GAFA, Geom. funct. anal.
Vol. 14 (2004) 913 – 940
1016-443X/04/050913-28
DOI 10.1007/s00039-004-0478-3

GAFA Geometric And Functional Analysis

GRAPH PRODUCTS, FOURIER ANALYSIS AND SPECTRAL TECHNIQUES

N. Alon, I. Dinur, E. Friedgut and B. Sudakov

Abstract. We consider powers of regular graphs defined by the weak graph product and give a characterization of maximum-size independent sets for a wide family of base graphs which includes, among others, complete graphs, line graphs of regular graphs which contain a perfect matching and Kneser graphs. In many cases this also characterizes the optimal colorings of these products.

We show that the independent sets induced by the base graph are the only maximum-size independent sets. Furthermore we give a qualitative stability statement: any independent set of size close to the maximum is close to some independent set of maximum size.

Our approach is based on Fourier analysis on Abelian groups and on Spectral Techniques. To this end we develop some basic lemmas regarding the Fourier transform of functions on $\{0, \ldots, r-1\}^n$, generalizing some useful results from the $\{0, 1\}^n$ case.

1 Introduction

Consider the following combinatorial problem:

Assume that at a given road junction there are n three-position switches that control the red-amber-green position of the traffic light. You are told that whenever you change the position of *all* the switches then the color of the light changes. Prove that in fact the light is controlled by only one of the switches.

The above problem is a special case of the problem we wish to tackle in this paper, characterizing the optimal colorings and maximal independent

N.A.'s research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University. E.F.'s research supported in part by the Israel Science Foundation. Part of this research was done while E.F. visited Microsoft Research and the Institute for Advanced Study in Princeton. B.S.'s research supported in part by NSF grants DMS-0106589, CCR-9987845 and by the State of New Jersey.

sets of products of regular graphs. The configuration space of the switches described above can be modeled by the n-fold product of K_3. Let us begin by defining the weak graph product of two graphs.

The *weak product* of G and H, denoted by $G \times H$ is defined as follows: the vertex set of $G \times H$ is the Cartesian product of the vertex sets of G and H. Two vertices (g_1, h_1) and (g_2, h_2) are adjacent in $G \times H$ if $g_1 g_2$ is an edge of G and $h_1 h_2$ is an edge of H. The "times" symbol, $\times$, is supposed to be reminiscent of the weak product of two edges: $| \times - = \times$. In this paper "graph product" will always mean the weak product.

In the first part of the paper we consider the interesting special case of the product of complete graphs on $r > 2$ vertices,

$$G = K_r^n = \times_{j=1}^n K_r \,.$$

We then discuss a more general setting, considering other r-regular graphs as well.

When $G = K_r^n$, we identify the vertices of G in the obvious way with the elements of $\mathbb{Z}_r^n$. Recalling the definition of the product, two vertices are adjacent in G iff the corresponding vectors differ in every coordinate. Clearly one can color G with r colors by choosing a coordinate i and coloring every vertex according to its ith coordinate. The following theorem asserts that if $r > 2$ then these are the only r-colorings. Here, and in what follows, we denote by $|H|$ the number of vertices of a graph H.

Theorem 1.1. *Let $G = K_r^n$, and assume $r \geq 3$. Let I be an independent set with $|I| = |G|/r$. Then there exists a coordinate $i \in \{1, \dots, n\}$ and $k \in \{0, \dots, r-1\}$ such that*

$$I = \{v : v_i = k\} \,.$$

Consequently, the only colorings of G by r colors are those induced by colorings of one of the factors K_r.

Greenwell and Lovász [GL] proved the above theorem (and actually, a somewhat stronger statement) more than a quarter of a century ago. See also [Mü] for a similar result. The novelty in this paper is the proof we supply that uses Fourier analysis on the group $\mathbb{Z}_r^n$. Our approach also allows us to deduce a stability version of the above theorem.

Theorem 1.2. *For every $r \geq 3$ there exists a constant $M = M(r)$ such that for any $\epsilon > 0$ the following is true. Let $G = K_r^n$. Let J be an independent set such that $\frac{|J|}{|G|} = \frac{1}{r} - \epsilon$. Then there exists an independent set I with $\frac{|I|}{|G|} = \frac{1}{r}$ such that $\frac{|J \triangle I|}{|G|} < M\epsilon$.*

Here "$\triangle$" denotes the symmetric difference. What the above theorem tells us is (in conjunction with Theorem 1.1) that any independent set that is *close* to being of maximum-size is close to being determined by one coordinate. We do not know of any purely combinatorial proof of this result.

The results in both theorems above can be extended to other base graphs. Let $\alpha(G)$ denote the maximum possible size of an independent set in a graph G. The following observation determines $\alpha(H^n)$ for any vertex transitive base graph H, in terms of $\alpha(H)$ and $|H|$.

PROPOSITION 1.3. *For any vertex transitive graph H and for any integer $n \geq 1$, if $G = H^n$ then*

$$\frac{\alpha(G)}{|G|} = \frac{\alpha(H)}{|H|} .$$

After the simple proof of this proposition (some special cases of which are proved in [BrNR]), we will provide some examples showing that the above equality does not necessarily hold without the transitivity assumption.

The relevance of graph eigenvalues to independent sets in graphs is well known and can be traced back to the old result that the independence number of any regular graph H on r vertices in which the eigenvalues of the adjacency matrix are $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r$, is at most $-r\mu_r/(\mu_1 - \mu_r)$. A proof of this fact, as well as of the related results on the connection between the Shannon capacity of a graph and its eigenvalues, can be found in [Lo]. This bound is tight for many graphs H including, for example, complete graphs and the Petersen graph. It turns out that the results in Theorem 1.1 and in Theorem 1.2 can be extended to any connected non-bipartite regular base graph H for which the above bound is tight.

Theorem 1.4. *Let H be a connected d-regular graph on r vertices and let $d = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_r$ be its eigenvalues. If*

$$\frac{\alpha(H)}{r} = \frac{-\mu_r}{d - \mu_r} \tag{1}$$

then for every integer $n \geq 1$,

$$\frac{\alpha(H^n)}{r^n} = \frac{-\mu_r}{d - \mu_r} .$$

Moreover, if H is also non-bipartite, and if I is an independent set of size $\frac{-\mu_r}{d-\mu_r} r^n$ in $G = H^n$, then there exists a coordinate $i \in \{1, 2, \ldots, n\}$ and a maximum independent set J in H, such that

$$I = \left\{ v \in V(H)^n : v_i \in J \right\} .$$

REMARK. Note that for any H and n, $\chi(H^n) = \chi(H)$. If H satisfies the conditions of the last theorem and if, in addition, $\chi(H) = r/\alpha(H)$ then every optimal coloring of H^n is induced by a coloring of one of the multiplicands, since it is a partition of H^n into maximum-size independent sets. Such a partition can only be consistent if each color class is induced by the same coordinate. The assumption $\chi(H) = r/\alpha(H)$ holds for many of the interesting classes of graphs to which Theorem 1.4 applies, see subsection 5.1.

