On the orthogonal rank of Cayley graphs and impossibility of quantum round elimination

Jop Briët* Jeroen Zuiddam†

August 22, 2016

Abstract

After Bob sends Alice a bit, she responds with a lengthy reply. At the cost of a factor of two in the total communication, Alice could just as well have given the two possible replies without listening and have Bob select which applies to him. Motivated by a conjecture stating that this form of "round elimination" is impossible in exact quantum communication complexity, we study the orthogonal rank and a symmetric variant thereof for a certain family of Cayley graphs. The orthogonal rank of a graph is the smallest number d for which one can label each vertex with a nonzero d-dimensional complex vector such that adjacent vertices receive orthogonal vectors.

We show an $\exp(n)$ lower bound on the orthogonal rank of the graph on $\{0,1\}^n$ in which two strings are adjacent if they have Hamming distance at least $n/2$. In combination with previous work, this implies an affirmative answer to the above conjecture.

1. Introduction

The orthogonal rank of Cayley graphs. In the following all graphs are simple and undirected. For a graph $G = (V, E)$, a map $\phi : V \to \mathbb{C}^d$ is an *orthogonal embedding in dimension* d if $\langle \phi(v), \phi(v) \rangle = 1$ for all $v \in V$ and $\langle \phi(v), \phi(w) \rangle = 0$ for all $(v, w) \in E$. The *orthogonal rank* of G, denoted $\xi(G)$, is the smallest positive integer d such that there is an orthogonal embedding of G in dimension d. Here we prove bounds on the orthogonal ranks of certain Cayley graphs.

For a finite group Γ and a subset $S \subseteq \Gamma$, the *Cayley graph* $G = \mathrm{Cay}(\Gamma, S)$ is the graph with vertex set $V = \Gamma$ and in which $g, h \in \Gamma$ are connected by

*QuSoft, CWI, Science Park 123, 1098 XG Amsterdam, Netherlands. Email: j.briet@cwi.nl. Supported by a VENI grant from the Netherlands Organisation for Scientific Research (NWO).

†QuSoft, CWI and University of Amsterdam, Science Park 123, 1098 XG Amsterdam, Netherlands. Email: j.zuiddam@cwi.nl. Supported by NWO through the research programme 617.023.116.

an edge if and only if $gh^{-1} \in S$ or $hg^{-1} \in S$. We shall also be interested in the following variant of the orthogonal rank. Say that an orthogonal embedding $\phi : \Gamma \to \mathbb{C}^d$ is *symmetric* if there exists a function $f : \Gamma \to \mathbb{C}$ such that $\langle \phi(g), \phi(h) \rangle = f(gh^{-1})$ for all $g, h \in \Gamma$. Then, the *symmetric orthogonal* rank of G, denoted $\xi_{\mathrm{sym}}(G)$, is defined as the smallest positive integer d such that there exists a symmetric orthogonal embedding in dimension d. Note that, clearly, $\xi(G) \leq \xi_{\mathrm{sym}}(G)$.

Our main results concern bounds on the orthogonal rank and symmetric orthogonal rank of certain Cayley graphs based on powers of cyclic groups. For a positive integer m, let $C_m = \{0, 1, \ldots, m-1\}$ be the cyclic group of m elements. For a positive integer k, let $[k]$ denote the set $\{1, 2, \ldots, k\}$. For a positive integer n and parameter $d \in [(m-1)n]$, define $H_m^n(d)$ to be the Cayley graph $\mathrm{Cay}(C_m^{\times n}, S)$ with $S = \{x \in C_m^{\times n} : \sum_i x_i \geq d\}$, where in the summation we consider the elements x_i as elements in $\mathbb{N}$. Then, the graph $H_2^n(d)$ is precisely the graph with vertex set $\{0,1\}^n$ where two strings form an edge if and only if they have Hamming distance at least d.

Quantum communication complexity. The orthogonal rank in general is a poorly understood parameter. Much impetus for its study has recently come from quantum information theory, in particular in the context of quantum entanglement [CMN$^+$07]. The problem of bounding the orthogonal rank of the above-mentioned Cayley graph $H_2^n(d)$ arose from a question in exact quantum communication complexity. Here two parties, Alice and Bob, receive inputs x, y from sets $\mathcal{X}, \mathcal{Y}$, respectively, and their goal is to compute a function $f(x, y)$ depending on both of their inputs, using as little communication as possible. In a *promise problem*, the inputs are guaranteed to be drawn from a subset $\mathcal{D} \subseteq \mathcal{X} \times \mathcal{Y}$ known to the parties in advance.

