Upper Bounds on Matching Families in $\mathbb{Z}_{pq}^n$

Yeow Meng Chee, *Senior Member, IEEE*, San Ling, Huaxiong Wang, and Liang Feng Zhang

***Abstract*—*Matching families* are one of the major ingredients in the construction of *locally decodable codes* (LDCs) and the best known constructions of LDCs with a constant number of queries are based on matching families. The determination of the largest size of any matching family in $\mathbb{Z}_m^n$, where $\mathbb{Z}_m$ is the ring of integers modulo m, is an interesting problem. In this paper, we show an upper bound of $O((pq)^{0.625n+0.125})$ for the size of any matching family in $\mathbb{Z}_{pq}^n$, where p and q are two distinct primes. Our bound is valid when n is a constant, $p \to \infty$, and $p/q \to 1$. Our result improves an upper bound of Dvir and coworkers.**

***Index Terms*—Locally decodable codes (LDCs), matching families, upper bound.**

I. Introduction

LOCALLY decodable codes: A classical error-correcting code C allows one to encode any message $\mathbf{x} = (\mathbf{x}(1), \ldots, \mathbf{x}(k))$ of k symbols as a codeword $C(\mathbf{x})$ of N symbols such that the message can be recovered even if $C(\mathbf{x})$ gets corrupted in a number of coordinates. However, to recover even a small fraction of the message, one has to consider all or most of the coordinates of the codeword. In such a scenario, more efficient schemes are possible. They are known as *locally decodable codes* (LDCs). Such codes allow the reconstruction of any symbol of the message by looking at a small number of coordinates of the codeword, even if a constant fraction of the codeword has been corrupted.

Let k, N be positive integers and let $\mathbb{F}$ be a finite field. For any $\mathbf{y}, \mathbf{z} \in \mathbb{F}^N$, we denote by $d_H(\mathbf{y}, \mathbf{z})$ the *Hamming distance* between $\mathbf{y}$ and $\mathbf{z}$.

Definition 1.1 (LDC): A code $C : \mathbb{F}^k \to \mathbb{F}^N$ is said to be (r, δ, ϵ)-locally decodable if there is a randomized decoding algorithm D such that

1) for every $\mathbf{x} \in \mathbb{F}^k, i \in [k]$ and $\mathbf{y} \in \mathbb{F}^N$ such that $d_H(C(\mathbf{x}), \mathbf{y}) \leq \delta N$, $\Pr[D^{\mathbf{y}}(i) = \mathbf{x}(i)] > 1 - \epsilon$, where the probability is taken over the random coins of D; and
2) D makes at most r queries to $\mathbf{y}$.

Manuscript received September 27, 2012; revised February 09, 2013; accepted March 03, 2013. Date of publication May 02, 2013; date of current version July 10, 2013. This work was supported by the National Research Foundation of Singapore under Research Grant NRF-CRP2–2007-03.

Y. M. Chee, S. Ling, and H. Wang are with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore (e-mail: ymchee@ntu.edu.sg; lingsan@ntu.edu.sg; hxwang@ntu.edu.sg).

L. F. Zhang was with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore. He is now with the University of Calgary, Calgary, AB T2N 1N4 Canada (e-mail: liangf.zhang@gmail.com).

Communicated by V. Guruswami, Associate Editor for Complexity and Cryptography.

Digital Object Identifier 10.1109/TIT.2013.2257918

The efficiency of C is measured by its *query complexity* r and *length* N (as a function of k). Ideally, one would like both r and N to be as small as possible.

Implicit discussion of the notion of LDCs dates back to [2], [30], [34]. Katz and Trevisan [24] were the first to formally define LDCs and prove (superlinear) lower bounds on their length. Kerenidis and de Wolf [26] showed a tight (exponential) lower bound for the length of 2-query LDCs. Woodruff [37] obtained superlinear lower bounds for the length of r-query LDCs, where $r \geq 3$. More lower bounds for specific LDCs can be found in [13], [16], [19], [29], [33], and [36]. On the other hand, many constructions of LDCs have been proposed in the past decade. These constructions can be classified into three generations based on their technical ideas. The first-generation LDCs [2], [6], [12], [24] are based on (low-degree) multivariate polynomial interpolation. In such a code, each codeword consists of evaluations of a low-degree polynomial in $\mathbb{F}[z_1, \ldots, z_n]$ at all points of $\mathbb{F}^n$, for some finite field $\mathbb{F}$. The decoder recovers the value of the unknown polynomial at a point by shooting a line in a random direction and decoding along it using noisy polynomial interpolation [5], [28], [35]. The second-generation LDCs [7], [38] are also based on low-degree multivariate polynomial interpolation but with a clever use of recursion. The third-generation LDCs, known as *matching vector codes* (MV codes), were initiated by Yekhanin [39] and developed further in [8], [10], [15], [17], [20], [22], [23], [25], and [31]. The constructions involve novel combinatorial and algebraic ideas, where the key ingredient is the design of large *matching families* in $\mathbb{Z}_m^n$. The interested reader may refer to Yekhanin [40] for a good survey of LDCs.

Matching families: Let m and n be positive integers. For any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{Z}_m^n$, we denote by $\langle \mathbf{u}, \mathbf{v} \rangle \triangleq \sum_{i=1}^k \mathbf{u}(i)\mathbf{v}(i) \mod m$ their *dot product*.

Definition 1.2 (Matching family): Let $S \subseteq \mathbb{Z}_m \setminus \{0\}$. Two families of vectors $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \mathbb{Z}_m^n$ form an S-matching family in $\mathbb{Z}_m^n$ if

1) $\langle \mathbf{u}_i, \mathbf{v}_i \rangle = 0$ for every $i \in [k]$; and
2) $\langle \mathbf{u}_i, \mathbf{v}_j \rangle \in S$ for every $i, j \in [k]$ such that $i \neq j$.

The matching family defined above is of *size* k. Dvir *et al.* [15] showed that if there is an S-matching family of size k in $\mathbb{Z}_m^n$, then there is an $(|S| + 1)$-query LDC encoding messages of length k as codewords of length m^n. Hence, large matching families are interesting because they result in short LDCs. For any $S \subseteq \mathbb{Z}_m \setminus \{0\}$, it is interesting to determine the largest size of any S-matching family in $\mathbb{Z}_m^n$. When $S = \mathbb{Z}_m \setminus \{0\}$, this largest size is often denoted by $k(m, n)$, which is clearly a *universal* upper bound for the size of any matching family in $\mathbb{Z}_m^n$.

Set systems: The study of matching families dates back to *set systems with restricted intersections* [3], whose study was initiated in [18].

Definition 1.3 (Set system): Let T and S be two disjoint subsets of $\mathbb{Z}_m$. A collection $\mathcal{F} = \{F_1, \ldots, F_k\}$ of subsets of $[n]$ is said to be a (T, S)- set system over $[n]$ if

1) $|F_i| \bmod m \in T$ for every $i \in [k]$; and
2) $|F_i \cap F_j| \bmod m \in S$ for every $i, j \in [k]$ such that $i \neq j$.

The set system defined above is of *size* k. When $T = \{0\}$ and $S \subseteq \mathbb{Z}_m \setminus \{0\}$, it is easy to show that the (T, S)-set system $\mathcal{F}$ yields an S-matching family of size k in $\mathbb{Z}_m^n$. To see this, let $\mathbf{u}_i = \mathbf{v}_i \in \mathbb{Z}_m^n$ be the characteristic vector of F_i for every $i \in [k]$, where $\mathbf{u}_i(j) = \mathbf{v}_i(j) = 1$ for every $j \in F_i$ and 0 otherwise. Clearly, $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ and $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ form an S-matching family of size k in $\mathbb{Z}_m^n$.

When m is a prime power and $n \geq m$, Deza *et al.* [14] and Babai *et al.* [4] showed that the largest size of any $(\{0\}, \mathbb{Z}_m \setminus \{0\})$-set systems over $[n]$ cannot be greater than $\binom{n}{m-1} + \cdots + \binom{n}{0}$. For any integer m, Sgall [32] showed that the largest size of any $(\{0\}, \mathbb{Z}_m \setminus \{0\})$-set system over $[n]$ is bounded by $O(2^{0.5n})$. On the other hand, Grolmusz [21] constructed a $(\{0\}, \mathbb{Z}_m \setminus \{0\})$-set system of (superpolynomial) size $\exp(O((\log n)^r/(\log\log n)^{r-1}))$ over $[n]$ when m has $r \geq 2$ distinct prime divisors. Grolmusz's set systems result in superpolynomial-sized matching families in $\mathbb{Z}_m^n$, which have been the key ingredient for Efremenko's LDCs [17].

Bounds: Due to the difficulty of determining $k(m, n)$ precisely, it is interesting to give both lower and upper bounds for $k(m, n)$. When $m \leq n$, a *simple lower bound* for $k(m, n)$ is $k \triangleq \binom{n}{m-1}$. To see this, let $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be the set of all 0–1 vectors of *Hamming weight* (i.e., the number of nonzero components) $m - 1$ in $\mathbb{Z}_m^n$. Let $\mathbf{v}_i = \mathbb{1} - \mathbf{u}_i$ for every $i \in [k]$, where $\mathbb{1}$ is the all-one vector. Then, $\mathcal{U}$ and $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ form a matching family of size k. When m is a composite number with $r \geq 2$ distinct prime factors, the $(\{0\}, \mathbb{Z}_m \setminus \{0\})$-set systems of [15], [21], and [27] result in superpolynomial-sized matching families in $\mathbb{Z}_m^n$. In particular, we have that $k(m, n) \geq \exp(O(\log^2 n/\log\log n))$ when $m = pq$ for two distinct primes p and q. On the other hand, Dvir *et al.* [15] obtained upper bounds for $k(m, n)$ for various settings of the integers m and n. More precisely, they showed that

1) $k(m, n) \leq m^{n-1+o_m(1)}$ for any integers m and n, where $o_m(1)$ is a term that tends to 0 as m approaches infinity;
2) $k(p, n) \leq \min\{1 + \binom{n+p-2}{p-1}, 4p^{0.5n} + 2\}$ for any prime p and integer n;
3) $k(m, n) \leq (m/q)^n k(q, n)$ for any integers m, n, and q such that $q|m$ and $\gcd(q, m/q) = 1$.

In particular, the latter two bounds imply that $k(m, n) \leq p^n(4q^{0.5n} + 2)$ when $m = pq$ for two distinct primes p and q such that $p \leq q$.

Our results: Dvir *et al.* [15] conjectured that $k(m, n) \leq O(m^{0.5n})$ for any integers m and n. A *special case* where the conjecture is open is when n is a constant, and $m = pq$ for two distinct primes p, q such that $p \to \infty$ and $p/q \to 1$. In this paper, we show that $k(m, n) \leq O(m^{0.625n+0.125})$ for this special case, which improves the best known upper bound that can be derived from results of Dvir *et al.* in [15], i.e., $k(m, n) \leq p^n(4q^{0.5n} + 2) = O(m^{0.75n})$.