Theorem 1.5. *Let H be a d-regular, connected, non-bipartite graph on r vertices, let $d = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_r$ be its eigenvalues and suppose its independence number satisfies* (1). *Then, there exists a constant $M = M(H)$ such that for any $\epsilon > 0$ the following holds. Let $G = H^n$ and let I be an independent set such that $\frac{|I|}{|G|} = \frac{\alpha(H)}{|H|} - \epsilon$. Then there exists an independent set I' with $\frac{|I'|}{|G|} = \frac{\alpha(H)}{|H|}$ such that $\frac{|I' \triangle I|}{|G|} < M\epsilon$.*

The rest of this paper is organized as follows. Section 2 is devoted to Fourier analysis of $\mathbb{Z}_r^n$, where we set up some of the necessary tools. We continue in section 3 with the proofs of several lemmas regarding Boolean functions on $\mathbb{Z}_r^n$. In section 4 we present the proofs of Theorem 1.1 and Theorem 1.2, and suggest a conjecture that strengthens Theorem 1.2 and offers a characterization of all independent sets that occupy a constant proportion of $G = K_r^n$. We also briefly consider the case of infinite dimensional products. Section 5 contains the proof of Proposition 1.3 and the proofs of Theorems 1.4 and 1.5, together with some related examples. These proofs are based on spectral techniques. Although the proofs of Theorems 1.1 and 1.2 can be obtained as special cases, as the eigenvectors of Cayley graphs of Abelian groups (and in particular, of complete graphs) are the characters of the corresponding groups, we believe that presenting them separately is instructive. It is interesting to note that the approach in section 5 suggests that many tools of Harmonic Analysis hold for general tensor products of some fixed orthogonal set of vectors, and not only for characters. Although this fact is not difficult and its proof essentially amounts to a change of basis, it seems very powerful and may well lead to additional interesting consequences. The final section 6 contains some concluding remarks.

2 Fourier Analysis on $\mathbb{Z}_r^n$

In this section we describe all the necessary background concerning discrete Fourier analysis. In addition, we present two extensions of known results for functions $f : \{0,1\}^n \to \{0,1\}$ to the equivalent results for functions $f : \mathbb{Z}_r^n \to \{0,1\}$. The first, Lemma 2.2, is a generalization of the Bonami–Beckner hyper-contractive estimate, and the second, Lemma 2.4, is a generalization of a result proved in [FKN] regarding functions whose Fourier representation is concentrated on subsets of size ≤ 1. We believe that these generalizations may find further uses. The proofs of these lemmas are presented in section 3.

2.1 Background and lemmas. Let $r > 2$ and $G = \{0,\dots,r-1\}^n = \mathbb{Z}_r^n$. Here is some notation we shall use: Let $\bar{0} = (0,0,\dots,0)$, $\bar{1} = (1,1,\dots,1)$, and $e_i = (0,\dots,1,\dots,0)$, the unit vector with 1 at the ith coordinate. For $S \in G$ let

$$|S| = \left|\{i : S_i \neq 0\}\right| .$$

We will occasionally refer to the "levels" of $\{0,\dots,r-1\}^n$: The 0-level contains only $\bar{0}$, the kth level all S such that $|S| = k$.

We think of G both as an Abelian group and as a probability space endowed with the uniform (product) measure μ. For any function $f : G \to \mathbb{C}$, let

$$\int_G f(S)dS \overset{def}{=} \frac{1}{|G|}\sum_{S\in G} f(S) .$$

For any such function f we define the p-norm of f :

$$\|f\|_p = \left(\int |f(S)|^p\right)^{1/p} .$$

The inner product between two functions f and g is

$$\langle f,g\rangle = \int_G f(S)\overline{g(S)}dS = \frac{1}{|G|}\sum_{S\in G} f(S)\overline{g(S)} .$$

We now wish to consider the dual group to G, the group of characters, whose elements form an orthonormal basis for the space of complex-valued functions on G. For any $S \in G$ let $u_S : G \to \mathbb{C}$ be defined by

$$u_S(T) = e^{2\pi i\langle S,T\rangle/r}$$

where $\langle S,T\rangle = \sum_{i=1}^n S_iT_i \bmod r$. It is easy to verify the following properties of the characters:

- They indeed form a group, and the mapping $S \mapsto u_S$ is a homomorphism from G to the multiplicative group of roots of unity:

  $$u_Su_T = u_{S+T}$$

  and

  $$u_{-T} = u_T^{-1} = \overline{u_T} .$$

- Since the sum of the roots of unity is 0 we have that for $S \neq \bar{0}$, $\sum_{T \in G} u_S(T) = 0$. Of course $u_{\bar{0}} \equiv 1$, hence $\frac{1}{|G|} \sum_{T \in G} u_{\bar{0}}(T) = 1$.
- It follows that the characters form an orthonormal basis:
$$\langle u_S, u_T \rangle = \delta_{S,T} .$$
- Every function $f : G \to \mathbb{C}$ has a unique expansion of the form $f = \sum \widehat{f}(S) u_S$, where $\widehat{f}(S) = \langle f, u_S \rangle = \frac{1}{|G|} \sum_{T \in G} f(T) \cdot \overline{u_S(T)}$.
- From orthonormality we get that, for any two functions f, g,
$$\langle f, g \rangle = \sum \widehat{f}(S) \overline{\widehat{g}(S)} ,$$
and in particular we get Parseval's identity,
$$\|f\|_2^2 = \langle f, f \rangle = \sum \left| \widehat{f}(S) \right|^2 .$$

The usage of Fourier analysis in the study of Boolean functions and various questions in combinatorics and computer science has seen much success in the last fifteen years, see for example [AKRS], [BenKS], [Bou2], [F],[FKN], [KKL], [KiS], [LiMN], [T]. All the above examples deal with problems set in $\{0,1\}^n$. In contrast there are relatively few combinatorial papers dealing with $\mathbb{Z}_r^n$ that use Fourier analysis. Some examples of the latter can be found in [BS], [ALM], [M].

A central tool that appears in almost every one of the aforementioned works that deal with $\{0,1\}^n$ is a hyper-contractive estimate due, independently, to Beckner and Bonami, [Be], [Bo]. One version of the Bonami–Beckner inequality that is often used is the following (see also [Bou1]):

LEMMA 2.1 (Beckner (1975), Bonami (1970)). *Let $f : \{0,1\}^n \to \mathbb{R}$ be a function that is a linear combination of $\{u_T : |T| \leq k\}$. Let $p > 2$. Then*
$$\|f\|_p \leq \left(\sqrt{p-1}\right)^k \|f\|_2 .$$

Bonami and Beckner actually proved the dual version of the above lemma, which is equivalent to it. Here we will need a more general version dealing with functions on $\mathbb{Z}_r^n$. On the other hand, as we do not need optimal constants or a statement regarding all p-norms, we will make do with the following statement. Recall that $|T| = |\{i : T_i \neq 0\}|$.