In the most general deterministic classical protocol the parties take turns sending each other binary sequences until both of them know the answer. The (exact) *communication complexity* is defined as the minimum number of bits sent back and forth in such a protocol under worst-case inputs. The *one-round* communication complexity is the length of the shortest string Alice can send Bob so that he can learn the answer under worst-case inputs.

In the quantum setting, the parties may send each other *qubits* instead of classical bits, which gives them at least as much power as in the classical setting and is well-known to sometimes lead to dramatic savings [BCW99] (see [NC10, Wil13] for an introduction to quantum information theory). The *quantum communication complexity* is defined analogously to its classical counterpart. The *one-round* quantum communication complexity turns out to be characterized precisely by the orthogonal rank of the graph $G = (V, E)$ with vertex set $V = \mathcal{X}$ and where $u, v \in V$ form an edge if there exists a $y \in \mathcal{Y}$ such that $(u, y) \in \mathcal{D}$ and $(v, y) \in \mathcal{D}$. De Wolf [dW01, Theorem 8.5.2] showed that the one-round quantum communication complexity equals $\lceil \log_2 \xi(G) \rceil$.

Impossibility of quantum round elimination. *Round elimination* is a basic procedure whereby the number of rounds of communication is reduced at the cost of some additional communication. For example, if Bob is supposed to send Alice a single bit after which she is supposed to reply with a k-bit string, she could just as well immediately send Bob the two k-bit strings corresponding to the bits he might have sent, after which he picks out the appropriate one. This removes one round of communication at the cost of roughly a factor of two increase in the total number of communicated bits.

It was conjectured in [BBL$^+$15] that a quantum analogue of round elimination is impossible, in the sense that removing a round can result in unexpectedly large increases in (quantum) communication. The following promise problem was suggested as a possible candidate to show this: Alice is given an n-bit string x and Bob is given a set $Y \subseteq \{0,1\}^n$ containing x such that for some integer $d \geq n/2$ the pairwise Hamming distances between the strings in Y equals d. The parties' goal is for Bob to learn x, that is $f(x, Y) = x$. The authors gave a two-round protocol for this problem in which Bob first sends Alice a single qubit, after which Alice replies with a $k = \lceil \log_2 n + 1 \rceil$-qubit sequence. The naive analogue of the above round-elimination example would say that there is a one-round $2k$-qubit protocol. However, in [BBL$^+$15] it was conjectured that the orthogonal rank of the graph associated to the problem, the graph $H_2^n(n/2)$, is of the order $n^{\omega(1)}$, implying that the one-round quantum communication complexity is in fact $\omega(k)$. It was shown that for n even, $2n \leq \xi(H_2^n(n/2)) \leq 2^{h(1/4)n+1} \approx 2^{0.81n}$, where $h(p) = -p \log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function.

Towards resolving this conjecture, it was suggested by Buhrman [Buh] to examine the potentially easier problem of determining the symmetric orthogonal rank of $H_2^n(n/2)$. Here we determine this parameter exactly for the more general class of Cayley graphs described above.

Theorem 1. *For all positive integers m, n and any $d \in [(m-1)n]$ that is divisible by $m-1$, we have*

$$\xi_{\mathrm{sym}}(H_m^n(d)) = m^{n - \frac{d}{m-1}}.$$

Observe that the above result improves on the upper bound of the orthogonal rank of [BBL$^+$15], since $\xi(H_2^n(n/2)) \leq \xi_{\mathrm{sym}}(H_2^n(n/2)) \leq 2^{n/2}$.

More importantly, the main conjecture of [BBL$^+$15] can be resolved in the affirmative.

Theorem 2. *There exist absolute constants $c, \varepsilon \in (0, \infty)$ such that for every positive integer n that is divisible by 8, we have*

$$\xi(H_2^n(n/2)) \geq 2^{\varepsilon n - c}.$$

Remark 3. Our proof of Theorem 2 indirectly establishes the result by bounding the Lovász theta number $\vartheta(\overline{G})$ of the complement of $G = H_2^n(n/2)$

(see Section 3) and using the fact that $\xi(G) \geq \vartheta(\overline{G})$ [BBL$^+$15]. While writing this note, Amir Yehudayoff brought to our attention an unpublished manuscript of Samorodnitsky's [Sam98] where he proves a slightly better lower bound on $\vartheta(\overline{G})$. He determines this value almost exactly with the use of cleverly-chosen orthogonal polynomials, giving $\vartheta(\overline{G}) \approx 2^{0.19n}$. Using more elementary methods, we prove that $\vartheta(\overline{G}) \geq 2^{0.0435n - \log_2(5/2)}$.

