Notes and Computations on Forbidden Differences

C. Dean, H. Havard, E. Hawkins, P. Heard,
A. Lott, and A. Rice*

Abstract - We explore from several perspectives the following question: given $X \subseteq \mathbb{Z}$ and $N \in \mathbb{N}$, what is the maximum size $D(X,N)$ of $A \subseteq \{1,2,\ldots,N\}$ before A is forced to contain two distinct elements that differ by an element of X? The set of forbidden differences, X, is called *intersective* if $D(X,N) = o(N)$, with the most well-studied examples being $X = S = \{n^2 : n \in \mathbb{N}\}$ and $X = \mathcal{P} - 1 = \{p - 1 : p \text{ prime}\}$. In addition to some new results, including exact formulas and estimates for $D(X,N)$ in some non-intersective cases like $X = \mathcal{P}$ and $X = S + k$, $k \in \mathbb{N}$, we also provide a comprehensive survey of known bounds and extensive computational data. In particular, we utilize an existing algorithm for finding maximum cliques in graphs to determine $D(S,N)$ for $N \leq 300$ and $D(\mathcal{P}-1,N)$ for $N \leq 500$. None of these exact values appear previously in the literature.

Keywords : difference set; intersectivity; Furstenburg-Sárközy theorem; maximum clique

Mathematics Subject Classification (2020) **:** 11B30

1 Introduction

Dating at least to the 1970s, a popular family of questions and results in arithmetic combinatorics concerns the existence of certain differences in dense sets of integers. For $X \subseteq \mathbb{Z}$ and $N \in \mathbb{N}$, let

$$D(X,N) = \max\{|A| : A \subseteq [N],\ (A-A) \cap X \subseteq \{0\}\},$$

where $[N] = \{1,2,\ldots,N\}$ and $A - A = \{a - b : a, b \in A\}$. In other words, $D(X,N)$ is the threshold such that a subset of $[N]$ with more than $D(X,N)$ elements necessarily contains two distinct elements that differ by an element of X. Related definitions in an infinite setting are as follows.

Definition 1.1 *For $A \subseteq \mathbb{N}$, we define the density of A by $\delta(A) = \lim_{N\to\infty} |A \cap [N]|/N$, provided this limit exists, and we define the upper density, $\bar{\delta}(A)$, by replacing the limit with limsup. For $X \subseteq \mathbb{Z}$, we define*

$$\mu(X) = \sup\left\{\bar{\delta}(A) : A \subseteq \mathbb{N},\ (A-A) \cap X \subseteq \{0\}\right\}.$$

We refer to $A \subseteq \mathbb{Z}$ satisfying $(A-A) \cap X \subseteq \{0\}$ as an X-set, and we say that X is intersective if $\mu(X) = 0$. This latter terminology is due to Ruzsa [28].

*All authors were supported by the Kinnaird Endowment, gifted to the Millsaps College Department of Mathematics. Alex Rice was partially supported by an AMS-Simons grant for PUI faculty.

Cantor and Gordon [4], Haralambis [11], and Gupta [10] have investigated $\mu(X)$ in some special cases where X is finite, and we borrow some notation and terminology developed in those papers. However, the majority of existing literature concerns infinite X, specifically the determination of whether X is intersective, and quantitative estimates for $D(X, N)$ in intersective cases. The most well-studied cases are the squares, i.e. $X = S = \{n^2 : n \in \mathbb{N}\}$, and $X = \mathcal{P} - 1 = \{p - 1 : p \in \mathcal{P}\}$, where $\mathcal{P}$ denotes the set of primes, originating from questions of Lovász and Erdős, respectively. The intersectivity of the squares was established independently by Furstenberg [7] and Sárközy [29], the latter of whom also established intersectivity of $\mathcal{P} - 1$ [30]. These results have spawned a great deal of refinements and extensions, and in Section 5, we provide, via tables and accompanying context, a survey of known upper and lower bounds on $D(X, N)$ in both classical and alternative intersective cases.

In addition to this summary of known results for intersective sets, we also explore $D(X, N)$ in some special nonintersective cases. In particular, we establish the following exact formulas in Section 3, which are compatible with a more general conjecture we state in Section 2. Throughout, we use 1_E to denote the characteristic function of a set E.

Theorem 1.2 *The formula* $D(\mathcal{P}, N) = \lceil N/4 \rceil + 1_E(N)$ *holds for all* $N \in \mathbb{N}$, *where* $E = \{2, 3, 4, 11, 12\}$.

Theorem 1.3 *The formula* $D(S+1, N) = \lceil N/3 \rceil + (1_E + 1_{\{9,24\}})(N)$ *holds for all* $N \in \mathbb{N}$, *where*

$$E = \{2, 3, 5, 6, 8, 9, 10, 11, 12, 17, 18, 20, 21, 23, 24, 25, 26, 27\}.$$

In Section 4, we shift to a computational approach to the well-studied quantities $D(S, N)$ and $D(\mathcal{P} - 1, N)$. Specifically, we leverage existing algorithms for finding maximum cliques in graphs to determine exact values of $D(S, N)$ for $N \leq 300$ and $D(\mathcal{P}-1, N)$ for $N \leq 500$, which have not previously appeared in the literature. We provide pseudocode descriptions of our computations, and also highlight some examples of notably dense sets lacking square or shifted prime differences.

2 Preliminaries

Before proceeding further, we establish the connection between the finitary and infinitary formulations given in Section 1. The following is similar to [28, Theorem 1], but with a different approach and construction.