LEMMA 2.2. *For every $r \geq 2$ there exists $C > 0$ such that the following holds. Let $G = \mathbb{Z}_r^n$ and $f : G \to \mathbb{C}$ be a function whose Fourier transform is concentrated on the first $k+1$ levels, that is, f is a linear combination of $\{u_T : |T| \leq k\}$. Then*
$$\|f\|_4 \leq C^k \|f\|_2 .$$

We will use this inequality in the proof of Lemma 2.4 below.

The following two lemmas give conditions on the Fourier transform of Boolean functions f under which f may be determined or approximated by a function depending on only one coordinate, a “dictator”.

LEMMA 2.3. *Let $f : \mathbb{Z}_r^n \to \{0,1\}$ be such that all its Fourier transform is concentrated on the first two levels,*

$$|S| > 1 \Rightarrow \widehat{f}(S) = 0 .$$

Then f is either constant or depends on precisely one coordinate.

The following lemma is a stability version of the above which ultimately enables us to deduce Theorem 1.2 and Theorem 1.5. The special case of this lemma with $G = \{0,1\}^n$ is stated and proven in [FKN].

LEMMA 2.4. *Let $r \geq 2$ and let $C = C(r)$ be the constant from Lemma 2.2. Let $K = 2+32C^8$. Then for any $\varepsilon > 0$ the following holds: let $f : \mathbb{Z}_r^n \to \{0,1\}$ be a function such that $Pr[f = 1] = \alpha$ and furthermore assume that*

$$\sum_{|S|>1} \left|\widehat{f}(S)\right|^2 = \varepsilon .$$

Then there exists a function $g : \mathbb{Z}_r^n \to \{0,1\}$ which depends on at most one coordinate such that

$$\|f - g\|_2^2 < \frac{K}{\alpha - \alpha^2 - \varepsilon}\varepsilon .$$

REMARK. Note that under the above assumptions $\varepsilon \leq \alpha - \alpha^2$, because $\alpha - \alpha^2 - \varepsilon = \sum_{|S|=1} |\widehat{f}(S)|^2$.

3 Generalizing the Boolean Inequalities

In this section we prove several lemmas regarding Boolean functions on $\mathbb{Z}_r^n$. The first, Lemma 2.2, generalizes a hyper-contractive estimate due to Beckner and Bonami [Be], [Bo]. Lemma 2.4 then generalizes a result of Friedgut, Kalai and Naor [FKN] regarding functions whose Fourier representation is concentrated on the levels 0 and 1. We begin by proving Lemma 2.2.

Proof of Lemma 2.2. Following the original proofs of Beckner and Bonami we will prove the dual statement formalized in Claim 3.1 below, using induction on the dimension. For $0 < \varepsilon < 1$ and

$$f(s) = \sum_{j=0}^{r-1} a_j e^{\frac{2\pi ijs}{r}} : \{0,\dots,r-1\} \to \mathbb{C} ,$$

define

$$T_\delta f(s) = a_0 + \delta\bigg(\sum_{j=1}^{r-1} a_j e^{\frac{2\pi ijs}{r}}\bigg).$$

By abuse of notation we will also denote $T_\delta^{\otimes n}$ by T_δ: for any function $f : \{0,\dots,r-1\}^n \to \mathbb{C}$ let $T_\delta f = \sum \widehat{f}(S)\delta^{|S|}u_S$.

CLAIM 3.1. *For every $r \geq 2$ there exists $0 < \delta < 1$ such that, for every $f : \{0,\dots,r-1\}^n \to \mathbb{C}$,*

$$\|T_\delta f\|_4 \leq \|f\|_2 \,.$$

REMARKS. This claim immediately implies Lemma 2.2 with $C = 1/\delta$. Indeed,

$$\|f\|_4 = \|T_\delta(T_{\delta^{-1}}f)\|_4 \leq \|T_{\delta^{-1}}f\|_2 = \Big(\sum_{|S|\leq k} \delta^{-2|S|}\big|\widehat{f}(S)\big|^2\Big)^{1/2}$$

$$\leq (1/\delta)^k\Big(\sum_{|S|\leq k} \big|\widehat{f}(S)\big|^2\Big)^{1/2} = (1/\delta)^k\|f\|_2 \,.$$

Furthermore, we interpret Claim 3.1 as saying that $\|T_\delta\| = 1$ as an operator from L_4 to L_2. As observed by Bonami and Beckner, for the class of operators defined by integration against a kernel (such as T_δ) one has multiplicativity of operator norms. (This can be proven rather easily using Minkowski's inequality.) Therefore one can now use induction on the dimension, and it suffices to prove the one-dimensional statement.

Proof of the claim. Fix $r \geq 2$. Let $\omega = e^{2\pi i/r}$ and let $f(s) = \sum_{j=0}^{r-1} a_j\omega^{sj}$. Clearly we may assume that $a_0 = 1$ and that $|a_1| \geq |a_j|$ for $j \geq 1$. Furthermore, since replacing a_i by $|a_i|$ does not change the norms of neither f nor $T_\delta f$, we may assume that the a_i's are non-negative reals.

Let $1 > \delta > 0$ and $T_\delta f(s) = \sum_{j=0}^{r-1} b_j\omega^{sj}$. Then $b_0 = a_0 = 1$ and $b_j = \delta a_j$ for $j \geq 1$. A simple calculation gives

$$[T_\delta f(s)]^2 = \sum_j \Big[\sum_k b_k b_{j-k}\Big]\omega^{sj}.$$

We then have the following:

$$\|T_\delta f\|_4^4 = \big\|[T_\delta f]^2\big\|_2^2 = \sum_j \Big[\sum_k b_k b_{j-k}\Big]^2$$

whereas

$$\|f\|_2^4 = \Big(\sum a_j^2\Big)^2.$$

Denote $a_1 = a$ and $b_1 = b = \delta a$. Then

$$1 + a^2 + a^4 \leq \|f\|_2^4 \,,$$

and, recalling that $b_1 = b \geq b_j$ for $j \geq 1$,

$$\|T_\delta f\|_4^4 \leq 1 + K[b^2 + b^3 + b^4] \leq 1 + K\delta[a^2 + a^3 + a^4]\,,$$

where K is some constant depending on r (but not on f.) Hence what we need to show is that for sufficiently small δ

$$1 + K\delta[a^2 + a^3 + a^4] \leq 1 + a^2 + a^4$$

holds for all values of a, or equivalently, that

$$(1 - K\delta)a^2 - K\delta a + (1 - K\delta) > 0\,.$$

The discriminant of this quadratic equation is

$$4K^2\delta^2 - (1 - K\delta)^2$$

which is negative for sufficiently small values of δ, which proves the assertion of the lemma. □

We now turn to the proof of Lemma 2.3, which is a softened version of Lemma 2.4.

