Minimal Symmetric Differences of Lines in Projective Planes

Paul Balister,[1] Béla Bollobás,[1,3] Zoltán Füredi,[3] and John Thompson[2]

[1]*Department of Mathematical Sciences, University of Memphis, Memphis Tennessee 38152, USA*

[2]*Department of Pure Mathematics and Mathematical Statistics, Cambridge CB3 0WB, United Kingdom*

[3]*Alfréd Rényi Institute of Mathematics, 1053 Budapest, Hungary*

Received March 17, 2013; revised October 29, 2013

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/jcd.21390

Abstract: **Let q be an odd prime power and let $f(r)$ be the minimum size of the symmetric difference of r lines in the Desarguesian projective plane $PG(2, q)$. We prove some results about the function $f(r)$, in particular showing that there exists a constant $C > 0$ such that $f(r) = O(q)$ for $Cq^{3/2} < r < q^2 - Cq^{3/2}$.** © 2014 Wiley Periodicals, Inc. J. Combin. Designs 00: 1–17, 2014

Keywords: *Desarguesian projective planes; codes; blocking*

1 INTRODUCTION

Let q be an odd prime power and consider the Desarguesian projective plane $PG(2, q)$. (For detailed definitions of lines, coordinates, conics, etc., see, e.g. the monograph Hirschfeld [11].) Write $\mathcal{P}$ and $\mathcal{L}$ for the set of points and lines of $PG(2, q)$, respectively. We shall consider the subsets of $\mathcal{P}$ or $\mathcal{L}$ as elements of a vector space isomorphic to $\mathbb{F}_2^N$, $N := q^2 + q + 1$, and will switch between the "subset" and "vector" interpretations without further comment. For example, for subsets A and B of $\mathcal{P}$ or $\mathcal{L}$, $A + B$ represents the symmetric difference of A and B.

Contract grant sponsor: Hungarian National Science Foundation OTKA 104343 (Z.F.); contract grant sponsor: European Research Council Advanced Investigators; contract grant number: 267195.

Define for $0 \le r \le N$,

$$f(r) = \min \left\{ \left| \sum_{i=1}^{r} \ell_i \right| : \ell_1, \dots, \ell_r \in \mathcal{L} \text{ distinct} \right\}, \tag{1}$$

that is the minimal symmetric difference of r lines in $PG(2, q)$.

The problem of determining $f(r)$ is motivated by the fact that it is an algebraic version of the Besicovitch-Kakeya [3] problem in a projective plane—determining the minimum size of a set that contains lines (or segments) in many directions. For more results on Kakeya's problem in the finite fields see [5], [10] and the references there.

Given a set R of lines in $PG(2, q)$, call a point *odd* if it is incident with an odd number of lines in R, and define the terms "even point," "single point," "double point," etc., analogously. Let $\mathcal{P}^o(R)$ be the set of odd points, and let $\mathcal{P}^e(R)$, $\mathcal{P}^k(R)$, $\mathcal{P}^{\ge k}(R)$ be defined analogously as the set of points that are even, multiplicity k, and multiplicity at least k, respectively.

Dually, for $S \subseteq \mathcal{P}$, define $\mathcal{L}^o(S)$ to be the set of lines $\ell \in \mathcal{L}$ such that $|\ell \cap S|$ is odd. Define $\mathcal{L}^e(S)$, $\mathcal{L}^k(S)$, and $\mathcal{L}^{\ge k}(S)$ analogously.

By duality of lines and points in the projective plane $PG(2, q)$ we can rewrite (1) in the equivalent forms

$$f(r) = \min_{R \subseteq \mathcal{L},\, |R| = r} |\mathcal{P}^o(R)| = \min_{S \subseteq \mathcal{P},\, |S| = r} |\mathcal{L}^o(S)|. \tag{2}$$

We shall therefore often switch the viewpoint and consider sets of points that have odd intersections with few lines.

The next observation, proved below, is that $\mathcal{P}^o(R)$ almost determines R, and $\mathcal{L}^o(S)$ almost determines S. Indeed, the N vectors specified by $\mathcal{L}$ span an $(N-1)$-dimensional subspace of $\mathbb{F}_2^{\mathcal{P}}$ and their only linear dependency is $\sum_{\ell \in \mathcal{L}} \ell = 0$. This gives that $\mathcal{P}^o(R) = \mathcal{P}^o(R')$ iff either $R = R'$ or $R' = \mathcal{L} \setminus R$. Indeed, it is well known that the $N \times N$ point line 0–1 incidency matrix A has rank $N-1$ (one can consider $AA^T = J + qI$ and this has rank $N-1$ over $\mathbb{F}_2$, see, e.g. Ryser [14]). The following useful lemma is based on this observation.

Lemma 1. *If $R = \mathcal{L}^o(S)$ then $|R|$ is even and either $S = \mathcal{P}^e(R)$ (if $|S|$ is odd) or $S = \mathcal{P}^o(R)$ (if $|S|$ is even). Dually, if $S = \mathcal{P}^o(R)$ then $|S|$ is even and either $R = \mathcal{L}^e(S)$ (if $|R|$ is odd) or $R = \mathcal{L}^o(S)$ (if $|R|$ is even).*

Proof. The maps $\mathcal{L}^o$ and $\mathcal{P}^o$ can be thought of as $\mathbb{F}_2$-linear maps between the set of subsets of $\mathcal{P}$ and $\mathcal{L}$, each regarded as a vector space isomorphic to $\mathbb{F}_2^N$. For $p \in \mathcal{P}$, $|\mathcal{L}^o(\{p\})| = |\{\ell \in \mathcal{L} : p \in \ell\}| = q + 1$ is even, so $|\mathcal{L}^o(S)|$ is even for all $S \subseteq \mathcal{P}$. Moreover

$$\mathcal{P}^o(\mathcal{L}^o(\{p\})) = \sum_{\ell \ni p} \ell = \mathcal{P} - \{p\} \in \mathbb{F}_2^{\mathcal{P}}$$

as the number $q + 1$ of lines through p is even and there is a unique line through p and p' for every $p' \neq p$. By linearity, $\mathcal{P}^o(\mathcal{L}^o(S)) = \sum_{p \in S}(\mathcal{P} - \{p\}) = S$ when $|S|$ is even, and so $\mathcal{P}^o$ has rank at least $N - 1$. Also, $\mathcal{P}^o(\mathcal{L}) = \emptyset$ as every point is in an even number of lines. Hence the kernel of $\mathcal{P}^o$ is $\{0, \mathcal{L}\}$. Similarly the kernel of $\mathcal{L}^o$ is $\{0, \mathcal{P}\}$. The result now follows as $\mathcal{P}^e(R) = \mathcal{P} \setminus \mathcal{P}^o(R)$ and $\mathcal{L}^e(R) = \mathcal{L} \setminus \mathcal{L}^o(R)$. □

Lemma 2. *For* $0 \le r \le N$, $f(N-r) = f(r)$.

Proof. Replacing any set $R = \{\ell_1, \dots, \ell_r\}$ by its complement $\mathcal{L} \setminus R$ and noting that $\sum_{\ell \notin R} \ell = \sum_{\ell \in R} \ell$, we find that $f(N-r) \le f(r)$. Reversing the roles of r and $N-r$ gives $f(N-r) \ge f(r)$. □

Lemma 3. *Let R be any set of r lines in $\mathcal{L}$. Then*

$$r(q+2-r) \le |\mathcal{P}^o(R)| \le rq+1$$

and

$$|\mathcal{P}^o(R)| \equiv r(q+2-r) \bmod 4.$$

In particular, $f(r) \ge r(q+2-r)$ *and* $f(r) \equiv r(q+2-r) \bmod 4$.