Theorem 2.1 *For every* $X \subseteq \mathbb{Z}$, *we have* $D(X, N)/N \to \mu(X)$ *as* $N \to \infty$. *Futher, there exists an* X*-set* $A \subseteq \mathbb{N}$ *with* $\bar{\delta}(A) = \mu(X)$.

Proof. Fix $X \subseteq \mathbb{Z}$. To first establish an upper bound on $\mu(X)$, suppose $A \subseteq \mathbb{N}$ is an X-set. By definition of $D(X, N)$, A satisfies $|A \cap [N]| \leq D(X, N)$ for all $N \in \mathbb{N}$. Dividing by N and passing to limits, we have

$$\bar{\delta}(A) = \limsup_{N\to\infty} \frac{|A \cap [N]|}{N} \leq \liminf_{N\to\infty} \frac{D(X, N)}{N}. \tag{1}$$

For the lower bound, let $\delta = \limsup_{N\to\infty} D(X,N)/N$, so it suffices to construct an X-set of upper density at least δ. For every $\varepsilon > 0$, we know by definition of lim sup that for every $M > 0$, there exists $N \geq M$ and an X-set $A \subseteq [N]$ with $|A| > (\delta - \varepsilon)N$. We claim something stronger, that for every $N \in \mathbb{N}$ there exists an X-set $A \subseteq [N]$ with $|A| \geq \delta N$. To verify the claim, first fix $\varepsilon > 0$, $N \in \mathbb{N}$, and an X-set $A' \subseteq [N']$ with $N' > 4N/\varepsilon$ and $|A'| > (\delta - \varepsilon/4)N'$. Let $m = \lfloor N'/N \rfloor$, noting that

$$|A' \cap [mN]| \geq |A'| - N > (\delta - \varepsilon/4)N' - N > (\delta - \varepsilon/2)N' \geq (\delta - \varepsilon/2)mN.$$

Partitioning $[mN]$ into m disjoint intervals of length N, the pigeonhole principle guarantees that A' has density greater than $\delta - \varepsilon$ on at least one interval. Since differences are invariant under translation, this high density interval corresponds to an X-set $A \subseteq [N]$ with $|A| > (\delta - \varepsilon)N$. However, since $|A|$ is an integer and ε was chosen independent of N, we can choose ε small enough that $|A| \geq \lceil (\delta - \varepsilon)N \rceil \geq \delta N$.

We now construct an infinite X-set as the union of an increasing sequence of finite sets. For each $N \in \mathbb{N}$, we let $G(N) = \{B \subseteq [N] : B \text{ is an } X\text{-set}, \ |B| \geq \delta N\}$, which the claim verified above ensures is nonempty. By the pigeonhole principle, for every $k \in \mathbb{N}$ and every $B \in G(2^k)$, there exist $B_j \in G(2^j)$ for $0 \leq j \leq k$, where $B_k = B$ and a translation of B_j is either the first or second half of B_{j+1} for $0 \leq j \leq k-1$. In this setting, if $i < j$, we say B_i *extends to* $[2^j]$.

We start the process with $A_0 = \{1\}$, which certainly extends to $[2^k]$ for all $k \in \mathbb{N}$. Inductively, suppose $j \geq 0$ and A_j is a translation containing 1 of an element of $G(2^j)$, which extends to $[2^k]$ for all $k > j$. We consider the infinitely many extensions of A_j, noting that, each time, a translation of A_j is half of a translation of a set $B \in G(2^{j+1})$. Since $G(2^{j+1})$ is finite, there must exist $\tilde{A}_{j+1} \in G(2^{j+1})$ which occurs as this set B infinitely often, so in other words $\tilde{A}_{j+1}$ also extends to $[2^k]$ for all $k > j+1$. We then define A_{j+1} as the translation of $\tilde{A}_{j+1}$ that makes $A_j \subseteq A_{j+1}$. In other words, A_{j+1} is obtained by adding a well-chosen block of length 2^j to either the left or right of A_j, maintaining density of at least δ and maintaining the property of extending to all higher levels. In particular, for each $j \geq 0$, A_j is contained in an interval of length 2^j containing 1, which may contain both positive and negative integers.

Finally, we define $A = \cup_{j=0}^{\infty} A_j$, so $A \subseteq \mathbb{Z}$ is an X-set, and for every $j \in \mathbb{N}$, A has density at least δ on an interval of length 2^j containing 1. This ensures that the lim sup of either $|A \cap [N]|/N$ or $|(-A) \cap [N]|/N$ is at least δ. Defining A' to be either $A \cap \mathbb{N}$ or $(-A) \cap \mathbb{N}$, we have $\mu(X) \geq \bar{\delta}(A') \geq \delta$, which combines with (1) to complete the proof. □

The following is a tangible, necessary condition for intersectivity.

Definition 2.2 *We say $X \subseteq \mathbb{Z}$ is locally intersective if X contains a nonzero multiple of every positive integer. This is equivalent to the statement that, for every infinite $A \subseteq \mathbb{N}$ and every $m \in \mathbb{N}$, the congruence $x - y \equiv z \pmod m$ is solvable with distinct $x, y \in A$ and $z \in X$, which is the motivation for the terminology.*

Clearly intersectivity implies local intersectivity, since if X contains no nonzero multiples of $m \in \mathbb{N}$, then the set $A = \{n \in \mathbb{N} : n \equiv 1 \pmod m\}$ is an X-set, hence $D(X,N) \geq$

$\lceil N/m \rceil$ and $\mu(X) \geq 1/m$. The converse, however, is false. For one large family of counterexamples, recall that a *lacunary sequence* $\{x_n\}$ of positive numbers satisfies $x_{n+1} \geq rx_n$ for some fixed $r > 1$. It is known that a finite union of lacunary sequences in $\mathbb{N}$ is *not* intersective (see [21], for example), but many lacunary sequences, such as the Fibonacci sequence or the sequence of factorials, are locally intersective. Further, the failure of the converse is not solely based on the sparseness of lacunary sequences. Another family of counterexamples is as follows: fix an irrational number $\alpha > 5$, let $A = \{\lfloor n\alpha \rfloor : n \in \mathbb{N}\}$, and let $X = \{n \in \mathbb{N} : |n - (k + 1/2)\alpha| < 1 \text{ for some } k \in \mathbb{N}\}$. By equidistribution, X is locally intersective with $\delta(X) = 2/\alpha$, while A is an X-set satisfying $\delta(A) = 1/\alpha$, hence X is not intersective.