Proof. The function f is of the form

$$f(S) = a_0 + \sum_{j=1}^{n} \sum_{k=1}^{r-1} a_{j,k} e^{\frac{2\pi i S_j k}{r}}\,.$$

Since f is Boolean we have $f = f^2$. From the uniqueness of the Fourier expansion we may compare the coefficients of $e^{2\pi i S_{j_1} k/r} e^{2\pi i S_{j_2} l/r}$ in f and f^2 and deduce that if $j_1 \neq j_2$ then

$$a_{j_1,k} a_{j_2,l} = 0\,.$$

Therefore there must exist one coordinate, say j_l, such that if $j \neq j_l$ then $a_{j,k} = 0$ for all k. Hence f is of the form

$$f(S) = a_0 + \sum_{k=1}^{r-1} a_k e^{\frac{2\pi i S_l k}{r}}\,.$$ □

Finally we turn to the proof of Lemma 2.4. Our proof follows the lines of one of the proofs presented in [FKN].

Proof of Lemma 2.4. To begin we may assume that $\varepsilon \leq 1/4C^8$ where C is the constant from Lemma 2.2. If this assumption fails then letting $K = 2 + 32C^8$, makes finding a $K\varepsilon$-approximation of f trivial. Let

$$f_S = \sum_{|T| \leq 1} \widehat{f}(T) u_T \quad \text{and} \quad f_L = \sum_{|T| > 1} \widehat{f}(T) u_T\,.$$

(S and L stand for small and large.)

If f_S were Boolean, then the lemma would follow easily. We will show that the same is true for f_S that is close to being Boolean. Let us consider the following function, h, that we use to measure, in a sense, how far f_S is from being Boolean:

$$h = f_S^2 - f_S\,.$$

Note that if f_S were Boolean then h would be identically 0.

We start by computing $\widehat{h}(S)$ on the second level. Let $i \neq j$. Then

$$\widehat{h}(k_1e_i + k_2e_j) = \widehat{f_S^2}(k_1e_i + k_2e_j) - \widehat{f_S}(k_1e_i + k_2e_j) = 2\widehat{f}(k_1e_i)\widehat{f}(k_2e_j)\,, \quad (2)$$

where we have used the fact that $\widehat{f_S}$ is 0 on the second level. The following lemma shows that the L_2 norm of h is small.

Lemma 3.2. *Let C be the constant from Lemma 2.2 and let $\lambda = 32C^8$. Then*

$$E(|h|^2) \leq \lambda\varepsilon\,.$$

Assuming Lemma 3.2 it is now easy to prove that the dictator-approximation of f guaranteed by Lemma 2.4 can be obtained from f by "truncating and rounding":

Corollary 3.3. *There exists $1 \leq j \leq n$ such that the for the function g', defined as follows,*

$$g'(x) = \widehat{f}(\bar{0}) + \sum_{k=1}^{r-1} \widehat{f}(ke_j)e^{\frac{2\pi ikx_j}{r}} \quad (3)$$

it is true that $\|g' - f\|_2^2 \leq \varepsilon\big(1 + \frac{\lambda}{2(\alpha-\alpha^2-\varepsilon)}\big)$.

This corollary now implies Lemma 2.4. It is not hard to see that if $g(y)$ is the function obtained by rounding g' to 0 or 1 according to the majority value of $f(x)$ over all $\{x|x_j = y_j\}$ then $\|g - f\|_2^2 \leq 2\|g' - f\|_2^2 \leq \frac{K}{\alpha-\alpha^2-\varepsilon}\varepsilon$.

Proof of the corollary. Denote $a_i = \sum_{k=1}^{r-1} |\widehat{f}(ke_i)|^2$. Since $E(|h|^2) = \sum |\widehat{h}(T)|^2$ we have, by summing only over T of the form $T = k_1e_i + k_2e_j$ and using (2)

$$\sum_{i<j} a_ia_j = \sum_{i<j} \left[\sum_{k=1}^{r-1} \big|\widehat{f}(ke_i)\big|^2\right]\left[\sum_{k=1}^{r-1} \big|\widehat{f}(ke_j)\big|^2\right] \leq \frac{E(|h|^2)}{4} \leq \frac{\lambda}{4}\varepsilon\,. \quad (4)$$

We also know that $\widehat{f}(\bar{0}) = \alpha$ and that $\sum |\widehat{f}(T)|^2 = \alpha$. This means that

$$\sum_i a_i = \alpha - \alpha^2 - \epsilon\,.$$

Let j be the index for which a_j is maximal.

$$\Big(\sum_i a_i\Big)^2 \le \sum_i a_i^2 + \tfrac{\lambda}{2}\varepsilon \le a_j \cdot \sum_i a_i + \tfrac{\lambda}{2}\varepsilon$$

which means, dividing by $\sum_i a_i{=}\alpha{-}\alpha^2{-}\varepsilon$, that $a_j{\ge}\alpha{-}\alpha^2{-}\varepsilon\big(1{+}\frac{\lambda}{2(\alpha-\alpha^2-\varepsilon)}\big)$. This means that $\|g'\|_2^2 = a_j + \alpha^2 \ge \alpha - \varepsilon\big(1 + \frac{\lambda}{2(\alpha-\alpha^2-\varepsilon)}\big)$ hence most of f's L_2 norm comes from the Fourier transform of g', implying the corollary. □

Proof of Lemma 3.2. We first show that h is 'mostly' smaller than $O(\sqrt{\varepsilon})$, and then use a concentration of measure argument to show that it cannot be much larger on the rest of the domain.

Recall that $f_S = f - f_L$ and that $f^2 - f = 0$, so

$$h = (f_S)^2 - f_S = (f - f_L)^2 - (f - f_L) = f_L^2 + f_L(1 - 2f)\,.$$

Let $k = 2C^4$, and set $Z = \{x \in \mathbb{Z}_r^n \mid |f_L(x)| \le k\sqrt{\varepsilon}\}$. Since $\|f_L\|_2^2 \le \varepsilon$, a Markov argument implies $\Pr_x[Z] \ge 1 - 1/k^2$. Additionally, for every $x \in Z$, $|h(x)| \le 2\,|f_L(x)| \le 2k\sqrt{\varepsilon}$. (Here we have used our assumption that $\varepsilon < 1/k^2 = C^{-8}/4$, see the first sentence of the proof.)

How can $E(|h|^2)$ be much larger than ε? Only if most of the contribution to this expectation comes from values of x outside Z, where h obtains values much larger than $\sqrt{\varepsilon}$. We next use the hyper-contractive estimate (Lemma 2.2) to show that this cannot happen; intuitively because having only low Fourier frequencies means $|h|$ is fairly 'concentrated' around its expectation.

