Strong blocking sets and minimal codes from expander graphs

Noga Alon*[1], Anurag Bishnoi[2], Shagnik Das†[3] and Alessandro Neri‡[4]

[1]Department of Mathematics, Princeton University, United States of America
[2]Delft Institute of Applied Mathematics, Delft University of Technology, Netherlands
[3]Department of Mathematics, National Taiwan University, Taiwan
[4]Department of Mathematics: Analysis, Logic and Discrete Mathematics, Ghent University, Belgium

May 24, 2023

Abstract

A strong blocking set in a finite projective space is a set of points that intersects each hyperplane in a spanning set. We provide a new graph theoretic construction of such sets: combining constant-degree expanders with asymptotically good codes, we explicitly construct strong blocking sets in the $(k-1)$-dimensional projective space over $\mathbb{F}_q$ that have size $O(qk)$. Since strong blocking sets have recently been shown to be equivalent to minimal linear codes, our construction gives the first explicit construction of $\mathbb{F}_q$-linear minimal codes of length n and dimension k, for every prime power q, for which $n = O(qk)$. This solves one of the main open problems on minimal codes.

1 Introduction

A blocking set in a finite projective or affine space is a set of points that intersects every hyperplane. The study of these objects is a classic topic in finite geometry [15, 17], with many applications in coding theory, combinatorics and computer science. One can strengthen this notion to that of a *strong blocking set* by requiring that the intersection with every hyperplane is not just nonempty, but forms a spanning set for that hyperplane. For example, in a projective plane, the set of all points on a single line is a blocking set, while the set of all points on three non-concurrent lines is a strong blocking set. This special kind of blocking set has also appeared in the literature under the names of generating sets [27, 31] and cutting blocking sets [1, 12, 16], but in this paper we follow the nomenclature of [24, 30].

Strong blocking sets have recently been shown to be in one-to-one correspondence with minimal codes [1, 47], a notion from coding theory. A linear code is simply a vector subspace of $\mathbb{F}_q^n$. A codeword in a linear code is called minimal if its support does not contain the support of any other codeword apart from its scalar multiples. Minimal codewords in a linear code have been studied for their applications in decoding algorithms [33] and cryptography [18, 40]. Determining the set of minimal codewords in a linear code is a difficult task that has only been achieved for a few families of linear codes, and this has led to the study of *minimal codes*, where *every* non-zero codeword is minimal (see, for example, [18]). Recently, minimal codes have also been linked to perfect hash families, and in particular the trifference problem [14], which have important applications in computer science (see for example [53] and the references therein).

*NA is partially supported by NSF grant DMS-2154082 and BSF grant 2018267.
†SD is supported by Taiwan NSTC grant 111-2115-M-002-009-MY2.
‡AN is supported by the Research Foundation - Flanders (FWO) grant 12ZZB23N.

The main problem is to construct minimal codes of a given dimension k of the shortest possible length n. It is known that any strong blocking set in the $(k-1)$-dimensional projective space obtained from $\mathbb{F}_q^k$, denoted by $\mathrm{PG}(k-1,q)$, must have size at least $(q+1)(k-1)$ [3]. Using the aforementioned connection, this implies that any minimal code of length n and dimension k over $\mathbb{F}_q$ must satisfy $n \geq (q+1)(k-1)$. Therefore, we would like to construct minimal codes of length close to this lower bound. Providing additional motivation for this problem is the fact that minimal codes whose length n is at most linear in k (for a fixed q) give rise to asymptotically good error-correcting codes [1]. While it is easy to show the existence of such short minimal codes using the probabilistic method, it is a challenging and central open problem to give good explicit constructions [21]. Many constructions of minimal codes have appeared in the last few years [1, 10, 21, 25, 27], but their lengths remain considerably larger than the theoretical lower bound.

Over the binary field, minimal codes are equivalent to linear intersecting codes [19, 22], which are codes with the property that the supports of any two non-zero codewords have non-empty intersection, but over larger fields it is a more restrictive notion than intersecting codes [21]. By this equivalence, we already have an explicit construction of minimal codes for $q = 2$ with n a linear function of k [22, Theorem 2.3]. Bartoli and Borello [11, Corollary 3.3] recently gave an explicit construction of strong blocking sets with size linear in the dimension, for any fixed $q \geq 3$, but the dependency on q in their construction is not linear: they proved that for every prime power q, there exists an infinite sequence of dimensions k for which they give an explicit construction of a strong blocking set in the projective space $\mathrm{PG}(k-1,q)$ of length roughly $q^4k/4$. The same construction appears in an earlier work of Cohen, Mesnager and Randriam [20], and the main idea is to concatenate algebraic geometric codes with the simplex code. The argument used in [11] also has the limitation that it can at best give an explicit construction of size approximately q^2k.

In this paper, we provide a novel graph-theoretic construction combining linear codes with graphs to produce minimal codes. By using explicit constructions of asymptotically good linear codes and constant-degree expander graphs, we then obtain, for some absolute constant c, the first explicit construction of strong blocking sets of size cqk in the projective space $\mathrm{PG}(k-1,q)$, and thus also of a k-dimensional minimal code over $\mathbb{F}_q$ of length at most cqk. By optimising the constant, we can show that our construction improves the previous best explicit constructions for every fixed $q \geq 7$.

There is a rich history of using expander graphs in the construction of asymptotically good linear codes [5, 45, 48], and we extend this line of research by showing that they can also be used to construct minimal codes. Central to our construction is the notion of vertex integrity of a graph, which measures how many vertices need to be removed from a graph to break it into small components (see Section 3 for a precise definition and references), and we prove a new lower bound on this parameter for d-regular graphs in terms of their eigenvalues. This in particular implies that the vertex integrity of constant-degree n-vertex expanders is linear in n.

Finite geometry has often been used to give extremal, or near extremal, constructions of graphs with respect to some property (for example, in Turán and Ramsey problems). We show that the other direction can also be fruitful, as in our construction we use extremal graphs to pick a subset of lines whose union has desirable intersection properties with hyperplanes in a finite projective space. This novel construction has also been used to give explicit constructions of certain affine blocking sets [14], and we expect that it will lead to many new results in finite geometry.

1.1 Outline

In Section 2, we give the necessary background on codes, blocking sets and expander graphs. We introduce the integrity of a graph in Section 3, and prove our lower bound for regular graphs. In Section 4, we describe our new explicit construction, proving the main result of this paper. In

Section 5, we optimize the size of our construction by using algebraic-geometric codes, almost Ramanujan graphs, and field reduction. Finally, in Section 6, we summarize our results and highlight some possible directions for further research.

2 Preliminaries

In this section we recall some basic notions and preliminary results from coding theory, with a focus on minimal linear codes and on how they can be viewed geometrically. We also recall the notion of expander graphs and some explicit constructions. For the rest of this paper, we shall assume that q is a prime power.

2.1 Error-correcting codes and minimal codes

Let us fix $\mathbb{F}_q$ to be the finite field with q elements and let $n \in \mathbb{N}$.

Definition 2.1. The **(Hamming) support** of a vector $v \in \mathbb{F}_q^n$ is the set

$$\sigma(v) := \{i \,:\, v_i \neq 0\} \subseteq [n].$$

The **(Hamming) weight** of v is

$$\mathrm{wt}(v) := |\sigma(v)|.$$

The Hamming weight induces a distance on $\mathbb{F}_q^n$, given by $d(u, v) := \mathrm{wt}(u - v)$. This is known as the **Hamming distance** and it is fundamental in the theory of error-correcting codes.

Definition 2.2. An $[n, k, d]_q$ code $\mathcal{C}$ is a k-dimensional subspace of $\mathbb{F}_q^n$, and

$$d := \min\{\mathrm{wt}(v) \,:\, v \in \mathcal{C} \setminus \{0\}\}$$

is called the **minimum distance** of $\mathcal{C}$. The elements of $\mathcal{C}$ are called **codewords**. Moreover, a **generator matrix** for $\mathcal{C}$ is a matrix $G \in \mathbb{F}_q^{k\times n}$ such that

$$\mathcal{C} = \{uG \,:\, u \in \mathbb{F}_q^k\};$$

that is, the rows of G span $\mathcal{C}$.