In contrast with the counterexamples provided above, it follows from a theorem of Kamae and Mendès France [13] that if $X = h(\mathbb{Z})$ for a polynomial $h \in \mathbb{Z}[x]$, then X is intersective if and only if it is locally intersective. For this reason, polynomials with locally intersective image, in other words nonzero polynomials with a root at every modulus, are called *intersective polynomials*. Similarly, the set $\mathcal{P} - 1$ (resp. $\mathcal{P} + 1$) was identified as a candidate for intersectivity because of its local intersectivity, as for every $m \in \mathbb{N}$ there are plenty of primes congruent to 1 (resp. -1) modulo m, while all other shifts of $\mathcal{P}$ fail to be locally intersective. More generally, for $h \in \mathbb{Z}[x]$, $h(\mathcal{P})$ being intersective is equivalent to h having a root modulo m that is coprime to m for every $m \in \mathbb{N}$, a condition known as $\mathcal{P}$-*intersectivity* (see [25]).

The underlying principle in the previous paragraph can be summarized as follows: for polynomial images of both $\mathbb{Z}$ and $\mathcal{P}$, local obstructions to intersectivity are the *only* obstructions. This idea can be naturally extended to nonintersective cases: the maximal density of an X-set, if X is built from polynomials and primes, should be determined by the maximal 'local avoidance' of X-differences. We summarize this philosophy with the following general conjecture.

Conjecture 2.3 For a nonzero polynomial $h \in \mathbb{Z}[x]$ and $X = h(\mathbb{Z})$ or $X = h(\mathcal{P})$, we have

$$\mu(X) = \sup_{m \in \mathbb{N}} \frac{d_X(m)}{m},$$

where $d_X(m) = \max\{|A| : A \subseteq \mathbb{Z}/m\mathbb{Z},\ (A - A) \cap X_m = \emptyset\}$, and X_m is the set of congruence class modulo m that intersect X.

We note that X is locally intersective if and only if $d_X(m) = 0$ for all $m \in \mathbb{N}$. The purported lower bound for $\mu(X)$ can be quickly established with X-sets formed from unions of congruence classes, so the content of Conjecture 2.3 lies in the upper bound.

3 Primes and Polynomials: Exact Formulas and Estimates

In this section we provide exact formulas or estimates for $D(X, N)$ in some special, nonintersective cases, which are compatible with Conjecture 2.3, including Theorems 1.2 and 1.3. We frequently rely on the translation invariance of differences, meaning

$x - y = (x + t) - (y + t)$, which in particular makes $D(X, N)$ a subadditive function of N. We include a proof of this standard fact below for completeness.

Lemma 3.1 *If $N, M \in \mathbb{N}$ and $X \subseteq \mathbb{Z}$, then $D(X, N + M) \leq D(X, N) + D(X, M)$.*

Proof. Suppose $N, M \in \mathbb{N}$ and $X \subseteq \mathbb{Z}$. Fix an X-set $A \subseteq [N + M]$ with $|A| = D(X, N + M)$. Letting $B = A \cap [N]$ and $C \cap [N + 1, N + M]$, we see that both B and $C - N = \{c - N : c \in C\} \subseteq [M]$ are X-sets, the latter by translation invariance. By definition of D, we have $|B| \leq D(X, N)$ and $|C| = |C - N| \leq D(X, M)$, so $D(X, N+M) = |A| = |B| + |C| \leq D(X, N) + D(X, M)$, as required. □

We now establish Theorems 1.2 and Theorem 1.3, the former by hand and the latter with computer assistance.

Proof. [Proof of Theorem 1.2] The following sets establish the lower bound for $N \in E$: $\{1, 2\}$ for $N = 2, 3, 4$, and $\{1, 2, 10, 11\}$ for $N = 11, 12$. The lower bound $D(\mathcal{P}, N) \geq \lceil N/4 \rceil$ for all $N \notin E$ is exhibited by $\{n \leq N : n \equiv 1 \pmod 4\}$. We next show that $D(\mathcal{P}, 8) = 2$, and hence also $D(\mathcal{P}, N) = 2$ for $3 \leq N \leq 7$, and $D(\mathcal{P}, N) = 4$ for $N = 11 \leq N \leq 16$. Suppose $A \subseteq [8]$ has no prime differences. By translation invariance, we can assume $1 \in A$ and, consequently, $3, 4, 6, 8 \notin A$. If any $2, 5$ or 7 is present in A, then neither of the other two can be, as the difference between each of them is prime. Therefore, $|A| \leq 2$. Now suppose $A \subseteq [10]$ has no prime differences. By translation invariance, we can assume $1 \in A$. If 2, 5, or 7 lie in A, then at most one of 9 and 10 can also be present in A, as each of $2, 5$, and 7 differ by a prime from at least one of 9 and 10. Therefore, $|A| \leq 3$ and $D(\mathcal{P}, 10) = D(\mathcal{P}, 9) = 3$.