Our techniques: Let $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \mathbb{Z}_m^n$ be a matching family of size $k = k(m, n)$, where $m = pq$ for two distinct primes p and q. We say that $\mathbf{u}, \mathbf{v} \in \mathbb{Z}_m^n$ are *equivalent* (and write $\mathbf{u} \sim \mathbf{v}$) if there is a $\lambda \in \mathbb{Z}_m^*$ such that $\mathbf{u}(i) = \lambda\mathbf{v}(i)$ for every $i \in [n]$, where $\mathbb{Z}_m^*$ is the set of units of $\mathbb{Z}_m$. Clearly, no two elements of $\mathcal{U}$ (resp. $\mathcal{V}$) can be equivalent to each other. Let $s, t \in \{1, p, q, m\}$. We say that $(\mathbf{u}_i, \mathbf{v}_i)$ is of *type* (s, t) if $\gcd(\mathbf{u}_i(1), \ldots, \mathbf{u}_i(n), m) = s$ and $\gcd(\mathbf{v}_i(1), \ldots, \mathbf{v}_i(n), m) = t$. We can partition the set $\{(\mathbf{u}_i, \mathbf{v}_i) : i \in [k]\}$ of pairs according to their types. Let $N_{s,t}$ be the number of pairs of type (s, t). Then, we have the following observations:

1) $N_{s,t} \leq 1$ when $m|st$ (see Lemma 3.9);
2) $N_{s,t} \leq k(q, n)$ when $(s, t) \in \{(1, p), (p, 1), (p, p)\}$ (see Lemma 3.10); and
3) $N_{s,t} \leq k(p, n)$ when $(s, t) \in \{(1, q), (q, 1), (q, q)\}$ (see Lemma 3.11).

These observations in turn imply that $k \leq 9 + N_{1,1} + 3k(p, n) + 3k(q, n)$ and enable us to reduce the problem of upper bounding k to that of establishing an upper bound for $N_{1,1}$.

As in [15], we establish an upper bound for $N_{1,1}$ by using an interesting relation between matching families and the *expanding properties* of the *projective graphs* (which will be explained shortly). Let

$$\mathbb{S}_{n,m} = \{\mathbf{u} \in \mathbb{Z}_m^n : \gcd(\mathbf{u}(1), \ldots, \mathbf{u}(n), m) = 1\} \text{ and}$$
$$\mathbb{P}_{n,m} = \mathbb{H}_{n,m} = \mathbb{S}_{n,m}/\sim . \quad (1)$$

We define the *projective* $(n-1)$*-space* over $\mathbb{Z}_m$ to be the pair $(\mathbb{P}_{n,m}, \mathbb{H}_{n,m})$. We call the elements of $\mathbb{P}_{n,m}$ *points* and the elements of $\mathbb{H}_{n,m}$ *hyperplanes*. We say that a point $\mathbf{u}$ *lies on* a hyperplane $\mathbf{v}$ if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. The projective graph $\mathbf{G}_{n,m}$ is defined to be a bipartite graph with classes of vertices $\mathbb{P}_{n,m} \cup \mathbb{H}_{n,m}$, where a point $\mathbf{u}$ and a hyperplane $\mathbf{v}$ are *adjacent* if and only if $\mathbf{u}$ lies on $\mathbf{v}$. Vertices that are adjacent to each other are called *neighbors*. A set $\mathcal{U}' \subseteq \mathbb{P}_{n,m}$ has the *unique neighbor property* if, for every $\mathbf{u} \in \mathcal{U}'$, there is a hyperplane $\mathbf{v}$ such that $\mathbf{v}$ is adjacent to $\mathbf{u}$ but to no other points in $\mathcal{U}'$ (see also Definition 3.1). Without loss of generality, let $\{(\mathbf{u}_i, \mathbf{v}_i) : i \in [k']\}$ be the set of pairs of type $(1, 1)$, where $k' = N_{1,1}$. Let $\mathcal{U}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_{k'}\} \subseteq \mathbb{P}_{n,m}$. It is straightforward to see that $\mathcal{U}'$ satisfies the unique neighbor property (Lemma 3.8). For any $X \subseteq \mathcal{U}'$, we denote by $N(X)$ the *neighborhood* of X, i.e., the collection of vertices in $\mathbb{H}_{n,m}$ that are adjacent to some vertex in X. Since every point in $\mathcal{U}' \setminus X$ must have a unique neighbor in $\mathbb{H}_{n,m} \setminus N(X)$, we have that

$$|\mathcal{U}'| \leq |X| + |\mathbb{H}_{n,m}| - |N(X)|. \quad (2)$$

We show that $\mathbf{G}_{n,m}$ has some kind of *expanding property* (see Theorem 3.1), meaning that $|N(X)|$ is large for certain choices of X, which allows us to obtain the expected upper bound for $k' = N_{1,1}$ (see Theorems 3.2 and 3.3). When m is a prime, such an expanding property of $\mathbf{G}_{n,m}$ was proved by Alon [1] using the *spectral method* and it says that

$$|N(X)| \geq |\mathbb{P}_{n,m}| - |\mathbb{P}_{n,m}|^{n/(n-1)}/|X| \quad (3)$$

where $X \subseteq \mathbb{P}_{n,m}$ is arbitrary.

Let $A_{n,m} = (a_{\mathbf{uv}})$ be the *adjacency matrix* of $\mathbf{G}_{n,m}$, where the rows are labeled by the points, the columns by the hyperplanes, and $a_{\mathbf{uv}} = 1$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are adjacent. Note that the matrix $A_{n,m}$ may take many different forms because the sets $\mathbb{P}_{n,m}$ and $\mathbb{H}_{n,m}$ are not *ordered*. However, from now on, we always assume that $\mathbb{P}_{n,m}$ and $\mathbb{H}_{n,m}$ are identical to each other as ordered sets. Hence, $A_{n,m}$ is *symmetric*. Let χ_X be the characteristic vector of X, where the components of χ_X are labeled by the elements $\mathbf{u} \in \mathbb{P}_{n,m}$ and $\chi_X(\mathbf{u}) = 1$ if and only if $\mathbf{u} \in X$. Alon [1] obtained both an upper bound and a lower bound for $\chi_X^t B_{n,m} \chi_X$ that jointly result in (3), where $B_{n,m} = A_{n,m} A_{n,m}^t$ with the superscript t denoting the *transpose* of a matrix. More precisely, Alon [1] determined the eigenvalues of $B_{n,m}$ and represented χ_X as a linear combination of the eigenvectors of $B_{n,m}$. In this paper, we develop their spectral method further and show a *tensor lemma* on $B_{n,m}$ (see Lemma 2.1), which says that $\mathbf{G}_{n,m}$ is a tensor product of $\mathbf{G}_{n,p}$ and $\mathbf{G}_{n,q}$ when $m = pq$, where p and q are two distinct primes. As in [1], we determine the eigenvalues of $B_{n,m}$ and represent χ_X as a linear combination of the eigenvectors of $B_{n,m}$. We obtain both an upper bound and a lower bound for $\chi_X^t B_{n,m} \chi_X$, which are then used to show that $\mathbf{G}_{n,m}$ has some kind of expanding property (see Theorem 3.1).

Subsequent work: Recently, in a follow-up work, Bhowmick *et al.* [9] obtained new upper bounds for $k(m,n)$. They used different techniques and showed that $k(m,n) \le m^{0.5n+14\log m}$ for any integers m and n. In particular, their upper bound translates into $k(m,n) \le m^{0.5n+O(1)}$ for the special case we consider in this paper.

Organization: In Section II, we study projective graphs over $\mathbb{Z}_m$ and matrices associated with such graphs. In Section III, we establish our upper bound for $k(pq,n)$ using the unique neighbor property in projective graphs. Section IV contains some concluding remarks.

II. Projective Graphs and Associated Matrices

Let d be a positive integer. We denote by $\mathbb{0}_d$, $\mathbb{1}_d$, I_d, and J_d the all-zero (either row or column) vector of dimension d, all-one (either row or column) vector of dimension d, identity matrix of order d, and all-one matrix of order d, respectively. We denote by O an all-zero matrix whose size is clear from the context. We also define

$$\begin{aligned} K_d &= I_d + J_d, \\ L_d &= \left((d+1)I_d - J_d - \mathbb{1}_d\right), \text{ and} \\ R_d &= \left(I_d - \mathbb{1}_d\right)^t. \end{aligned} \tag{4}$$

Let $A = (a_{ij})$ and B be two matrices. We define their *tensor product* to be the block matrix $A \otimes B = (a_{ij} \cdot B)$. We say that $A \simeq B$ if A can be obtained from B by *simultaneously* permuting the rows and columns (i.e., apply the same permutation to both the rows and columns). Clearly, A and B have the same eigenvalues if $A \simeq B$.

In this section, we study the projective graph $\mathbf{G}_{n,m}$ defined in Section I. We also follow the notation there. Let $\theta_{n,m} = |\mathbb{P}_{n,m}|$ be the number of points (or hyperplanes) in the projective $(n-1)$-space over $\mathbb{Z}_m$. Chee and Ling [11] showed that

$$\theta_{n,m} = m^{n-1} \prod_{p|m} (1 + 1/p + \cdots + 1/p^{n-1}) \tag{5}$$

and $|N(\mathbf{u})| = |N(\mathbf{v})| = \theta_{n-1,m}$ for every point $\mathbf{u}$ and hyperplane $\mathbf{v}$. When m is prime, Alon [1] showed that $\theta_{n-1,m}^2$ is an eigenvalue of $B_{n,m}$ of multiplicity 1 and m^{n-2} is an eigenvalue of $B_{n,m}$ of multiplicity $\theta_{n,m} - 1$. Furthermore, an eigenvector of $B_{n,m}$ with eigenvalue $\theta_{n-1,m}^2$ is $\mathbb{1}$ and linearly independent eigenvectors of $B_{n,m}$ with eigenvalue m^{n-2} can be chosen to be the columns of R_d, where $d = \theta_{n,m} - 1$. However, the eigenvalues of $B_{n,m}$ have not been studied when m is composite. Here, we determine the eigenvalues of $B_{n,m}$ when $m = pq$ for two distinct primes p and q.

Lemma 2.1 (Tensor lemma): Let $n > 1$ be an integer and let $m = pq$ for two distinct primes p and q. Then, $B_{n,m} \simeq B_{n,p} \otimes B_{n,q}$.