Definition 2.3. Let $\{n_i\}_{i\geq 1}$ be an increasing sequence of lengths and suppose that there exist sequences $\{k_i\}_{i\geq 1}$ and $\{d_i\}_{i\geq 1}$ such that for all $i \geq 1$ there is an $[n_i, k_i, d_i]_q$ code $\mathcal{C}_i$. Then the sequence $\{\mathcal{C}_i\}_{i\geq 1}$ is called an $(R, \delta)_q$**-family of codes**, where the **rate** of this family is defined as

$$R := \liminf_{i\to\infty} \frac{k_i}{n_i},$$

and the **relative distance** is defined as

$$\delta := \liminf_{i\to\infty} \frac{d_i}{n_i}.$$

One of the central problems on error-correcting codes is to understand the trade-off between the rate and the relative distance of codes. A family of codes for which $R > 0$ and $\delta > 0$, is known as an *asymptotically good code.* An easy probablistic argument shows the existence of such codes for every $\delta \in [0, 1 - 1/q)$ and $R = 1 - H_q(\delta)$, where

$$H_q(x) := x \log_q(q - 1) - x \log_q(x) - (1 - x) \log_q(1 - x),$$

is the q-ary entropy function, defined on the domain $0 \leq x \leq 1 - 1/q$. This is known as the Gilbert-Varshamov bound. The first explicit construction of asymptotically good codes was given by Justesen [35], who showed that for every $0 < R < 1/2$, there is an explicit family of codes with rate R and relative distance $\delta \geq (1-2R)H_q^{-1}\left(\frac{1}{2}\right)$. Note that for any prime power q, $H_q^{-1}\left(\frac{1}{2}\right) \geq H_2^{-1}\left(\frac{1}{2}\right) > 0.11$, and thus there are absolute constants $R, \delta > 0$, not depending on q, for which we have an explicit construction of a family of codes with rate R and relative distance δ. Improving the values of the rate R and the relative distance δ for which there is an explicit construction, and reducing the computational complexity of these constructions, has been an active area of research in coding theory since the 1970s (see for example [5,46,50]). One of the most significant developments in the area was the use of modular curves to show that, for $q \geq 49$, there are explicit constructions of linear codes over $\mathbb{F}_q$ that are even better than the probabilistic ones (see [23,49] for some recent surveys on these constructions).

In this paper, we study a special class of codes called *minimal (linear) codes*. These are codes with interesting features from a combinatorial point of view.

Definition 2.4. Let $\mathcal{C}$ be an $[n,k,d]_q$ code. A nonzero codeword $v \in \mathcal{C}$ is said to be **minimal** (in $\mathcal{C}$) if $\sigma(v)$ is minimal with respect to the inclusion in the set

$$\sigma(\mathcal{C}) \coloneqq \{\sigma(c) \,:\, u \in \mathcal{C} \setminus \{0\}\}.$$

The code $\mathcal{C}$ is a **minimal linear code** if all its nonzero codewords are minimal.

Minimal codewords were first studied by Hwang for decoding purposes [33]. Later, they were analyzed by Massey in connection with secret sharing schemes [40]. Since then, minimal codewords and minimal codes attracted renewed interest within the coding theory community (see for example [1,6,10,21]). These concepts were further studied from a combinatorial point of view, since they correspond to circuits in the matroid associated to the dual code [26]. Recently, minimal codes have also been linked to linear trifferent codes [14], which are a special case of perfect hash families [53].

2.2 Projective systems and strong blocking sets

In this section we briefly describe the geometric dual approach to coding theory, where linear codes can be identified with set of points in a suitable projective space. For $k > 1$, the finite projective space of dimension $k-1$ over the finite field $\mathbb{F}_q$ is defined as

$$\mathrm{PG}(k-1,q) \coloneqq \left(\mathbb{F}_q^k \setminus \{0\}\right) / \sim,$$

where $u \sim v$ if and only if $u = \lambda v$ for some non-zero $\lambda \in \mathbb{F}_q$ (in some circles the same object will be denoted by $\mathbb{P}^{k-1}(\mathbb{F}_q)$). The equivalence class that a non-zero vector v belongs to is denoted by $[v]$. The 1-dimensional, 2-dimensional, ..., $(k-1)$-dimensional vector subspaces of $\mathbb{F}_q^k$ correspond to the points, lines, ..., hyperplanes of $\mathrm{PG}(k-1,q)$. We denote the span of a subset S of points in a projective space by $\langle S \rangle$ and the dimension $\dim(\langle S \rangle)$ is one less than the vector space dimension of the corresponding vector subspace. For example, the span of two distinct points P, Q in a projective space, which we will also denote by $\langle P, Q \rangle$, is a 1-dimensional projective subspace that we refer to as the line joining P and Q.

Definition 2.5. A **projective** $[n,k,d]_q$ **system** is a (multi)set of n points, $\mathcal{M} \subseteq \mathrm{PG}(k-1,q)$, such that $\langle \mathcal{M} \rangle = \mathrm{PG}(k-1,q)$ and

$$d = n - \max\{|H \cap \mathcal{M}| : H \text{ is a hyperplane}\}.$$

The term projective system and the notation used come from the correspondence with linear codes. Indeed, a projective $[n,k,d]_q$ system is simply a dual interpretation of a nondegenerate $[n,k,d]_q$ code. More precisely, an $[n,k,d]_q$ code $\mathcal{C}$ is **nondegenerate** if there is no identically zero coordinate in $\mathcal{C}$. In other words, $\mathcal{C}$ is not contained in any principal hyperplane $H_i := \{v \in \mathbb{F}_q^n : v_i = 0\}$.

The aforementioned correspondence – up to equivalence – comes from putting representatives of the points of the projective system as columns of a $k \times n$ matrix, and considering the code generated by (the rows of) this matrix. Vice versa, given a generator matrix of a nondegenerate code, we can obtain the associated projective system by taking the columns of such a matrix as a multiset of points in $\mathrm{PG}(k-1,q)$. There is more work required to make this correspondence well-defined, and we refer the reader to [49, Theorem 1.1.6] for a formal treatment of the correspondence. Due to this correspondence, a sequence $\{\mathcal{M}_i\}_{i\in\mathbb{N}}$ of projective systems is called an $(R,\delta)_q$**-family of projective systems** if the corresponding family of codes is an $(R,\delta)_q$-family of codes.

We now define the main finite geometric object studied in this paper.

Definition 2.6. A set $\mathcal{M} \subseteq \mathrm{PG}(k-1,q)$ is said to be a **strong blocking set** if

$$\langle H \cap \mathcal{M} \rangle = H,$$

for every hyperplane H of $\mathrm{PG}(k-1,q)$.

Remark 2.7. In the vector space notation, a strong blocking set in $\mathbb{F}_q^k$ is a collection of 1-dimensional vector subspaces that intersects every $(k-1)$-dimensional vector subspace in a spanning set.

Theorem 2.8 (see [1], [47])**.** Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_q$ code and let $G = (g_1 \mid \ldots \mid g_n) \in \mathbb{F}_q^{k\times n}$ be any of its generator matrices. The following are equivalent:

1. $\mathcal{C}$ is a minimal code;
2. $\mathcal{M} = \{[g_1],\ldots,[g_n]\}$ is a strong blocking set in $\mathrm{PG}(k-1,q)$.

The main and most relevant problem – from both a coding theoretic and geometric point of view – is the construction of small strong blocking sets, or, equivalently, of short minimal codes. The first step is to ask how small a strong blocking set can be. Answers to this question are partial and given by the following results. The first one is a general lower bound observed in [1], proved using the Combinatorial Nullstellensatz (see [14, 30] for alternative proofs using the results from [17, 34]).

Theorem 2.9. For any prime power q, every strong blocking set in $\mathrm{PG}(k-1,q)$ has size at least $(q+1)(k-1)$.

Recently this lower bound has been improved by using Delsarte's linear programming bound in coding theory, which is also known as the MRRW bound.

Theorem 2.10 (see [14, Theorem 1.4], [44, Theorem 3.3])**.** For any prime power q, there is a constant $c_q > 1$ such that every strong blocking set in $\mathrm{PG}(k-1,q)$ has size at least $(c_q - o(1))(q+1)(k-1)$.

We also have the following existence result shown using the probabilistic method that provides the best-known upper bounds.

Theorem 2.11 (see [41] for $q = 2$ and [2, 14] for $q > 2$)**.** The size of the smallest strong blocking set in $\mathrm{PG}(k-1,q)$ is at most

$$\begin{cases} \frac{2k-1}{\log_2(4/3)} & \text{if } q = 2, \\ (q+1)\frac{2k}{\log_q\left(\frac{q^4}{q^3-q+1}\right)} & \text{otherwise.} \end{cases}$$

This is an existence result that does not provide any explicit constructions. We now recall some of the most relevant general explicit constructions of small strong blocking sets that are known in the literature.

Rational normal tangents: Assume that $q \geq 2k-3$ and that $\mathrm{char}(\mathbb{F}_q) > k$. Fancsali and Sziklai [27] showed that under these hypothesis, one can take any distinct $2k-3$ points on a rational normal curve, and then take the union of the tangent lines to this curve at those points. The resulting set is a strong blocking set of size $(2k-3)(q+1)$. In the same paper, they also showed how to get rid of the hypothesis on the characteristic of the field, by using what they call the *diverted tangents* method. However, the hypothesis on the field size must be kept, implying that such a construction provides only finitely many strong blocking sets for a given field size.

Tetrahedron: This construction is probably the most natural one. It is obtained by selecting any k points in $\mathrm{PG}(k-1,q)$ in general position, and then taking the union of the lines spanned by every pair of these points. It works over every field, but its size $\binom{k}{2}(q-1)+k$ is quadratic in k, while we know by Theorem 2.11 about the existence of strong blocking sets whose size is linear in k. The tetrahedron was first observed by Davydov, Giulietti, Marcugini and Pambianco [24] and then rediscovered by several authors.