Suppose $A \subseteq [20]$ has no prime differences, and assume $1 \in A$. The interval [20] can be split into two sets of eight ($[8], [9, 16]$) and one set of four ($[17, 20]$). As $D(\mathcal{P}, 8) = 2$, we know $|A| \leq 2 + 2 + |A \cap [17, 20]|$. Since $1 \in A$, we know $18, 20 \notin A$. This leaves only 17 and 19 as possible elements of $|A \cap [17, 20]|$, and since they are separated by a prime difference of 2, only one of these numbers can be in A, hence $|A| \leq 5$ and $D(\mathcal{P}, 20) = 5$, and also $D(\mathcal{P}, N) = 5$ for $N = 17, 18, 19$. Finally, for $N > 20$, $[N]$ can be partitioned into $[8], [9, 16], \ldots, [N - k - 7, N - k], [N - k + 1, N]$, where $k \in \{0, 1, 10, 19, 20, 5, 6, 7\}$ with $k \equiv N \pmod 8$, and Lemma 3.1 is applied to establish $D(\mathcal{P}, N) = \lceil N/4 \rceil$. □

To clarify the compatibility of Theorem 1.2 with Conjecture 2.3, we note that $d_{\mathcal{P}}(4) = 1$ because there are no primes divisible by 4, so $1/4 = \mu(\mathcal{P}) \geq \sup_{m \in \mathbb{N}} d_{\mathcal{P}}(m)/m \geq d_{\mathcal{P}}(4)/4 = 1/4$, and hence all quantities involved are equal to $1/4$. Similarly for Theorem 1.3, $h(x) = x^2 + 3$ has no root modulo 3, so $d_{S+3}(3) = 1$ and $\mu(S + 3) \geq \sup_{m \in \mathbb{N}} d_{S+3}(m)/m \geq 1/3$, and the formula proven below establishes equality.

Proof. [Proof of Theorem 1.3] Maximum clique computations, as described in Section 4, establish the formula for $N \leq 50$, the most exceptional cases being $D(S + 1, 9) = 5$ and $D(S + 1, 24) = 10$, with lower bounds exhibited by the sets $\{1, 2, 5, 8, 9\}$ and $\{1, 2, 5, 8, 9, 16, 17, 20, 23, 24\}$, respectively. The lower bound $D(S + 1, N) \geq \lceil N/3 \rceil$ for all N is exhibited by $\{n \leq N : n \equiv 1 \pmod 3\}$. Then, for $N > 42$, $[N]$ can be partitioned into $[15], [16, 30], \ldots, [N - k - 14, N - k], [N - k + 1, N]$, where $k \in \{0, 1, 32, 33, 4, 35, 36, 7, 8, 39, 40, 41, 42, 13, 14\}$ with $k \equiv N \pmod{15}$, and Lemma 3.1 is applied to establish $D(S + 1, N) \leq \lceil N/3 \rceil$. □

In the two preceeding results, the formula $D(X,N) = \lceil N/m^* \rceil$ holds for sufficiently large N, where $m^* = \min\{m \in \mathbb{N} : (X \cap m\mathbb{Z}) = \emptyset\}$. We note that this does not hold in general for $X = h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$, as the supremum in Conjecture 2.3 is not necessarily $1/m^*$. For example, the set $X = S+3$ intersects $2\mathbb{Z}$, $3\mathbb{Z}$, and $4\mathbb{Z}$, but $\mu(S+3) \geq 1/4$ as exhibited by the set $A = \{n \in \mathbb{N} : n \equiv 1 \text{ or } 3 \pmod 8\}$, in other words $d_{S+3}(8) = 2$. Further, even when numerical evidence suggests $\mu(X) = 1/m^*$, we can still have $D(X,N) > \lceil N/m^* \rceil$ infinitely often, as shown in the following example.

Theorem 3.2 *The lower bound $D(S+2,N) > \lceil N/4 \rceil$ holds unless $N = 4k^2+5$ for some $k \in \mathbb{N}$. More specifically, letting $j = \lceil N/4 \rceil$, we have*

$$D(S+2,N) \geq j + \begin{cases} 1 & N \equiv 2 \pmod 4 \text{ or } (N \equiv 0 \text{ or } 3 \pmod 4 \text{ and } 4j-2 \in S+2) \\ & \text{or } (N \equiv 1 \pmod 4 \text{ and } 4j-6 \notin S+2) \\ 2 & N \equiv 0 \text{ or } 3 \pmod 4 \text{ and } 4j-2 \notin S+2 \\ 0 & N \equiv 1 \pmod 4 \text{ and } 4j-6 \in S+2 \end{cases}.$$

Proof. Let $X = S+2$ and note that elements of X are positive and congruent to 2 or 3 modulo 4. Fix $N \in \mathbb{N}$, let $j = \lceil N/4 \rceil$, and let $A = \{1,2,6,10,\ldots,4j-6\}$, so $|A| = j$. Further, differences in A are either 0 modulo 4, positive and 1 modulo 4, or negative and 3 modulo 4, hence none lie in X, so A is an X-set. If $N \equiv 2 \pmod 4$, then $4j-2 = N$, and $A \cup \{4j-2\}$ is an X-set for the same reasons as A, completing the proof in this case. If $N \equiv 1 \pmod 4$, then $A' = A \cup \{4j-5\} \subseteq [N-2]$, and differences in A' are either differences in A, positive and 1 modulo 4, negative and 3 modulo 4, or $4j-6$. The only one of these that can possibly lie in X is $4j-6$, so A' is an X-set unless $4j-6 \in X$, completing the proof for $N \equiv 1 \pmod 4$. Moreover, if $4j-6 = n^2+2$, n must be even, so $4j-6 = (2k)^2+2$, hence $N = 4j-3 = 4k^2+5$.