Proof: Let $\pi : \mathbb{P}_{n,p} \times \mathbb{P}_{n,q} \to \mathbb{P}_{n,m}$ be the mapping defined by $\pi(\mathbf{u},\mathbf{v}) = \mathbf{w}$, where

$$\mathbf{w}(i) \equiv \mathbf{u}(i) \bmod p \quad \text{and} \quad \mathbf{w}(i) \equiv \mathbf{v}(i) \bmod q \tag{6}$$

for every $i \in [n]$. Then, π is well defined. To see this, let $\mathbf{w}' = \pi(\mathbf{u}',\mathbf{v}')$ and $\mathbf{w} = \pi(\mathbf{u},\mathbf{v})$ for $\mathbf{u},\mathbf{u}' \in \mathbb{S}_{n,p}$ and $\mathbf{v},\mathbf{v}' \in \mathbb{S}_{n,q}$. If $\mathbf{u} \sim \mathbf{u}'$ and $\mathbf{v} \sim \mathbf{v}'$, then there are integers $\lambda \in \mathbb{Z}_p^*$ and $\mu \in \mathbb{Z}_q^*$ such that

$$\mathbf{u}'(i) \equiv \lambda \mathbf{u}(i) \bmod p \quad \text{and} \quad \mathbf{v}'(i) \equiv \mu \mathbf{v}(i) \bmod q \tag{7}$$

for every $i \in [n]$. Let $\delta \in \mathbb{Z}_m^*$ be an integer such that

$$\delta \equiv \lambda \bmod p \quad \text{and} \quad \delta \equiv \mu \bmod q. \tag{8}$$

By (6)–(8), we have that $\mathbf{w}'(i) \equiv \delta \mathbf{w}(i) \bmod m$ for every $i \in [n]$. Hence, $\mathbf{w} \sim \mathbf{w}'$.

Let $\mathbb{P}_{n,p} = \{\mathbf{u}_1, \ldots, \mathbf{u}_{\ell_1}\}$ and $\mathbb{P}_{n,q} = \{\mathbf{v}_1, \ldots, \mathbf{v}_{\ell_2}\}$, where $\ell_1 = \theta_{n,p}$ and $\ell_2 = \theta_{n,q}$. It is clear that π is injective and $\theta_{n,m} = \ell_1 \ell_2$ (this is clear from (5)). It follows that π is bijective and

$$\mathbb{P}_{n,m} = \{\pi(\mathbf{u}_1,\mathbf{v}_1), \ldots, \pi(\mathbf{u}_1,\mathbf{v}_{\ell_2}), \ldots, \pi(\mathbf{u}_{\ell_1},\mathbf{v}_{\ell_2})\}. \tag{9}$$

Let $\mathbf{w}$ and $\mathbf{w}'$ be as above. Then, $\langle \mathbf{w},\mathbf{w}' \rangle \equiv 0 \bmod m$ if and only if $\langle \mathbf{u},\mathbf{u}' \rangle \equiv 0 \bmod p$ and $\langle \mathbf{v},\mathbf{v}' \rangle \equiv 0 \bmod q$. Hence, the $(\mathbf{w},\mathbf{w}')$-entry of $A_{n,m}$ is equal to 1 if and only if the $(\mathbf{u},\mathbf{u}')$-entry of $A_{n,p}$ and the $(\mathbf{v},\mathbf{v}')$-entry of $A_{n,q}$ are both equal to 1. Hence, $A_{n,m} \simeq A_{n,p} \otimes A_{n,q}$. It follows that

$$\begin{aligned} B_{n,m} &= A_{n,m} A_{n,m}^t \\ &\simeq (A_{n,p} \otimes A_{n,q})(A_{n,p} \otimes A_{n,q})^t \\ &= (A_{n,p} A_{n,p}^t) \otimes (A_{n,q} A_{n,q}^t) \\ &= B_{n,p} \otimes B_{n,q} \end{aligned}$$

as desired. □

In fact, we could have concluded that $A_{n,m} = A_{n,p} \otimes A_{n,q}$ and therefore $B_{n,m} = B_{n,p} \otimes B_{n,q}$ in Lemma 2.1. The sole

$\mathbb{P}_{3,2}$	$\mathbb{P}_{3,3}$	$\mathbb{P}_{3,6}$						
(0, 0, 1)	(0, 0, 1)	(0, 0, 1)	(0, 3, 4)	(0, 3, 1)	(3, 0, 4)	(3, 0, 1)	(3, 3, 4)	(3, 3, 1)
(0, 1, 0)	(0, 1, 0)	(0, 4, 3)	(0, 1, 0)	(0, 1, 3)	(3, 4, 0)	(3, 4, 3)	(3, 1, 0)	(3, 1, 3)
(0, 1, 1)	(0, 1, 1)	(0, 4, 1)	(0, 1, 4)	(0, 1, 1)	(3, 4, 4)	(3, 4, 1)	(3, 1, 4)	(3, 1, 1)
(1, 0, 0)	(0, 1, 2)	(0, 4, 5)	(0, 1, 2)	(0, 1, 5)	(3, 4, 2)	(3, 4, 5)	(3, 1, 2)	(3, 1, 5)
(1, 0, 1)	(1, 0, 0)	(4, 0, 3)	(4, 3, 0)	(4, 3, 3)	(1, 0, 0)	(1, 0, 3)	(1, 3, 0)	(1, 3, 3)
(1, 1, 0)	(1, 0, 1)	(4, 0, 1)	(4, 3, 4)	(4, 3, 1)	(1, 0, 4)	(1, 0, 1)	(1, 3, 4)	(1, 3, 1)
(1, 1, 1)	(1, 0, 2)	(4, 0, 5)	(4, 3, 2)	(4, 3, 5)	(1, 0, 2)	(1, 0, 5)	(1, 3, 2)	(1, 3, 5)
	(1, 1, 0)	(4, 4, 3)	(4, 1, 0)	(4, 1, 3)	(1, 4, 0)	(1, 4, 3)	(1, 1, 0)	(1, 1, 3)
	(1, 1, 1)	(4, 4, 1)	(4, 1, 4)	(4, 1, 1)	(1, 4, 4)	(1, 4, 1)	(1, 1, 4)	(1, 1, 1)
	(1, 1, 2)	(4, 4, 5)	(4, 1, 2)	(4, 1, 5)	(1, 4, 2)	(1, 4, 5)	(1, 1, 2)	(1, 1, 5)
	(1, 2, 0)	(4, 2, 3)	(4, 5, 0)	(4, 5, 3)	(1, 2, 0)	(1, 2, 3)	(1, 5, 0)	(1, 5, 3)
	(1, 2, 1)	(4, 2, 1)	(4, 5, 4)	(4, 5, 1)	(1, 2, 4)	(1, 2, 1)	(1, 5, 4)	(1, 5, 1)
	(1, 2, 2)	(4, 2, 5)	(4, 5, 2)	(4, 5, 5)	(1, 2, 2)	(1, 2, 5)	(1, 5, 2)	(1, 5, 5)

Fig. 1. Ordered point sets.

reason that we did not do so is that those matrices may take different forms, as noted in Section I. To facilitate further analysis, we make the matrices unique such that $A_{n,m} = A_{n,p} \otimes A_{n,q}$. This can be achieved by making the sets $\mathbb{P}_{n,p}, \mathbb{P}_{n,q}$ and $\mathbb{P}_{n,m}$ unique. To do so, we first make $\mathbb{P}_{n,p} = [\mathbf{u}_1, \ldots, \mathbf{u}_{\ell_1}]$ and $\mathbb{P}_{n,q} = [\mathbf{v}_1, \ldots, \mathbf{v}_{\ell_2}]$ unique as ordered sets, where $\ell_1 = \theta_{n,p}$ and $\ell_2 = \theta_{n,q}$.

For example, as shown in Fig. 1, we may set $\mathbb{P}_{3,2} = [(0,0,1),\ (0,1,0),\ (0,1,1),(1,0,0),\ (1,0,1),\ (1,1,1)]$ and $\mathbb{P}_{3,3} = [(0,0,1), (0,1,0), (0,1,1), (0,1,2), (1,0,0), (1,0,1), (1,0,2), (1,1,0), (1,1,1), (1,1,2), (1,2,0), (1,2,1), (1,2,2)]$. Then, both $\mathbb{P}_{3,2}$ and $\mathbb{P}_{3,3}$ have been made unique as ordered sets. (Here, each equivalence class in $\mathbb{P}_{n,p}$ is represented by the first element when its elements are arranged in lexicographical order, and these representatives are subsequently also arranged in lexicographical order.) Once $\mathbb{P}_{n,p}$ and $\mathbb{P}_{n,q}$ have been made unique as ordered sets, we can simply set $\mathbb{P}_{n,m} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_\ell] = [\pi(\mathbf{u}_1, \mathbf{v}_1), \pi(\mathbf{u}_1, \mathbf{v}_2), \ldots, \pi(\mathbf{u}_{\ell_1}, \mathbf{v}_{\ell_2})]$, where $\ell = \ell_1 \ell_2$ and $\mathbf{w}_1 = \pi(\mathbf{u}_1, \mathbf{v}_1), \mathbf{w}_2 = \pi(\mathbf{u}_1, \mathbf{v}_2), \ldots, \mathbf{w}_\ell = \pi(\mathbf{u}_{\ell_1}, \mathbf{v}_{\ell_2})$. For example, as shown in Fig. 1, $\mathbb{P}_{3,6}$ consists of $\ell_1 (= 7)$ columns and the ith column corresponds to $\pi(\mathbf{u}_i, \mathbf{v}_1), \ldots, \pi(\mathbf{u}_i, \mathbf{v}_{\ell_2})$ for every $i \in [\ell_1]$. From now on, we suppose that the point sets $\mathbb{P}_{n,p}, \mathbb{P}_{n,q}$, and $\mathbb{P}_{n,m}$ have always been made unique, such as in the way illustrated above. Then, we have

$$A_{n,m} = A_{n,p} \otimes A_{n,q} \quad \text{and} \quad B_{n,m} = B_{n,p} \otimes B_{n,q}. \tag{10}$$

Let $d_1 = 1,\ d_2 = \ell_1 - 1,\ d_3 = \ell_2 - 1$, and $d_4 = (\ell_1 - 1)(\ell_2 - 1)$. We define an $\ell \times \ell$ matrix

$$\begin{aligned} Y &= (\, Y_1 \quad Y_2 \quad Y_3 \quad Y_4 \,) \\ &= (\, \mathbb{1}_\ell \quad R_{d_2} \otimes \mathbb{1}_{\ell_2} \quad \mathbb{1}_{\ell_1} \otimes R_{d_3} \quad R_{d_2} \otimes R_{d_3} \,). \end{aligned} \tag{11}$$

Lemma 2.2: For every $s \in \{1,2,3,4\}$, the d_s columns of Y_s are linearly independent eigenvectors of $B_{n,m}$ with eigenvalue λ_s, where $\lambda_1 = \theta_{n-1,m}^2,\ \lambda_2 = p^{n-2}\theta_{n-1,q}^2,\ \lambda_3 = q^{n-2}\theta_{n-1,p}^2$, and $\lambda_4 = m^{n-2}$.