Proof. Each line of R contains at least $q+1-(r-1) = q+2-r$ points that do not lie on any other line of R. Thus there are at least $r(q+2-r)$ points lying on a single line, and so in particular $|\mathcal{P}^o(R)| \ge r(q+2-r)$. On the other hand, one line contains $q+1$ points and the symmetric difference of two lines contains exactly $2q$ points. Thus $|\mathcal{P}^o(R)| \le rq+1$ for $r \le 2$. For $r > 2$ write $R = R' \cup \{\ell, \ell'\}$. Then by induction

$$\begin{aligned} |\mathcal{P}^o(R)| &= |\mathcal{P}^o(R') + \mathcal{P}^o(\{\ell, \ell'\})| \\ &\le |\mathcal{P}^o(R')| + |\mathcal{P}^o(\{\ell, \ell'\})| \\ &\le ((r-2)q+1) + 2q = rq+1. \end{aligned}$$

Now let $t_i = |\mathcal{P}^i(R)|$ be the set of points of multiplicity i. Then $\sum i t_i = r(q+1)$ is the number of points in all the lines counted with multiplicity, and $\sum i(i-1)t_i = r(r-1)$ is the number of intersection points between ordered pairs of lines counted with multiplicity. Subtracting gives $\sum i(2-i)t_i = r(q+2-r)$. But $i(2-i) \equiv 0 \bmod 4$ when i is even and $i(2-i) \equiv 1 \bmod 4$ when i is odd. Thus $r(q+2-r) \equiv \sum_{i \text{ odd}} t_i = |\mathcal{P}^o(R)|$ mod 4. □

The function $f(r)$ is easily determined for $0 \le r \le q+1$ (and hence by Lemma 2 also for $N-q-1 \le r \le N$).

Theorem 4. *For* $0 \le r \le q+1$, $f(r) = r(q+2-r)$.

Proof. Lemma 3 implies $f(r) \ge r(q+2-r)$, so it remains by (2) to construct a set S of points with $|S| = r$ and $|\mathcal{L}^o(S)| = r(q+2-r)$.

Let $C = \{[s^2 : st : t^2] : [s:t] \in PG(1,q)\}$ be the conic $XZ = Y^2$. We note that all lines ℓ intersect C in at most 2 points, and $|\ell \cap C| = 1$ if and only if ℓ is one of the $q+1$ tangent lines to C.

Let S be any subset of C of size r. No line intersects S in more than two points and so for any $p \in S$ exactly $r-1$ lines through p meet C at another point of S, while $(q+1)-(r-1) = q+2-r$ lines through p fail to meet C at any other point of S. Thus there are exactly $r(q+2-r)$ lines that meet S in an odd number of points and so $|\mathcal{L}^o(S)| = r(q+2-r)$ as required. □

The function $f(r)$ cannot vary too rapidly; trivially we have $|f(r+1) - f(r)| \le q+1$. In fact, we can say slightly more.

Theorem 5. *For* $0 < r < N - 2$, $|f(r+1) - f(r)| \leq q - 1$.

Note that $f(0) = f(N) = 0$ and $f(1) = f(N-1) = q+1$, so this result fails for $r = 0, N-1$. On the other hand, the inequality can be sharp. For example, $f(2) - f(1) = f(q+1) - f(q) = q - 1$ by Theorem 4. There are other examples, e.g. $f(2q-1) = q+1$ and $f(2q) = 2$ (see Theorem 13 below).

Proof. Assume $|R| = r$ and $\mathcal{P}^o(R) = S$ with $|S| = f(r)$. Note that $S \neq \emptyset$ as $R \neq \emptyset, \mathcal{L}$. Pick $p \in S$. Assume every line ℓ through p intersects S in an odd number of points. Then every line through p intersects $S \setminus p$ is an even number of points. Since distinct lines through p partition $S \setminus p$, we see that $|S \setminus p|$ is even and hence $|S|$ is odd, contradicting Lemma 1. Thus there exists a line ℓ_e that meets S in an even (and positive) number of points. If all $\ell \in \mathcal{L}$ met S in an even number of points then $\mathcal{L}^o(S) = \emptyset$ and so $S = \emptyset$ or $\mathcal{P}$, a contradiction. Thus there exists a line ℓ_o that meets S in an odd number of points. As $R = \mathcal{L}^o(S)$ or $\mathcal{L}^e(S)$, either ℓ_e or ℓ_o fails to lie in R. Adding such a line to R increases r by one and increases S by at most $q - 1$, implying $f(r+1) - f(r) \leq q - 1$.

Replacing r by $N - r - 1$ and applying Lemma 2 gives $f(r+1) - f(r) = -(f(N-r) - f(N-r-1)) \geq -(q-1)$, completing the proof of Theorem 5. □

2 THE CASE OF $q + 2$ LINES

Our next aim is to prove that the jump $f(q+2) - f(q+1) = f(q+2) - (q+1)$ is not too small.

Theorem 6. $f(q+2) = 2q - 2$ *for* $q \leq 13$. *More generally, for* $q \geq 7$ *we have* $\frac{3}{2}(q+1) \leq f(q+2) \leq 2q - 2$.

To prove this we shall use several lemmas, some classical results of this topic. Most of their proofs use either Rédei's method (see e.g. [13]) or some version of Combinatorial Nullstellensatz (see e.g. [1, Theorem 1.2]). Arrangements of $q+2$ lines are the most investigated part of finite geometries. In the following, a *triple point* with respect to a set of lines R will refer to a point that lies on *at least* three lines.

Lemma 7 (Bichara and Korchmáros [2]). *Let R be a set of $q+2$ lines in $PG(2, q)$. Then there are at most two lines without triple points.*

A *blocking set* in the affine plane $AG(2, q)$ or in the projective plane $PG(2, q)$ is a set B of points such that each line is incident with at least one point of B.

Lemma 8 (Brouwer and Schrijver [6] and Jamison [12]). *Let B be a blocking set in $AG(2, q)$. Then B consists of at least $2q - 1$ points.*

Lemma 9 (Szőnyi [15]). *Let B be a minimal blocking set in $PG(2, q)$ of size less than $3(q+1)/2$ where $q = p^h$ for some prime p. Then all lines meet B in 1 mod p points.*

The following lemma is contained in [5] (top of page 211) as a part of a more complex argument. For completeness we reproduce its proof here.

Lemma 10 (Blokhuis and Mazzocca [5]). *Let R be a set of $q+2$ lines with at least one of the lines containing no triple points. Then the number of odd points is at least $2q$ minus the number of lines in R without triple points.*

Proof. Without loss of generality, we may assume that R contains the line at infinity and that this line has no triple point. Let L be the set of $q+1$ lines in $AG(2, q)$ obtained by restricting the remaining lines of R to $AG(2, q)$. As the line at infinity contains no triple point, no two lines in L are parallel. Then as $|L| = q + 1$, every line ℓ in $AG(2, q)$ is parallel to precisely one line of L.

Claim. In $AG(2, q)$ the odd points block all lines in $AG(2, q)$, except those in L that have no triple points.

Indeed, assume first that $\ell \notin L$. Then ℓ intersects q of the lines in L; indeed it intersects all but the unique line in L parallel to ℓ. Since q is odd, ℓ has an odd point.

Now assume $\ell \in L$ and has a triple point. As there are q points in L and only q other lines in L, the fact that some point in ℓ meets at least two of these lines implies that there is a point of ℓ that meets no other line of L. Such a point is a single (and hence odd) point.

Adding one point from each line without a triple point (except the line at infinity) we obtain a blocking set of the affine plane, which by Lemma 8 contains at least $2q - 1$ points. The result follows. □

Proof of the lower bound in Theorem 6. Let R be a set of $q+2$ lines with $f(q+2) = |\mathcal{P}^o(R)|$, $S := \mathcal{P}^o(R)$, and let T_3 be the set of triple points. We will show that $|S| \geq 3(q+1)/2$.

First, suppose that R has a line without a triple point. Then by Lemmas 7 and 10 there are at least $2q - 2$ odd points.

Second, suppose all $q+2$ lines in R have triple points and $|S| < 2q - 2$. Since $f(q+2) \equiv 0 \bmod 4$ by Lemma 3 we may suppose that $|S| \leq 2q - 6$.

Claim. S is a minimal blocking set in $PG(2, q)$.

Indeed, every line ℓ in $PG(2, q)$ is either in our set (in which case it contains a single point), or intersects all $q+2$ lines of R. As $q+2$ is odd, ℓ must contain an odd point.