Finally, if $N \equiv 0$ or $3 \pmod 4$, we apply the previous case for $N+1$ or $N+2$, respectively, which are congruent to 1 mod 4, noting that $\lceil (N+1)/4 \rceil$ and $\lceil (N+2)/4 \rceil$ are equal to $j+1$ in the respective cases. □

While we do not provide a proof that $\mu(S+2) = 1/4$, maximum clique computations, as discussed in the next section, yield that the lower bound in Theorem 3.2 holds with equality for all $51 \leq N \leq 150$. In particular, $\mu(S+2) \leq D(S+2,149)/149 = 38/149 \approx 0.255$.

4 Maximum Cliques and Computational Data

In an effort to collect computational data for the quantities $D(S,N)$ and $D(\mathcal{P}-1,N)$, we make a connection between forbidden differences and graph theory, as was done previously in [15].

Definition 4.1 *For a graph G, a clique of G is a complete subgraph $C \subseteq G$. Further, C is called a maximum clique of G if $|C| \geq |C'|$ for all cliques $C' \subseteq G$.*

Note the distinction between a maximum clique and a *maximal clique*, which is a clique that is not contained in any strictly larger clique. The problem of developing algorithms that take a graph as input and produce a single maximum clique, a list of all maximum cliques, or just the size of a maximum clique, is a well-studied problem in computer science that is known to be NP-complete. Maximum cliques are directly connected with the forbidden differences problems discussed in this paper in the following way: given $X \subseteq \mathbb{Z}$ and $N \in \mathbb{N}$, we can build a graph G with vertices in $[N]$ by connecting vertex i with vertex j if and only if neither $i-j$ nor $j-i$ lie in X. Then, $D(X,N)$ is precisely the size of a maximum clique in G. Further, by translation invariance, we can instead identify the vertices of the graph with $\{0,1,\ldots,N-1\}$, and restrict our search to maximum cliques containing 0. In particular, instead of working with the full graph G, we can restrict to the neighbors of 0, and add 0 to the returned maximum clique. A pseudocode description of the construction of this 'ZeroNeighborhoodGraph' is given below.

ZeroNeighborhoodGraph

Input: $N \in \mathbb{N}, X \subseteq \mathbb{Z}$
$V \leftarrow \{\}$
$E \leftarrow \{\}$
for $1 \leq i \leq N-1$ **do**
  if $i \notin X$ and $-i \notin X$ **then**
    $V \leftarrow V \cup \{i\}$
    for $1 \leq j < i$ **do**
      if $j \in V$ and $j-i \notin X$ and $i-j \notin X$ **then** $E \leftarrow E \cup \{\{i,j\}\}$
      end if
    end for
  end if
  $i \leftarrow i+1$
end for
return (V,E)

We use the existing MaxCliqueDyn algorithm [20], developed in [14], implemented in C++, which takes in a graph $G = (V,E)$ as input and outputs (M,C), where $C \subseteq G$ is a maximum clique and $M = |C|$. Rather than run the algorithm for each N, which would result in a great deal of redundancy, we instead take a 'top-down' approach. For example, suppose we input $N = 280$ and $X = S$, build the ZeroNeighborhood graph G, compute $(53, C) = \text{MaxCliqueDyn}(G)$, and add 0 to C to get a maximum clique C' of the full graph G', with $|C'| = 54$. Suppose further that, with elements sorted in increasing order, $C' = \{0,\ldots,268\}$. In this case, we know that C' is a maximum clique not only for $N = 280$, but in fact for all $269 \leq N \leq 280$, meaning $D(S,N) = 54$ for $269 \leq N \leq 280$, and we can skip over 11 values of N and continue the top-down progression by inputting $N = 268$. A pseudocode description of this 'cascade' process, for N in a chosen range, is provided below.

MaxCliqueCascade

Input: Min, Max$\in \mathbb{N}$, $X \subseteq \mathbb{Z}$
$N \leftarrow$ Max
while $N >$ Min **do**
  Print N
  $G \leftarrow$ ZeroNeighborhoodGraph(N, X)
  $(M, C) \leftarrow$ MaxCliqueDyn(G)
  Print M and C
  $N' \leftarrow \max(C) \triangleright D(X, n) = M + 1$ with optimal example $A = \{1\} \cup (C + 1)$ for all $N' + 1 \leq n \leq N$
  $N \leftarrow N'$
end while

With computing time totaling a few days on personal laptops, we collected the following values for $D(S, N)$.

Table 1: Exact values for $D(S, N)$ computed with MaxCliqueCascade.