Based on Theorem 1 and Samorodnitsky's results, which cover the graphs $H_2^n(d)$ for all $d \in [n]$, the best bounds on the orthogonal rank of $G = H_2^n(d)$ can be summarized as follows:

$$2^{(1-h(d/(2n)))n - o(n)} \leq \vartheta(\overline{G}) \leq \xi(G) \leq \xi_{\mathrm{sym}}(G) = 2^{n-d},$$

where h denotes the binary entropy function defined above.

Connection with k-wise independence. Before going into the proofs, we would like to mention a connection between the Lovász theta number of the graph $H_2^n(d)$ and $(d-1)$-wise independent distributions on $\{0,1\}^n$. A probability distribution P on $\{0,1\}^n$ is *k-wise independent* if for any k indices $i_1 < i_2 < \cdots < i_k$ and any string $v \in \{0,1\}^k$

$$\Pr_{x \sim P}[x_{i_1} x_{i_2} \cdots x_{i_k} = v] = 2^{-k}.$$

In words, the restriction of P to any k indices is a uniform distribution. Now view P as a function $\{0,1\}^n \to \mathbb{R}$ such that $P(x) \geq 0$ for all $x \in \{0,1\}^n$ and $\sum_x P(x) = 1$. It is a standard and easy fact that P being k-wise independent is equivalent to the Fourier coefficients $\widehat{P}(z)$ being zero for all $z \in \{0,1\}^n$ with Hamming weight $|z| \in \{1, 2, \ldots, k\}$. Let $G = H_2^n(d)$. With the above observation and Equation (3) on page 11 one can prove that the value of

$$\begin{array}{lll} \max & P[00\cdots0] & \text{s.t.} \quad 1.\ P \text{ is a prob. distr. on } \{0,1\}^n \\ & & \quad\ 2.\ P \text{ is } (d-1)\text{-wise independent} \end{array}$$

is exactly $2^{-n}\,\vartheta(\overline{G})$. The maximal probability that all bits are zero was studied in [PYY11] and [BGGP12] and it was stated as an open problem in [PYY11] to determine this value for all $d \in [n/2, n) \cap \mathbb{N}$. Our proof of Theorem 2 gives a nontrivial lower bound for $d = n/2$. The results of [Sam98] yield the asymptotically tight value $2^{-h(d/(2n))n}$ for any $d \in [n]$.

A concise way to phrase the above in Fourier analytic terms is as follows. For a function $f : \{0,1\}^n \to \mathbb{R}$ and $p \in [1, \infty)$, the ℓ_p-norm of f is defined as

$$\|f\|_p \coloneqq \Big(2^{-n} \sum_{x \in \{0,1\}^n} |f(x)|^p\Big)^{1/p}.$$

Define $\|f\|_\infty \coloneqq \max_{x \in \{0,1\}^n} |f(x)|$. The above lower is then equivalent to the assertion that for any $d \in [n]$ and any function $f : \{0,1\}^n \to \mathbb{R}$ of polynomial

degree d, we have

$$\|f\|_\infty \leq 2^{(1-h(d/(2n)))n+o(n)} \|f\|_1 .$$

This may be compared with the following standard consequence of the hypercontractive inequality [BLM13, Corollary 5.16], which says that for any function $f : \{0,1\}^n \to \mathbb{R}$ of polynomial degree d and for all $1 < p < q < \infty$,

$$\|f\|_q \leq \left(\frac{q-1}{p-1}\right)^{d/2} \|f\|_p .$$

2. The symmetric orthogonal rank

In this section we prove Theorem 1. Let us first review some results on the character group of a finite group. Let Γ be a finite group and let $\mathbb{C}^\times$ be the multiplicative group $\mathbb{C} \setminus \{0\}$. The character group $\widehat{\Gamma}$ of Γ is the group consisting of all homomorphisms $G \to \mathbb{C}^\times$, that is, maps $f : G \to \mathbb{C}^\times$ such that $f(gh) = f(g)f(h)$ for any $g, h \in \Gamma$. Now consider the complex vector space $\mathbb{C}^\Gamma$ consisting of all maps $\Gamma \to \mathbb{C}$. Endow this space with the inner product defined by $\langle f, f' \rangle \coloneqq |\Gamma|^{-1} \sum_{g \in G} f(g)f'(g)$. Then the characters $\widehat{\Gamma}$ form an orthonormal basis of $\mathbb{C}^\Gamma$. We can thus write every map $f : \Gamma \to \mathbb{C}$ in the form

$$f(g) = \sum_{\chi \in \widehat{\Gamma}} \widehat{f}(\chi)\, \chi(x)$$

with $\widehat{f}(\chi) \in \mathbb{C}$. The complex numbers $\widehat{f}(\chi)$ are called Fourier coefficients and the map $f \mapsto \widehat{f}$ is called the Fourier transform.