That S is minimal can be seen as follows: Let $v \in S$ and suppose on the contrary that $S \setminus \{v\}$ meets all lines. Since v is an odd point, there are $2m+1$ lines of R containing it. Each of these lines contains at least $2m - 1$ additional odd (single) points of S. Moreover, every line ℓ not in R has an odd number of odd points. Then if $\ell \notin R$ is a line through v, we have $|S \cap \ell| \geq 2$ and hence $|S \cap \ell| \geq 3$. In total we find at least $(2m+1)(2m-1) + 2(q-2m) \geq 2q - 1$ odd points beside v. This contradiction completes the proof of the Claim.

We count multiplicities of intersections as in the proof of Lemma 3. If we let t_i be the number of points that occur in exactly i of our lines, then $\sum_i i t_i = \sum_i i(i-1)t_i = (q+2)(q+1)$. Thus $\sum_i i(i-2)t_i = 0$, rearranging

$$|S| = \sum_{i \text{ odd}} t_i = \sum_{i \geq 3} (i(i-2) + (i \bmod 2))\, t_i = 4t_3 + 8t_4 + 16t_5 + 24t_6 + \cdots \qquad (3)$$

Let $R_3 \subseteq R$ be the set of lines having a single triple point, and that point has degree three, and let $R_4 \subseteq R$ be the set of lines having a single triple point, and that point has degree at least four. Every line in R has at least one triple point, the members of $R \setminus (R_3 \cup R_4)$ have at least two. So adding up the degrees of triple points we obtain

$\sum_{i\geq 3} it_i = \sum_{\ell\in R} |\ell \cap T_3| \geq 2|R| - |R_3| - |R_4|$. Consider $\sum_{i\geq 4} it_i$, it is an upper bound for $|R_4|$. Summarizing we obtain

$$3t_3 + \sum_{i\geq 4} 2it_i \geq 2|R| - |R_3|.$$

This and (3) yield $|S| \geq 2q + 4 - |R_3|$. Every R_3 line meets S in two elements, so actually $R_3 = \emptyset$ by Lemma 9 for $|S| < 3(q+1)/2$. This contradiction completes the proof of $|S| \geq 3(q+1)/2$. For $q \leq 13$ we note that $3(q+1)/2 > 2q - 6$, so $f(q+2) = 2q - 2$. □

Finally, to show $f(q+2) \leq 2q - 2$ recall that $f(q+2) \leq f(q+1) + (q-1) = 2q$ by Theorems 5 and 4, while $f(q+2) \equiv 0 \bmod 4$ by Lemma 3. Thus $f(q+2) \leq 2q - 2$.

This upper bound on $f(q+2)$ can also be seen in the following way. There is an action of $SL(2,q)$ on $PG(2,q)$ in which the orbits are A, B, and C, where C is the conic described above, A is the set of points that lie on no tangent of C and B is the set of points that lie on two tangents of C. Now $|\mathcal{L}^o(C)| = q+1$, so if $p \in A$ then $|\mathcal{L}^o(C \cup \{p\})| = (q+1) + (q+1)$ as all lines through p change from having an even intersection with C to having an odd intersection with $C \cup \{p\}$. On the other hand, if $p \in B$ then $|\mathcal{L}^o(C \cup \{p\})| = (q+1) + (q-1) - 2 = 2q - 2$ as there are $q - 1$ lines thorough p with an even intersection with C and an odd intersection with $C \cup \{p\}$, while there are two lines through p that are tangent to C and so have odd intersection with C and even intersection with $C \cup \{p\}$. The result now follows from (2).

We conjecture that in fact the upper bound is correct in Theorem 6.

Conjecture 11. $f(q+2) = 2q - 2.$

3 EXACT VALUES NEAR $2q$

A few more values of $f(r)$ are known when r is small. To derive these we shall make use of the following result.

Lemma 12. *For even s, $f(s)$ is the minimum even r such that there exists a set R of lines with $|R| = r$ and $|\mathcal{P}^o(R)| = s$.*

Proof. Assume R is a set of lines with $|R| = r$ and $\sum_{\ell\in R} \ell = S$ with $|S| = s$. Now $|\mathcal{L}^o(S)|$ is even while $|\mathcal{L}^e(S)|$ is odd. Hence $R = \mathcal{L}^o(S)$ as r is even. Thus, by (2), $f(s) \leq r$. Conversely, if $f(s) = r$ and $|S| = s$ with $|\mathcal{L}^o(S)| = r$, then r is even and, setting $R = \mathcal{L}^o(S)$, we have $|R| = r$ and $|\mathcal{P}^o(R)| = |S| = s$ as s is even. □

Theorem 13. $f(2q-1) = q+1,\ f(2q) = 2,\ f(2q+1) = q-1.$

Proof. If $|R| = 2$ then $|\mathcal{P}^o(R)| = 2q$, so $f(2q) \leq 2$ by Lemma 12. However $f(r) > 0$ and $f(r)$ is even for $0 < r < N$, so $f(2q) = 2$. Thus $f(2q-1), f(2q+1) \leq q+1$ by Theorem 5. Also $f(2q+1) \equiv (2q+1)(-q+1) \equiv q - 1 \bmod 4$ and $f(2q-1) \equiv (2q-1)(-q+3) \equiv q+1 \bmod 4$ by Lemma 3. Thus it is sufficient to show that $f(2q \pm 1) > q - 3$. As $2q \pm 1$ is odd, there exists a R with $|R| = f(2q \pm 1)$ and $|\mathcal{P}^o(R)| = N - (2q \pm 1) \geq q^2 - q$. But $|\mathcal{P}^o(R)| \leq q|R| + 1$ by Lemma 3, so $|R| > q - 3$. □

4 A GRAPH CLIQUE DECOMPOSITION LEMMA

The values of $f(r)$ for $q+2<r<2q-1$ remain to be determined, and indeed $f(r)$ is unknown for many values of $r<Cq^{3/2}$, although some non-trivial bounds are given by Lemmas 19 and 20 below. For larger r, between $Cq^{3/2}$ and $N-Cq^{3/2}$, we shall show much more. Indeed it seems that $f(r)$ can be determined for most values of r in this range, although an explicit description of these values seems difficult.

Suppose that s is even (the case when s is odd follows by considering $f(N-s)$). By Lemma 12 and duality it is enough to determine for each even r in turn whether or not there exists a set S of points such that $|\mathcal{L}^o(S)|=s$. Any set of points S induces an edge-decomposition of the complete graph K_S with vertex set S into cliques on the sets $\ell\cap S$, $\ell\in\mathcal{L}$. Indeed, every pair of points of S lie in a unique line $\ell\in\mathcal{L}$ so each edge K_S lies in a unique clique $K_{\ell\cap S}$. We show that $s=|\mathcal{L}^o(S)|$ can be determined in terms of the sizes of these cliques.

Lemma 14. *Suppose $r=|S|$ is even and $|\mathcal{L}^o(S)|=rq-4t$. For $\ell\in\mathcal{L}$ write $r_\ell=|S\cap\ell|$. Then $\sum_{\ell\in\mathcal{L}}\left\lfloor\frac{r_\ell}{2}\right\rfloor=\frac{r}{2}+2t$.*

Proof. As there are $q+1$ lines through each point of S, $\sum_{\ell\in\mathcal{L}}r_\ell=r(q+1)$. Thus

$$rq-4t=|\mathcal{L}^o(S)|=\sum_{r_\ell\ \text{odd}}1=\sum_{\ell}\left(r_\ell-2\left\lfloor\frac{r_\ell}{2}\right\rfloor\right)=rq+r-2\sum_{\ell}\left\lfloor\frac{r_\ell}{2}\right\rfloor.$$

Hence $\sum\left\lfloor\frac{r_\ell}{2}\right\rfloor=\frac{r}{2}+2t$. □

Note that by Lemma 3 $s=|\mathcal{L}^o(S)|$ must be of the form $rq-4t$ with $0\le t\le\binom{r}{2}$. Since we are interested in the smallest r for which a suitable set S exists, typically we expect t to be relatively small and r not much bigger that s/q. We can therefore reduce the problem to the question of (a) whether there is *any* clique decomposition of K_r into cliques of size $r_1,\dots,r_n$ with a given value of $\sum\left\lfloor\frac{r_i}{2}\right\rfloor$, and (b) whether such a decomposition can be realized by a set of points inside $PG(2,q)$.