N	$D(S,N)$	N	$D(S,N)$
$1 \le N \le 2$	1	$120 \le N \le 124$	29
$3 \le N \le 5$	2	$125 \le N \le 130$	30
$6 \le N \le 7$	3	$131 \le N \le 132$	31
$8 \le N \le 10$	4	$133 \le N \le 137$	32
$11 \le N \le 12$	5	$138 \le N \le 143$	33
$13 \le N \le 15$	6	$144 \le N \le 145$	34
$16 \le N \le 17$	7	$146 \le N \le 150$	35
$18 \le N \le 20$	8	$151 \le N \le 156$	36
$21 \le N \le 22$	9	$157 \le N \le 158$	37
$23 \le N \le 34$	10	$159 \le N \le 163$	38
$35 \le N \le 37$	11	$164 \le N \le 188$	39
$38 \le N \le 42$	12	$189 \le N \le 198$	40
$43 \le N \le 47$	13	$199 \le N \le 202$	41
$48 \le N \le 52$	14	$203 \le N \le 205$	42
$53 \le N \le 57$	15	$206 \le N \le 207$	43
$58 \le N \le 65$	16	$208 \le N \le 218$	44
$66 \le N \le 67$	17	$219 \le N \le 222$	45
$68 \le N \le 70$	18	$223 \le N \le 235$	46
$71 \le N \le 72$	19	$236 \le N \le 241$	47
$73 \le N \le 80$	20	$242 \le N \le 247$	48
$81 \le N \le 85$	21	$248 \le N \le 252$	49
$86 \le N \le 91$	22	$253 \le N \le 257$	50
$92 \le N \le 96$	23	$258 \le N \le 262$	51
$97 \le N \le 101$	24	$263 \le N \le 265$	52
$102 \le N \le 106$	25	$266 \le N \le 268$	53
$107 \le N \le 111$	26	$269 \le N \le 282$	54
$112 \le N \le 117$	27	$283 \le N \le 284$	55
$118 \le N \le 119$	28	$285 \le N \le 287$	56
$120 \le N \le 124$	29	$288 \le N \le 292$	57
$293 \le N \le 300$	58		

While we have examples of optimal square-difference free sets for all $N \le 300$, we omit them here for the sake of brevity. However, we provide one notable example below, fore-

shadowed prior to the pseudocode description of MaxCliqueCascade: a square difference free set $A \subseteq [269]$ with $|A| = 54$, which remains optimal until N reaches 283.

$$\begin{aligned} A = \{&1, 4, 6, 9, 11, 14, 16, 21, 28, 33, 38, 48, 51, 59, 66, 72, 79, 86, 89, 94, 96, 107, \\ &113, 118, 124, 126, 131, 139, 144, 146, 152, 157, 163, 174, 176, 181, 184, 191, \\ &204, 211, 214, 219, 222, 232, 237, 242, 249, 254, 256, 259, 261, 264, 266, 269\} \end{aligned}$$

In addition, with computing time totaling about one day on a personal laptop, we collected the following values for $D(\mathcal{P}-1, N)$. The algorithm runs faster for $X = \mathcal{P}-1$ because the corresponding graphs are substantially sparser.

Table 2: Exact values for $D(\mathcal{P}-1, N)$ computed with MaxCliqueCascade.

N	$D(\mathcal{P}-1, N)$	N	$D(\mathcal{P}-1, N)$
$1 \le N \le 3$	1	$136 \le N \le 152$	14
$4 \le N \le 8$	2	$153 \le N \le 155$	15
$9 \le N \le 11$	3	$156 \le N \le 208$	16
$12 \le N \le 32$	4	$209 \le N \le 211$	17
$33 \le N \le 35$	5	$212 \le N \le 216$	18
$36 \le N \le 48$	6	$217 \le N \le 219$	19
$49 \le N \le 51$	7	$220 \le N \le 242$	20
$52 \le N \le 64$	8	$243 \le N \le 245$	21
$65 \le N \le 67$	9	$246 \le N \le 298$	22
$68 \le N \le 104$	10	$299 \le N \le 301$	23
$105 \le N \le 107$	11	$302 \le N \le 488$	24
$108 \le N \le 132$	12	$489 \le N \le 491$	25
$133 \le N \le 135$	13	$492 \le N \le 500$	26

As with the squares, we include the most notably dense example of a $(p-1)$-difference free set, in this case $A \subseteq [302]$ with $|A| = 24$, which is optimal until N reaches 489:

$$\begin{aligned} A = \{&1, 4, 57, 60, 65, 68, 91, 94, 141, 144, 155, 158, 175, \\ &178, 189, 192, 209, 212, 265, 268, 273, 276, 299, 302\} \end{aligned}$$

5 Survey of Known Bounds

The following is a (to our knowledge) comprehensive survey of known upper and lower bounds on $D(X, N)$ for intersective X, with results ranging from the mid 1970s to 2025. We separate the bounds into three tables: squares, $\mathcal{P}-1$, and others, the last of which

predominantly consists of more general polynomial images. An asterisk (*) on a citation indicates the current best-known upper bound of its type, while a double asterisk (**) indicates the current best-known lower bound of its type. We use the Vinogradov symbol $\ll$ to denote 'less than a constant times', and implied constants may depend on parameters *other than* N. We use c to denote a positive constant, which must also be independent of N.

Table 3: Bounds on $D(S, N)$, where S is the set of squares

Bound on $D(S,N)$	Due to
$o(N)$	Furstenberg [7]
$\leq N(\log N)^{-\frac{1}{3}+o(1)}$	Sárközy [29]
$\gg N^{\frac{1+\frac{\log 7}{\log 65}}{2}} = N^{0.7331...}$	Ruzsa [27]
$\ll N(\log N)^{-\frac{1}{12}\log\log\log\log N}$	Pintz, Steiger, Szemerédi [22]
$\gg N^{\frac{1+\frac{\log 12}{\log 205}}{2}} = N^{0.7334...}$	Lewko [15]**
$\ll N(\log N)^{-c\log\log\log N}$	Bloom, Maynard [3]
$\ll N\exp(-c\sqrt{\log N})$	Green, Sawhney [8]*

Table 4: Bounds on $D(\mathcal{P}-1, N)$, where $\mathcal{P}-1 = \{p-1 : p \text{ prime}\}$.