Let $m \in \mathbb{N}$ and let $\zeta_m \in \mathbb{C}^\times$ be an mth primitive root of unity. Let Γ be the cyclic group $C_m = \{0, 1, \ldots, m-1\}$. Then the character group $\widehat{\Gamma}$ consists of the maps

$$\chi_z : C_m \to \mathbb{C}^\times : x \mapsto (\zeta_m^z)^x \qquad \text{with } z \in C_m.$$

Let $n \in \mathbb{N}$ and let Γ be the direct power $C_m^{\times n}$. Then the character group $\widehat{\Gamma}$ consists of the maps

$$\chi_z : C_m^{\times n} \to \mathbb{C}^\times : x \mapsto (\zeta_m^{z_1})^{x_1} \cdots (\zeta_m^{z_n})^{x_n} \qquad \text{with } z \in C_m^{\times n}.$$

We will write the product $(\zeta_m^{z_1})^{x_1} \cdots (\zeta_m^{z_n})^{x_n}$ as $\zeta_m^{z \cdot x}$, and $\widehat{f}(\chi_z)$ as $\widehat{f}(z)$.

We will use Bochner's theorem for finite groups. Let f be a map $\Gamma \to \mathbb{C}$. Let e be the unit element of Γ. We say f is *normalized* if $f(e) = 1$. We say that f is *positive semidefinite* (PSD) if for any $k \in \mathbb{N}$ and any $g_1, \ldots, g_k \in \Gamma$ the matrix $(f(g_i g_j^{-1}))_{i,j \in [k]}$ is PSD.

Theorem 4 (Bochner's theorem for finite groups)**.** *Let Γ be a finite abelian group. Let $\widehat{\Gamma}$ be the character group of Γ. Let f be a map $\Gamma \to \mathbb{C}$. Then, the following two statements are equivalent:*

1. *The map f is normalized and PSD.*

2. *The map f satisfies $\widehat{f}(\chi) \in \mathbb{R}_{\geq 0}$ for all $\chi \in \widehat{\Gamma}$, and $\sum_{\chi \in \widehat{\Gamma}} \widehat{f}(\chi) = 1$.*

Proof. Assume f is normalized and PSD. Consider the Fourier decomposition $f = \sum_{\chi \in \widehat{\Gamma}} \widehat{f}(\chi)\chi$. Then $f(xy^{-1}) = \sum_{\chi} \widehat{f}(\chi)\chi(x)\overline{\chi(y)}$. Define the matrix $M = (f(gh^{-1}))_{g,h \in \Gamma}$. Then for any vector $v \in \mathbb{C}^{\Gamma}$, we have $v^* M v \in \mathbb{R}_{\geq 0}$ and

$$\begin{aligned} v^* M v &= v^* \Big(\sum_{\chi} \widehat{f}(\chi)(\chi(g)\overline{\chi(h)})_{g,h} \Big) v \\ &= \sum_{\chi} \widehat{f}(\chi) v^* (\chi\chi^*) v, \end{aligned}$$

where in the last line we used χ to denote the complex vector $(\chi(g))_g$. By taking $v = \chi$ and using the orthogonality of the characters, we get $\widehat{f}(\chi) \in \mathbb{R}_{\geq 0}$ for all χ. Also $1 = f(e) = \sum_{\chi} \widehat{f}(\chi)\chi(e) = \sum_{\chi} \widehat{f}(\chi)$.

Assume f has real nonnegative Fourier coefficients summing to 1. Then $f(e) = \sum_{\chi} \widehat{f}(\chi)\chi(e) = \sum_{\chi} \widehat{f}(\chi) = 1$. Let $g_1, \ldots, g_k \in \Gamma$. Define the matrix $M = (f(g_i g_j^{-1}))_{i,j \in [k]}$. Then $M = \sum_{\chi} \widehat{f}(\chi)(\chi(g_i)\overline{\chi(g_j)})_{i,j \in [k]} = \sum_{\chi} \widehat{f}(\chi) N_{\chi}$, where each N_{χ} is a submatrix of the PSD matrix $\chi\chi^*$. Therefore, the matrix M is PSD. □

The following proposition relates symmetric orthogonal embeddings to maps $f : \Gamma \to \mathbb{C}$ with restrictions on the Fourier coefficients.