We call an edge-decomposition Π of K_r into cliques of orders $r_1,\dots,r_n$ a *simple decomposition* if there is at most one value of i with $r_i>3$. In other words, K_r is decomposed as single edges, triangles, and at most one larger clique. We write $M(\Pi)$ for the sum $\sum_{i=1}^n\left\lfloor\frac{r_i}{2}\right\rfloor$.

Lemma 15. *Suppose we are given an edge-decomposition Π of K_r with $M(\Pi)<\frac{1}{4}r(\sqrt{4r-3}-1)$. Then there exists a simple edge-decomposition Π' of K_r with $M(\Pi')=M(\Pi)$.*

Proof. Assume Π decomposes K_r into cliques of orders $r_1,\dots,r_n$ with $r_1\ge r_2\ge\dots\ge r_n$. Let C_i be the i'th clique. Then there are $r_1(r-r_1)$ edges from $V(C_1)$ to $V(K_r)\setminus V(C_1)$. Moreover, each clique C_i, $i>1$, can meet C_1 in at most one vertex and hence covers at most r_i-1 of these edges. Thus $\sum_{i>1}(r_i-1)\ge r_1(r-r_1)$ and hence

$$M(\Pi)\ge\sum_{i=1}^n\frac{r_i-1}{2}\ge\frac{r_1-1}{2}+\frac{r_1(r-r_1)}{2}.\tag{4}$$

On the other hand there are $\binom{r}{2}$ edges to be covered in total, so

$$M(\Pi) \geq \sum_{i=1}^{n} \frac{r_i - 1}{2} = \sum_{i=1}^{n} \frac{1}{r_i}\binom{r_i}{2} \geq \frac{1}{r_1}\binom{r}{2}. \tag{5}$$

For $r_1 < r/2$, the bound in (4) is increasing and the bound in (5) is decreasing as r_1 increases, so the smallest bound on $M(\Pi)$ occurs when the two bounds are equal. It can be checked that this occurs when $r = r_1^2 - r_1 + 1$ with a common bound $M(\Pi) \geq \frac{1}{2}r(r_1 - 1) = \frac{1}{4}r(\sqrt{4r-3} - 1)$. This contradicts the assumption on $M(\Pi)$, so we may assume $r_1 \geq r/2$.

Let E_1 be the set of $r_1(r - r_1)$ edges joining C_1 to the rest of K_r and E_2 be the set of $\binom{r-r_1}{2}$ edges of K_r not meeting C_1. For each clique C_i, $i > 1$, we note that for all $r_i \geq 2$,

$$|E_1 \cap E(C_i)| - |E_2 \cap E(C_i)| \leq \left\lfloor \frac{r_i}{2} \right\rfloor \leq |E_1 \cap E(C_i)| + |E_2 \cap E(C_i)|.$$

Indeed, the right hand side is just $\binom{r_i}{2}$, while the left hand side is either $(r_i - 1) - \binom{r_i - 1}{2}$ or $-\binom{r_i}{2}$ depending on whether or not C_i meets some vertex of C_1. Note that the lower bound is achieved if $r_i \in \{2, 3\}$ and C_i meets C_1. Summing over all cliques gives

$$\left\lfloor \frac{r_1}{2} \right\rfloor + |E_1| - |E_2| \leq M(\Pi) \leq \left\lfloor \frac{r_1}{2} \right\rfloor + |E_1| + |E_2|. \tag{6}$$

Also note that $\lfloor \frac{r_i}{2} \rfloor \equiv \binom{r_i}{2} \bmod 2$, so that $M(\Pi)$ is equivalent to either bound modulo 2.

As $r_1 \geq r/2$, the graph on $E_1 \cup E_2$ can be packed with $|E_2|$ triangles each meeting C_1. Indeed, it is enough to decompose K_{r-r_1} completely into at most r_1 partial matchings $M_1, \ldots, M_{r_1}$ and then join each matching to a distinct vertex of C_1 to obtain sets of edge-disjoint triangles. For even $r - r_1$, it is well-known that K_{r-r_1} can be decomposed into $r - r_1 - 1 < r_1$ perfect matchings. For odd $r - r_1$ decompose K_{r-r_1+1} into $r - r_1 \leq r_1$ perfect matchings and remove a single vertex to give a decomposition of K_{r-r_1} into $r - r_1$ partial matchings. Completing the packing of $E_1 \cup E_2$ by including K_2s covering the remaining edges of E_1 gives a decomposition Π'' of K_r that achieves the lower bound $M_0 = \lfloor r_1/2 \rfloor + |E_1| - |E_2|$ in (6). Now replacing $(M(\Pi) - M_0)/2 \leq |E_2|$ of the triangles of this packing with three K_2s, allows us to increase $M(\Pi'')$ in steps of 2 until we get to a packing Π' of C_1, edges, and triangles, with $M(\Pi') = M(\Pi)$. □

Lemma 16. *Let $m = \lceil \sqrt{r-3} \rceil - 1$. Then for any integer s with $s \leq \binom{r}{2}$, $s \equiv \binom{r}{2} \bmod 2$, and $s \geq \lfloor \frac{r-m}{2} \rfloor + \frac{m}{2}(2r - 3m + 1)$ there exists a simple decomposition Π of K_r with $M(\Pi) = s$.*

Proof. From the proof of Lemma 15 we know that we can construct a simple a decomposition for any $s \equiv \binom{r}{2}$ and

$$\left\lfloor \frac{r_1}{2} \right\rfloor + r_1(r - r_1) - \binom{r - r_1}{2} \leq s \leq \left\lfloor \frac{r_1}{2} \right\rfloor + r_1(r - r_1) + \binom{r - r_1}{2}$$

with $r_1 \geq \frac{r}{2}$. It is a simple but tedious exercise to show that the intervals for $r_1 = \lceil \frac{r}{2} \rceil, \ldots, r - m$ cover every $s \equiv \binom{r}{2}$ in the range from $\lfloor \frac{r-m}{2} \rfloor + \frac{m}{2}(2r - 3m + 1)$ to $\frac{3}{4}\binom{r}{2}$.

For $s > \frac{3}{4}\binom{r}{2}$ it is enough to show that one can pack $(\binom{r}{2} - s)/2 \le \binom{\lfloor r/2 \rfloor}{2}$ triangles into K_r. This also follows from the proof of Lemma 15 where it was shown that one can pack $\binom{\lfloor r/2 \rfloor}{2}$ triangles into $K_r \setminus E(K_{\lceil r/2 \rceil})$. □

Lemmas 15 and 16 show that if there exists a decomposition with $M(\Pi) = s$ then there exists a simple decomposition with $M(\Pi) = s$ except possibly in the range between about $\frac{1}{2}r^{3/2}$ and about $r^{3/2}$. There can exist non-simple decompositions in this range for which there is no simple decomposition. For example, the lines of a projective plane of order q', q' odd, give rise to a decomposition Π of K_r when $r = q'^2 + q' + 1$ with $M(\Pi) = (q'^2 + q' + 1)(q' + 1)/2$ (exactly the bound in Lemma 15). One can check that for a simple decomposition to have the same value of $M(\Pi)$ would require $\frac{q'-1}{2} < r_1 < \frac{q'+1}{2}$ for large q', an impossibility, so no corresponding simple decomposition exists.

5 REALIZING CLIQUE DECOMPOSITIONS OF THE PROJECTIVE PLANE

We now turn to the question of whether a simple decomposition can be realized by a set of points in $PG(2, q)$. One needs a set S formed by taking a large number r_1 of points in one line, and the remaining points only on lines intersecting S in at most three points. The proof of the following lemma provides a construction that realizes this in most relevant cases.