Proof: The proof consists of simple verification. For example, when $s = 4$, we have that $B_{n,m} \cdot Y_4 = (B_{n,p} \otimes B_{n,q}) \cdot (R_{d_2} \otimes R_{d_3}) = (B_{n,p} \cdot R_{d_2}) \otimes (B_{n,q} \cdot R_{d_3}) = \left(p^{n-2} \cdot R_{d_2}\right) \otimes \left(q^{n-2} \cdot R_{d_3}\right) = \lambda_4 \cdot Y_4$, where the first equality is due to (10). Similarly, we can verify for $s \in \{1,2,3\}$. The linear independence of the columns of Y_s can be checked using the linear independence of the columns of R_{d_2} and R_{d_3}. □

Lemma 2.3: We have that

$$Y^{-1} = \ell^{-1} \cdot \begin{pmatrix} \mathbb{1}_\ell \\ L_{d_2} \otimes \mathbb{1}_{\ell_2} \\ \mathbb{1}_{\ell_1} \otimes L_{d_3} \\ L_{d_2} \otimes L_{d_3} \end{pmatrix} \text{ and}$$

$$Y^t \cdot Y = \begin{pmatrix} \ell & O & O & O \\ O & \ell_2 K_{d_2} & O & O \\ O & O & \ell_1 K_{d_3} & O \\ O & O & O & K_{d_2} \otimes K_{d_3} \end{pmatrix}.$$

Proof: Note that $L_d \cdot R_d = (d+1) \cdot I_d$ and $\mathbb{1}_d \cdot R_d = O$ and $R_d^t \cdot R_d = K_d$ for every integer d. Both equalities follow from simple calculations. □

III. Main Result

In this section, we present our main result, i.e., a new upper bound for $k(m,n)$, where $m = pq$ is the product of two distinct primes p and q. As noted in Section I, our arguments consist of a series of reductions. First of all, we reduce the problem of finding an upper bound for $k(m,n)$ to one of establishing an upper bound for $N_{1,1}$, the number of pairs $(\mathbf{u}_i, \mathbf{v}_i)$ of type $(1,1)$. The latter problem is in turn reduced to the study of the projective graph $\mathbf{G}_{n,m}$. More precisely, we follow the techniques of [15] and use the unique neighbor property of $\mathbf{G}_{n,m}$. However, the validity of the technique depends on some expanding property of $\mathbf{G}_{n,m}$.

A. Expanding Property

We follow the notations of Section II. In this section, we show that the projective graph $\mathbf{G}_{n,m}$ has some kind of expanding property (see Theorem 3.1), in the sense that $|N(X)|$ is large for certain choices of X. Expanding properties of the projective graph $\mathbf{G}_{n,p}$, where p is a prime, have been studied by Alon [1] using the well-known spectral method. In Section II, we made the observation that the graph $\mathbf{G}_{n,m}$ is a tensor product of the graphs $\mathbf{G}_{n,p}$ and $\mathbf{G}_{n,q}$. This observation enables us to obtain interesting properties (see Lemmas 2.2 and 2.3) which in turn facilitate our proof that $\mathbf{G}_{n,m}$ has some kind of expanding property.

Let $\mathbb{N}$ be the set of nonnegative integers and let $\mathbb{R}$ be the field of real numbers. For any vectors $\phi = (\phi_1, \ldots, \phi_\ell)^t, \psi = (\psi_1, \ldots, \psi_\ell)^t \in \mathbb{R}^\ell$, we let $\langle \phi, \psi \rangle = \sum_{i=1}^{\ell} \phi_i \cdot \psi_i$ and $\|\phi\|^2 = \langle \phi, \phi \rangle$. Furthermore, we define the weight of ϕ to be $\mathrm{wt}(\phi) = \sum_{i=1}^{\ell} \phi_i$. For a set $X \subseteq \mathbb{P}_{n,m}$, we denote by $\chi_X \in \mathbb{R}^\ell$ its characteristic vector whose components are labeled by the elements $\mathbf{u} \in \mathbb{P}_{n,m}$ and $\chi_X(\mathbf{u}) = 1$ if $\mathbf{u} \in X$ and 0 otherwise. Due to Lemmas 2.2 and 2.3, the column vectors of Y form a basis of the vector space $\mathbb{R}^\ell$. Therefore, there is a real vector

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}$$

where

$$\alpha_1 = \alpha_{11}, \alpha_2 = \begin{pmatrix} \alpha_{21} \\ \vdots \\ \alpha_{2d_2} \end{pmatrix}, \alpha_3 = \begin{pmatrix} \alpha_{31} \\ \vdots \\ \alpha_{3d_3} \end{pmatrix}, \alpha_4 = \begin{pmatrix} \alpha_{41} \\ \vdots \\ \alpha_{4d_4} \end{pmatrix}$$

such that χ_X can be written as a linear combination of the columns of Y, say

$$\chi_X = Y\alpha = \sum_{s=1}^{4} Y_s \alpha_s. \quad (12)$$

Let $\psi = A_{n,m}^t \chi_X$. The main idea of Alon's spectral method in [1] is to establish both a lower bound and an upper bound for the following number:

$$\begin{aligned}\|\psi\|^2 = \chi_X^t \cdot B_{n,m}\chi_X &= \sum_{r=1}^{4} \alpha_r^t Y_r^t \cdot \sum_{s=1}^{4} \lambda_s Y_s \alpha_s \\ &= \sum_{s=1}^{4} \lambda_s \|Y_s \alpha_s\|^2 \quad (13)\end{aligned}$$

where the second equality is due to Lemma 2.2, and the third equality follows from the second part of Lemma 2.3. For every $s \in \{1, 2, 3, 4\}$, we set

$$\Delta_s = \|Y_s \alpha_s\|^2. \quad (14)$$

Lemma 3.1: The quantities Δ_1, Δ_2, and Δ_3 can be written as

$$\begin{aligned}\Delta_1 = \ell\alpha_{11}^2, \quad \Delta_2 &= \ell_2(\|\alpha_2\|^2 + \mathrm{wt}(\alpha_2)^2) \quad \text{and} \\ \Delta_3 &= \ell_1(\|\alpha_3\|^2 + \mathrm{wt}(\alpha_3)^2). \quad (15)\end{aligned}$$

Proof: Lemma 2.3 shows that $Y_2^t Y_2 = \ell_2 K_{d_2}$. Then, we have

$$\begin{aligned}\Delta_2 = \|Y_2\alpha_2\|^2 = \alpha_2^t \cdot Y_2^t Y_2 \cdot \alpha_2 &= \alpha_2^t \cdot \ell_2 K_{d_2} \cdot \alpha_2 \\ &= \ell_2(\|\alpha_2\|^2 + \mathrm{wt}(\alpha_2)^2)\end{aligned}$$

which is the second equality. The first and third equalities can be proved similarly. □

Lemma 3.1 allows us to represent $\|\psi\|^2$ as an explicit function of α_1, α_2, and α_3. Let

$$\begin{aligned}S_1 &= \|\alpha_2\|^2 + \mathrm{wt}(\alpha_2)^2 \quad \text{and} \\ S_2 &= \|\alpha_3\|^2 + \mathrm{wt}(\alpha_3)^2. \quad (16)\end{aligned}$$

Lemma 3.2: We have that $\|\psi\|^2 = \lambda_4|X| + \ell(\lambda_1 - \lambda_4)\alpha_{11}^2 + \ell_2(\lambda_2 - \lambda_4)S_1 + \ell_1(\lambda_3 - \lambda_4)S_2$.

Proof: From Lemma 2.3, we have that $|X| = \|\chi_X\|^2 = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$. It follows that $\Delta_4 = |X| - \Delta_1 - \Delta_2 - \Delta_3$. Along with (13)–(16), this implies the expected equality. □

Although Lemma 3.2 gives us a representation of $\|\psi\|^2$ in terms of $|X|$, α_{11}, S_1, and S_2, it can be more explicit if we know how the quantities α_{11}, S_1 and S_2 are connected to X. Note that $\alpha = Y^{-1}\chi_X$ according to (12). Let $Z_1 = \ell^{-1} \cdot \mathbb{1}_\ell$, $Z_2 = \ell^{-1} \cdot L_{d_2} \otimes \mathbb{1}_{\ell_2}$, and $Z_3 = \ell^{-1} \cdot \mathbb{1}_{\ell_1} \otimes L_{d_3}$. Then

$$\alpha_s = Z_s \chi_X \quad (17)$$

for every $s \in \{1, 2, 3\}$, by Lemma 2.3. As an immediate consequence, we then have that

$$\alpha_{11} = \alpha_1 = Z_1\chi_X = \ell^{-1}|X|. \quad (18)$$

On the other hand, recall that $\mathbb{P}_{n,p}$, $\mathbb{P}_{n,q}$, and $\mathbb{P}_{n,m}$ have been made unique as ordered sets in Section II. For every $h \in [\ell]$, there exists $(i, j) \in [\ell_1] \times [\ell_2]$ such that $\mathbf{w}_h = \pi(\mathbf{u}_i, \mathbf{v}_j)$. Let $\sigma : \mathbb{P}_{n,m} \to [\ell_1]$ be the mapping defined by

$$\sigma(\mathbf{w}_h) = \left\lfloor \frac{h-1}{\ell_2} \right\rfloor + 1 \quad (19)$$

and let $\tau : \mathbb{P}_{n,m} \to [\ell_2]$ be the mapping defined by

$$\tau(\mathbf{w}_h) = h - (\sigma(\mathbf{w}_h) - 1)\ell_2. \quad (20)$$

Lemma 3.3: We have that $\mathbf{w}_h = \pi(\mathbf{u}_{\sigma(\mathbf{w}_h)}, \mathbf{v}_{\tau(\mathbf{w}_h)})$ for every $h \in [\ell]$.

Proof: Suppose that $\mathbf{w}_h = \pi(\mathbf{u}_i, \mathbf{v}_j)$ for $(i, j) \in [\ell_1] \times [\ell_2]$. Then, the representation of $\mathbb{P}_{n,m}$ in Section II shows that $h = (i-1)\ell_2 + j$. It is easy to see that $i = \sigma(\mathbf{w}_h)$ and $j = \tau(\mathbf{w}_h)$. □

For every $i \in [\ell_1]$ and $j \in [\ell_2]$, let $\sigma^{-1}(i)$ be the preimage of i under σ and let $\tau^{-1}(j)$ be the preimage of j under τ. Let $\mathbf{a} \in \mathbb{R}^{\ell_1}$ and $\mathbf{b} \in \mathbb{R}^{\ell_2}$ be two real vectors defined by

$$\mathbf{a}(i) = |\sigma^{-1}(i) \cap X| \quad \text{and} \quad \mathbf{b}(j) = |\tau^{-1}(j) \cap X| \quad (21)$$

where $i \in [\ell_1]$ and $j \in [\ell_2]$. Then, we clearly have that $\mathrm{wt}(\mathbf{a}) = \mathrm{wt}(\mathbf{b}) = |X|$.

Lemma 3.4: We have that $S_1 = \ell^{-2}\ell_1(\ell_1\|\mathbf{a}\|^2 - |X|^2)$ and $S_2 = \ell^{-2}\ell_2(\ell_2\|\mathbf{b}\|^2 - |X|^2)$.