Line subspreads: This is a slight improvement on the size of the tetrahedron. It works whenever $k=2t$ is even, and it consists of carefully choosing t^2 points in $\mathrm{PG}(t-1,q^2)$, and then using the field reduction map to obtain t^2 lines in $\mathrm{PG}(k-1,q)$ whose union is a strong blocking set. This construction has size $\frac{k^2}{4}(q+1)$ and was recently pointed out in [3].

All these constructions are obtained as unions of lines in the projective space. This is mainly due to the fact that with such a structure it is easy to control their intersections with subspaces. In particular, the main feature that these constructions possess is the following property, which is stronger than being a strong blocking set.

Definition 2.12. A set $\mathcal{L}$ of lines in a projective space satisfies the **avoidance property** if there is no codimension-2 space meeting every line $\ell \in \mathcal{L}$.

The relation between these sets of lines and strong blocking sets is the following observation of Fancsali and Sziklai [27, Theorem 11], whose proof we include for the sake of convenience.

Theorem 2.13. If a set $\mathcal{L}$ of lines satisfies the avoidance property, then the point-set $\mathcal{B} = \cup_{\ell \in \mathcal{L}} \ell$ is a strong blocking set.

Proof. Let $\mathcal{L}$ be a set of lines and let $\mathcal{B} = \cup_{\ell \in \mathcal{L}} \ell$. Assume that $\mathcal{B}$ is not a strong blocking set. Then there exists a hyperplane Π such that $\mathcal{B} \cap \Pi$ does not span Π. In particular, $\mathcal{B} \cap \Pi$ is contained in a hyperplane H of Π.

Since Π is a hyperplane, it meets every line of the projective space. Thus, $\ell \cap \Pi \neq \emptyset$ for all $\ell \in \mathcal{L}$, but since $\ell \subseteq \mathcal{B}$ and $\mathcal{B} \cap \Pi \subseteq H$, it follows that $\ell \cap H \neq \emptyset$. That is, H is a codimension-2 subspace meeting every line of $\mathcal{L}$, and so $\mathcal{L}$ does not satisfy the avoidance property. □

Remark 2.14. As shown in [27, Lemma 13], any collection of lines that satisfy the avoidance property must have size at least $k-1+\lfloor (k-1)/2 \rfloor$, thus giving a lower bound of roughly $1.5(q+1)(k-1)$ on the smallest possible size of a strong blocking set that can be constructed using such a set of lines.

2.3 Expander graphs

In our construction, we will make use of explicit constructions of constant-degree expander graphs. Informally, the edges of expander graphs are very well spread out, ensuring that there are many outgoing edges from all vertex subsets that are not too large. We refer the reader to the survey [32] for a formal definition and for various applications of expanders.

Expansion in graphs can be measured by their spectral properties. Given an n-vertex graph G, we denote the eigenvalues of its adjacency matrix by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. These eigenvalues encode a lot of information about the graph; for instance, if G is connected and d-regular, then $\lambda_1 = d$ and $\lambda_2 < d$. A graph G is called an (n, d, λ)-graph if it is a d-regular graph on n vertices with $|\lambda_i| \leq \lambda$ for all $i \geq 2$. The following lemma is one of the central tools for studying such graphs.

Lemma 2.15 (Expander-Mixing Lemma)**.** Let G be an (n, d, λ)-graph and S, T be two vertex-subsets of G. Denote by $e(S, T)$ the number of pairs $(x, y) \in S \times T$ such that xy is an edge of G. Then

$$\left| e(S,T) - \frac{d|S||T|}{n} \right| \leq \lambda \sqrt{|S||T| \left(1 - \frac{|S|}{n}\right)\left(1 - \frac{|T|}{n}\right)}.$$

A proof of this lemma can be found in [51, Lemma 4.15]. Note that the error term on the right-hand side is directly proportional to λ, and so it is natural to try to make this parameter as small as possible. The Alon–Bopanna bound [43] limits how far one can go, and motivates the definition of Ramanujan graphs, which are the ultimate expanders.

Theorem 2.16 (Alon-Bopanna)**.** Let G be an (n, d, λ)-graph. Then $\lambda \geq 2\sqrt{d-1} - o(1)$ as $n \to \infty$.

Definition 2.17. Let G be a d-regular graph with the eigenvalues $d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. If $\max\{|\lambda_i| : |\lambda_i| < d\} \leq 2\sqrt{d-1}$, then G is said to be a **Ramanujan graph**.

Lubotsky, Phillips and Sarnak [38] and Margulis [39] gave explicit constructions of d-regular Ramanujan graphs for $d = p + 1$, where p is prime. We denote by H_d the d-regular Ramanujan graph constructed by Lubotsky, Phillips and Sarnak. For the convenience of the reader, we briefly describe this construction.

Fix a prime $p \equiv 1 \pmod 4$. By Jacobi's four square theorem, there exist exactly $p+1$ integer solutions to the equation

$$p = b_1^2 + b_2^2 + b_3^2 + b_4^2, \qquad b_1 > 0, b_2, b_3, b_4 \equiv 0 \pmod 2 \tag{1}$$

Now, let $r \equiv 1 \pmod 4$ be a distinct prime. To each solution of (1) we associate the matrix

$$\begin{pmatrix} b_1 + ib_2 & b_3 + ib_4 \\ -b_3 + ib_4 & b_1 - ib_2 \end{pmatrix} \in \mathbb{F}_r^{2\times 2}, \tag{2}$$

where i is a square root of -1 in $\mathbb{F}_r$. If p is a quadratic residue modulo r, we define H_{p+1} to be the Cayley graph of $\mathrm{PGL}(2, \mathbb{F}_r)$ with the $p + 1$ generators given in (2), which has $r(r^2 - 1)$ vertices. If p is not a quadratic residue modulo r, then we define H_{p+1} to be the Cayley graph of $\mathrm{PSL}(2, \mathbb{F}_r)$ with the $p + 1$ generators given in (2), which has $\frac{r(r^2-1)}{2}$ vertices.

It was shown in [38] that the graphs H_{p+1} are Ramanujan graphs, and hence $\lambda_2(H_{p+1}) \leq 2\sqrt{p}$. A few years later, this construction was adapted in [42] to produce $(q+1)$-regular graphs for prime powers q. However, both of these constructions have the disadvantage that they only produce (n, d, λ)-graphs with very restricted choices of n and d. If we slightly relax the requirement that $\lambda \leq 2\sqrt{d-1}$, we can find explicit constructions for every degree d and every large enough n.

Theorem 2.18 (see [4, Theorem 1.3])**.** For every positive integer d, and every $\varepsilon > 0$, there is an $n_0(d, \varepsilon)$ such that, for all $n \geq n_0(d, \varepsilon)$ with nd even, there is an explicit construction of an (n, d, λ)-graph $G^{\varepsilon}_{n,d}$ with $\lambda \leq 2\sqrt{d-1} + \varepsilon$

3 Integrity of a graph

Crucial to our work is the following graph parameter, known as the (vertex) *integrity* of a graph, which was originally introduced in late 1980s as a measure of the robustness of a network under vertex deletion [7, 9].

Definition 3.1. Let $G = (V, E)$ be a simple connected graph. For any subgraph H, let $\kappa(G)$ denote the largest size of a connected component in H. The **integrity** of G is the integer

$$\iota(G) := \min_{S\subseteq V} \left(|S| + \kappa(G - S)\right).$$

It is a challenging problem to determine the integrity of graphs precisely, or even asymptotically (see [7] for an old survey and [8, 13] for some recent bounds on different families of graphs). We prove a new lower bound on the vertex integrity of (n, d, λ)-graphs. First, we introduce another graph parameter and show that it is closely related to the integrity of a graph.

Definition 3.2. For a graph G, let $z(G)$ denote the largest integer z such that there are two disjoint sets of vertices in G, each of size z, with no edge between them.

Proposition 3.3. For every graph $G = (V, E)$ on n vertices,

$$n - 2z(G) \leq \iota(G) \leq n - z(G).$$

Proof. For the upper bound, let A, B be two disjoint sets of size z with no edges between them. Put $S = V - (A \cup B)$. Then any connected component in $G - S$ is either contained in A or in B, and thus has size at most z. Therefore $\iota(G) \leq |S| + z = (n - 2z) + z = n - z$.