Lemma 17. *Fix r, $0 \le r \le q + 1$ and assume $r_1 \ge \max\{\frac{1}{3}(2r - 3), (2r - 3) - (q + 1)\}$. Then any simple decomposition Π of K_r with maximal clique of order r_1 can be realized by a set of points in $PG(2, q)$.*

Proof. Consider sets of points that are subsets of $C \cup L$, where $C = \{XZ = Y^2\}$ is the conic used in the proof of Theorem 4 and $L = \{X = dZ\}$ is a line that does not intersect C (so d is chosen to be a quadratic nonresidue in the field $\mathbb{F}_q$). A simple calculation shows that the secant line joining $[s^2 : st : t^2]$ and $[s'^2 : s't' : t'^2]$ on C meets L at the point $[d(st' + s't) : dtt' + ss' : st' + s't]$ on L. This mapping of pairs of points on C to L is more easily described by introducing the norm group $G = \mathbb{F}_{q^2}^\times / \mathbb{F}_q^\times$. The points $p = [s^2 : st : t^2] \in C$ correspond to the coset $\phi(p) = (s + t\sqrt{d})\mathbb{F}_q^\times$ and the coset $\alpha = (a + b\sqrt{d})\mathbb{F}_q^\times$ corresponds to the point $\psi(\alpha) = [db : a : b] \in L$. The secant line through $p, p' \in C$ then meets L at $\psi(\phi(p)\phi(p'))$. The key point is that G is cyclic of order $q + 1$. Hence by taking a subset $P = \{p_1, p_2, \ldots, p_s\}$ of C with $2s - 3 \le q + 1$ such that $\phi(p_i)$ form a suitable geometric progression, the secants through these points meet L in only $2s - 3$ points (assuming $s \ge 2$). Indeed, we can take $\phi(p_i) = \alpha^i$ where α is a generator of G so that the secants meet L at the points $\psi(\alpha^3), \psi(\alpha^4), \ldots, \psi(\alpha^{2s-1})$. Moreover there are 4 points $(\psi(\alpha^3), \psi(\alpha^4), \psi(\alpha^{2s-2}), \psi(\alpha^{2s-1}))$ on L that meet just one secant, four which meet exactly two secants, etc., with one or three points meeting $\lfloor s/2 \rfloor$ secants (depending on the parity of s). Now let $P' = \{p'_1, \ldots, p'_t\}$ be a set of t points on the line L and suppose there are k secants through two points of P meeting P'. then $P \cup P'$ induces a simple edge decomposition of $K_{P \cup P'}$ with one clique of order $|P'|$ and k triangles, the remaining cliques being single edges.

We now consider the conditions on the parameter that allow us to vary k between the minimum of zero and the maximum of $\binom{s}{2}$, where $s \ge 2$. To achieve $k = 0$ requires $t \le (q + 1) - (2s - 3)$ as P' must avoid all the secant lines through P. To achieve $k = \binom{s}{2}$

requires $t \geq 2s - 3$ as P' must meet all secants through P. All values of k between the minimum and maximum can be achieved one step at a time by moving some point of P' so that it meets one more secant line. Now $s = r - r_1$ and $t = r_1$ so these conditions become

$$r_1 \leq q + 1 - (2r - 2r_1 - 3) \qquad \text{and} \qquad r_1 \geq 2r - 2r_1 - 3,$$

or equivalently $r_1 \geq (2r - 3) - (q + 1)$ and $r_1 \geq \frac{1}{3}(2r - 3)$. For $s < 2$ there are no secant lines and the only restriction is $t = r_1 \leq q + 1$ that follows from $r_1 \leq r \leq q + 1$. □

Corollary 18. *There exists an absolute constant $C > 0$ such that $w/q \leq f(w) \leq w/q + C(w^{3/2}/q^{5/2} + 1)$ for all even w with $Cq^{3/2} \leq w \leq N - Cq^{3/2}$.*

Note that for odd w, $N - w$ is even and so $(N - w)/q \leq f(w) = f(N - w) \leq (N - w)/q + C((N - w)^{3/2}/q^{5/2} + 1)$.

Proof. By choosing C sufficiently large we may assume that q is also large. The lower bound follows from Lemmas 12 and 3. For the upper bound choose r minimal such that $r > w/q + 2w^{3/2}/q^{5/2}$ and $r \equiv qw \bmod 4$. Write $w = rq - 4t$, so that $r^{3/2} \leq 4t \ll r^2$ and $r > \sqrt{q}$. By Lemma 16 there exists a simple decomposition of K_r with $M(\Pi) = r/2 + 2t$ and indeed, this decomposition must have maximal clique size $r_1 = r - O(\sqrt{r})$. Then by Lemma 17 this decomposition can be realized by a subset S of $PG(2, q)$. Now $|\mathcal{L}^o(S)| = qr - 4t = w$ by Lemma 14 and so $f(w) \leq r \leq w/q + C(w^{3/2}/q^{5/2} + 1)$. □

6 FURTHER CONSTRUCTIONS FROM BLOCKING SETS AND THE MAXIMUM OF $f(r)$

We shall now provide some constructions that give at least some reasonable bounds on $f(r)$ for $r < Cq^{3/2}$ or $r > N - Cq^{3/2}$.

Let $Q^+ \subseteq \mathbb{F}_q$ be the set of nonzero quadratic residues and $Q^- \subseteq \mathbb{F}_q$ be the set of quadratic nonresidues. Both sets have $(q - 1)/2$ elements. Define $Q_i \subseteq \mathcal{P}$, $i = 0, 1$ by

$$Q_0 = \{[x\!:\!0\!:\!1] : x \in Q^+\} \cup \{[1\!:\!x\!:\!0] : x \in Q^+\} \cup \{[0\!:\!1\!:\!x] : x \in Q^-\},$$

and

$$Q_1 = \{[x\!:\!0\!:\!1] : x \in Q^+\} \cup \{[1\!:\!x\!:\!0] : x \in Q^+\} \cup \{[0\!:\!1\!:\!x] : x \in Q^+\}.$$

Given any line $\ell : \alpha X + \beta Y + \gamma Z = 0$ that does not go through the points $O_x := [1\!:\!0\!:\!0]$, $O_y := [0\!:\!1\!:\!0]$, $O_z := [0\!:\!0\!:\!1]$, we have $|\ell \cap Q_i| \equiv i \bmod 2$. Indeed, ℓ intersects $\{[x\!:\!0\!:\!1] : x \in Q^+\}$ iff $\alpha/\gamma \in Q^+$ and similarly for the others. But for any $\alpha, \beta, \gamma \neq 0$ an odd number of the conditions $\alpha/\gamma \in Q^+$, $\beta/\gamma \in Q^+$, and $\gamma/\alpha \in Q^+$ hold.

The example Q_0 is due to J. di Paola. By a famous result of Blokhuis [4] the set $Q_0 \cup \{O_x, O_y, O_z\}$ is the smallest nontrivial blocking set on $PG(2, q)$ when q is prime.

Lemma 19.

$$f\left(\frac{3}{2}(q - 1) + kq + j\right) \leq 3q + j(q + 2 - j)$$

for $0 \leq k \leq (q - 1)/2$ *and* $0 \leq j \leq q + 1$.

Proof. Let V be the set of kq points that lie in one of k "vertical" lines of the form $X = \alpha Z$, $\alpha \in Q^-$, not including the point O_y at infinity. Let C be any set of j points on the conic $XZ = Y^2$. Note that V, Q_i, and C are pairwise disjoint for $i = 0, 1$. Let $S = V \cup Q_{k \bmod 2} \cup C$ so that $|S| = \frac{3}{2}(q-1) + kq + j$. Consider a line ℓ that does not meet $\{O_x, O_y, O_z\}$. Then $|\ell \cap V| = k$ and $|\ell \cap Q_{k \bmod 2}| \equiv k \bmod 2$. Thus $|\ell \cap S| \equiv |\ell \cap C| \bmod 2$. From the proof of Theorem 4 there are $j(q+2-j)$ lines that meet C in an odd number of points, and there are only $3q$ lines that meet $\{O_x, O_y, O_z\}$, so $f(|S|) \le |\mathcal{L}^o(S)| \le 3q + j(q+2-j)$ as required. □

Lemma 20.

$$f(kq + j) \le k + j(q + 2 - j)$$

for $0 \le k \le (q-1)/2$, *k even, and* $0 \le j \le q+1$.