Proof: For every $i \in [d_2]$, the ith row of Z_2 is

$$\begin{aligned}Z_2[i] &= \ell^{-1} \left(-\mathbb{1}_{i-1} \quad \ell_1 - 1 \quad -\mathbb{1}_{\ell_1 - i} \right) \otimes \mathbb{1}_{\ell_2} \\ &= \ell^{-1}\ell_1 \left(\mathbb{0}_{i-1} \quad 1 \quad \mathbb{0}_{\ell_1 - i} \right) \otimes \mathbb{1}_{\ell_2} - \ell^{-1}\mathbb{1}_\ell.\end{aligned}$$

Let $T = \ell^{-1}\ell_1 \left(\mathbb{0}_{i-1} \quad 1 \quad \mathbb{0}_{\ell_1 - i} \right) \otimes \mathbb{1}_{\ell_2}$ so that

$$Z_2[i] = T - \ell^{-1}\mathbb{1}_\ell.$$

The nonzero components of T are labeled by $\sigma^{-1}(i)$. It follows that $T \cdot \chi_X = \ell^{-1}\ell_1\mathbf{a}(i)$ and therefore

$$\alpha_{2i} = Z_2[i] \cdot \chi_X = T \cdot \chi_X - \ell^{-1}\mathbb{1}_\ell \cdot \chi_X = \ell^{-1}(\ell_1 \cdot \mathbf{a}(i) - |X|).$$

Note that $\mathrm{wt}(\mathbf{a}) = \mathbf{a}(1) + \cdots + \mathbf{a}(\ell_1) = |X|$ and $d_2 = \ell_1 - 1$. Due to (16), we have that

$$\begin{aligned}S_1 = \|\alpha_2\|^2 + \mathrm{wt}(\alpha_2)^2 &= \sum_{i=1}^{d_2} \alpha_{2i}^2 + \Big(\sum_{i=1}^{d_2} \alpha_{2i} \Big)^2 \\ &= \ell^{-2}\ell_1(\ell_1 \cdot \|\mathbf{a}\|^2 - |X|^2)\end{aligned}$$

which is the first equality.

For every $j \in [d_3]$, the jth row of Z_3 is

$$\begin{aligned}Z_3[j] &= \ell^{-1}\mathbb{1}_{\ell_1} \otimes \left(-\mathbb{1}_{j-1} \quad \ell_2 - 1 \quad -\mathbb{1}_{\ell_2 - j} \right) \\ &= \ell^{-1}\ell_2\mathbb{1}_{\ell_1} \otimes \left(\mathbb{0}_{j-1} \quad 1 \quad \mathbb{0}_{\ell_2 - j} \right) - \ell^{-1}\mathbb{1}_\ell.\end{aligned}$$

Let $T' = \ell^{-1}\ell_2\mathbb{1}_{\ell_1} \otimes \left(\mathbb{0}_{j-1} \quad 1 \quad \mathbb{0}_{\ell_2 - j} \right)$ so that

$$Z_3[j] = T' - \ell^{-1}\mathbb{1}_\ell.$$

The nonzero components of T' are labeled by $\tau^{-1}(j)$. It follows that $T' \cdot \chi_X = \ell^{-1}\ell_2 \mathbf{b}(j)$ and therefore

$$\alpha_{3j} = Z_3[j] \cdot \chi_X = T' \cdot \chi_X - \ell^{-1}\mathbb{1}_\ell \cdot \chi_X = \ell^{-1}(\ell_2 \mathbf{b}(j) - |X|).$$

Note that $\mathrm{wt}(\mathbf{b}) = \mathbf{b}(1) + \cdots + \mathbf{b}(\ell_2) = |X|$ and $d_3 = \ell_2 - 1$. Due to (16), we have that

$$\begin{aligned} S_2 = \|\alpha_3\|^2 + \mathrm{wt}(\alpha_3)^2 &= \sum_{i=1}^{d_3} \alpha_{3i}^2 + \Big(\sum_{i=1}^{d_3} \alpha_{3i}\Big)^2 \\ &= \ell^{-2}\ell_2(\ell_2 \cdot \|\mathbf{b}\|^2 - |X|^2) \end{aligned}$$

which is the second equality. □

Lemmas 3.2 and 3.4, together with (18), result in an explicit representation of $\|\psi\|^2$ in terms of X

$$\begin{aligned} \|\psi\|^2 = \lambda_4|X| &+ \ell^{-1}(\lambda_1 - \lambda_4)|X|^2 \\ &+ (\lambda_2 - \lambda_4)\ell^{-1}(\ell_1\|\mathbf{a}\|^2 - |X|^2) \\ &+ (\lambda_3 - \lambda_4)\ell^{-1}(\ell_2\|\mathbf{b}\|^2 - |X|^2). \quad (22) \end{aligned}$$

For simplicity, we denote by $F(\mathbf{a}, \mathbf{b})$ the right-hand side of (22). We aim to deduce an upper bound for $F(\mathbf{a}, \mathbf{b})$ in terms of $|X|$. Clearly, this also provides an upper bound for $\|\psi\|^2$ and is crucial for establishing that $\mathbf{G}_{n,m}$ has some kind of expanding property. Let

$$\kappa_p = \lfloor 4p^{0.5n} + 2 \rfloor \quad \text{and} \quad \kappa_q = \lfloor 4q^{0.5n} + 2 \rfloor. \qquad (23)$$

Dvir *et al.* [15] showed that $k(p, n) \le \kappa_p$ and $k(q, n) \le \kappa_q$. Let $\mathcal{U}, \mathcal{V} \subseteq \mathbb{P}_{n,m}$ form a matching family. From now on, we suppose that $X \subseteq \mathcal{U}$ and, furthermore, that its cardinality $|X| = x \le \min\{\kappa_q \ell_1, \kappa_p \ell_2\}$ is fixed. We remark that this assumption does not affect our proof adversely (see Theorem 3.3).

Lemma 3.5: Let $\mathbf{a}, \mathbf{b}$ be the real vectors defined by (21). Then, we have that $\mathbf{a}(i) \le \kappa_q$ for every $i \in [\ell_1]$ and $\mathbf{b}(j) \le \kappa_p$ for every $j \in [\ell_2]$.

Proof: Suppose that $\mathbf{a}(i) > \kappa_q$ for some $i \in [\ell_1]$. Let $\mathcal{U}' = \sigma^{-1}(i) \cap X \triangleq \{\mathbf{u}'_s : s \in [\mathbf{a}(i)]\} \subseteq \mathcal{U}$. Then, by the definition of matching families, there is a subset of $\mathcal{V}$, say $\mathcal{V}' = \{\mathbf{v}'_s : s \in [\mathbf{a}(i)]\}$, such that $\mathcal{U}'$ and $\mathcal{V}'$ form a matching family. It follows that

- $\langle \mathbf{u}'_s, \mathbf{v}'_s \rangle \equiv 0 \bmod m$ for every $s \in [\mathbf{a}(i)]$;
- $\langle \mathbf{u}'_s, \mathbf{v}'_t \rangle \not\equiv 0 \bmod m$ whenever $s, t \in [\mathbf{a}(i)]$ and $s \ne t$.

  On the one hand, we immediately have that
- $\langle \mathbf{u}'_s, \mathbf{v}'_s \rangle \equiv 0 \bmod q$ for every $s \in [\mathbf{a}(i)]$.

  On the other hand, Lemma 3.3 shows that any two elements in $\mathbb{P}_{n,m}$ are equivalent to each other as elements of $\mathbb{Z}_p^n$ as long as they have the same image under σ. Therefore, $\mathbf{u}'_s \sim \mathbf{u}'_t$ as elements of $\mathbb{Z}_p^n$ for any $s, t \in [\mathbf{a}(i)]$. It follows that $\langle \mathbf{u}'_s, \mathbf{v}'_t \rangle \equiv \langle \mathbf{u}'_t, \mathbf{v}'_t \rangle \equiv 0 \bmod p$. Recall that $\langle \mathbf{u}'_s, \mathbf{v}'_t \rangle \not\equiv 0 \bmod m$ whenever $s \ne t$. It follows that
- $\langle \mathbf{u}'_s, \mathbf{v}'_t \rangle \not\equiv 0 \bmod q$ whenever $s, t \in [\mathbf{a}(i)]$ and $s \ne t$.

Therefore, $\mathcal{U}'$ and $\mathcal{V}'$ form a matching family in $\mathbb{Z}_q^n$ of size $\mathbf{a}(i) > \kappa_q$, which contradicts Dvir *et al.* [15]. Hence, we must have that $\mathbf{a}(i) \le \kappa_q$ for every $i \in [\ell_1]$.

Similarly, we must have that $\mathbf{b}(j) \le \kappa_p$ for every $j \in [\ell_2]$. □

Lemma 3.5 shows that the components of $\mathbf{a}$ and $\mathbf{b}$ cannot be too large when $X \subseteq \mathcal{U}$. In fact, several conditions must be satisfied by the real vectors $\mathbf{a}$ and $\mathbf{b}$, which can be summarized as follows:

while $\mathbf{c} \ne \mathbf{a}^*$ do

- set $i_0 = \min\{i \in [\ell_1] : \mathbf{c}(i) \ne \mathbf{a}^*(i)\}$;
- set $a = \begin{cases} \min\{\kappa_q, \mathbf{c}(i_0) + \mathbf{c}(i_0+1)\}, \text{ if } i_0 \le \mu_q, \\ \min\{\nu_q, \mathbf{c}(i_0) + \mathbf{c}(i_0+1)\}, \text{ if } i_0 = \mu_q + 1; \end{cases}$
- set $b = \mathbf{c}(i_0) + \mathbf{c}(i_0+1) - a$;
- set $j_0 = \min\left(\{j : j \in \{i_0+2, \ldots, \ell_1\} \wedge \mathbf{c}(j) \le b\} \cup \{0\}\right)$;
  - if $j_0 = i_0 + 2$, set $\mathbf{c}' = (\mathbf{c}(1), \ldots, \mathbf{c}(i_0-1), a, b, \mathbf{c}(i_0+2), \ldots, \mathbf{c}(\ell_1))$;
  - if $j_0 = 0$ or ℓ_1, set $\mathbf{c}' = (\mathbf{c}(1), \ldots, \mathbf{c}(i_0-1), a, \mathbf{c}(i_0+2), \ldots, \mathbf{c}(\ell_1), b)$;
  - otherwise, set $\mathbf{c}' = (\mathbf{c}(1), \ldots, \mathbf{c}(i_0-1), a, \mathbf{c}(i_0+2), \ldots, \mathbf{c}(j_0-1), b, \mathbf{c}(j_0), \ldots, \mathbf{c}(\ell_1))$;
- set $\mathbf{c} = \mathbf{c}'$.

Fig. 2. Algorithm.

- $0 \le \mathbf{a}(i) \le \kappa_q$ for every $i \in [\ell_1]$, due to Lemma 3.5 and (21);
- $0 \le \mathbf{b}(j) \le \kappa_p$ for every $j \in [\ell_2]$, due to Lemma 3.5 and (21);
- $\mathrm{wt}(\mathbf{a}) = \mathrm{wt}(\mathbf{b}) = |X| = x$, due to (21).