We now prove the lower bound. Let $z = z(G)$ and let S be a subset of size σ such that the maximum size of a connected component in $G - S$ is κ, with $\sigma + \kappa = \iota(G)$. Let $C_1, \dots, C_t$ be the connected components in $G - S$ of sizes $\kappa = c_1 \geq \cdots \geq c_t$. Note that $n - \iota(G) = \sum_{i=2}^{t} c_i$, and thus it suffices to upper bound this sum by $2z$. Also note that there are no edges between C_i and C_j for any $i \neq j$. If $c_1 \geq z + 1$, then by the maximality of z the size of $C_2 \cup \cdots \cup C_t$ is at most z, and we are done. Therefore we have $c_1 \leq z$ and, for the sake of contradiction, we assume that $\sum_{i=2}^{t} c_i \geq 2z + 1$. Let $2 \leq s \leq t$ be the largest index s for which $c_s + \cdots + c_t \geq z + 1$. Since $c_s \leq c_1$, it follows that $c_s + \cdots + c_t \leq z + c_1$. Therefore, $c_2 + \cdots + c_{s-1} \geq 2z + 1 - (z + c_1) = z + 1 - c_1$. Let $X = C_1 \cup \cdots \cup C_{s-1}$ and $Y = C_s \cup \cdots \cup C_t$. Then both X and Y have size at least $z + 1$, which is a contradiction since they do not have any edges between them. □

Corollary 3.4. For any (n, d, λ)-graph G, we have $\iota(G) \geq \left(\frac{d-\lambda}{d+\lambda}\right) n$.

Proof. Let $z(G)$ be as in Definition 3.2. For any two sets S, T of vertices with $e(S, T) = 0$ and $|S| = |T| = z(G)$, Lemma 2.15 implies that

$$z(G) \leq \frac{\lambda n}{d + \lambda}.$$

Applying the lower bound $\iota(G) \geq n - 2z(G)$ from Proposition 3.3 gives $\iota(G) \geq n - 2\frac{\lambda}{d+\lambda}n = \frac{d-\lambda}{d+\lambda}n$. □

Remark 3.5. A lower bound on the integrity of cubic graphs was proved in [52, Theorem 8]. The argument there, along with Cheeger's inequality [32, Theorem 2.4], can be used to prove the weaker bound of $\iota(G) \geq n \min\{1/2, (d - \lambda)/(3d - \lambda)\}$.

Remark 3.6. When applied to d-regular Ramanujan graphs, Corollary 3.4 yields a lower bound of $\iota(G) = \left(1 - O\left(d^{-1/2}\right)\right) n$. In Appendix A, we show that the largest possible integrity of n-vertex graphs with average degree at most d is in fact of the form $\iota(G) = \left(1 - \Theta\left(d^{-1} \log d\right)\right) n$.

4 Constructing Strong Blocking Sets from Graphs

In this section, we will provide a new general construction inspired by the tetrahedron (see Section 2). We will use the data from a projective $[n, k, d]_q$ system and a graph on n vertices in order to construct a set of lines with the avoidance property, whose union, in light of Theorem 2.13, forms a strong blocking set.

Definition 4.1. Let $\mathcal{M} = \{P_1, \dots, P_n\}$ be a set of n points in $\mathrm{PG}(k-1, q)$ and let $G = (\mathcal{M}, E)$ be a graph with vertex set equal to $\mathcal{M}$. We define the set of lines

$$\mathcal{L}(\mathcal{M}, G) \coloneqq \{\langle P_i, P_j \rangle : P_i P_j \in E\}$$

and the set of points

$$\mathcal{B}(\mathcal{M}, G) \coloneqq \bigcup_{\ell \in \mathcal{L}(\mathcal{M}, G)} \ell$$

Remark 4.2. The size of $\mathcal{B}(\mathcal{M}, G)$ is at most $n + (q-1)|E|$, since there are $|E|$ lines, each of which contains at most $q-1$ points not in $\mathcal{M}$.

The following result lies at the heart of our construction as it gives a sufficient condition for the line-set $\mathcal{L}(\mathcal{M}, G)$ to satisfy the avoidance property.

Proposition 4.3. Let $\mathcal{M} = \{P_1, \dots, P_n\}$ be a set of points in $\mathrm{PG}(k-1, q)$ and let $G = (\mathcal{M}, E)$ be a graph whose set of vertices is $\mathcal{M}$. If for every $S \subseteq \mathcal{M}$ there exists a connected component C in $G - S$ such that

$$\langle S \cup C \rangle = \mathrm{PG}(k-1, q),$$

then the set $\mathcal{L}(\mathcal{M}, G) = \{\langle P_i, P_j \rangle : P_i P_j \in E\}$ satisfies the avoidance property; that is, no codimension-2 subspace of $\mathrm{PG}(k-1, q)$ meets every line of $\mathcal{L}(\mathcal{M}, G)$.

Proof. Say G satisfies the property and, for the sake of contradiction, let H be a codimension-2 subspace that meets every line in $\mathcal{L} = \mathcal{L}(\mathcal{M}, G)$. Let $S = H \cap \mathcal{M}$ and let C be a connected component of $G - S$ such that S and C together span the whole space.

For every edge $e = P_i P_j$ whose endpoints P_i, P_j lie in $\mathcal{M} \setminus S$, there is a corresponding line $\ell_{ij} = \langle P_i, P_j \rangle \in \mathcal{L}$, which by our assumption intersects H. Since $P_i, P_j \notin H$, there must be a unique point $Q_{ij} \in \ell_{ij} \cap H$. Thus, writing $\mathcal{Q}$ for the set $\{Q_{ij} : P_i P_j \in E, P_i, P_j \notin S\}$, we have $S \cup \mathcal{Q} \subseteq H$.

Now observe that for an edge $P_i P_j \in E$ with endpoints $P_i, P_j \notin S$, if a subspace contains both P_i and Q_{ij}, then it must also contain P_j, which lies on the line spanned by P_i and Q_{ij}. Fixing some point P_r in the component $C \subseteq \mathcal{M} \setminus S$, since every point in C is connected by a path to P_r, the previous observation implies that any subspace containing $\mathcal{Q} \cup \{P_r\}$ must contain all of C. Hence,

$$\langle H \cup \{P_r\} \rangle \supseteq \langle S \cup \mathcal{Q} \cup \{P_r\} \rangle \supseteq \langle S \cup C \rangle = \mathrm{PG}(k-1, q).$$

This is a contradiction, as H is a codimension-2 subspace, and thus $\langle H \cup \{P_r\} \rangle$ has codimension at least 1. □

Proposition 4.3 provides a general method of constructing strong blocking sets by combining a graph G with a set $\mathcal{M}$ of points in a projective space. However, the construction requires nontrivial interplay between G and $\mathcal{M}$ and their local properties, and it seems quite difficult to design them simultaneously. For this reason, we will simplify the approach by assuming the worst-case global parameters.

Lemma 4.4. Let $\mathcal{M}$ be a projective $[n, k, d]_q$ system and let $G = (\mathcal{M}, E)$ be a graph of integrity $\iota(G) \geq n - d + 1$. Then $\mathcal{L}(\mathcal{M}, G)$ satisfies the avoidance property, and thus $\mathcal{B}(\mathcal{M}, G)$ is a strong blocking set in $\mathrm{PG}(k-1, q)$ of size at most $n + (q-1)|E|$.

Proof. Let S be an arbitrary subset of $\mathcal{M}$. Since $\iota(G) \geq n - d + 1$, there exists a connected component C in G such that $|S| + |C| \geq n - d + 1$. From the definition of projective systems (see Section 2), it follows that every hyperplane meets $\mathcal{M}$ in at most $n - d$ points. Therefore, $S \cup C \subseteq \mathcal{M}$ is not contained in any hyperplane of $\mathrm{PG}(k-1, q)$, thus implying $\langle S \cup C \rangle = \mathrm{PG}(k-1, q)$. From Proposition 4.3, we conclude that $\mathcal{L}(\mathcal{M}, G)$ satisfies the avoidance property and thus, by Theorem 2.13, $\mathcal{B}(\mathcal{M}, G)$ is a strong blocking set. As per Remark 4.2, $|\mathcal{B}(\mathcal{M}, G)| \leq n + (q-1)|E|$. □

We now prove the main result of our paper by giving an explicit construction of strong blocking sets in $\mathrm{PG}(k-1, q)$ with size linear in qk.

Theorem 4.5. There is an absolute constant c such that for every prime power q, there exists an explicit construction of strong blocking sets of size at most cqk_i in $\mathrm{PG}(k_i - 1, q)$, for some infinite increasing sequence $\{k_i\}_{i \in \mathbb{N}}$.

Proof. Let R be any constant satisfying $0 < R < 1/2$ and let $\delta = 0.11(1-2R)$. Let $\mathcal{M}_i$ be the projective $[n_i, k_i, d_i]_q$ systems given by the Justesen construction [35], which exist for an infinite increasing sequence $\{k_i\}_{i \in \mathbb{N}}$. Then $\lim_{i \to \infty} k_i/n_i = R$ and $\lim_{i \to \infty} d_i/n_i \geq (1-2R)H_q^{-1}(1/2) > \delta$. Therefore, there exists an i_0, which we may assume to be sufficiently large for all subsequent calculations, such that for all $i \geq i_0$, we have $d_i/n_i \geq \delta$ and $k_i/n_i \geq R/2$. Let $\{G_i\}_{i \geq i_0}$ be an explicit family of (n_i, d, λ)-graphs, where d and λ are positive constants for which $(d-\lambda)/(d+\lambda) \geq 1 - \delta + 1/n_i$. From Theorem 2.18, it follows that such an explicit construction of graphs is always possible. By Corollary 3.4, we have $\iota(G_i) \geq (1-\delta)n_i + 1 \geq n_i - d_i + 1$. Therefore, by Lemma 4.4, $\mathcal{B}(\mathcal{M}_i, G_i)$ is a strong blocking set in $\mathrm{PG}(k_i - 1, q)$ of size at most

$$n_i + (q-1)\frac{dn_i}{2} < \frac{d}{2}qn_i \leq \frac{d}{R}qk_i.$$

This concludes the proof with $c = \frac{d}{R}$. □

5 Strong blocking sets from expander graphs and AG codes

Using the construction of Theorem 4.5, the best constant c that we get is quite large; for the optimal choice of R, it is approximately $c \simeq 8276$. However, we can reduce it substantially by replacing the Justesen codes with some families of AG codes and – depending on the field – by using field reduction. In this section we optimize the value of the constant c in our construction for all values of q.