Proof. Let V and C be as in the proof of Lemma 19. Then the number of lines meeting C in an odd number of points is $j(q+2-j)$ while the number of lines meeting V in an odd number of points is just k (the lines of V). As $|V \cup C| = kq + j$, $f(kq+j) \le k + j(q+2-j)$. □

Lemma 21.

$$f(q + 1 + kq + j) \le q + 1 + k + j(q + 2 - j)$$

for $0 \le k \le (q-1)/2$, *k even, and* $0 \le j \le q-1$,

Proof. Let V and C be as in the proof of Lemma 19 except that we shall now insist that $O_x, O_z \notin C$. Let C' be the conic $XZ = 4Y^2$. Note that C' could only meet C at the points O_x, O_z, which we have assumed do not lie in C. Also $C' \cap V = \emptyset$. There are $q+1$ lines that meet C' in an odd number of points, $j(q+2-j)$ lines that meet C in an odd number of points, and k lines that meet V in an odd number of points. The result follows since $|V \cup C \cup C'| = q + 1 + kq + j$. □

Corollary 22. *For large q, the maximum value of* $f(r)$ *is* $(q^2 + 4q + 3)/4$ *and occurs only at* $r = (q+1)/2$, $r = (q+3)/2$, $r = N - (q+1)/2$, *and* $r = N - (q+3)/2$.

Proof. The result follows when r is restricted to the range $0 \le r \le q+1$ and $N - (q+1) \le r \le N$ by Theorem 4 and Lemma 2, so it is enough by Lemma 2 to bound $f(r)$ in the range $r \in [q+2, N/2]$. For $r \in [q+2, (\frac{3}{2} - \varepsilon)q]$ we can apply Lemma 21 with $k = 0$ to obtain $f(r) \le (\frac{1}{4} - \varepsilon^2)q^2 + O(q)$. For $r \in [(\frac{3}{2} - \varepsilon)q, \frac{3}{2}(q-1)]$ we can apply Lemma 19 with $k = j = 0$ and Theorem 5 to obtain $f(r) \le 3q + (q-1)\varepsilon q$. Thus we may assume $r \ge \frac{3}{2}(q-1)$.

If $|r/q - t| \ge \frac{1}{4}$ for every integer t, then we write $r = \frac{3}{2}(q-1) + kq + j$, where either $0 \le j \le \frac{3}{2} + \frac{q}{4}$ or $\frac{3}{2} + \frac{3q}{4} \le j < q$. In either case Lemma 19 implies

$$f(r) \le 3q + \frac{q+5}{4} \cdot \frac{3q+3}{4} = \frac{1}{16}(3q^2 + 66q + 15).$$

If $|r/q - t| < \frac{1}{4}$ and $\lfloor (r-1)/q \rfloor$ is even, we write $r = kq + j$ with $1 \le j < \frac{q}{4}$ or $\frac{3q}{4} < j \le q$. In either case Lemma 20 gives

$$f(r) \le k + \frac{3q+1}{4} \cdot \frac{q+7}{4} \le \frac{1}{16}(3q^2 + 30q - 1).$$

Finally, if $|r/q - t| < \frac{1}{4}$ and $\lfloor (r-1)/q \rfloor$ is odd, we write $r = q + 1 + kq + j$ with $0 \le j < \frac{q}{4} - 1$ or $\frac{3q}{4} - 1 < j \le q$. In either case Lemma 21 gives

$$f(r) \le q + 1 + k + \frac{3q-3}{4} \cdot \frac{q+11}{4} \le \frac{1}{16}(3q^2 + 38q + 24).$$

Thus in all cases

$$f(r) \le \frac{1}{16}(3q^2 + 66q + 15) < \frac{1}{4}(q^2 + 4q + 3).$$

for q sufficiently large. □

7 EXACT VALUES FROM THE BAER SUBPLANE

A subset of points $S \subseteq \mathcal{P}$ is a *subplane of order k* if $|S| = k^2 + k + 1$ and the sets $\{\ell \cap S : \ell \in \mathcal{L},\ |\ell \cap S| > 1\}$ form the line system of a finite projective plane of order k. In the case when $k = \sqrt{q}$, we call S a *Baer subplane*. It is well known that such Baer subplanes exists whenever q is a perfect square (see Bruck [7]). Even more (see, e.g. Yff [16]) $\mathcal{P}$ can be partitioned into $q - \sqrt{q} + 1$ Baer subplanes.

Consider a Baer sublane B and let $R_B \subseteq \mathcal{L}$ be the set of lines meeting it in exactly $\sqrt{q} + 1$ points. Then $|R_B| = q + \sqrt{q} + 1$. The lines of R_B cover every point of B exactly $\sqrt{q} + 1$ times, and every other point exactly once. Thus $\mathcal{P}^o(R_B) = \mathcal{P} \setminus B$, which is very large. However, consider an arbitrary point $p \notin B$ and let R be the symmetric difference of R_B and $\mathcal{L}(\{p\})$ (these two families contain only one common line $\ell_p \in R_B$ through p). Then $\mathcal{P}^o(R) = B \cup \{p\}$. We obtain

$$f(2q + \sqrt{q}) \le q + \sqrt{q} + 2. \tag{7}$$

Considering $p \in B$ and the set of even lines of $B \setminus \{p\}$ (it is again the symmetric difference of R_B and $\mathcal{L}(\{p\})$, now they have $\sqrt{q} + 1$ common lines) we obtain

$$f(2q - \sqrt{q}) \le q + \sqrt{q}. \tag{8}$$

Considering two disjoint Baer subplanes we get

$$f(2q + 2\sqrt{q} + 2) \le 2q + 2\sqrt{q} + 2. \tag{9}$$

Theorem 23. *Equality holds in (7) and (8) for* $q \ge 81$.

We also **conjecture** that equality holds in (9), too (at least for large enough q). For the proof of Theorem 23 we need the following classical results and a few lemmata.

Lemma 24 (Bruen [8], sharpening by Bruen and Thas [9]). *Suppose that $S \subseteq \mathcal{P}$ is a nontrivial blocking set (i.e. it meets every line but does not contain any) then $|S| \geq q + \sqrt{q} + 1$. Moreover, if $|S| = q + \sqrt{q} + 2$, and $q \geq 9$ is of square order, then there exists a point $x \in S$ such that $S \setminus \{x\}$ is the point set of a Baer subplane.*

Let $\mathcal{U} \subseteq \mathcal{L}$ be a set of lines. A set $C \subseteq \mathcal{P}$ is called a *near-blocker of $\mathcal{U}$* if it meets exactly all but one member of $\mathcal{U}$.

Lemma 25. *Let $\mathcal{U}$ be a set of lines in $PG(2, q)$.*

(a) Suppose that $\cap_{\ell \in \mathcal{U}} \ell = \emptyset$. Then there exists a near-blocker of size at most $|\mathcal{U}|/2$.
(b) Suppose that $q \geq 5$ is odd and $\mathcal{U}$ cannot be blocked by a 2-element set. Then there exists a near-blocker of size at most $|\mathcal{U}|/3 + (q + 1)/6$.

Proof. (a) Let us apply induction on the size of $|\mathcal{U}|$. The cases $|\mathcal{U}| = 1, 2, 3$ are trivial. If $\mathcal{U}$ cannot be covered by two points then select any point $p \in \mathcal{P}$ covered at least twice by the lines of $\mathcal{U}$ and use induction from $\mathcal{U} \setminus \mathcal{L}(\{p\})$. Otherwise, some two points x_1, x_2 cover all lines. Assuming that $\deg_{\mathcal{U}}(x_1) \geq \deg_{\mathcal{U}}(x_2)$, select x_1 and one element from all but one of the lines of $\mathcal{U}$ going through x_2 and avoiding x_1.

(b) For $|\mathcal{U}| \leq q + 2$ we have $\lfloor |\mathcal{U}|/2 \rfloor \leq |\mathcal{U}|/3 + (q + 1)/6$ and we can apply case (a). (If $|\mathcal{U}| = q + 2$ we make use of the fact that q is odd.) We may now suppose $|\mathcal{U}| \geq q + 3$, so $\max_p \deg_{\mathcal{U}}(p) \geq 3$. Consider first the case when $\mathcal{U}$ cannot be covered by three vertices. Chose a maximum degree vertex p and apply the induction hypothesis to $\mathcal{U} \setminus \mathcal{L}(\{p\})$. Finally, if some set $\{x_1, x_2, x_3\}$ meets every member of $\mathcal{U}$ we choose the two highest degree vertices among them and one element from all but one of the lines of $\mathcal{U}$ going through the third, avoiding the other two. In this way we obtain a near-cover of size at most $2 + (|\mathcal{U}|/3 - 1)$. □

The following lemma will be useful when $|\mathcal{L}^e(A)|$, t_1, and t_2 are all small.