Clearly, when x is fixed, the problem of establishing an upper bound for $F(\mathbf{a}, \mathbf{b})$ can be reduced to that of determining the maximum value of $F(\mathbf{a}, \mathbf{b})$ subject to the conditions above. Let

$$\begin{aligned} \mu_q &= \left\lfloor \frac{x}{\kappa_q} \right\rfloor, \quad \nu_q = x - \kappa_q \mu_q \\ \mathbf{a}^* &= (\ \kappa_q \cdot \mathbb{1}_{\mu_q} \quad \nu_q \quad \mathbb{0}_{\ell_1 - 1 - \mu_q}\ ) \\ \mu_p &= \left\lfloor \frac{x}{\kappa_p} \right\rfloor, \quad \nu_p = x - \kappa_p \mu_p \\ \mathbf{b}^* &= (\ \kappa_p \cdot \mathbb{1}_{\mu_p} \quad \nu_p \quad \mathbb{0}_{\ell_2 - 1 - \mu_p}\ ). \qquad (24) \end{aligned}$$

We show below that $F(\mathbf{a}^*, \mathbf{b}^*)$ is the maximum value of $F(\mathbf{a}, \mathbf{b})$ subject to the conditions above.

Lemma 3.6: Let $a, b, c, d \in \mathbb{N}$ be such that $a \ge b$, $c \ge d$, and $a + b = c + d$. If $a \ge c$, then $a^2 + b^2 \ge c^2 + d^2$.

Proof: Clearly, we have that $a^2 + b^2 - c^2 - d^2 = (a-c)(a+c) + (b-d)(b+d) = (a-c)(a+c) - (a-c)(b+d) = (a-c)(a+c-b-d) \ge 0$, where the second equality follows from $a + b = c + d$ and the last inequality follows from $a \ge b$, $c \ge d$ and $a \ge c$. □

Lemma 3.7: We have that $\|\psi\|^2 = F(\mathbf{a}, \mathbf{b}) \le F(\mathbf{a}^*, \mathbf{b}^*)$.

Proof: First, we note that the vectors $\mathbf{a}^*$ and $\mathbf{b}^*$ satisfy the three conditions above. In order to show that $F(\mathbf{a}, \mathbf{b}) \le F(\mathbf{a}^*, \mathbf{b}^*)$, in view of (22), it suffices to show that $\|\mathbf{a}\|^2 \le \|\mathbf{a}^*\|^2$ and $\|\mathbf{b}\|^2 \le \|\mathbf{b}^*\|^2$. We only show the first inequality. The second one can be proved similarly.

Without loss of generality, we can suppose that $\mathbf{a}(1) \ge \mathbf{a}(2) \ge \cdots \ge \mathbf{a}(\ell_1)$. From Lemma 3.5, we have that $\mathbf{a}(i) \le \kappa_q$ for every $i \in [\ell_1]$. We claim that the algorithm in Fig. 2 takes as input the original vector $\mathbf{a}_0 = \mathbf{a}$ and produces a sequence of vectors, say $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_h$, such that

- $\kappa_q \ge \mathbf{a}_s(1) \ge \mathbf{a}_s(2) \ge \cdots \ge \mathbf{a}_s(\ell_1) \ge 0$ and $\mathrm{wt}(\mathbf{a}_s) = x$ for every $s \in \{0, 1, \ldots, h\}$; and
- $\mathbf{a}_0 = \mathbf{a}$, $\mathbf{a}_h = \mathbf{a}^*$ and $\|\mathbf{a}_s\|^2 \le \|\mathbf{a}_{s+1}\|^2$ for every $s \in \{0, 1, \ldots, h-1\}$.

Clearly, if the algorithm does have the above functionality, then we must have that $\|\mathbf{a}\|^2 \le \|\mathbf{a}^*\|^2$.

Consider the algorithm in Fig. 2. In order to get the expected sequence, i.e., $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_h$, it will be run with an initial input $\mathbf{c} = \mathbf{a}_0 = \mathbf{a}$. In every iteration, the algorithm outputs a $\mathbf{c}'$ and then checks whether $\mathbf{c}' = \mathbf{a}^*$. It halts once the equality holds.

We must show that this algorithm does halt in a finite number of steps and achieves the promised functionality. The algorithm starts with $\mathbf{c}$ and checks whether $\mathbf{c} = \mathbf{a}^*$. If the equality holds, it halts. Otherwise, it constructs a new vector $\mathbf{c}'$ such that $\mathrm{wt}(\mathbf{c}') = x$, $\kappa_q \geq \mathbf{c}'(1) \geq \mathbf{c}'(2) \geq \cdots \geq \mathbf{c}'(\ell_1) \geq 0$ and $\|\mathbf{c}\|^2 \leq \|\mathbf{c}'\|^2$. More concretely, the algorithm finds the first coordinate (say $i_0 \in [\ell_1]$) where $\mathbf{c}$ and $\mathbf{a}^*$ differ. Clearly, we have that $i_0 \leq \mu_q + 1$, $\mathbf{c}(i) = \mathbf{a}^*(i)$ for every $i \in \{1, 2, \ldots, i_0 - 1\}$ and $\mathbf{c}(i_0) < \mathbf{a}^*(i_0)$. Next, the algorithm affects a carry from $\mathbf{c}(i_0 + 1)$ to $\mathbf{c}(i_0)$. This is done by setting $\mathbf{c}'(i_0) = a$. Finally, the algorithm must determine $\mathbf{c}'(i)$ for every $i \in \{i_0 + 1, \ldots, \ell_1\}$. This is done by rearranging the $\ell_1 - i_0$ numbers $b, \mathbf{c}(i_0 + 2), \ldots, \mathbf{c}(\ell_1)$ such that they are in descending order. By the description above, it is clear that

- $\mathrm{wt}(\mathbf{c}') = \sum_{i=1}^{i_0-1} \mathbf{c}(i) + a + b + \sum_{i=i_0+2}^{\ell_1} \mathbf{c}(i) = \mathrm{wt}(\mathbf{c}) = x$;
- $\mathbf{c}'(1) = \mathbf{c}'(2) = \cdots = \mathbf{c}'(i_0 - 1) = \kappa_q \geq \mathbf{c}'(i_0) = a > \mathbf{c}(i_0) \geq \mathbf{c}'(i_0 + 1) \geq \cdots \geq \mathbf{c}'(\ell_1)$;
- $0 \leq \mathbf{c}'(i) \leq \kappa_q$ for every $i \in [\ell_1]$;
- $\|\mathbf{c}'\|^2 - \|\mathbf{c}\|^2 = a^2 + b^2 - \mathbf{c}(i_0)^2 - \mathbf{c}(i_0 + 1)^2 \geq 0$, due to Lemma 3.6.

In each iteration, either i_0 becomes greater than it was in the previous iteration or i_0 does not change but the new $\mathbf{c}'(i_0)$ obtained is strictly greater than $\mathbf{c}(i_0)$. However, since $\mathbf{c}'(i_0)$ must be bounded by κ_q, in the latter case, $\mathbf{c}'(i_0)$ will eventually become $\mathbf{a}^*(i_0)$ in a finite number of iterations. Then, in the next iteration, i_0 will be increased by at least 1. Therefore, we can get a sequence $\mathbf{a}_0 = \mathbf{a}, \mathbf{a}_1, \ldots, \mathbf{a}_h = \mathbf{a}^*$, where h is the number of iterations. □

Lemma 3.7 shows that $F(\mathbf{a}^*, \mathbf{b}^*)$ is a valid upper bound for $\|\psi\|^2$. This bound is nice because both $\mathbf{a}^*$ and $\mathbf{b}^*$ depend only on x, which will facilitate our analysis. More precisely, we have that

$$\|\psi\|^2 \leq \ell^{-1}\lambda_1 x^2 + \Delta \tag{25}$$

where $\Delta = \lambda_4 x - \ell^{-1}\lambda_4 x^2 + (\lambda_2 - \lambda_4)\ell^{-1}(\ell_1\|\mathbf{a}^*\|^2 - x^2) + (\lambda_3 - \lambda_4)\ell^{-1}(\ell_2\|\mathbf{b}^*\|^2 - x^2)$.

We proceed to develop an explicit lower bound for $\|\psi\|^2$ in terms of x and $|N(X)|$. Recall that the components of ψ are labeled by all the hyperplanes. It is easy to see that

$$\psi(\mathbf{v}) = |N(\mathbf{v}) \cap X| \tag{26}$$

is the number of neighbors of $\mathbf{v}$ in X for every $\mathbf{v} \in \mathbb{H}_{n,m}$. Hence, $\psi(\mathbf{v}) = 0$ whenever $\mathbf{v} \notin N(X)$. It follows that

$$\begin{aligned}\sum_{\mathbf{v}\in\mathbb{H}_{n,m}} \psi(\mathbf{v}) &= \sum_{\mathbf{v}\in N(X)} \psi(\mathbf{v}) \\ &= \sum_{\mathbf{u}\in X} |N(\mathbf{u})| = x \cdot \theta_{n-1,m} \end{aligned}\tag{27}$$

where the last equality follows from Chee and Ling [11]. It follows from the Cauchy–Schwarz inequality that

$$\begin{aligned}\|\psi\|^2 &= \sum_{\mathbf{v}\in N(X)} \psi(\mathbf{v})^2 \geq \frac{1}{|N(X)|}\left(\sum_{\mathbf{v}\in N(X)} \psi(\mathbf{v})\right)^2 \\ &= \frac{x^2\theta_{n-1,m}^2}{|N(X)|} = \frac{\lambda_1 x^2}{|N(X)|}\end{aligned}\tag{28}$$

where the second equality follows from (27) and the last equality follows from Lemma 2.2.

Both the upper bound (see (25)) and the lower bound (see (28)) for $\|\psi\|^2$ involve only x and $|N(X)|$. Together, they demonstrate that the projective graph $\mathbf{G}_{n,m}$ has some kind of expanding property.

Theorem 3.1 (Expanding property): Let $\mathcal{U}, \mathcal{V} \subseteq \mathbb{P}_{n,m}$ form a matching family and let $X \subseteq \mathcal{U}$ be of cardinality $x \leq \min\{\kappa_q\ell_1, \kappa_p\ell_2\}$. Then, we have that $|N(X)| \geq \lambda_1 x^2/(\ell^{-1}\lambda_1 x^2 + \Delta)$.

B. On the Largest Matching Family in $\mathbb{P}_{n,m}$

In this section, we deduce an upper bound on the largest matching family in $\mathbb{P}_{n,m}$. As in [15], our analysis depends on the unique neighbor property defined in Definition 3.1 and the fact that the projective graph $\mathbf{G}_{n,m}$ has some kind of expanding property (see Theorem 3.1).