Proposition 5. *Let Γ be a finite abelian group and let S be a subset of Γ. Let $f : \Gamma \to \mathbb{C}$ be a map such that $f(e) = 1$ and $f(g) = 0$ for all $g \in S$. Then, there exists a map $\phi : \Gamma \to \mathbb{C}^d$ such that $\langle \phi(g), \phi(h) \rangle = f(gh^{-1})$ if and only if all Fourier coefficients $\widehat{f}(\chi)$ are real and nonnegative and $|\mathrm{supp}(\widehat{f})| \leq d$.*

Proof. Let $\phi : \Gamma \to \mathbb{C}^d$ be a map such that $\langle \phi(g), \phi(h) \rangle = f(gh^{-1})$. We have the following equality of matrices

$$(\langle \phi(g), \phi(h) \rangle)_{g,h \in \Gamma} = (f(gh^{-1}))_{g,h \in \Gamma} = \sum_{\chi} \widehat{f}(\chi)\chi\chi^*.$$

The left-hand side is a Gram matrix and therefore PSD. Bochner's theorem (Theorem 4) says that the Fourier coefficients $\widehat{f}(\chi)$ are then real and nonnegative. Moreover, the rank of the left-hand side is at most d while the rank of the right-hand side equals $|\mathrm{supp}(\widehat{f})|$.

On the other hand, suppose $\widehat{f}(\chi) \in \mathbb{R}_{\geq 0}$ for all χ. Let S be the set $\{\chi \in \widehat{\Gamma} : \widehat{f}(\chi) \neq 0\}$. For any $g \in \Gamma$, define the vector

$$\phi(g) \coloneqq \Big(\sqrt{\widehat{f}(\chi)}\, \chi(g) \Big)_{\chi \in S} \in \mathbb{C}^S.$$

We claim that ϕ satisfies $\langle \phi(g), \phi(h) \rangle = f(gh^{-1})$ for all $g, h \in \Gamma$. Indeed, we have, for any $g, h \in \Gamma$,

$$\langle \phi(g), \phi(h) \rangle = \sum_{\chi \in S} \widehat{f}(\chi)\chi(g)\overline{\chi(h)} = f(gh^{-1}) = 0, \quad \text{when } gh^{-1} \in S,$$
$$\langle \phi(g), \phi(g) \rangle = \sum_{\chi \in S} \widehat{f}(\chi)\chi(g)\overline{\chi(g)} = f(e) = 1,$$

which proves the claim. □

We also use the following well-known result on the number of roots of multivariate polynomials, the particular form of which is taken from [CT15]. View the cyclic group C_m as a multiplicative subgroup of $\mathbb{C}$ by mapping a generator to a primitive mth root of unity. For any map $f : C_m^{\times n} \to \mathbb{C}$, we define the *polynomial degree* $\deg(f)$ to be the smallest number d such that there is a polynomial $p \in \mathbb{C}[x_1, \ldots, x_n]$ of degree d that interpolates f, that is, $p(z) = f(z)$ for all $z \in C_m^{\times n}$. We write $U(f) := \{z \in C_m^{\times n} \mid f(z) \neq 0\}$ for the set of nonzeros of f in $C_m^{\times n}$.

Theorem 6 (DeMillo-Lipton-Schwartz-Zippel)**.** *Let $f : C_m^{\times n} \to \mathbb{C}$ be a nonzero map of polynomial degree d. Then,*

$$|U(f)| \geq \frac{m^n}{m^{d/(m-1)}}.$$

Proof. By viewing C_m as a multiplicative subgroup of $\mathbb{C}$ we can identify f with a nonzero polynomial in $\mathbb{C}[x_1, \ldots, x_n]$ of degree d such that each variable in f has degree at most $m-1$. We induce on n. For the base case $n = 1$, f is a nonzero univariate polynomial of degree at most $m-1$ and f thus has at most $m-1$ zeros. Therefore, $|U(f)| \geq 1 \geq m/m^{d/(m-1)}$.

Assume the theorem statement is proven for polynomials in $n-1$ variables. We can write f in the form

$$f(t, y_1, \ldots, y_{n-1}) = \sum_{i=1}^{d} t^i g_i(y_1, \ldots, y_{n-1}),$$

with $g_i \in \mathbb{C}[y_1, \ldots, y_{n-1}]$ a polynomial of degree at most $d - i$. Let k be the maximum i for which g_i is nonzero. By the induction hypothesis, the polynomial g_k satisfies

$$|U(g_k)| \geq m^{n-1}/m^{(d-k)/(m-1)}.$$

For each $y \in U(g_k)$, let $h_y \in \mathbb{C}[t]$ be the univariate polynomial defined by $h_y(t) = f(t, y_1, \ldots, y_{n-1})$. We know that each h_y is nonzero and has degree k. Therefore, $|U(h_y)| \geq m/m^{k/(m-1)}$. We conclude that

$$|U(f)| \geq \sum_{y \in U(g_k)} |U(h_y)| \geq m^n/m^{d/(m-1)},$$

which proves the theorem. □

Lower bound proof for Theorem 1. For $x \in C_m^{\times n}$, define the weight $|x|$ to be $\sum_i x_i$ where the sum is taken in $\mathbb{N}$. Denote the unit element in $C_m^{\times n}$ by 0. Let $f : C_m^{\times n} \to \mathbb{C}$ be a map satisfying the three properties

1. $f(0) = 1$,
2. $f(x) = 0$ when $|x| \geq d$,
3. $\widehat{f}(z) \geq 0$ for all $z \in C_m^{\times n}$.