Lemma 26.

(a) Let $A = (\ell \setminus T_1) \cup T_2$ where ℓ is a line, $T_1 \subseteq \ell$, $T_2 \cap \ell = \emptyset$, and $t_i = |T_i|$. Then $|\mathcal{L}^e(A)| \geq (t_1 + t_2)q - t_2(2t_1 + t_2 - 2)$.
(b) Let $A = (B \setminus T_1) \cup T_2$ where B is a Baer subplane, $T_1 \subseteq B$, $T_2 \cap B = \emptyset$, and $t_i = |T_i|$. Then $|\mathcal{L}^e(A)| \geq (t_1 + t_2)q - t_2(2t_1 + t_2 - 1) - t_1\sqrt{q}$.

Proof. (a) Consider the lines through a point $x \in T_2$. Exactly $q + 1 - t_1$ of them meet $\ell \setminus T_1$. At most $t_2 - 1$ of these lines contain a further point of A (namely a point from T_2). Thus we have obtained at least $t_2(q + 1 - t_1 - (t_2 - 1))$ 2-point lines. Next consider the q lines through a point $y \in T_1$ other than ℓ. All but t_2 avoids T_2, too, thus giving at least $t_1(q - t_2)$ zero-point lines. The total number of these lines gives the desired lower bound.

(b) Every point $x \in T_2$ is incident to at least $(q - t_1) - (t_2 - 1)$ 2-point lines, and every point $y \in T_1$ is incident to at least $q - \sqrt{q} - t_2$ zero-point lines. □

Proof of equality in (7). Suppose, on the contrary, that we have a set of lines R, $|R| = 2q + \sqrt{q}$, such that for $S = \sum_{\ell \in R} \ell$ we have $|S| < q + \sqrt{q} + 2$. Since $|S|$ is even, we have $|S| \leq q + \sqrt{q}$. Since R is odd we have $R = \mathcal{L}^e(S)$. Thus S meets every line from $\mathcal{L} \setminus R$. Let $\mathcal{U}$ be the set of lines avoiding S, we have $\mathcal{U} \subseteq R$.

First consider the case when there is a set V, $|V| \leq 2$, meeting all points of $\mathcal{U}$. (This includes the case $\mathcal{U} = \emptyset$.) Then $S \cup V$ meets all lines, so is a blocking set.

We claim that $S \cup V$ does not contain a line, so is a non-trivial blocking set. Suppose, on the contrary, that there is a line $\ell \subseteq S \cup V$. Apply Lemma 26 (a) with $A = S = (\ell \setminus T_1) \cup T_2$ where $T_1 = \ell \cap V$, $|T_1| \le 2$ and $T_2 = S \setminus \ell$, $|T_2| \le |S \cup V| - |\ell| \le \sqrt{q} + 1$. We obtain that

$$|\mathcal{L}^e(S)| \ge t_1 q + t_2(q + 2 - 2t_1 - t_2) \ge t_1 q + t_2(q - \sqrt{q} - 3).$$

Since $|\mathcal{L}^e(S)| = 2q + \sqrt{q}$ we obtain that $|T_1| + |T_2| \le 2$ for $q \ge 49$.

We finish the proof of our claim by observing that for $|T_1| + |T_2| \le 2$, $T_1 \subseteq \ell$, the number of even lines $|\mathcal{L}^e((\ell \setminus T_1) \cup T_2)|$ cannot be $2q + \sqrt{q}$. Indeed, in the case $T_1 = \emptyset$ we have $|\mathcal{L}^e(S)| \le t_2 q + 2 < 2q + \sqrt{q}$. In the case $t_2 = 0$ we have $|\mathcal{L}^e(S)| \le 1 + t_1 q < 2q + \sqrt{q}$. Finally, in the case $t_1 = t_2 = 1$ we have $|\mathcal{L}^e(S)| = 2q - 1 < 2q + \sqrt{q}$.

Consider $S \cup V$, which is a nontrivial blocking set of size at most $q + \sqrt{q} + 2$. By the Bruen-Thas theorem (Lemma 24) there is a Baer subplain $B \subseteq S \cup V$. Thus we know a lot about the structure of S, we can write $S = (B \setminus T_1) \cup T_2$ where $T_1 = B \setminus S$ (it is a subset of V, so $t_1 \le 2$) and $T_2 = S \setminus B \subseteq (S \cup V) \setminus B$ so $t_2 \le 1$.

We finish the proof of the case $|V| \le 2$ by checking all possible values of t_1 and t_2. In case of $t_1 = 2$, $t_2 = 1$, Lemma 26 (b) applied to $A = S$ gives $|\mathcal{L}^e(S)| \ge 3q - 4 - 2\sqrt{q}$. This exceeds $2q + \sqrt{q}$ for $q \ge 25$. We obtain that $t_1 + t_2 \le 2$. Since $|S|$ is even and $|B|$ is odd their symmetric difference (i.e. $T_1 \cup T_2$) is odd, we get $t_1 + t_2 = 1$. So S should be one of the examples discussed in the beginning of this section and we are done.

From now on suppose that there is no set V, $|V| \le 2$, meeting all points of $\mathcal{U}$. Apply Lemma 25 (b) to $\mathcal{U}$ to obtain a near-blocker C of $\mathcal{U}$ of size at most $|\mathcal{U}|/3 + (q+1)/6$ and a line $\ell_C \in \mathcal{U}$ missed by C. We proceed as in the proof of Theorem 6.

The set $S \cup C$ meets all lines except ℓ_C, so it is a blocking set of the *affine* plane $PG(2, q) \setminus \ell_C$. Then Lemma 8 yields $|S \cup C| \ge 2q - 1$. We obtain

$$2q - 1 \le |S| + |C| \le (q + \sqrt{q}) + |\mathcal{U}|/3 + (q+1)/6.$$

Here $|\mathcal{U}| \le |R| = 2q + \sqrt{q}$ so the right hand side is at most $(11q + 8\sqrt{q} + 1)/6$. This cannot hold for $q \ge 81$. This final contradiction implies that $|S| \le q + \sqrt{q}$ is not possible for $q \ge 81$ and we are done. □

Proof of equality in (8). This proof is similar to the previous proof, but simpler. Suppose, on the contrary, that we have a set of lines R, $|R| = 2q - \sqrt{q}$ such that for $S = \sum_{\ell \in R} \ell$ we have $|S| < q + \sqrt{q}$. As $|S|$ is even, we have $|S| \le q + \sqrt{q} - 2$. Since R is odd we have $R = \mathcal{L}^e(S)$. Thus S meets every line from $\mathcal{L} \setminus R$. Let $\mathcal{U}$ be the set of lines avoiding S, so that $\mathcal{U} \subseteq R$.

If there is a set V, $|V| \le 2$, meeting all points of $\mathcal{U}$ (including the case $\mathcal{U} = \emptyset$) then $S \cup V$ meets all lines, it is a blocking set of size at most $q + \sqrt{q}$. By the Bruen theorem (Lemma 24) it must contain a line ℓ. Apply Lemma 26 (a) with $A = S = (\ell \setminus T_1) \cup T_2$ where $T_1 = \ell \cap V$, $|T_1| \le 2$ and $T_2 = S \setminus \ell$, $|T_2| \le |S \cup V| - |\ell| \le \sqrt{q} - 1$. We obtain that

$$|\mathcal{L}^e(S)| \ge t_1 q + t_2(q + 2 - 2t_1 - t_2) \ge t_1 q + t_2(q - \sqrt{q} - 1).$$

Since $|\mathcal{L}^e(S)| = 2q - \sqrt{q}$ we obtain that $|T_1| + |T_2| \le 2$ for $q \ge 25$.