Definition 3.1 (Unique neighbor property): We say that $\mathcal{U} \subseteq \mathbb{P}_{n,m}$ satisfies the unique neighbor property if, for every $\mathbf{u} \in \mathcal{U}$, there is a $\mathbf{v} \in N(\mathbf{u})$ such that $\mathbf{v}$ is not adjacent to any $\mathbf{w} \in \mathcal{U} \setminus \{\mathbf{u}\}$.

As noted by Dvir *et al.* [15], there is a set $\mathcal{U} \subseteq \mathbb{P}_{n,p}$ of cardinality k that satisfies the unique neighbor property in $\mathbf{G}_{n,p}$ if and only if there is a matching family in $\mathbb{Z}_p^n$ of size k. As an analog, the following lemma is true for $\mathbf{G}_{n,m}$.

Lemma 3.8: A set $\mathcal{U} \subseteq \mathbb{P}_{n,m}$ satisfies the unique neighbor property if and only if there is a $\mathcal{V} \subseteq \mathbb{H}_{n,m}$ such that $\mathcal{U}$ and $\mathcal{V}$ form a matching family.

Proof: Suppose that $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. If it satisfies the unique neighbor property in $\mathbf{G}_{n,m}$, then, for every $i \in [k]$, there is a $\mathbf{v}_i \in N(\mathbf{u}_i)$ such that $\mathbf{v}_i \notin N(\mathbf{u}_j)$ for every $j \in [k] \setminus \{i\}$. Equivalently, we have that $\langle \mathbf{u}_i, \mathbf{v}_i\rangle = 0$ and $\langle \mathbf{u}_j, \mathbf{v}_i\rangle \neq 0$ for every $j \in [k] \setminus \{i\}$. Let $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Then, $\mathcal{U}$ and $\mathcal{V}$ form a matching family.

Conversely, let $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \mathbb{H}_{n,m}$ be such that $\mathcal{U}$ and $\mathcal{V}$ form a matching family. For every $i \in [k]$, we have that $\langle \mathbf{u}_i, \mathbf{v}_i\rangle = 0$ and $\langle \mathbf{u}_j, \mathbf{v}_i\rangle \neq 0$ whenever $j \in [k]$ and $j \neq i$. Equivalently, $\mathbf{v}_i \in N(\mathbf{u}_i)$ and $\mathbf{v}_i \notin N(\mathbf{u}_j)$ when $j \in [k] \setminus \{i\}$. Hence, $\mathcal{U}$ satisfies the unique neighbor property. □

Theorem 3.2: Let $\mathcal{U}, \mathcal{V} \subseteq \mathbb{P}_{n,m}$ form a matching family and let $X \subseteq \mathcal{U}$ be of cardinality x. Then, we have that $|\mathcal{U}| \leq x + \ell\Delta/(\ell^{-1}\lambda_1 x^2 + \Delta)$.

Proof: By Lemma 3.8, $\mathcal{U}$ satisfies the unique neighbor property in $\mathbf{G}_{n,m}$. Hence, every element in $\mathcal{U} \setminus X$ must have a unique neighbor in $\mathbb{H}_{n,m} \setminus N(X)$. It follows that $|\mathcal{U} \setminus X| \leq |\mathbb{H}_{n,m} \setminus N(X)| = \ell - |N(X)|$, which implies that $|\mathcal{U}| \leq |X| + \ell - |N(X)|$. Along with Theorem 3.1, the inequality desired follows. □

The following theorem gives an explicit upper bound for the largest matching family in $\mathbb{P}_{n,m}$.

Theorem 3.3: Let $\mathcal{U}, \mathcal{V} \subseteq \mathbb{P}_{n,m}$ form a matching family. Then, $|\mathcal{U}| \leq (8 + \epsilon)m^{0.625n+0.125}$ for any constant $\epsilon > 0$ as $p \to \infty$ and $p/q \to 1$, where n is a constant.

Proof: Suppose that $|\mathcal{U}| > (8+\epsilon)m^{0.625n+0.125}$. Then, we can take a set of points $X \subseteq \mathcal{U}$ of cardinality $x = \lfloor \ell^{0.625} \rfloor \leq \min\{\kappa_q\ell_1, \kappa_p\ell_2\}$. By Theorem 3.2, we have that $|\mathcal{U}| \leq x + \ell\Delta/(\ell^{-1}\lambda_1 x^2 + \Delta) \approx 8\ m^{0.625n+0.125}$ when $p \to \infty$ and $p/q \to 1$, with n a constant. This is a contradiction. □

C. On the Largest Matching Family in $\mathbb{Z}_m^n$

Let $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}, \mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a matching family of size $k = k(m, n)$ in $\mathbb{Z}_m^n$. In order to establish the final upper bound for $k(m,n)$, we have to classify the pairs $\{(\mathbf{u}_i,\mathbf{v}_i) : i \in [k]\}$ into types and establish an upper bound for each type.

Definition 3.2 (Type of pairs): For every $i \in [k]$, the pair $(\mathbf{u}_i,\mathbf{v}_i)$ is said to be of type (s,t) if $\gcd(\mathbf{u}_i(1),\ldots,\mathbf{u}_i(n),m) = s$ and $\gcd(\mathbf{v}_i(1),\ldots,\mathbf{v}_i(n),m) = t$, where s,t are positive divisors of m.

Let $s,t \in \{1,p,q,m\}$. We define $\Omega_{s,t}$ to be the set of pairs $(\mathbf{u}_i,\mathbf{v}_i)$ of type (s,t) and $N_{s,t} = |\Omega_{s,t}|$. Clearly, the elements of the set $\{(\mathbf{u}_i,\mathbf{v}_i) : i \in [k]\}$ fall into 16 different classes as s and t vary.

Lemma 3.9: If $m|st$, then $N_{s,t} \le 1$.

Proof: Suppose that $N_{s,t} > 1$. Then, we can take two pairs, say $(\mathbf{u}_1,\mathbf{v}_1)$ and $(\mathbf{u}_2,\mathbf{v}_2)$, from $\Omega_{s,t}$. Clearly, we have that $\langle \mathbf{u}_1,\mathbf{v}_2\rangle = \langle \mathbf{u}_2,\mathbf{v}_1\rangle = 0$, which is a contradiction. □

Lemma 3.9 covers nine of the 16 classes. More precisely, we have that $N_{s,t} \le 1$ when $(s,t) \in \{(p,q),(q,p),(m,1),(m,p),(m,q),(m,m),(1,m),(p,m),(q,m)\}$.

Lemma 3.10: If $(s,t) \in \{(1,p),(p,1),(p,p)\}$, then $N_{s,t} \le \kappa_q$.

Proof: We prove for $(s,t) = (1,p)$. The other cases can be treated similarly. Without loss of generality, we can suppose that $\{(\mathbf{u}_1,\mathbf{v}_1),\ldots,(\mathbf{u}_c,\mathbf{v}_c)\}$ are the pairs of type (s,t), where $c = N_{1,p}$. Let

- $\mathbf{u}'_i = (\mathbf{u}_i(1) \bmod q,\ldots,\mathbf{u}_i(n) \bmod q)$ and
- $\mathbf{v}'_i = (\mathbf{v}_i(1)/p \bmod q,\ldots,\mathbf{v}_i(n)/p \bmod q)$

for every $i \in [c]$. Then, $\mathcal{U}' = \{\mathbf{u}'_1,\ldots,\mathbf{u}'_c\}$ and $\mathcal{V}' = \{\mathbf{v}'_1,\ldots,\mathbf{v}'_c\}$ form a matching family of size c in $\mathbb{Z}_q^n$. This implies that $N_{s,t} = c \le \kappa_q$, using results of Dvir *et al.* in [15]. □

Similarly, we have

Lemma 3.11: If $(s,t) \in \{(1,q),(q,1),(q,q)\}$, then $N_{s,t} \le \kappa_p$.

Finally, we have the main result of this paper.

Theorem 3.4: Let n be a constant and let $m = pq$ for two distinct primes p and q. Then, we have that $k(m,n) \le O(m^{0.625n+0.125})$ when $p \to \infty$ and $p/q \to 1$.

Proof: Theorem 3.3 gives an upper bound for $N_{1,1}$. By Theorem 3.3 and Lemmas 3.9–3.11, $k(m,n) = k = \sum_{s|m,t|m} N_{s,t} \le 9+3\kappa_p + 3\kappa_q + O\big(m^{0.625n+0.125}\big)$, which is asymptotically bounded by $O\big(m^{0.625n+0.125}\big)$ when $p \to \infty$ and $p/q \to 1$. □

IV. Concluding Remarks

It would be attractive if the method in this paper could be extended to work for any integer m. To this end, we would need to show that the projective graph $\mathbf{G}_{n,m}$, for a general integer m, has some kind of expanding property. In fact, a generalized version of the tensor lemma (see Lemma 4.1) does exist for the matrix $B_{n,m}$, and it is also possible to determine the eigenvalues of $B_{n,m}$ for a general integer m (see Theorems 4.1 and 4.2).

Lemma 4.1 (Tensor lemma): Let $m = m_1\cdots m_r = p_1^{e_1}\cdots p_r^{e_r}$ for distinct primes $p_1,\ldots,p_r$ and positive integers $e_1,\ldots,e_r$, where $m_s = p_s^{e_s}$ for every $s \in [r]$. Then, we have that

$$B_{n,m} \simeq B_{n,m_1} \otimes \cdots \otimes B_{n,m_r}. \tag{29}$$

Theorem 4.1 (Eigenvalues of $B_{n,m}$ when m is a prime power): Let $m = p^e$, where p is a prime and e is a positive integer, and let n be a positive integer. Then, $\lambda_1 = p^{2(e-1)(n-2)} \cdot \theta_{n-1,p}^2$ is an eigenvalue of $B_{n,m}$ of multiplicity $d_1 = 1$, $\lambda_2 = p^{(2e-1)(n-2)}$ is an eigenvalue of $B_{n,m}$ of multiplicity $d_2 = \theta_{n,p} - 1$, and $\lambda_s = p^{(2e+1-s)(n-2)}$ is an eigenvalue of $B_{n,m}$ of multiplicity $d_s = (p^{n-1}-1)\theta_{n,p^{s-2}}$ for every $s \in \{3,\ldots,e+1\}$.

Theorem 4.2 (Eigenvalues of $B_{n,m}$ when m is any positive integer): Let $m = m_1\cdots m_r = p_1^{e_1}\cdots p_r^{e_r}$ for distinct primes $p_1,\ldots,p_r$ and positive integers $e_1,\ldots,e_r$, where $m_s = p_s^{e_s}$ for every $s \in [r]$. Let λ_s be an eigenvalue of B_{n,m_s} of multiplicity d_s for every $s \in [r]$. Then, $\lambda_1\cdots\lambda_r$ is an eigenvalue of $B_{n,m}$ of multiplicity $d_1\cdots d_r$.