Bound on $D(\mathcal{P}-1,N)$	Due to
$\leq N(\log\log N)^{-2+o(1)}$	Sárközy [30]
$\geq N^{\left(\frac{\log 2}{2}-o(1)\right)/\log\log N}$	Ruzsa [28]**
$\ll N(\log\log N)^{-c\log\log\log\log\log N}$	Lucier [18]
$\ll N\exp(-c(\log N)^{1/4})$	Ruzsa, Sanders [26]
$\ll N\exp(-c(\log N)^{1/3})$	Wang [32]
$\ll N^{1-c}$	Green [9]*

All of the best upper bounds in Tables 3-5 are obtained through Fourier analytic density increment arguments, with the exception of Green's [9] breakthrough bound $D(\mathcal{P}-1, N) \ll N^{1-c}$, which follows from a quantitative improvement of the *van der Corput property* for shifted primes. All of the lower bounds in Tables 3-5, with the exception of the greedy algorithm, have their roots in constructions of Ruzsa [27, 28], which bridge the modular version of the problem (i.e. forbidden differences in $\mathbb{Z}/m\mathbb{Z}$) and the integer version. In Table 5, we use $d_X(m)$ as defined in Conjecture 2.3. Further, the terms *strongly Deligne* and *$\mathcal{P}$-Deligne* refer to large families of intersective and $\mathcal{P}$-intersective multivariate polynomials, respectively, satisfying appropriate nonsingularity conditions, as defined in [5] and [6].

Table 5: Bounds on $D(X, N)$ for other X

X	Bound on $D(X, N)$	Due to
any $X \subseteq \mathbb{N}$	$\geq \dfrac{N-1}{\lvert X \cap [N] \rvert + 1}$	greedy algorithm (see [19])
$h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$ intersective	$o(N)$	Kamae, Mendés-France [13]
$\{n^k : n \in \mathbb{N}\}$	$\gg N^{\frac{k-1+\frac{\log d_X(m)}{\log m}}{k}}$, m squarefree	Ruzsa [27]**
$\{n^k : n \in \mathbb{N}\}$	$\ll N(\log N)^{-\frac{1}{4}\log\log\log\log N}$	Balog, Pelikán, Pintz, Szemerédi [2]
$h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$, $h(0) = 0$	$\ll N/\log\log\log N$	Slijepčević [31]
$h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$ intersective	$\leq N(\log N)^{-\frac{1}{\deg(h)-1}+o(1)}$	Lucier [17]
$h(\mathcal{P})$, $h \in \mathbb{Z}[x]$, $h(1) = 0$	$\ll N/\log\log\log N$	Li, Pan [16]
$h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$ intersective, $\deg(h) = 2$	$\leq N(\log N)^{\left(-\frac{1}{\log 3}+o(1)\right)\log\log\log\log N}$	Hamel, Lyall, Rice [12]
$h(\mathcal{P})$, $h \in \mathbb{Z}[x]$, $\mathcal{P}$-intersective	$\leq N(\log N)^{-\frac{1}{2(\deg(h)-1)}+o(1)}$	Rice [25]
$h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$ intersective, $\deg(h) = k$	$\leq N(\log N)^{\left(-\frac{1}{\log\left(\frac{k^2+k}{2}\right)}+o(1)\right)\log\log\log\log N}$	Rice [23]
$\{am^2 + bmn + cn^2 : m, n \in \mathbb{N}\}$ $a, b, c \in \mathbb{Z}$, $b^2 - 4ac \neq 0$	$\ll N\exp(-c\sqrt{\log N})$	Rice [24]
$\{m^2 + n^2 : m, n \in \mathbb{N}\}$	$\gg \sqrt{N}$	Younis [33]**
$h(\mathbb{Z}^\ell)$, $h \in \mathbb{Z}[x_1, \ldots, x_\ell]$, $\ell \geq 2$, strongly Deligne	$\ll N\exp(-c(\log N)^\mu)$, $\mu = \begin{cases} [(\deg(h)-1)^2+1]^{-1} & \text{if } \ell = 2 \\ 1/2 & \text{if } \ell \geq 3 \end{cases}$	Doyle, Rice [5]*
$h(\mathbb{Z})$, $h \in \mathbb{Z}[x]$ intersective	$\ll N(\log N)^{-c\log\log\log N}$	Arala [1]*
$h(\mathcal{P})$, $h \in \mathbb{Z}[x]$, $\mathcal{P}$-intersective	$\ll N(\log N)^{-c\log\log\log N}$	Doyle, Rice [6]*
$h(\mathcal{P}^\ell)$, $h \in \mathbb{Z}[x_1, \ldots, x_\ell]$, $\ell \geq 2$, $\mathcal{P}$-Deligne	$\ll N\exp(-c(\log N)^\mu)$, $\mu = \begin{cases} [2(\deg(h)-1)^2+6]^{-1} & \text{if } \ell = 2 \\ 1/4 & \text{if } \ell \geq 3 \end{cases}$	Doyle, Rice [6]*

Acknowledgments

This research was initiated during the Summer 2024 Kinnaird Institute Research Experience at Millsaps College. All authors were supported during the summer by the Kinnaird Endowment, gifted to the Millsaps College Department of Mathematics. At the time of completion, all authors except Alex Rice and Andrew Lott were Millsaps College undergraduate students. Alex Rice was partially supported by an AMS-Simons research grant for PUI faculty. The authors would like to thank Drewrey Lupton and Philipp Birklbauer for their code-related assistance, and John Griesmer for his helpful insights and references.

References

[1] N. Arala, A maximal extension of the Bloom-Maynard bound for sets with no square differences, *Funct. Approx. Comment. Math.*, **71** (2024), 271–296.