To this end, we use the asymptotically good Algebraic-Geometry (AG) codes, explicit constructions of which can be found in [28, 29, 50]. In particular, for every square prime power q and $R, \delta > 0$ satisfying $R + \delta \geq 1 - (\sqrt{q} - 1)^{-1}$, we can construct an $(R, \delta)_q$-family of $[n_i, k_i, d_i]_q$ codes for some increasing sequences $\{n_i\}_{i \in \mathbb{N}}, \{k_i\}_{i \in \mathbb{N}}$ and $\{d_i\}_{i \in \mathbb{N}}$.

Definition 5.1. Given a square prime power q, for every $R \in (0, 1)$, set $\delta = 1 - R - (\sqrt{q} - 1)^{-1}$. Given the $(R, \delta)_q$-family of $[n_i, k_i, d_i]_q$ codes described above, we denote by $\{\mathcal{A}_{n_i, R}\}_{i \in \mathbb{N}}$ the associated $(R, 1 - R - (\sqrt{q} - 1)^{-1})_q$-family of projective $[n_i, k_i, d_i]_q$ systems.

With this notation in place, we can now proceed to describe our improved constructions.

5.1 Quadratic Fields

We start with a simple result, obtained by combining Lemma 4.4 and Corollary 3.4 with the explicit construction of expander graphs given by Theorem 2.18.

Theorem 5.2. Let $d \geq 3$, let $q > 4$ be a square prime power such that $(\sqrt{q}-1)^{-1} < \frac{d-2\sqrt{d-1}}{d+2\sqrt{d-1}}$, and let $\varepsilon > 0$. Then there is an increasing sequence $\{k_i\}_{i\in\mathbb{N}}$ for which we can explicitly construct strong blocking sets in $\mathrm{PG}(k_i-1,q)$ of size at most

$$\left(\frac{d(d+2\sqrt{d-1})(\sqrt{q}-1)}{2\left(d(\sqrt{q}-2)-2\sqrt{q(d-1)}\right)} + \varepsilon \right) k_i q.$$

Before we proceed with the proof, let us explore what this result implies about that constant in the bound on the size of strong blocking sets in $\mathrm{PG}(k-1,q)$ when q is a square. For each such q, we can choose an optimal value for d to minimize the bound. This amounts to finding the minimum values of the function

$$F_q(d) = \frac{d(d+2\sqrt{d-1})(\sqrt{q}-1)}{2\left(d(\sqrt{q}-2)-2\sqrt{q(d-1)}\right)},$$

where we can extend the domain to $\mathbb{R}_{>2}$. To simplify the calculations, we can make the substitution $y = \sqrt{d-1}$, and then find the local extrema by setting the derivative equal to zero. This amounts to finding the zeros of the polynomial

$$\psi_q(y) = \sqrt{q}(y-1)(y^3-2y^2-y-2) - 2(y^2+1)^2.$$

As q grows, the roots of this polynomial converge to those of $(y-1)(y^3-2y^2-y-2)$, and the unique root in our domain of interest $(y > 1)$ is

$$y_0 = \frac{1}{3}\left(2 + (44-3\sqrt{177})^{\frac{1}{3}} + (44+3\sqrt{177})^{\frac{1}{3}}\right).$$

Hence, for large values of q, $F_q(d)$ will be minimized for

$$d \approx d_0 = 1 + y_0^2 = 3 + \frac{1}{3}(459-12\sqrt{177})^{\frac{1}{3}} + \frac{1}{3}(459+12\sqrt{177})^{\frac{1}{3}} \approx 8.0701,$$

and so one should take $d = 8$ or 9. It is straightforward to verify that, for large enough q, we have $F_q(8) < F_q(9)$, and hence, as q tends to infinity, the optimal constant this construction provides is

$$\lim_{q\to\infty} F_q(8) = \frac{4}{9}(23+8\sqrt{7}) \approx 19.63,$$

a very significant saving compared to the construction from the previous section. For smaller values of q, we can compute the optimal choice of d and the corresponding constant, and these are given in Table 1.

Sufficiently motivated, we now prove the theorem.

Proof of Theorem 5.2. Let $\varepsilon_1 = \varepsilon_1(d,q) > 0$ be sufficiently small, and set $\lambda = 2\sqrt{d-1}+\varepsilon_1$. We have $(\sqrt{q}-1)^{-1} < \frac{d-\lambda}{d+\lambda} - 2\varepsilon_1$, and set $R = \frac{d-\lambda}{d+\lambda} - \varepsilon_1 - (\sqrt{q}-1)^{-1}$ and $\delta = 1 - R - (\sqrt{q}-1)^{-1} = 1 - \frac{d-\lambda}{d+\lambda} + \varepsilon_1$.

Let $\{\mathcal{A}_{n_i,R}\}_{i\in\mathbb{N}}$ be the $(R,\delta)_q$-family of projective $[n_i,k_i,d_i]_q$ systems from Definition 5.1. Theorem 2.18 shows that there is some M for which we obtain an explicit sequence $\{G_i\}_{i>M}$ of (n_i,d,λ)-graphs. Corollary 3.4 gives

$$\iota(G_{n_i}) \geq n_i \frac{d-\lambda}{d+\lambda} = n_i\left(1-\delta+\varepsilon_1\right).$$

Since $\lim_{i\to\infty} \frac{d_i}{n_i} = \delta$, we have $\iota(G_{n_i}) \geq n_i - d_i + 1$ for sufficiently large i. Thus, by Lemma 4.4, $\mathcal{B}(\mathcal{A}_{n_i,R}, G_{n_i})$ is a strong blocking set in $\mathrm{PG}(k_i-1,q)$. Since G_{n_i} has $\frac{1}{2}n_i d$ edges, we have

$$|\mathcal{B}(\mathcal{A}_{n_i,R}, G_{n_i})| \leq n_i + (q-1)\frac{n_i d}{2} < \frac{n_i d}{2} q.$$

q	argmin $F_q(d)$	upper bound$/k(q+1)$
9	$d=85$	292.68
16	$d=37$	104.60
25	$d=26$	66.86
49	$d=18$	43.91
64	$d=16$	39.07
81	$d=15$	35.83
121	$d=13$	31.76
169	$d=12$	29.31
$256 \le q \le 361$	$d=11$	27.06
$529 \le q \le 1024$	$d=10$	24.44
$1369 \le q \le 11881$	$d=9$	22.46
$q \ge 12769$	$d=8$	20.52

Table 1: For given ranges of square prime powers, this table provides the values of d that minimize the size of the strong blocking sets obtained by Theorem 5.2, and upper bounds on the corresponding sizes.

Now, since $\lim_{i\to\infty} \frac{k_i}{n_i} = R$, we have $n_i \le \frac{k_i}{R-\varepsilon_1}$ for sufficiently large i. Making this substitution, and recalling our choice of $R = \frac{d-\lambda}{d+\lambda} - \varepsilon_1 - (\sqrt{q}-1)^{-1}$, our upper bound becomes

$$\frac{n_i d}{2} q \le \frac{d}{2(R-\varepsilon_1)} k_1 q = \frac{d(d+\lambda)(\sqrt{q}-1)}{2\left((d-\lambda)(\sqrt{q}-1)-(d+\lambda)-2\varepsilon_1(d+\lambda)(\sqrt{q}-1)\right)} k_i q.$$

If we choose ε_1 to be sufficiently small, we obtain the upper bound

$$|\mathcal{B}(\mathcal{A}_{n_i,R}, G_{n_i})| \le \left(\frac{d(d+\lambda)(\sqrt{q}-1)}{2\left((d-\lambda)(\sqrt{q}-1)-(d+\lambda)\right)} + \frac{\varepsilon}{2} \right) k_i q.$$

Recalling that $\lambda = 2\sqrt{d-1} + \varepsilon_1$, we have

$$\frac{d(d+\lambda)(\sqrt{q}-1)}{2\left((d-\lambda)(\sqrt{q}-1)-(d+\lambda)\right)} = \frac{d(d+2\sqrt{d-1})(\sqrt{q}-1)+\varepsilon_1 d(\sqrt{q}-1)}{2\left(d(\sqrt{q}-2)-2\sqrt{q(d-1)}-\varepsilon_1\sqrt{q}\right)},$$

and the result follows provided ε_1 is suitably small. □

5.2 Non-Quadratic Fields

Theorem 5.2 shows that replacing the Justesen codes with AG codes in our construction can greatly reduce the size of the strong blocking sets we obtain. However, the one drawback is that the construction is only possible over quadratic fields. In this section we show how to use one final trick — field reduction — to take a strong blocking set over $\mathbb{F}_{q^2}$ and build from it a strong blocking set over $\mathbb{F}_q$ that is not much larger.

