], whose running time is polynomial in n and $\binom{t}{d} = 2^{O(d)}$. When $d = O(\log n)$, this algorithm runs in polynomial time. A different and parallel algorithm was given by Arvind and Srinivasan [AS10b], but they also use the structural result of Theorem 4.1. We remark that [APY09] state their result for the field $\mathbb{F}_2$, but the proof can be easily modified for other finite fields, and was extended by Arvind and Srinivasan [AS10b] further to finite abelian groups.

We also use Theorem 4.1 in order to prove:

Theorem 4.2. *Let $\mathbb{F}$ be a fixed finite field and $L \subseteq \mathbb{F}^n$ be a linear subspace of dimension $k \leq n/2$. There exists an algorithm that runs in time polynomial in n and finds a point that is $\Omega(\frac{n}{\max\{k, \log n\}} \cdot \log n)$ far from L.*

Proof. Let $b = \max\{2k, 20 \log n\}$. Given a subspace L of dimension k, partition the n coordinates into n/b blocks, denoted $B_1, \ldots, B_{n/b}$, of size b each. The projection of L on each block B_i has dimension at most k.

For every $i \in [n/b]$, use Theorem 4.1 (with ambient dimension b) to find a point $v_i \in \mathbb{F}^b$ that is d-far from the projection of L on B_i, for $d = \log n$ and $t = 4 \log n$. This is possible, since the number of subspaces in the collection $\mathcal{S}$ in Theorem 4.1 is at most $\binom{t}{d} \leq 2^{4 \log n} = n^4$ and each subspace has dimension at most

$$k + \frac{d}{t} b \leq \frac{b}{2} + \frac{b}{4} = \frac{3b}{4},$$

where the inequality uses the fact that $k \leq \frac{b}{2}$ and $\frac{d}{t} = \frac{1}{4}$. Hence, the total number of points covered by the union of subspaces in $\mathcal{S}$ is at most

$$n^4 \cdot |\mathbb{F}|^{\frac{3b}{4}} < |\mathbb{F}|^b,$$

where we use the facts that $b \geq 20 \log n$ and $|\mathbb{F}| \geq 2$, so $n^4 < |\mathbb{F}|^{b/4}$. Hence, there exists a point that is not in the union of subspaces in $\mathcal{S}$, which will be found by the algorithm from Theorem 4.1.

The concatenation of $v_1, ..., v_{n/b}$ gives a point that is $\frac{n}{b} \cdot \log n$ far from L since for every $u \in L$, the restriction of u to B_i differs from v_i in at least $\log n$ coordinates. □

The difference comes from the fact that the analogous theorem in [APY09] uses the same strategy but on every block finds a point that is $(\log k)$-far. The observation is that when k is very small, we can run the algorithm from Theorem 4.1 with improved remoteness parameters since we only want the running time to be polynomial in n.

5 Open Problems

This work raises several natural open problems.

- Can the algorithm from Theorem 1.1 be adapted to obtain a deterministic algorithm for RPP over finite fields, for any remoteness parameter $d = \omega(\log n)$? Small fields like $\mathbb{F}_2$ seem like the hardest case; perhaps a more viable intermediate goal is to consider fields with size growing polynomially or even exponentially with n, noting that even when the field size is exponential in n, representing each element in it requires only polynomially many bits.
- Can the algorithm from Theorem 1.1 be adapted to obtain an explicit construction of rigid matrices over $\mathbb{Q}$? It is possible to obtain "semi-explicit" rigid matrices over $\mathbb{Q}$ in which the entries are integers of doubly-exponential magnitude (and hence exponential bit complexity) [Lok09, Ram20]. The techniques for proving that such matrices are rigid are somewhat reminiscent of our proof of Theorem 1.1. Nevertheless, we are not able to obtain matrices whose entries are integers (or rationals) with polynomial bit complexity.
- A natural intermediate goal is the construction of *rigid sets*: these are sets that are guaranteed to contain, for every subspace L, a vector far from L. A rigid matrix corresponds to such a set of size n (the set is the rows of the matrix), but one could relax that requirement and try to obtain sets that are as small as possible. [APY09] construct such sets of size $2^{O(d)}n/d$ for remoteness parameter d (see also [AC15] for alternative constructions with roughly the same size). Constructions of size polynomial in n for $d = \omega(\log n)$ are of great interest, even over $\mathbb{Q}$. They do not seem to imply any circuit lower bounds.

Acknowledgements

The author thanks Mrinal Kumar and Ramprasad Saptharishi for various useful discussions on the remote point problem. ChatGPT was used for help in proof organization and simplification, and proofreading. The paper was written by the human author.

References

[AC15] Noga Alon and Gil Cohen. On Rigid Matrices and U-Polynomials. *Comput. Complex.*, 24(4):851–879, 2015.

[APY09] Noga Alon, Rina Panigrahy, and Sergey Yekhanin. Deterministic Approximation Algorithms for the Nearest Codeword Problem. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop, APPROX 2009, and 13th International Workshop, RANDOM 2009, 2009. Proceedings*, volume 5687 of *Lecture Notes in Computer Science*, pages 339–351. Springer, 2009.

[AS10a] Vikraman Arvind and Srikanth Srinivasan. Circuit Lower Bounds, Help Functions, and the Remote Point Problem. In *Innovations in Computer Science - ICS 2010, Tsinghua University. Proceedings*, pages 383–396. Tsinghua University Press, 2010.

[AS10b] Vikraman Arvind and Srikanth Srinivasan. The Remote Point Problem, Small Bias Spaces, and Expanding Generator Sets. In *27th International Symposium on Theoretical Aspects of Computer Science, STACS 2010*, volume 5 of *LIPIcs*, pages 59–70. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2010.

[Lok09] Satyanarayana V. Lokam. Complexity Lower Bounds using Linear Algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1-2):1–155, 2009.

[Ram20] C. Ramya. Recent Progress on Matrix Rigidity - A Survey. *CoRR*, abs/2009.09460, 2020. Pre-print available at `arXiv:2009.09460`.

[Val77] Leslie G. Valiant. Graph-Theoretic Arguments in Low-Level Complexity. In *Mathematical Foundations of Computer Science 1977, 6th Symposium*, volume 53 of *Lecture Notes in Computer Science*, pages 162–176. Springer, 1977.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il