Write f in the Fourier basis, $f = \sum_{z \in C_m^{\times n}} \widehat{f}(z)\chi_z(x)$. Define $g : C_m^{\times n} \to \mathbb{C}$ by $g(z) = \widehat{f}(z)$. Then $\widehat{g}(z) = m^{-n} f(z)$. The Fourier expansion of g is thus $g(x) = m^{-n} \sum_{z \in C_m^{\times n}} f(z)\chi_z(x)$. Since $f(x) = 0$ when $|x| \geq d$, the polynomial degree of g is at most d. By Theorem 6 there are at least $m^n/m^{d/(m-1)}$ points where g is nonzero. The map f thus has at least $m^n/m^{d/(m-1)}$ nonzero Fourier coefficients, which means by Proposition 5 that $\xi_{\mathrm{sym}}(G) \geq m^{n-d/(m-1)}$. □

Upper bound proof for Theorem 1. Define $k = n - d/(m-1)$. Let $h : C_m^{\times k} \to \mathbb{C}$ be the indicator function of the zero element in $C_m^{\times k}$, so $h(0) = 1$ and $h(x) = 0$ for all $x \neq 0$. Let $g : C_m^{\times n} \to \mathbb{C}$ map x to $h(x_1, \ldots, x_k)$. We see that $g(0) = 1$ and that $g(x) = 0$ whenever $|x| \geq d$. For the Fourier coefficients of g we get

$$\langle g, \chi_z \rangle = \frac{1}{m^n} \sum_{x \in C_m^{\times n}} g(x)\chi_z(x) = \frac{1}{m^n} \sum_{\substack{x \in C_m^{\times n}: \\ x = 0^k x'}} \zeta_m^{z' \cdot x'},$$

where $z' \in C_m^{\times d/(m-1)}$ is the vector consisting of the last $d/(m-1)$ entries in z, and similarly for x'. If $z' = 0$, then the above expression is positive. If $z' \neq 0$, then the above expression is zero. We conclude that g has m^k nonzero Fourier coefficients and moreover all nonzero Fourier coefficients are positive. Therefore, by Proposition 5, there is a symmetric orthogonal embedding of the graph $H_m^n(d)$ in dimension m^k. □

3. Lower bounds on the orthogonal rank

In this section we prove Theorem 2. The workhorse in this proof is the following special case of [Sam98, Lemma 3.3], for which we give an alternative, more elementary proof, albeit with slightly worse constants.

Proposition 7. *There exist absolute constants $c, \varepsilon \in (0, \infty)$ such that for every $n \in \mathbb{N}$ that is divisible by 8 and any $p \in \mathbb{R}[x]$ of degree $< n/2$, we have*

$$p(0) \leq 2^{-\varepsilon n + c} \sum_{i=0}^{n} \binom{n}{i} |p(i)| .$$

Proof. Let $S \subseteq \{0, 1, \ldots, n\}$ be any set of size at least $n/2$. Writing p as an interpolation polynomial in Lagrange form with respect to the set S, gives

$$p(x) = \sum_{i \in S} p(i) \prod_{\ell \in S \setminus \{i\}} \frac{\ell - x}{\ell - i}.$$

By the triangle inequality, $p(0)$ is therefore at most

$$\begin{aligned} p(0) &\leq \sum_{i \in S} |p(i)| \prod_{\ell \in S \setminus \{i\}} \frac{\ell}{|\ell - i|} \\ &= \sum_{i \in S} |p(i)| \binom{n}{i} \binom{n}{i}^{-1} \prod_{\ell \in S \setminus \{i\}} \frac{\ell}{|\ell - i|} \\ &\leq \left(\sum_{i=0}^{n} \binom{n}{i} |p(i)| \right) \max_{i \in S} \binom{n}{i}^{-1} \prod_{\ell \in S \setminus \{i\}} \frac{\ell}{|\ell - i|}. \end{aligned}$$