We finish the investigation of this case by observing that for $|T_1| + |T_2| \le 2$, $T_1 \subseteq \ell$, the number of even lines $|\mathcal{L}^e((\ell \setminus T_1) \cup T_2)|$ cannot be $2q - \sqrt{q}$. Since both S and ℓ are even sets, their symmetric difference (i.e. $T_1 \cup T_2$) is even. We have four cases to check according to the value of $(t_1, t_2) \in \{(2, 0), (1, 1), (0, 2), (0, 0)\}$. The sizes of $|\mathcal{L}^e(S)|$ are $2q + 1$, $2q - 1$, again $2q + 1$, and 1, respectively. None of these is equal to $2q - \sqrt{q}$.

From now on suppose that $\mathcal{U} \neq \emptyset$ and there is no set V, $|V| \leq 2$, meeting all points of $\mathcal{U}$. Apply Lemma 25 (b) to $\mathcal{U}$ to obtain a near-blocker C of $\mathcal{U}$ of size at most $|\mathcal{U}|/3 + (q + 1)/6$ and a line $\ell_C \in \mathcal{U}$ missed by C. We proceed as in the proof of Theorem 6.

The set $S \cup C$ meets all lines except ℓ_C, so it can be considered as a blocking set of the affine plane $PG(2, q) \setminus \ell_C$. Then Lemma 8 yields $|S \cup C| \geq 2q - 1$. We obtain

$$2q - 1 \leq |S| + |C| \leq (q + \sqrt{q} - 2) + |\mathcal{U}|/3 + (q + 1)/6.$$

Here $|\mathcal{U}| \leq |R| = 2q - \sqrt{q}$ so the right-hand-side is at most $(11q + 4\sqrt{q} - 11)/6$. This cannot hold for $q \geq 49$ implying that $|S| \leq q + \sqrt{q}$ is not possible for $q \geq 49$ and we are done. $\square$

With some more work we can see that only the examples from the Baer subplane give equalities in (7) and (8) (for $q > q_0$).

Many questions remain open. What is $f(q + 2)$, and $f(q + 3)$? The least we should be able to do is to prove better bounds on these. Also, any information about $f(r)$ for $r \leq 2q^{3/2}$ would be great.

ACKNOWLEDGMENTS

The authors are indebted to the referees for their helpful comments and suggestions.

REFERENCES

[1] N. Alon, Combinatorial Nullstellensatz, Combin Prob Comput 8 (1999), 7–29.

[2] A. Bichara and G. Korchmáros, Note on $(q + 2)$-sets in a Galois plane of order q, Combinatorial and geometric structures and their applications (Trento, 1980), pp. 117–121, Ann. Discrete Math. 14, North-Holland, Amsterdam-New York, 1982.

[3] S. Besicovitch, The Kakeya problem, Amer Math Monthly 70 (1963), 697–706.

[4] A. Blokhuis, On the size of a blocking set in $PG(2, p)$, Combinatorica 14 (1994), 111–114.

[5] A. Blokhuis and F. Mazzocca, The finite field Kakeya problem, Building bridges 205–218, Bolyai Soc Math Stud 19, Springer, Berlin, 2008.

[6] A. E. Brouwer and A. Schrijver, The blocking number of an affine space, J Combin Theory Ser A 24 (1978), 251–253.

[7] R. H. Bruck, Quadratic extensions of cyclic planes, Proc Sympos Appl Math, Vol. 10, pp. 15–44. American Mathematical Society, Providence, RI, 1960.

[8] A. Bruen, Blocking sets in finite projective planes, SIAM J Appl Math 21 (1971), 380–392.

[9] A. A. Bruen and J. A. Thas, Blocking sets, Geometriae Dedicata 6 (1977), 193–203.

[10] X. W. C. Faber, On the finite field Kakeya problem in two dimensions, J Number Theory 124 (2007), 248–257.

[11] J. W. P. Hirschfeld, Projective geometries over finite fields, 2nd edn. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1998.

[12] R. Jamison, Covering finite fields with cosets of subspaces, J Combin Theory Ser A 22 (1977), 253–266.

[13] L. Lovász and A. Schrijver, Remarks on a theorem of Rédei, Studia Scient Math Hungar 16 (1981), 449–454.

[14] H. J. Ryser, Geometries and incidence matrices, Amer Math Monthly 62 (1955), 25–31.

[15] T. Szőnyi, Blocking sets in Desarguesian affine and projective planes, Finite Fields Appl 3 (1997), 187–202.

[16] P. Yff, On subplane partitions of a finite projective plane, J Combinatorial Theory Ser A 22 (1977), 118–122.

APPENDIX: VALUES OF $f(r)$ FOR SMALL q:

TABLE A1. $q = 3$

r	$f(r)$	r	$f(r)$
1	4	4	4
2	6	5	4
3	6	6	2

TABLE A2. $q = 5$

r	$f(r)$	r	$f(r)$	r	$f(r)$
1	6	6	6	11	4
2	10	7	8	12	4
3	12	8	8	13	6
4	12	9	6	14	6
5	10	10	2	15	4

TABLE A3. $q = 7$

r	$f(r)$	r	$f(r)$	r	$f(r)$	r	$f(r)$
1	8	8	8	15	6	22	6
2	14	9	12	16	8	23	6
3	18	10	10	17	8	24	4
4	20	11	10	18	6	25	8
5	20	12	12	19	10	26	6
6	18	13	8	20	4	27	6
7	14	14	2	21	8	28	4

TABLE A4. $q = 9$

r	*f(r)*	*r*	*f(r)*	*r*	*f(r)*	*r*	*f(r)*		
1	10	10	10	19	8	28	4	37	6
2	18	11	16	20	12	29	10	38	6
3	24	12	12	21	10	30	6	39	8
4	28	13	14	22	10	31	8	40	8
5	30	14	14	23	12	32	4	41	10
6	30	15	12	24	8	33	10	42	6
7	28	16	16	25	10	34	6	43	8
8	24	17	10	26	10	35	8	44	8
9	18	18	2	27	12	36	4	45	6

TABLE A5. $q = 11$

r	*f(r)*	*r*	*f(r)*	*r*	*f(r)*	*r*	*f(r)*	*r*	*f(r)*	*r*	*f(r)*
1	12	12	12	23	10	34	10	45	8	56	8
2	22	13	20	24	16	35	14	46	6	57	8
3	30	14	14–26	25	16	36	4	47	10	58	6
4	36	15	14–18	26	14	37	12	48	8	59	10
5	40	16	16	27	14	38	10	49	12	60	8
6	42	17	16	28	12	39	10	50	6	61	8
7	42	18	14–18	29	16	40	4	51	10	62	10
8	40	19	14–26	30	10	41	12	52	8	63	10
9	36	20	16–20	31	14–18	42	6	53	12	64	8
10	30	21	12	32	12	43	14	54	6	65	8
11	22	22	2	33	16	44	4	55	10	66	6

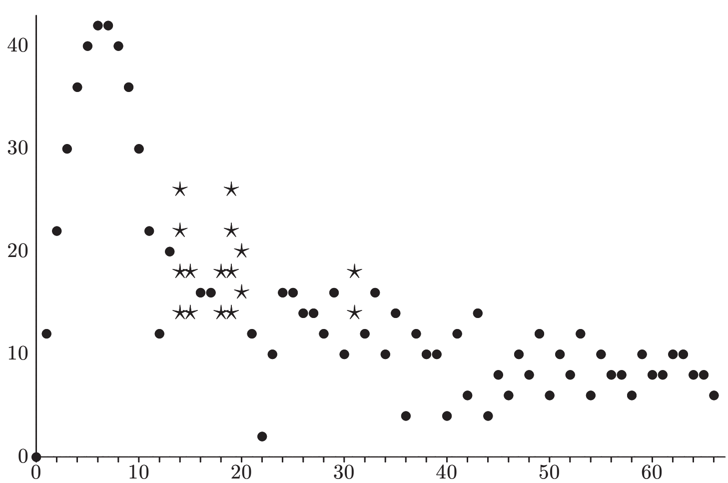

FIGURE A1. Graph of $f(r)$ for $q = 11$. Dots represent known values, and stars represent possible values for the values of r for which $f(r)$ is unknown.