We first recall the **field reduction map**, which we denote by $\mathcal{F}_{q,r}$. This map uses the fact that points of $\mathrm{PG}(K-1,q^r)$ are 1-dimensional $\mathbb{F}_{q^r}$-subspaces of $\mathbb{F}_{q^r}^K$, which in turn can be viewed as r-dimensional $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^{rK}$. Hence, $\mathcal{F}_{q,r}$ sends points of $\mathrm{PG}(K-1,q^r)$ to $(r-1)$-spaces of $\mathrm{PG}(rK-1,q)$; see [37] for a survey on field reduction.

As shown in [3], the field reduction map also preserves some key properties related to strong blocking sets. We begin with a definition.

Definition 5.3. Let $\mathcal{L} = \{\ell_1, \ell_2, \ldots, \ell_t\}$ be a collection of lines in $\mathrm{PG}(K-1,q^2)$. We say a set of points $\Lambda \subseteq \mathrm{PG}(K-1,q^2)$ is **viable for** $\mathcal{L}$ if $\Lambda = \cup_{i=1}^t \Lambda^{(i)}$, where each $\Lambda^{(i)} =$

$\left\{\lambda_1^{(i)},\lambda_2^{(i)},\lambda_3^{(i)},\lambda_4^{(i)}\right\} \subseteq \ell_i$ is a set of four points that do not lie on a common $\mathbb{F}_q$-subline of ℓ_i.

Given a viable set Λ, we define the **derived set** to be the set

$$\mathcal{F}_{q,2}(\Lambda) = \left\{\mathcal{F}_{q,2}(\lambda_j^{(i)}) : i \in [t], j \in [4]\right\}$$

of lines in $\mathrm{PG}(2K-1,q)$.

The following result, obtained by combining [3, Theorem 4.2] and [3, Proposition 4.5], allows us to turn strong blocking sets in $\mathrm{PG}(k-1,q^2)$ into strong blocking sets in $\mathrm{PG}(2k-1,q)$. This was also highlighted in [2], where viable sets are shown to be **outer strong blocking sets**.

Theorem 5.4 (see [3])**.** Let $\mathcal{L} = \{\ell_1,\ldots,\ell_t\}$ be a set of lines in $\mathrm{PG}(K-1,q^2)$ whose union forms a strong blocking set. If Λ is a viable set for $\mathcal{L}$, then the union of the lines in the derived set $\mathcal{F}_{q,2}(\Lambda)$ is a strong blocking set in $\mathrm{PG}(2K-1,q)$.

This field reduction process is especially effective when used on our strong blocking sets $\mathcal{B}(\mathcal{M},G)$ constructed from graphs, as the points of $\mathcal{M}$ belong to several lines.

Lemma 5.5. Let $\mathcal{M}$ be an $[n,K,d]_{q^2}$ projective system and let $G=(\mathcal{M},E)$ be a graph. Then we can find a viable set Λ of size at most $n+2|E|$ for the associated set of lines $\mathcal{L}(\mathcal{M},G)$ in $\mathrm{PG}(K-1,q^2)$.

Proof. Let us enumerate the edges of G as $E=\{e_1,e_2,\ldots,e_m\}$. If $e_i = P_aP_b$, then the corresponding line $\ell_i \in \mathcal{L}(\mathcal{M},G)$ is given by $\ell_i = \langle P_a,P_b\rangle$. We then take $\lambda_1^{(i)} = P_a$ and $\lambda_2^{(i)} = P_b$, and let $\lambda_3^{(i)} = Q_3^{(a,b)}$ be an arbitrary third point on ℓ_i. Since any three points on an $\mathbb{F}_{q^2}$-line define a unique $\mathbb{F}_q$-subline, we can then choose a fourth point $\lambda_4^{(i)} = Q_4^{(a,b)}$ that avoids this subline.

Thus,

$$\Lambda = \bigcup_{i=1}^{m}\left\{\lambda_1^{(i)},\lambda_2^{(i)},\lambda_3^{(i)},\lambda_4^{(i)}\right\} = \mathcal{M} \cup \left(\bigcup_{P_aP_b\in E}\left\{Q_3^{(a,b)},Q_4^{(a,b)}\right\}\right)$$

is viable for $\mathcal{L}(\mathcal{M},G)$, and $|\Lambda| \le n+2|E|$. □

We can now apply Lemma 5.5 and Theorem 5.4 to the construction from Theorem 5.2 in order to build small strong blocking sets even when q is not a square.

Theorem 5.6. Let $d\ge 3$, let $q>2$ be such that $(q-1)^{-1} < \frac{d-2\sqrt{d-1}}{d+2\sqrt{d-1}}$, and let $\varepsilon>0$. Then there is an increasing sequence $\{k_i\}_{i\in\mathbb{N}}$ for which we can explicitly construct strong blocking sets in $\mathrm{PG}(k_i-1,q)$ of size at most

$$\left(\frac{(d+1)(d+2\sqrt{d-1})(q-1)}{2\left(d(q-2)-2q\sqrt{d-1}\right)}+\varepsilon\right)k_i(q+1).$$

Proof. This proof follows the same lines as that of Theorem 5.2, and so we will mainly highlight the changes. As before, we let $\varepsilon_1 = \varepsilon_1(d,q) > 0$ be sufficiently small and set $\lambda = 2\sqrt{d-1}+\varepsilon_1$. By assumption, $(q-1)^{-1} < \frac{d-\lambda}{d+\lambda} - 2\varepsilon_1$, and we set $R = \frac{d-\lambda}{d+\lambda} - \varepsilon_1 - (q-1)^{-1}$ and $\delta = 1-R-(q-1)^{-1} = 1-\frac{d-\lambda}{d+\lambda}+\varepsilon_1$.

We then take $\{\mathcal{A}_{n_i,R}\}_{i\in\mathbb{N}}$ to be an $(R,\delta)_{q^2}$-family of projective $[n_i,K_i,d_i]_{q^2}$ systems, and $\{G_i\}_{i>M}$ a sequence of (n_i,d,λ)-graphs. As before, our choice of parameters ensures that for sufficiently large i, we have $\iota(G_{n_i}) \ge n_i-d_i+1$. Applying Lemma 4.4, we deduce that the set of lines $\mathcal{L} = \mathcal{L}(\mathcal{A}_{n_i,R},G)$ has the avoidance property in $\mathrm{PG}(K-1,q^2)$, and hence, by Theorem 2.13, the union $\mathcal{B}(\mathcal{A}_{n_i,R},G)$ of those lines is a strong blocking set.

By Lemma 5.5, we can find a set Λ that is viable for $\mathcal{L}$ of size at most $|\mathcal{A}_{n_i,R}| + 2e(G_{n_i}) = n_i(d+1)$. Theorem 5.4 shows that the union of the lines in the derived set $\mathbb{F}_{q,2}(\Lambda)$ is then a strong blocking set in $\mathrm{PG}(2K_i - 1, q)$. Since each line in $\mathrm{PG}(2K_i - 1, q)$ has $q+1$ points, this strong blocking set has size at most $n_i(d+1)(q+1)$. Since $R = \lim_{i\to\infty} \frac{K_i}{n_i}$, recalling our choices for the parameters R and λ and setting $k_i = 2K_i$ then yields the claimed bound. □

As before, we can determine the optimal degree d to use by minimizing the quantity

$$R_q(d) := \frac{(d+1)(d+2\sqrt{d-1})(q-1)}{2\left(d(q-2) - 2q\sqrt{d-1}\right)}.$$

It is again advisable to make the substitution $d = 1 + y^2$, following which we find the zeros of the derivative coincide with those of $\phi_q(y) = q(y-1)(y^3 - 2y^2 - y - 4) - 2(y^2 - y + 1)(y^2 + y + 2)$. As q grows, the roots of ϕ_q converge to the roots of $(y-1)(y^3 - 2y^2 - y - 4)$. This polynomial has a unique root y_0 that is larger than 1, and this corresponds to

$$d_0 = 1 + y_0^2 = 3 + (31 - 2\sqrt{58})^{\frac{1}{3}} + (31 + 2\sqrt{58})^{\frac{1}{3}} \approx 9.0967.$$

Hence the asymptotically optimal degree must be either 9 or 10, and inspection shows $R_q(10) > R_q(9) \to \frac{5}{49}\left(113 + 72\sqrt{2}\right) \approx 21.92$ as $q \to \infty$.