However, the method used in this paper may become weaker as the number of distinct prime factors of m increases. As in many other classic applications, the performance of our method depends on the difference between the two largest eigenvalues of $B_{n,m}$. Roughly speaking, a larger difference corresponds to better performance. However, Theorems 4.1 and 4.2 show that this difference becomes less significant as the number of distinct prime factors of m increases. On the other hand, this does not rule out the possibility of applying Theorems 4.1 and 4.2 in a different way. Currently, an exponential gap still exists between the best lower bound and the best upper bound for $k(m,n)$. We hope that these general theorems (Theorems 4.1 and 4.2) can be used in some way to close this gap in the future.

Acknowledgment

The authors thank the anonymous referees for very detailed and helpful comments that significantly improved the organization and exposition of the paper.

References

[1] N. Alon, "Eigenvalues, geometric expanders, sorting in rounds, and Ramsey theory," *Combinatorica*, vol. 6, no. 3, pp. 207–219, 1986.

[2] L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy, "Checking computations in polylogarithmic time," in *Proc. 33rd ACM Symp. Theory Comput.*, 2001, pp. 21–31.

[3] L. Babai and P. Frankl, Linear algebra methods in combinatorics Univ. Chicago. Chicago, IL, USA, 1998.

[4] N. Alon, L. Babai, P. Frankl, S. Kutin, and D. Stefankovic, "Set systems with restricted intersections modulo prime powers," *J. Combin. Theory, A*, vol. 95, no. 1, pp. 39–73, 2001.

[5] D. Beaver and J. Feigenbaum, "Hiding instances in multioracle queries," in *Proc. 7th Annu. Symp. Theor. Aspects Comput. Sci.*, 1990, vol. 415, pp. 37–48.

[6] A. Beimel, Y. Ishai, and E. Kushilevitz, "General constructions for information-theoretic private information retrieval," *J. Comput. Syst. Sci.*, vol. 71, no. 2, pp. 213–247, 2005.

[7] A. Beimel, Y. Ishai, E. Kushilevitz, and J. Raymond, "Breaking the $O(n^{1/(2k-1)})$ barrier for information-theoretic private information retrieval," in *Proc. 43rd Annu. IEEE Symp. Found. Comput. Sci.*, 2002, pp. 261–270.

[8] A. Ben-Aroya, K. Efremenko, and A. Ta-Shma, “A note on amplifying the error-tolerance of locally decodable codes,” *Electron. Colloq. Comput. Complex.*, vol. 17, 2010.
[9] A. Bhowmick, Z. Dvir, and S. Lovett, “New lower bounds for matching vector codes,” *Electron. Colloq. Comput. Complex.*, vol. 19, 2012.
[10] Y. M. Chee, T. Feng, S. Ling, H. Wang, and L. F. Zhang, “Query-efficient locally decodable codes of subexponential length,” *Comput. Complex.*, vol. 22, no. 1, pp. 159–189, 2013.
[11] Y. M. Chee and S. Ling, “Projective covering designs,” *Bull. Lond. Math. Soc.*, vol. 25, no. 3, pp. 231–239, 1993.
[12] B. Chor, E. Kushilevitz, and O. G. Sudan, “Private information retrieval,” *J. ACM*, vol. 45, no. 6, pp. 965–981, 1998.
[13] A. Deshpande, R. Jain, T. Kavitha, J. Radhakrishnan, and S. V. Lokam, “Better lower bounds for locally decodable codes,” in *Proc. 17th Annu. IEEE Conf. Comput. Complex.*, Jun. 2002, pp. 184–193.
[14] M. Deza, P. Frankl, and N. M. Singh, “On functions of strength t,” *Combinatotica*, vol. 3, no. 3–4, pp. 331–339, 1983.
[15] Z. Dvir, P. Gopalan, and S. Yekhanin, “Matching vector codes,” in *Proc. 51st Annu. IEEE Symp. Found. Comput. Sci.*, Oct. 2010, pp. 705–714.
[16] Z. Dvir and A. Shpilka, “Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits,” in *Proc. 37th ACM Symp. Theory Comput.*, May 2005, pp. 592–601.
[17] K. Efremenko, “3-query locally decodable codes of subexponential length,” in *Proc. 41st ACM Symp. Theory Comput.*, May 2009, pp. 39–44.
[18] P. Erdos, C. Ko, and R. Rado, “Intersection theorems for systems of finite sets,” *Quart. J. Math.*, vol. 12, pp. 313–320, 1961.
[19] O. Goldreich, H. J. Karloff, L. J. Schulman, and L. Trevisan, “Lower bounds for linear locally decodable codes and private information retrieval,” in *Proc. 17th Annu. IEEE Conf. Comput. Complex.*, Jun. 2002, pp. 175–183.
[20] P. Gopalan, “A note on Efremenko’s locally decodable codes,” *Electron. Colloq. Comput. Complex.*, vol. 16, 2009.
[21] V. Grolmusz, “Superpolynomial size set-systems with restricted intersections mod 6 and explicit Ramsey graphs,” *Combinatorica*, vol. 20, no. 1, pp. 71–86, 2000.
[22] T. Itoh and Y. Suzuki, “New constructions for query-efficient locally decodable codes of subexponential length,” in 2008 [Online]. Available: abs/0810.4576
[23] T. Itoh and Y. Suzuki, “Improved constructions for query-efficient locally decodable codes of subexponential length,” *IEICE Trans.*, vol. 93-D, no. 2, pp. 263–270, 2010.
[24] J. Katz and L. Trevisan, “On the efficiency of local decoding procedures for error-correcting codes,” in *Proc. 32nd ACM Symp. Theory Comput.*, May 2000, pp. 80–86.
[25] K. S. Kedlaya and S. Yekhanin, “Locally decodable codes from nice subsets of finite fields and prime factors of Mersenne numbers,” *SIAM J. Comput.*, vol. 38, no. 5, pp. 1952–1969, 2009.
[26] I. Kerenidis and R. de Wolf, “Exponential lower bound for 2-query locally decodable codes via a quantum argument,” in *Proc. 35th ACM Symp. Theory Comput.*, Jul. 2003, pp. 106–115.
[27] S. Kutin, “Constructing large set systems with given intersection sizes modulo composite numbers,” *Combinat., Probab. Comput.*, vol. 11, no. 5, pp. 475–486, 2002.
[28] R. J. Lipton, “Efficient checking of computations,” in *Proc. 7th Annu. Symp. Theor. Aspects Comput. Sci.*, Feb. 1990, vol. 415, pp. 207–215.
[29] K. Obata, “Optimal lower bounds for 2-query locally decodable linear codes,” in *Proc. 6th Int. Workshop Random. Approx. Tech.*, Sep. 2002, vol. 2483, pp. 39–50.
[30] A. Polishchuk and D. A. Spielman, “Nearly-linear size holographic proofs,” in *Proc. 26th ACM Symp. Theory Comput.*, May 1994, pp. 194–203.
[31] P. Raghavendra, “A note on Yekhanin’s locally decodable codes,” *Electron. Colloq. Comput. Complex.*, vol. 14, 2007.
[32] J. Sgall, “Bounds on pairs of families with restricted intersections,” *Combinatorica*, vol. 19, no. 4, pp. 555–566, 1999.
[33] D. Shiowattana and S. V. Lokam, “An optimal lower bound for 2-query locally decodable linear codes,” *Inf. Process. Lett.*, vol. 97, no. 6, pp. 244–250, 2006.
[34] M. Sudan, *Efficient Checking of Polynomials and Proofs and the Hardness of Approximation Problems*, ser. Lecture Notes in Computer Science. New York, NY, USA: Springer-Verlag, 1995, vol. 1001.
[35] M. Sudan, L. Trevisan, and S. P. Vadhan, “Pseudorandom generators without the XOR lemma,” in *Proc. 31st ACM Symp. Theory Comput.*, May 1999, pp. 537–546.
[36] S. Wehner and R. de Wolf, “Improved lower bounds for locally decodable codes and private information retrieval,” in *Proc. 32nd Int. Colloq. Autom., Lang. Program.*, Jul. 2005, vol. 3580, pp. 1424–1436.
[37] D. P. Woodruff, “New lower bounds for general locally decodable codes,” *Electron. Colloq. Comput. Complex.*, vol. 14, 2007.
[38] D. P. Woodruff and S. Yekhanin, “A geometric approach to information-theoretic private information retrieval,” in *Proc. 20th Annu. IEEE Conf. Comput. Complex.*, Jun. 2005, pp. 275–284.
[39] S. Yekhanin, “Towards 3-query locally decodable codes of subexponential length,” in *Proc. 39th ACM Symp. Theory Comput.*, Jun. 2007, pp. 266–274.
[40] S. Yekhanin, “Locally decodable codes,” *Found. Trends Theor. Comput. Sci.*, vol. 6, no. 3, pp. 139–255, 2012.

Yeow Meng Chee (SM’08) received the B.Math. degree in computer science and combinatorics and optimization and the M.Math. and Ph.D. degrees in computer science, from the University of Waterloo, Waterloo, ON, Canada, in 1988, 1989, and 1996, respectively.

Currently, he is an Associate Professor at the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. Prior to this, he was Program Director of Interactive Digital Media R&D in the Media Development Authority of Singapore, Postdoctoral Fellow at the University of Waterloo and IBM’s Zürich Research Laboratory, General Manager of the Singapore Computer Emergency Response Team, and Deputy Director of Strategic Programs at the Infocomm Development Authority, Singapore. His research interest lies in the interplay between combinatorics and computer science/engineering, particularly combinatorial design theory, coding theory, extremal set systems, and electronic design automation.

San Ling received the B.A. degree in mathematics from the University of Cambridge, Cambridge, U.K., in 1985, and the Ph.D. degree in mathematics from the University of California, Berkeley, in 1990. Since April 2005, he has been a Professor with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. Prior to that, he was with the Department of Mathematics, National University of Singapore. His research fields include arithmetic of modular curves and application of number theory to combinatorial designs, coding theory, cryptography, and sequences.

Huaxiong Wang obtained a Ph.D. in Mathematics from the University of Haifa, Israel (1996) and a Ph.D. in Computer Science from the University of Wollongong, Australia (2001). He is currently with Nanyang Technological University, Singapore. His research interests include cryptography, information security, coding theory, combinatorics and theoretical computer science. He is on the editorial boards of Designs, Codes and Cryptography, Journal of Communications and Journal of Communications and Networks and was the Program Co-Chair of 9th Australasian Conference on Information Security and Privacy (ACISP’04), Sydney, Australia, July, 2004 and 4th International Conference on Cryptology and Network Security (CANS05), Xiamen, Fujian, China, December, 2005. He won the inaugural Prize of Research Contribution awarded by the Computer Science Association of Australasia in 2004.

Liang Feng Zhang received the B.S. degree in applied mathematics from the Tongji University, Shanghai, China, in 2004, the M.S. degree in applied mathematics from the Shanghai Jiao Tong University, Shanghai, China, in 2007, and the Ph.D. degree in cryptography from the Nanyang Technological University, Singapore, in 2012. He is currently a postdoctoral fellow at the University of Calgary, Canada. His research interests include coding theory, cryptography, and discrete mathematics.