Hence,

$$p(0) \leq \left(\sum_{i=0}^{n} \binom{n}{i} |p(i)| \right) \min_{S} \max_{i \in S} \binom{n}{i}^{-1} \prod_{\ell \in S \setminus \{i\}} \frac{\ell}{|\ell - i|},$$

where the minimum is taken over all sets $S \subseteq \{0, 1, \ldots, n\}$ of size at least $n/2$. To prove the result it thus suffices to exhibit a set S for which the maximum is exponentially small in n. To this end, define the sets $S_1 \coloneqq (\frac{n}{8}, \frac{3n}{8}] \cap \mathbb{N}$ and $S_2 \coloneqq [\frac{5n}{8}, \frac{7n}{8}) \cap \mathbb{N}$ and let $S = S_1 \cup S_2$. Define

$$f(i) \coloneqq \binom{n}{i}^{-1} \prod_{\ell \in S \setminus \{i\}} \frac{\ell}{|\ell - i|}. \tag{1}$$

Since f is symmetric about $n/2$, we have $\max_{i \in S_1} f(i) = \max_{i \in S_2} f(i)$ and it follows that to bound $\max_{i \in S} f(i)$, it is sufficient to maximize f over S_1.

Define $k = n/8$. Let $j \in [1, 3]$ such that $i \coloneqq jk \in S_1$. We claim that

$$f(i) \leq \frac{\binom{3k}{jk}\binom{7k-1}{jk}\binom{jk}{k}}{\binom{5k-1}{jk}\binom{8k}{jk}}. \tag{2}$$

Indeed, in (1) we can split the product over S into a product over S_1 and a product over S_2 to obtain

$$\begin{aligned} f(jk) &= \binom{8k}{jk}^{-1} \prod_{\ell \in S_1 \setminus \{jk\}} \frac{\ell}{|\ell - jk|} \prod_{\ell \in S_2} \frac{\ell}{|\ell - jk|} \\ &\leq \binom{8k}{jk}^{-1} \frac{(3k)!}{k!\,((3-j)k)!\,((j-1)k)!} \frac{(7k-1)!\,((5-j)k-1)!}{(5k-1)!\,((7-j)k-1)!}. \end{aligned}$$

Multiplying with $(jk)!^2/(jk)!^2$ and grouping appropriately, one recognizes the required binomial coefficients.

Observe that the product of the first and third coefficients in the numerator of (2), namely $\binom{3k}{jk}\binom{jk}{k}$, counts the number of ways to choose a jk-subset in a $3k$-set and then a k-subset in this jk-subset. We get the same count by first choosing a k-subset in a $3k$-set and then choosing a jk-subset in the $3k$-set which includes the k-set. Therefore, $\binom{3k}{jk}\binom{jk}{k} = \binom{3k}{k}\binom{3k-k}{jk-k} = \binom{3k}{k}\binom{2k}{jk-k}$. Next, it follows from the Cauchy–Vandermonde identity [Juk11, Exercise 1.9],

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k},$$

that $\binom{3k}{k}\binom{2k}{jk-k} \leq \binom{5k}{jk} = \frac{5}{5-j}\binom{5k-1}{jk}$. Hence, $f(i) \leq \frac{5}{5-j}\binom{7k-1}{jk}\binom{8k}{jk}^{-1}$. Finally, since $j \leq 3$ and for any integers $b \leq b+c < a$, we have $\binom{a-c}{b}\binom{a}{b}^{-1} \leq (\frac{a-b}{a})^c$ [Juk11, Exercise 1.18], it follows that

$$\begin{aligned} f(i) &\leq \frac{5}{5-j}\binom{7k-1}{jk}\binom{8k}{jk}^{-1} \\ &\leq \frac{5}{2}\binom{7k}{jk}\binom{8k}{jk}^{-1} \\ &\leq \frac{5}{2}\Big(\frac{8k-jk}{8k}\Big)^k \leq \frac{5}{2}\Big(\frac{7}{8}\Big)^k, \end{aligned}$$

which establishes the result. □

For a matrix $X \in \mathbb{R}^{n\times n}$, we write $X \succcurlyeq 0$ if X is symmetric and PSD.

Proof of Theorem 2. Let $G = H_2^n(d)$. We lower bound $\xi(G)$ by lower bounding $\vartheta(\overline{G})$. This we do by looking at the value of $\vartheta(G)$. By definition,

$$\vartheta(G) = \max \sum_{i,j\in[2^n]} X_{ij} \quad \text{s.t.}$$

1. X is a real $2^n \times 2^n$ matrix
2. $X \succcurlyeq 0$
3. $\sum_{i=1}^{2^n} X_{ii} = 1$
4. $X_{ij} = 0 \quad \forall (i,j) \in E(G)$.