For convenience of notation, let $X = E(|h|^2)$ and $Y = E(|h|^4)$, and denote $p = \Pr_x[x \notin Z] \le 1/k^2$. For every function h, it is true that $X \le \sqrt{Y}$. Since the Fourier support of h is only on the first two levels, Lemma 2.2 implies that $\sqrt{Y} \le C^4 X$. Thus,

$$\begin{aligned} X = E\big(|h|^2\big) &= (1-p)\cdot E\big(|h(x)|^2 \bigm| x \in Z\big) + p \cdot E\big(|h(x)|^2 \bigm| x \notin Z\big) \\ &\le (1-p)4k^2\varepsilon + p\sqrt{E(|h(x)|^4 \mid x \notin Z)} \\ &\le 4k^2\varepsilon + p\sqrt{\tfrac{Y}{p}} \\ &\le 4k^2\varepsilon + \sqrt{p}C^4 X \le 4k^2\varepsilon + \tfrac{1}{2}X \end{aligned}$$

Thus, $X \le 8k^2\varepsilon = 32C^8\varepsilon$. □

4 Independent Sets in K_r^n

In this section we prove Theorems 1.1 and 1.2. Let $G = K_r^n$. Identify the vertices of G with the elements of $\mathbb{Z}_r^n$. Let $D = (\mathbb{Z}_r \setminus 0)^n$. Recall that two vertices $u, v \in \mathbb{Z}_r^n$ are adjacent iff $u - v \in D$. G is thus the Cayley graph of $\mathbb{Z}_r^n$ with the elements of D as generators, and letting $d = |D| = (r-1)^n$ we have that the graph G is d regular. Aside from its group structure we will also think of $\mathbb{Z}_r^n$ as a measure space with the uniform (product) measure, which we shall denote by μ.

For any i between 1 and n and any $0 \leq k \leq r-1$, the set $V_i^k = \{ v \mid v_i = k \}$ is an example of a large independent set, $\mu(V_i^k) = 1/r$. (Although the measure notation is unusual in the combinatorial setting, it will blend in well with our analytical methods in the sequel.) We first observe that this is the maximum size that an independent set can attain.

CLAIM 4.1. *If I is an independent set in G then $\mu(I) \leq 1/r$.*

We now prove the claim by a simple combinatorial argument. However, it is interesting to note that the claim will also follow easily from the analytical proof of Theorem 1.1 that we present below.

Proof of the claim. We will partition the graph into disjoint cliques of size r. The claim will then follow from the fact that I can intersect every clique in at most one vertex.

Let $\bar{i} = (i, i, \ldots, i) \in \mathbb{Z}_r^n$, and let $K = \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{(r-1)}\}$. The claim is proven by observing that $v + K$ is a clique for any $v \in \mathbb{Z}_r^n$ and that

$$\mathbb{Z}_r^n = \bigcup_{(v_2, \ldots, v_n) \in \mathbb{Z}_r^{n-1}} \big((0, v_2, \ldots, v_n) + K \big) .$$

(Note that this is the partition of $\mathbb{Z}_r^n$ into cosets of the subgroup K.) □

Proposition 1.3, proved in section 5, generalizes the above claim for arbitrary regular vertex transitive graphs.

Theorems 1.1 and 1.2 will follow from the next lemma, that is inspired by observations of Håstad [Ha], and is also a special case of the basic spectral approach described in section 5.

LEMMA 4.2. *Let $I \subset \mathbb{Z}_r^n$ be an independent set in G, and let $f : \mathbb{Z}_r^n \to \{0,1\}$ be its indicator function, i.e. $f(x) = 1$ iff $x \in I$. Then,*

$$\sum_S \big| \widehat{f}(S) \big|^2 \left(\frac{-1}{r-1} \right)^{|S|} = 0 .$$

Proof. For any $\tau \in D$ define $f_\tau(x) = f(x + \tau)$. Since I is an independent set, always $f_\tau(x) \cdot f(x) = 0$, so $\langle f, f_\tau \rangle = 0$. Define

$$A(f) = \tfrac{1}{d} \sum_{\tau \in D} f_\tau .$$

A is the averaging operator, replacing $f(x)$ by the average of f on the neighbors of x. The above implies that (and indeed is equivalent to)

$$\langle f, A(f) \rangle = 0\,. \tag{5}$$

Let us now compute the Fourier transform of $A(f)$ in terms of the Fourier transform of f. Recall that for $S \in \mathbb{Z}_r^n$ we denote $|S| = |\{ i \mid S_i \neq 0 \}|$,

Claim 4.3.

$$\widehat{A(f)}(S) = \widehat{f}(S) \left(\frac{-1}{r-1} \right)^{|S|}.$$

Proof. We first claim that $\widehat{f_\tau}(S) = \widehat{f}(S) u_\tau(S)$. Indeed,

$$\begin{aligned}\widehat{f_\tau}(S) = \int f(x+\tau)\overline{u_S(x)}dx &= \int f(x)\overline{u_S(x-\tau)}dx \\ &= \int f(x)\overline{u_S(x)}u_S(\tau)dx = \widehat{f}(S)u_\tau(S)\,.\end{aligned}$$

Recall that $D = \{1, \dots, r-1\}^n$, $d = |D|$, and denote $\omega = e^{2\pi i/r}$. We have

$$\begin{aligned}\widehat{A(f)}(S) &= \tfrac{1}{d} \sum_{\tau \in D} \widehat{f_\tau}(S) = \tfrac{1}{d} \widehat{f}(S) \sum_{\tau \in D} u_\tau(S) \\ &= \tfrac{1}{d} \widehat{f}(S) \prod_{j=1}^{n} \sum_{k=1}^{r-1} \omega^{kS_j} \\ &= \frac{1}{d} \widehat{f}(S) \prod_{j:S_j=0} (r-1) \prod_{j:S_j \neq 0} (-1) = \widehat{f}(S) \left(\frac{-1}{r-1} \right)^{|S|}\end{aligned}$$

where in the last equality we have used $d = (r-1)^n$. The proof of the lemma is now completed by orthogonality, writing

$$0 = \langle f, A(f) \rangle = \sum_S \widehat{f}(S) \overline{\widehat{A(f)}(S)} = \sum_S \left| \widehat{f}(S) \right|^2 \left(\frac{-1}{r-1} \right)^{|S|}. \qquad \square$$

Proof of Theorems 1.1 and 1.2. Denote $\alpha = \mu(I)$. Now, from Parseval we have

$$\sum \left| \widehat{f}(S) \right|^2 = \alpha\,.$$

Also,

$$\widehat{f}(\bar{0}) = \int f(x)dx = \alpha\,.$$

Hence

$$\sum_{S \neq \bar{0}} \left| \widehat{f}(S) \right|^2 = \alpha - \alpha^2 \tag{6}$$

and subtracting $|\widehat{f}(0)|^2 = \alpha^2$ from the identity of Lemma 4.2,

$$\sum_{S\neq\bar{0}} \left|\widehat{f}(S)\right|^2 \left(\frac{-1}{r-1}\right)^{|S|} = -\alpha^2 \,. \tag{7}$$

Comparing (6) and (7) now yields valuable information on the Fourier transform of f. It is convenient to interpret this information via the following probability distribution: Let T be a random variable taking values in $G \setminus \bar{0}$ with

$$Pr[T = S] = \frac{|\widehat{f}(S)|^2}{\alpha - \alpha^2} \,.$$

Let

$$X = X(T) = \left(\frac{-1}{r-1}\right)^{|T|} \,.$$

Then (7) now gives that

$$E(X) = \frac{-\alpha^2}{\alpha - \alpha^2} = \frac{-\alpha}{1-\alpha} \,.$$

We now break our analysis into three cases

- $\alpha > 1/r$. We will see that this cannot happen, which implies Claim 4.1.
- $\alpha = 1/r$. The analysis of this case will imply Theorem 1.1.
- $\alpha = 1/r - \varepsilon$. This case leads to the proof of Theorem 1.2.