For large q, then, Theorem 5.6 yields a larger strong blocking set than Theorem 5.2. However, aside from the fact that Theorem 5.6 works over any field, not just quadratic ones, it also outperforms Theorem 5.2 for small values of q. More precise estimates are given in Table 2, and in comparison to Table 1, we find that the field reduction can lead to significantly smaller constants when q is small.

q	argmin $R_q(d)$	upper bound$/k(q+1)$
3	$d = 85$	296.12
4	$d = 38$	107.35
5	$d = 27$	69.41
7	$d = 19$	46.32
8	$d = 17$	41.45
9	$d = 16$	38.18
11	$d = 14$	34.08
13	$d = 13$	31.62
$16 \leq q \leq 19$	$d = 12$	29.36
$23 \leq q \leq 32$	$d = 11$	26.73
$37 \leq q \leq 109$	$d = 10$	24.75
$q \geq 113$	$d = 9$	22.81

Table 2: For given ranges of prime powers, this table provides the degrees d that minimize the sizes of the strong blocking sets constructed in Theorem 5.6, and upper bounds on their sizes.

6 Conclusion

In this paper, we describe a general machinery for constructing strong blocking sets in finite projective spaces starting from a graph and a linear code. In particular, taking explicit constructions of constant-degree expanders and asymptotically good linear codes, we provide the first explicit construction of strong blocking sets in $\mathrm{PG}(k-1, q)$ whose size is linear in both k and q. As a consequence, as highlighted in Theorem 4.5, this also provides an explicitly constructed family of asymptotically good minimal codes over $\mathbb{F}_q$ of rate at least $(cq)^{-1}$, for some absolute constant c.

These constructions are based on our new results on the vertex integrity of a graph. Concretely, Corollary 3.4 bounds the vertex integrity of d-regular graphs from below by a quantity only depending on their eigenvalues. Finally, in Section 5 we optimize the constant c: we make use of almost Ramanujan graphs and asymptotically good families of AG codes (Theorem 5.2), and combine this construction with the field reduction on a viable set of points, obtaining a derived strong blocking set (Theorem 5.6).

It must be noted that the strong blocking sets from Theorem 5.6 are themselves the unions of lines in $\mathrm{PG}(k-1,q)$, and if q is a square, we can again apply Theorem 5.4 to derive strong blocking sets in $\mathrm{PG}(2k-1,\sqrt{q})$. We can then repeat this process further until we reach a field of non-square order. Thus, we can start with a strong blocking set over $\mathbb{F}_{q^{2^r}}$, and then get a strong blocking set over $\mathbb{F}_q$ in r steps. The calculations, which we omit, are essentially the same as those in the proof of Theorem 5.6. The only exception is that since our intermediate blocking sets will not be coming from graphs, we cannot apply Lemma 5.5 each time to produce very economical viable sets. Instead, since we choose four points from each line to make a viable set, we shall bound the number of lines in each iteration as being at most four times larger than in the previous step. This allows us to derive an explicit construction of strong blocking sets in $\mathrm{PG}(k_i-1,q)$ of size at most

$$\left(\frac{2^{r-1}(d+1)(d+2\sqrt{d-1})(q^{2^r}-1)}{d(q^{2^r}-2)-2q^{2^r}\sqrt{d-1}}+\varepsilon\right)k_i(q+1).$$

for some increasing sequence $\{k_i\}_{i\in N}$, provided that $(q^{2^r}-1)^{-1}<\frac{d-2\sqrt{d-1}}{d+2\sqrt{d-1}}$.

Performing the optimization reveals that the third derivation ($r=3$) minimizes the size of the blocking sets for $q=2$ and that the second derivation ($r=2$) is optimal for $3\leq q\leq 5$. For $q\geq 7$, however, the strong blocking sets from Theorems 5.2 and 5.6 are already so small — they are better than the previous best-known constructions [11,20] — that repeated derivations offer no improvement. For the convenience of the reader, we summarize in Table 3 the smallest strong blocking sets we obtained using our constructions.

Our construction of strong blocking sets can be used to give explicit constructions of affine blocking sets with respect to codimension-2 subspaces [14]. Motivated by this problem, it will be interesting to explore a generalization of our construction to r-uniform hypergraphs, which would lead to a construction of affine blocking sets with respect to codimension-r subspaces.

References

[1] G. N. Alfarano, M. Borello, and A. Neri. A geometric characterization of minimal codes and their asymptotic performance. *Advances in Mathematics of Communications*, 16(1):115–133, 2022.

[2] G. N. Alfarano, M. Borello, and A. Neri. Outer strong blocking sets. *preprint arXiv:2301.09590*, 2023.

[3] G. N. Alfarano, M. Borello, A. Neri, and A. Ravagnani. Three combinatorial perspectives on minimal codes. *SIAM Journal on Discrete Mathematics*, 36(1):461–489, 2022.

[4] N. Alon. Explicit expanders of every degree and size. *Combinatorica*, pages 1–17, 2021.

[5] N. Alon, J. Bruck, J. Naor, M. Naor, and R. M. Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 38(2):509–516, 1992.

[6] A. Ashikhmin and A. Barg. Minimal vectors in linear codes. *IEEE Transactions on Information Theory*, 44(5):2010–2017, 1998.

q	Construction	upper bound/$k(q+1)$
2	3rd derivation	118
3	2nd derivation	77
4	2nd derivation	59
5	2nd derivation	54
7	1st derivation	47
8	1st derivation	42
9	1st derivation	39
11	1st derivation	35
13	1st derivation	32
16	1st derivation	30
17	1st derivation	29
19	1st derivation	28
$23 \leq q \leq 25$	1st derivation	27
$27 \leq q \leq 32$	1st derivation	26
$37 \leq q \leq 49$	1st derivation	25
$53 \leq q \leq 109$	1st derivation	24
$113 \leq q \leq 1217$	1st derivation	23
$1223 \leq q \leq 12763$	1st derivation	22
$q \geq 12769$, q non-square	1st derivation	22
$12769 \leq q < 70603$, q square	original	21
$q > 70603$, q square	original	20

Table 3: For given ranges of prime powers q, this table provides the best upper bound on the size of the constructed strong blocking sets in finite projective spaces over the finite field $\mathbb{F}_q$, together with an indication of which construction achieve this: Original (Theorem 5.2), 1st derivation (Theorem 5.6) or rth derivation.

[7] K. S. Bagga, L. W. Beineke, W. D. Goddard, M. J. Lipman, and R. E. Pippert. A survey of integrity. *Discrete Applied Mathematics*, 37:13–28, 1992.

[8] J. Balogh, T. Mészáros, and A. Z. Wagner. Two results about the hypercube. *Discrete Applied Mathematics*, 247:322–326, 2018.

[9] C. A. Barefoot, R. Entringer, and H. Swart. Vulnerability in graphs-a comparative survey. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 1(38):13–22, 1987.

[10] D. Bartoli and M. Bonini. Minimal linear codes in odd characteristic. *IEEE Transactions on Information Theory*, 65(7):4152–4155, 2019.

[11] D. Bartoli and M. Borello. Small strong blocking sets by concatenation. *SIAM Journal on Discrete Mathematics*, 37(1):65–82, 2023.

[12] D. Bartoli, A. Cossidente, G. Marino, and F. Pavese. On cutting blocking sets and their codes. *Forum Mathematicum*, 34(2):347–368, 2022.

[13] D. Benko, C. Ernst, and D. Lanphier. Asymptotic bounds on the integrity of graphs and separator theorems for graphs. *SIAM Journal on Discrete Mathematics*, 23(1):265–277, 2009.

[14] A. Bishnoi, J. D'haeseleer, D. Gijswijt, and A. Potukuchi. Blocking sets, minimal codes and trifferent codes. *arXiv:2301.09457*, 2023.

[15] A. Blokhuis, P. Sziklai, and T. Szonyi. Blocking sets in projective spaces. *Current research topics in Galois geometry*, pages 61–84, 2011.

[16] M. Bonini and M. Borello. Minimal linear codes arising from blocking sets. *Journal of Algebraic Combinatorics*, 53:327–341, 2021.

[17] A. E. Brouwer and A. Schrijver. The blocking number of an affine space. *Journal of Combinatorial Theory, Series A*, 24(2):251–253, 1978.

[18] H. Chabanne, G. Cohen, and A. Patey. Towards secure two-party computation from the wire-tap channel. In *International Conference on Information Security and Cryptology*, pages 34–46. Springer, 2013.

[19] G. Cohen and A. Lempel. Linear intersecting codes. *Discrete Mathematics*, 56(1):35–43, 1985.

[20] G. Cohen, S. Mesnager, and H. Randriam. Yet another variation on minimal linear codes. *Advances in Mathematics of Communications*, 10(1):53–61, 2016.

[21] G. D. Cohen, S. Mesnager, and A. Patey. On minimal and quasi-minimal linear codes. In *IMA International Conference on Cryptography and Coding*, pages 85–98. Springer, 2013.