[2] A. Balog, J. Pelikán, J. Pintz, E. Szemerédi, Difference sets without κ-th powers, *Acta. Math. Hungar.*, **65** (1994), 165–187.

[3] T. Bloom, J. Maynard, A new upper bound for sets with no square differences, *Comp. Math.*, **158** (2022), 1777–1798.

[4] D.G. Cantor, B. Gordon, Sequences of integers with missing differences, *J. Combinatorial Theory Ser. A*, **14** (1973), 281–287.

[5] J.R. Doyle, A. Rice, Multivariate polynomial values in difference sets, *Discrete Anal.*, 2021:11, 46pp.

[6] J.R. Doyle, A. Rice, The Furstenberg-Sárközy theorem for polynomials in one or more prime variables, *Ramanujan J.*, **67** (2025), Paper No. 64, 33pp.

[7] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math.*, **71** (1977), 204–256.

[8] B. Green, M. Sawhney, Improved bounds for the Furstenberg-Sárközy theorem, available online at the URL: `https://arxiv.org/abs/2411.17448v1`.

[9] B. Green, On Sárközy's theorem for shifted primes, *J. Amer. Math. Soc.*, **37** (2024), 1121–1201.

[10] S. Gupta, Sets of integers with missing differences, *J. Combinatorial Theory Ser. A*, **89** (2000), 55–69.

[11] N.M. Haralambis, Sets of integers with missing differences, *J. Combinatorial Theory Ser. A*, **23** (1977), 22–33.

[12] M. Hamel, N. Lyall, A. Rice, Improved bounds on Sárközy's theorem for quadratic polynomials, *Int. Math. Res. Not.*, (2013), 1761–1782

[13] T. Kamae, M. Mendès France, van der Corput's difference theorem, *Israel J. Math.*, **31**, (1978), 335–342.

[14] J. Konc, D. Janezic, An improved branch and bound algorithm for the maximum clique problem, *MATCH Commun. Math. Comput. Chem.*, **58** (2007), 569–590.

[15] M. Lewko, An improved lower bound related to the Sárközy-Furstenberg Theorem, *Elect. J. Comb.*, **22** (2015), 1–6.

[16] H.-Z. Li, H. Pan, Difference sets and polynomials of prime variables, *Acta. Arith.*, **138** (2009), 25–52.

[17] J. Lucier, Intersective sets given by a polynomial, *Acta Arith.*, **123** (2006), 57–95.

[18] J. Lucier, Difference sets and shifted primes, *Acta. Math. Hungar.*, **120** (2008), 79–102.

[19] N. Lyall, A new proof of Sárközy's theorem, *Proc. Amer. Math. Soc.*, **141** (2013), 2253–2264.

[20] MaxCliqueDyn algorithm, available online at the URL: http://insilab.org/maxclique/

[21] Y. Peres, W. Schlag, Two Erdős problems on lacunary sequences: chromatic number and diophantine approximation, *Bull. Lond. Math. Soc.*, **42** (2010), 295–300.

[22] J. Pintz, W. L. Steiger, E. Szemerédi, On sets of natural numbers whose difference set contains no squares, *J. London Math. Soc.*, **37** (1988), 219–231.

[23] A. Rice, A maximal extension of the best-known bounds for the Furstenberg-Sárközy Theorem, *Acta Arith.*, **187** (2019), 1–41.

[24] A. Rice, Binary quadratic forms in difference sets, *Combinatorial and Additive Number Theory III*, Springer Proc. of Math. and Stat., **297** (2020), 175–196.

[25] A. Rice, Sárközy's theorem for $\mathcal{P}$-intersective polynomials, *Acta Arith.*, **157** (2013), 69–89.

[26] I. Ruzsa, T. Sanders, Difference sets and the primes, *Acta. Arith.*, **131** (2008), 281–301.

[27] I. Ruzsa, Difference sets without squares, *Period. Math. Hungar.*, **15** (1984), 205–209.

[28] I. Ruzsa, On measures on intersectivity, *Acta Math. Hungar.*, **43** (1984), 335–340.

[29] A. Sárközy, On difference sets of sequences of integers I, *Acta. Math. Hungar.*, **31** (1978), 125–149.

[30] A. Sárközy, On difference sets of sequences of integers III, *Acta. Math. Hungar.*, **31** (1978), 355–386.

[31] S. Slijepčević, A polynomial Sárközy-Furstenberg theorem with upper bounds, *Acta Math. Hungar.*, **98** (2003), 275–280.

[32] R. Wang, On a theorem of Sárközy for difference sets and shifted primes, *J. of Number Theory*, **211** (2020), 220–234.

[33] K. Younis, Lower bounds in the polynomial Szemerédi theorem, available online at the URL: https://arxiv.org/abs/1908.06058.

Christian Dean
Millsaps College
1701 N State St.
Jackson, MS 39210
E-mail: deanmchris@gmail.com

Haley Havard
University of Georgia
Department of Mathematics
Athens, GA 30602
E-mail: heh25323@uga.edu

Elizabeth Hawkins
Millsaps College
1701 N State St.
Jackson, MS 39210
E-mail: hawkie@millsaps.edu

Patch Heard
Millsaps College
1701 N State St.
Jackson, MS 39210
E-mail: patchh@icloud.com

Andrew Lott
University of Georgia
Department of Mathematics
Athens, GA 30602
E-mail: andrew.lott@uga.edu

Alex Rice
Millsaps College
1701 N State St
Jackson, MS 39210
E-mail: riceaj@millsaps.edu

Received: November 20, 2025 **Accepted**: March 5, 2026
Communicated by Scott Annin