For all three cases it is important to note that for all T, $X(T) = X \geq \frac{-1}{r-1}$ with equality iff $|T| = 1$.

1. Let $\alpha > \frac{1}{r}$. Then $E(X) = \frac{-\alpha}{1-\alpha} < \frac{-1}{r-1}$. However this is impossible, because $X \geq \frac{-1}{r-1}$. Hence if I is an independent set then $\mu(I) \leq \frac{1}{r}$. Thus we have just proven Claim 4.1 (again).
2. Next, if $\alpha = 1/r$ then $E(X) = \frac{-1}{r-1}$ which implies that $X \equiv \frac{-1}{r-1}$ meaning that all the support of $|\widehat{f}(S)|^2$ must be concentrated on u_S with $|S| = 1$. Hence f has all its Fourier transform concentrated on the first two levels. According to Lemma 2.3 this implies that f depends on one coordinate only, and hence, since f is not constant, it is the indicator function of $S_l = k$ for some l and k, as asserted by Theorem 1.1.
3. Finally we analyze the case where α is slightly less than $1/r$. This means that $E(X)$ is very close to $\frac{-1}{r-1}$, its minimal value. Observing that for S, $|S| > 1$,

$$X(S) \geq \frac{-1}{(r-1)^3} > \frac{-1}{r-1},$$

this implies that almost all the weight of $|\widehat{f}(S)|^2$ is concentrated on S, $|S| \le 1$.
Formally, let

$$Y = X + \frac{1}{r-1} \ge 0\,.$$

When $Y > 0$ then $Y \ge \frac{-1}{(r-1)^3} + \frac{1}{r-1} = \frac{r(r-2)}{(r-1)^3}$. Hence, by Markov's inequality,

$$\text{Prob}[Y > 0] \le E(Y) \cdot \frac{(r-1)^3}{r(r-2)}\,. \tag{8}$$

Now, if $\alpha > 1/r - \varepsilon$ then

$$E(Y) = E(X) + \frac{1}{r-1} < \frac{\varepsilon r^2}{(r-1)(r\varepsilon + r - 1)}\,,$$

which, using (8) together with $r \ge 3$, implies

$$\text{Prob}[Y > 0] \le \frac{\varepsilon r}{r\varepsilon + r - 1} \cdot \frac{(r-1)^2}{(r-2)} \le 2r\varepsilon\,. \tag{9}$$

Recalling that

$$Y(S) > 0 \ \text{ iff } \ X(S) > \frac{-1}{r-1} \ \text{ iff } \ |S| > 1\,,$$

the upshot of (9) is that $\forall\varepsilon > 0$ and for every independent set I with $\mu(I) > 1/r - \varepsilon$ we have

$$\sum_{|S|>1} \left|\widehat{f}(S)\right|^2 = (\alpha - \alpha^2)Prob[Y > 0] \le 2\varepsilon\,,$$

where f is the characteristic function of I. We now deduce directly from Lemma 2.4 that there is some 'dictator' function g that depends on at most one coordinate, and such that $\|f-g\|_2^2 < \frac{K}{\alpha-\alpha^2-\varepsilon} \cdot 2\varepsilon = O(\varepsilon)$. (Here, if ε is not sufficiently small, $g \equiv 0$ is also a possibility, but if ε is such that $\|f\|_2^2 - \frac{K}{\alpha-\alpha^2-\varepsilon} \cdot 2\varepsilon > 0$ then g must be a dictator.) □

4.1 Infinite weak products. Let A be an infinite set and consider $G = \times^A K_r$, the infinite weak product of K_r. Let μ be the product measure on A, the measure generated by the cylinders of finite codimension. The set $\{u_S\}_{|S|<\infty}$ is a basis for the measurable functions, hence the same proof technique as before yields the following:

Theorem 4.4. *Let $I \subseteq \times^A K_r$ be a measurable independent set with $\mu(I) = 1/r$. Then there exists $j \in A$ and a set $J \subset G$ of the form*

$$J = \{S : S_j = k\}$$

with

$$\mu(J \triangle I) = 0\,.$$

REMARKS. • Alternatively one can deduce the infinite dimensional statement from the finite case using the stability statement of Theorem 1.2. It seems impossible to make this deduction directly without the stability theorem.

• In [GL], an example due to Pósa is presented that shows, using ultrafilters, an optimal coloring of G which is not of the form $\sigma(S) = S_j$. Of course, the color classes thus formed must be non-measurable.

4.2 Large independent sets in weak products. Theorem 1.1 gives a complete characterization of independent sets of measure $1/r$. Theorem 1.2 also characterizes independent sets of size close to $1/r$. A natural question is whether smaller independent sets, of size, say, $1/(100r)$ still have some simple structure. It seems quite hopeless to approach this question using traditional combinatorial arguments. We now present a conjecture regarding such sets, that may be susceptible to Fourier methods.

Let us start with an example. Consider $G = \times^n K_3$ over vertices $\{0,1,2\}^n$. Let

$$I = \{S : \text{ at least two of the three coordinates } S_1, S_2, S_3 \text{ are equal to } 0\}\,.$$

Clearly, if $S, T \in I$ then they must agree on one of the first three coordinates, hence I is an independent set, with $\mu(I) = 7/27$. A naive conjecture would be that any reasonably large independent set is similarly determined by few coordinates. However, this is clearly too strong an assumption, since a random subset of I is also an independent set. Moreover, this problem does not go away if we consider only maximal independent sets (with respect to containment). Indeed consider any 'nice' independent set I_0, e.g. the one above with the following incremental modification. At each step select some random vertex $v \in \{0,1,2\}^n$ and let I_t be an arbitrarily chosen maximal independent set containing $(I_{t-1} \cup \{v\}) \setminus \Gamma(v)$ where $\Gamma(v)$ are the neighbors of v in G. By definition, at each step I_t remains a maximal independent set, and after sufficiently many steps its size is still about half that of I, and it certainly is not determined by only a few coordinates.

Thus, we conjecture that every large independent set is close to *being contained* in a set determined by few coordinates. Sets of the form $\{S : S_j = k\}$ are sometimes referred to as dictatorships (dictated by the jth coordinate.) Analogously, we call S a k-junta if its characteristic function depends on at most k different coordinates.

CONJECTURE 4.5. *There exists a function $M = M(r,\alpha,\varepsilon)$ (note that M does not depend on n!), such that for every integer $r \geq 3$ and for every $\alpha > 0$ if $I \subseteq K_r^n$ is an independent set with $\mu(I) > \alpha$ then for any $\varepsilon > 0$ there is an M-junta J which is itself an independent set such that*

$$\mu(I \setminus J) < \varepsilon ,$$

i.e. I is almost contained in a junta-type independent set.