[22] G. D. Cohen and G. Zémor. Intersecting codes and independent families. *IEEE Transactions on Information Theory*, 40(6):1872–1881, 1994.

[23] A. Couvreur and H. Randriambololona. Algebraic geometry codes and some applications. In *Concise Encyclopedia of Coding Theory*, pages 307–362. Chapman and Hall/CRC, 2021.

[24] A. A. Davydov, M. Giulietti, S. Marcugini, and F. Pambianco. Linear nonbinary covering codes and saturating sets in projective spaces. *Advances in Mathematics of Communications*, 5(1):119–147, 2011.

[25] C. Ding. Linear codes from some 2-designs. *IEEE Transactions on Information Theory*, 61(6):3265–3275, 2015.

[26] G. Dósa, I. Szalkai, C. Laflamme, et al. The maximum and minimum number of circuits and bases of matroids. *Pure Mathematics and Applications*, 15(4):383–392, 2004.

[27] S. Fancsali and P. Sziklai. Lines in higgledy-piggledy arrangement. *Electronic Journal of Combinatorics*, 21, 2014.

[28] A. Garcia and H. Stichtenoth. A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound. *Inventiones mathematicae*, 121(1):211–222, 1995.

[29] A. Garcia and H. Stichtenoth. On the asymptotic behaviour of some towers of function fields over finite fields. *Journal of Number Theory*, 61(2):248–273, 1996.

[30] T. Héger and Z. L. Nagy. Short minimal codes and covering codes via strong blocking sets in projective spaces. *IEEE Transactions on Information Theory*, 68(2):881–890, 2021.

[31] T. Héger, B. Patkós, and M. Takáts. Search problems in vector spaces. *Designs, Codes and Cryptography*, 76(2):207–216, 2015.

[32] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[33] T.-Y. Hwang. Decoding linear block codes for minimizing word error rate (corresp.). *IEEE Transactions on Information Theory*, 25(6):733–737, 1979.

[34] R. E. Jamison. Covering finite fields with cosets of subspaces. *Journal of Combinatorial Theory, Series A*, 22(3):253–266, 1977.

[35] J. Justesen. Class of constructive asymptotically good algebraic codes. *IEEE Transactions on Information Theory*, 18(5):652–656, 1972.

[36] T. Kövári, V. T. Sós, and P. Turán. On a problem of Zarankiewicz. *Colloquium Mathematicum*, 3:50–57, 1954.

[37] M. Lavrauw and G. Van de Voorde. Field reduction and linear sets in finite geometry. *Topics in finite fields*, 632:271–293, 2015.

[38] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[39] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[40] J. L. Massey. Minimal codewords and secret sharing. In *Proceedings of the 6th joint Swedish-Russian international workshop on information theory*, pages 276–279, 1993.

[41] D. Miklós. Linear binary codes with intersection properties. *Discrete Applied Mathematics*, 9(2):187–196, 1984.

[42] M. Morgenstern. Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory, Series B*, 62(1):44–62, 1994.

[43] A. Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

[44] M. Scotti. On the lower bound for the length of minimal codes. *arXiv preprint arXiv:2302.05350*, 2023.

[45] M. Sipser and D. A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996.

[46] A. Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 238–251, 2017.

[47] C. Tang, Y. Qiu, Q. Liao, and Z. Zhou. Full characterization of minimal linear codes as cutting blocking sets. *IEEE Transactions on Information Theory*, 67(6):3690–3700, 2021.

[48] R. Tanner. A recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, 27(5):533–547, 1981.

[49] M. Tsfasman and S. G. Vladut. *Algebraic-geometric codes*, volume 58. Springer Science & Business Media, 2013.

[50] M. A. Tsfasman, S. G. Vlădut, and T. Zink. Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982.

[51] S. P. Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.

[52] A. Vince. The integrity of a cubic graph. *Discrete Applied Mathematics*, 140(1-3):223–239, 2004.

[53] H. Wang and C. Xing. Explicit constructions of perfect hash families from algebraic curves over finite fields. *Journal of Combinatorial Theory, Series A*, 93(1):112–124, 2001.

A Tight bounds for integrity

In this appendix, we show that the maximum possible integrity of an n-vertex graph of average degree d is $\left(1-\Theta\left(\frac{\log d}{d}\right)\right)n$, complementing the bounds given in Section 3.

Proposition A.1. Let $d \geq 2$.

(i) If $n \geq 48d$, the integrity of any n-vertex graph G of average degree at most d satisfies

$$\iota(G) \leq \left(1-\frac{\log d}{4d}\right)n.$$

(ii) For all $n \geq d$, there are n-vertex graphs G of average degree at most d with

$$\iota(G) \geq \left(1-\frac{4\log d}{d}\right)n.$$

Proof. For both parts, we shall appeal to Proposition 3.3, which asserts that $n-2z(G) \leq \iota(G) \leq n-z(G)$, where $z(G)$ is the largest z such that G contains two disjoint sets of z vertices that have no edges between them.

For part (i), we need to show that $z(G) \geq \frac{n\log d}{4d}$ for all such graphs G. This follows from the Kővári–Sós–Turán Theorem [36], applied to the complement of G. For the sake of completeness, though, we provide a simple probabilistic proof.

Let A be a random subset of $V = V(G)$ obtained by selecting each vertex of G independently with probability $p = \frac{\log d}{2d}$. The size of A is then a binomial random variable, and the Chernoff bound shows that $|A| \geq \frac{n\log d}{4d}$ with probability at least $1-e^{-n\log d/(16d)}$, which is at least $1-\frac{1}{d^3}$.

We now define B to be the set of all vertices in $V \setminus A$ that have no neighbors in A. Note that if a vertex v has degree d_v, then $\mathbb{P}(v \in B) = (1-p)^{d_v+1}$, as we need that neither v nor any of its d_v neighbors belong to A. Thus, the expected size of B is $\sum_{v\in V}(1-p)^{d_v+1}$. Since $(1-p)^x$ is a convex function, and the average degree is at most d, we have

$$\mathbb{E}[|B|] = \sum_{v\in V}(1-p)^{d_v+1} \geq n(1-p)^{d+1} = n\left(1-\frac{\log d}{2d}\right)^{d+1}.$$

Computation shows that this is at least $\frac{3n}{4d^{1/2}}$. Since $|B|$ cannot be larger than n, we have

$$\frac{3n}{4d^{1/2}} \leq \mathbb{E}[|B|] \leq n\mathbb{P}\left(|B| \geq \frac{n}{2d^{1/2}}\right) + \frac{n}{2d^{1/2}}\mathbb{P}\left(|B| \leq \frac{n}{2d^{1/2}}\right) \leq n\mathbb{P}\left(|B| \geq \frac{n}{2d^{1/2}}\right) + \frac{n}{2d^{1/2}},$$

whence it follows that $\mathbb{P}\left(|B| \geq \frac{n}{2d^{1/2}}\right) \geq \frac{1}{4d^{1/2}} > \frac{1}{d^3}$.

Hence, with positive probability, we have both $|A| \geq \frac{n\log d}{4d}$ and $|B| \geq \frac{n}{2d^{1/2}} \geq \frac{n\log d}{4d}$, and the existence of such a pair of sets shows $z(G) \geq \frac{n\log d}{4d}$, as required.

For part (ii), we need to show the existence of a graph G, of average degree at most d, for which $z(G) \leq \frac{2n\log d}{d}$. Note that we always have $z(G) \leq \frac{n}{2}$, and so this is trivial if $d \leq 8$.

Now consider the random graph $G\left(n, \frac{d}{n}\right)$, where every edge is present independently with probability $p = \frac{d}{n}$. The number of edges is a binomial random variable, whose median is at most $\left\lceil \frac{d(n-1)}{2} \right\rceil$, and hence $\mathbb{P}\left(e(G) > \frac{dn}{2}\right) < \frac{1}{2}$.

We can use a straightforward union bound to show that there are no large subsets without any edges between them. Indeed, the expected number of pairs of sets of z vertices such that all z^2 cross-edges are missing is at most

$$\binom{n}{z}^2(1-p)^{z^2} \leq \left(\frac{ne}{z}\right)^{2z}e^{-pz^2} = \left(\frac{n^2e^2}{z^2e^{pz}}\right)^z.$$

Substituting our choice of $z = \frac{2n \log d}{d}$, this simplifies to $\left(\frac{e^2}{4\log^2 d}\right)^{\frac{2n \log d}{d}}$. Since $d \geq 8$, this is at most $2^{-\frac{2n \log d}{d}}$, which is less than $\frac{1}{2}$.

Thus, with positive probability, G is such that $e(G) \leq \frac{dn}{2}$ and $z(G) \leq \frac{2n \log d}{d}$, as required. □

Remark A.2. We have dealt with graphs of bounded average degree for simplicity, so that we could use the binomial random graph in part (ii). If one is primarily interested in d-regular graphs, as we have been using in this paper, then the upper bound in part (i) naturally still applies. For the lower bound in part (ii), one must replace the binomial random graph with the random d-regular graph. At the expense of more complicated calculations, a similar bound can be shown, provided d is not too small.