If X is a feasible solution to the above maximisation, then for any element $a \in G$ (we implicitly identify G with the group $C_2^{\times n}$ here) the matrix Y^a defined by $Y^a_{x,y} = X_{x+a,y+a}$ is a feasible solution with the same value. Let Y be the average $\frac{1}{2^n}\sum_{a\in G} Y^a$. This is again feasible with the same value. Moreover $Y_{x,y}$ depends only on $x - y$; namely, if $x - y = x' - y'$, then $Y_{x,y} = Y_{x-y,0} = Y_{x'-y',0} = Y_{x',y'}$. With this observation and Bochner's Theorem (Theorem 4) we obtain

$$\vartheta(G) = \max \sum_{x\in\{0,1\}^n} f(x) \qquad \text{s.t.} \qquad \begin{array}{l} 1.\ f : \{0,1\}^n \to \mathbb{R} \\ 2.\ \widehat{f}(z) \geq 0 \text{ for all } z \in \{0,1\}^n \\ 3.\ f(0) = 1 \\ 4.\ f(x) = 0 \text{ for } d \leq |x|. \end{array}$$

Through the Fourier transform we get

$$\vartheta(G) = \max \quad 2^n g(0) \qquad \text{s.t.} \qquad \begin{array}{l} 1.\ g : \{0,1\}^n \to \mathbb{R} \\ 2.\ g(z) \geq 0 \text{ for all } z \\ 3.\ 2^n\,\widehat{g}(0) = 1 \\ 4.\ \widehat{g}(x) = 0 \text{ for } d \leq |x|. \end{array} \tag{3}$$

(See [Sch79] for a similar description.) Because $\vartheta(\overline{G})\,\vartheta(G) \geq 2^n$ (see [Lov79]), we have

$$\vartheta(\overline{G}) \geq \frac{2^n}{\max\ 2^n g(0)} \qquad \text{s.t.} \qquad \begin{array}{l} 1.\ g : \{0,1\}^n \to \mathbb{R} \\ 2.\ g(z) \geq 0 \text{ for all } z \\ 3.\ 2^n\,\widehat{g}(0) = 1 \\ 4.\ \widehat{g}(x) = 0 \text{ for } d \leq |x|. \end{array}$$

Now let $d = n/2$. According to Proposition 7, the value of $g(0)$ is at most $2^{-\varepsilon n+c}$, so the value of $\vartheta(\overline{G})$ is at least $2^{\varepsilon n-c}$. □

Acknowledgements. The authors thank Harry Buhrman, Teresa Piovesan, Oded Regev, Ronald de Wolf, and Amir Yehudayoff for helpful discussions.

References

[BBL+15] J. Briët, H. Buhrman, D. Leung, T. Piovesan, and F. Speelman. Round elimination in exact communication complexity. In *Proceedings of the 10th Conference on the Theory of Quantum Computation and Cryptography*, 2015.

[BCW99] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *STOC '98*, pages 63–68. ACM, New York, 1999.

[BGGP12] I. Benjamini, O. Gurel-Gurevich, and R. Peled. On k-wise independent distributions and boolean functions. *arXiv preprint arXiv:1201.3261*, 2012.

[BLM13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence.

[Buh] H. Buhrman. Personal communication.

[CMN+07] P. J. Cameron, A. Montanaro, M. W. Newman, S. Severini, and A. Winter. On the quantum chromatic number of a graph. *Electron. J. Combin.*, 14(1):Research Paper 81, 15 pp. (electronic), 2007.

[CT15] G. Cohen and A. Tal. Two Structural Results for Low Degree Polynomials and Applications. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, volume 40 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 680–709, 2015.

[dW01] R. de Wolf. *Quantum Computing and Communication Complexity*. PhD thesis, Universiteit van Amsterdam, 2001.

[Juk11] S. Jukna. *Extremal combinatorics: with applications in computer science*. Springer Science & Business Media, 2011.

[Lov79] L. Lovász. On the Shannon capacity of a graph. *Information Theory, IEEE Transactions on*, 25(1):1–7, 1979.

[NC10] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[PYY11] R. Peled, A. Yadin, and A. Yehudayoff. The maximal probability that k-wise independent bits are all 1. *Random Structures & Algorithms*, 38(4):502–525, 2011.

[Sam98] A. Samorodnitsky. Extremal properties of solutions for Delsarte's linear program. Preliminary version. URL: `http://www.cs.huji.ac.il/~salex/papers/old_sq_measure.ps`, 1998.

[Sch79] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *Information Theory, IEEE Transactions on*, 25(4):425–429, 1979.

[Wil13] M. M. Wilde. *Quantum information theory*. Cambridge University Press, 2013.