We believe that this conjecture, if proven, may be quite useful in resolving the hardness of approximate-coloring 3-colorable graphs, an infamously open hardness of approximation problem. We outline the connection roughly as follows. Following the PCP approach we begin with some NP-hard constraint satisfaction problem (CSP) and construct from a given CSP instance a 3-colorable graph. This graph would consist of many copies of $(K_3)^n$ that are further interconnected, such that each copy represents a variable in the CSP taking values in $\{1,\dots,n\}$. One would use the above conjecture to prove that a given coloring of this graph corresponds, in each copy of $(K_3)^n$, to a 'commitment' assigning some constant number of values in $[n]$ to that variable (these are the coordinates of the junta). Then, the structure of the graph would be used to show that these values must satisfy the constraints of the CSP we started out with.

This scheme has been successfully used with similar approximation problems such as vertex-cover [DS], where similarly structured graphs containing many interconnected copies of $\{0,1\}^n$ have been constructed, and combinatorial statements regarding the 'juntativity' of subsets of $\{0,1\}^n$ were essential. One must add that it is not altogether clear how to construct such a reduction and such a graph in the 3-colorability case, so proving the above conjecture is by no means the last step of completing such a proof.

Just as the generalization of the results from [FKN], which characterize Boolean functions that are close to dictatorships, are essential for proving Theorem 1.2, we believe that for the proof of the above conjecture it will be necessary to generalize the junta-type results of [Bou2] and [KiS]. First steps towards this goal appear in [DFK].

5 Independent Sets in H^n

We begin this section with the simple proof of Proposition 1.3. Let H be a vertex-transitive graph on a set $V = \{v_1,\dots,v_r\}$ of r vertices, let A be the automorphism group of H, and let S be an independent set of maximum size in $G = H^n$. Let $a_1, a_2, \ldots, a_n$ be n random elements of A, where each a_i is chosen, randomly, independently and uniformly, among the members of A. For each i, $1 \leq i \leq r$, let x_i be the random vertex of G given by $x_i = (a_1(v_i), a_2(v_i), \ldots, a_n(v_i))$. It is easy to see that x_i is uniformly distributed among all vertices of G. Therefore, the expected size of the intersection $S \cap \{x_1, \ldots, x_r\}$ is precisely $|S|\frac{r}{|G|}$. Note, however, that the induced subgraph of G on $\{x_1, \ldots, x_r\}$ is a copy of H, and therefore the size of the intersection is at most $\alpha(H)$ for every choice of $a_1, a_2, \ldots, a_n$. We conclude that $|S|\frac{r}{|G|} \leq \alpha(H)$, and as $|S| = \alpha(G)$ this implies that

$$\frac{\alpha(G)}{|G|} \leq \frac{\alpha(H)}{|H|},$$

as needed. □

Remark. The transitivity assumption is needed in the above proposition, that is, its assertion fails for some connected regular base graphs H. Indeed, consider, for example, the following d-regular graph H. Its vertex set is the disjoint union of three sets, A, B and C, where $|A| = |B| = d$ and $|C| = d + 1$. Let a, b, c_1, c_2 be distinguished elements of the corresponding sets, and let the edge set of H consist of all edges between A and B except ab, all internal edges in C except c_1c_2, and the two additional edges ac_1 and bc_2. It is easy to check that $\alpha(H) = d+1$. For each $n \geq 1$, let $m_n = \alpha(H^n)/|H^n|$ denote the maximum possible measure of an independent set in H^n. Thus $m_1 = \frac{d+1}{3d+1}$. We claim that for $n > 1$,

$$m_n \geq \frac{d}{3d+1} + \frac{d}{3d+1} m_{n-1}.$$

To prove this claim, note that the set of vertices of H^n whose first coordinate is a member of A (with all possible extensions to the rest of the coordinates), together with the set of all vertices of H^n whose first coordinate is in $C - \{c_1\}$, with the last $n-1$ coordinates belonging to a maximum size independent set in H^{n-1}, form an independent set in H^n. This proves the claim and implies that

$$m_n \geq \frac{d}{3d+1} + \left(\frac{d}{3d+1}\right)^2 + \cdots + \left(\frac{d}{3d+1}\right)^{n-1} + \left(\frac{d}{3d+1}\right)^{n-1} \frac{d+1}{3d+1}.$$

For any fixed d, the above lower bound tends to $\frac{d}{2d+1}$ as n tends to infinity, showing that the graph H does not satisfy the assertion of Proposition 1.3.

We proceed with the proofs of Theorems 1.4 and 1.5. These are based on some basic facts about graph eigenvalues and eigenvectors. The relation between the spectral properties of a graph and its structural properties has been studied extensively, cf., e.g. [A1], [KrS] and its references.

The *adjacency matrix* of a graph $H=(V,E)$ is the matrix $A=(a_{u,v})_{u,v\in V}$, in which $a_{u,v}=1$ if $uv\in E$ and $a_{u,v}=0$ otherwise. This matrix is symmetric and hence has real eigenvalues $\mu_1\geq\ldots\geq\mu_r$ (called the eigenvalues of H) and an orthonormal basis of eigenvectors $u_1,\ldots,u_r$. If H is connected and d-regular, then $\mu_1=d$, and the corresponding (normalized) eigenvector is $u_1=\frac{1}{\sqrt{r}}(1,1,\ldots,1)^t$. By the Perron–Frobenius theorem $d\geq|\mu_i|$ for all i, and $\mu_2<d$ iff H is connected, whereas $\mu_r>-d$ iff H is non-bipartite. Note that the trace of A is 0, as H has no loops, and therefore if H has edges, then μ_1 is strictly positive and μ_r is strictly negative.

The adjacency matrix of the weak power H^n of H is the tensor n-th power of A. Its eigenvalues are all products $\prod_{i=1}^n\mu_{g(i)}$, as g ranges over all r^n functions mapping $\{1,2,\ldots,n\}$ into $\{1,2,\ldots,r\}$. The corresponding orthonormal basis of eigenvectors is obtained by taking all tensor products of the eigenvectors u_i.

LEMMA 5.1. *Let $G=(V,E)$ be a D-regular, connected graph on m vertices, let $D=\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_m$ be its eigenvalues, and let $v_1,v_2,\ldots,v_m$ be the corresponding orthonormal basis of eigenvectors, where $v_1=\frac{1}{\sqrt{m}}(1,1,\ldots,1)^t$. Let I be an independent set in G, let f be its characteristic function, and suppose $f=\sum_{i=1}^m c_iv_i$.*

(i) *The size of I satisfies*
$$|I|\leq\frac{-\lambda_m m}{D-\lambda_m}\,,\tag{10}$$
and equality holds if and only if f is a linear combination of the eigenvectors corresponding to the eigenvalues λ_1 and λ_m.

(ii) *If $\sum_{i:\lambda_i\neq\lambda_1,\lambda_m}c_i^2=\delta m$ and $\lambda=\min\{\lambda_i:\lambda_i\neq\lambda_m\}$ then*
$$\delta\leq\frac{(D-\lambda_m)}{(\lambda-\lambda_m)}\frac{|I|}{m}\left(\frac{-\lambda_m}{D-\lambda_m}-\frac{|I|}{m}\right).\tag{11}$$

Proof. Let A be the adjacency matrix of G. Since f is the characteristic vector of an independent set, $f^tAf=0$. As the vectors v_i form an orthonormal basis, $f^tAf=\sum_{i=1}^m c_i^2\lambda_i$, $\sum_{i=1}^m c_i^2=f^tf=\sum_{v\in V}f^2(v)=|I|$, and $c_1=f^tv_1=|I|/\sqrt{m}$. It thus follows that
$$\begin{aligned}0=f^tAf=\sum_{i=1}^m c_i^2\lambda_i=Dc_1^2+\sum_{i=2}^m c_i^2\lambda_i&\geq D\tfrac{|I|^2}{m}+\lambda_m\sum_{i=2}^m c_i^2\\&=D\tfrac{|I|^2}{m}+\lambda_m\left(|I|-\tfrac{|I|^2}{m}\right).\end{aligned}$$

Therefore,

$$-\lambda_m|I| \geq \tfrac{|I|^2}{m}(D-\lambda_m)\,,$$

implying (10). Moreover, equality holds if and only if $c_i = 0$ whenever λ_i is neither $D = \lambda_1$ nor λ_m. This establishes the assertion of part (i).

The proof of part (ii) is very similar. Here

$$\begin{aligned}0 = f^tAf = \sum_{i=1}^{m} c_i^2\lambda_i &= D\tfrac{|I|^2}{m} + \lambda_m\left(|I| - \tfrac{|I|^2}{m} - \delta m\right) + \sum_{i:\lambda_i\neq\lambda_1,\lambda_m} c_i^2\lambda_i \\ &\geq D\tfrac{|I|^2}{m} + \lambda_m\left(|I| - \tfrac{|I|^2}{m} - \delta m\right) + \delta m\lambda\,.\end{aligned}$$

Therefore,

$$(D-\lambda_m)|I|\left(\frac{-\lambda_m}{D-\lambda_m} - \frac{|I|}{m}\right) = -\lambda_m|I| - \frac{D-\lambda_m}{m}|I|^2 \geq \delta m(\lambda-\lambda_m)\,,$$

implying (11). □

Proof of Theorem 1.4. Put $G = H^n$ and let $m = r^n$ be the number of vertices of G. By the discussion preceding Lemma 5.1, the eigenvalues of G are all products of eigenvalues of H and the corresponding eigenvectors are the tensor products of the eigenvectors of H. In particular, the largest eigenvalue of G is its degree $\lambda_1 = D = d^n$, and its smallest (most negative) eigenvalue is $\lambda_m = \mu_r d^{n-1}$. Therefore, by Lemma 5.1, part (i),

$$\alpha(G) \leq \frac{-\lambda_m m}{D-\lambda_m} = \frac{-\mu_r d^{n-1}r^n}{d^n - \mu_r d^{n-1}} = \frac{-\mu_r r^n}{d-\mu_r}\,,$$

as needed. Moreover, if I is an independent set of size $\frac{-\mu_r r^n}{d-\mu_r}$ in G then, by Lemma 5.1 part (i), the characteristic function f of I is a linear combination of the eigenvectors corresponding to the eigenvalues λ_1 and λ_m. Since H is connected and non-bipartite the multiplicity of the eigenvalue $\lambda_1 = d^n$ of G is one, and its only eigenvector is the vector $\frac{1}{r^{n/2}}(1,1,\dots,1)^t$ obtained by taking the tensor n-th power of the first eigenvector of H. Similarly, the only eigenvectors corresponding to the smallest eigenvalue $\lambda_m = \mu_r d^{n-1}$ are the tensor products of one eigenvector of H corresponding to its smallest eigenvalue, with $n-1$ copies of the vector $\frac{1}{\sqrt{r}}(1,1,\dots,1)^t$. Let $u_1, u_2, \dots, u_r$ be an orthonormal basis of eigenvectors of H, where u_i corresponds to the eigenvalue μ_i. In particular, $u_1 = \frac{1}{\sqrt{r}}(1,1,\dots,1)^t$. For each sequence $S \in \{1,2,\dots,r\}^n$, let v_S denote the tensor product $\prod_{i=1}^n u_{S_i}$. Define, also $|S| = |\{i : S_i \neq 1\}|$. In addition, if $|S| = 1$, let $i(S)$ denote the unique index i such that $S_i \neq 1$. The vectors v_S form an orthonormal basis of eigenvectors of G, and by the above discussion the vector f is a linear combination of vectors v_S with $|S| \leq 1$. Let $f = \sum_{|S|\leq 1} \alpha_S v_S$ be this combination.

Imitating the proof of Lemma 2.3 we next show that since f is Boolean, all the vectors S for which $|S| = 1$ and $\alpha_S \neq 0$ have the same value of $i(S)$. Indeed, since f is Boolean, $f^2 = f$. However, $f^2 = \sum_{S,S'} \alpha_S \alpha_{S'} v_S v_{S'}$, where here $v_S v'_S$ denotes the vector whose coordinates are the products of the corresponding coordinates of v_S and $v_{S'}$. It is easy to check that if $|S| = |S'| = 1$ and $i(S) = i(S')$, then $v_S v_{S'}$ is a linear combination of the vectors v_T with $i(T) = i(S)$ and the constant vector. It is also easy to check that for the unique sequence $S = \overline{1} = (1, 1, \ldots, 1)$, and for any S' with $|S'| = 1$, $v_{\overline{1}} v_{S'}$ is a scalar multiple of $v_{S'}$. On the other hand, if $|S| = |S'| = 1$ and $i(S) \neq i(S')$, then $v_S v_{S'}$ is a multiple of the vector v_T, where T coincides with S in coordinate number $i(S)$, with S' in coordinate number $i(S')$, and with both in all other coordinates (in which they are both 1). In particular, $|T| = 2$. By the unique representation of f as a linear combination of the members of the basis v_S, it thus follows that all vectors S for which $\alpha_S \neq 0$ and $|S| = 1$ have the same value of $i(S)$. In other words, f is determined by a single coordinate, that is, there is a coordinate i such that for each vertex $x = (x_1, x_2, \ldots, x_n)$ of H^n, x lies in I depending only on the value of x_i. It follows that all the vertices that appear in the i-th coordinate of members of I form an independent set of H, and as this set has to be of maximum size, the assertion of the second part of Theorem 1.4 follows. □

Proof of Theorem 1.5. Let f be the characteristic vector of I, put $D = d^n$, $m = r^n$ and let $\lambda_1 = D \geq \lambda_2 \geq \cdots \geq \lambda_m$ be the eigenvalues of G. Define also $\lambda = \min\{\lambda_i : \lambda_i \neq \lambda_m\}$. Let $u_1, \ldots, u_r$ as well as $\{v_S : S \in \{1, 2, \ldots, r\}^n\}$ be, as in the previous proof, the eigenvectors of H and of G, respectively. Let $f = \sum c_S v_S$ be a representation of f as a linear combination of the vectors v_S, and define δ by

$$\delta m = \sum_{|S|>1} c_S^2 \,.$$

By Lemma 5.1, part (ii), we conclude that

$$\delta \leq \frac{(D - \lambda_m)}{(\lambda - \lambda_m)} \frac{|I|}{m} \epsilon \,. \tag{12}$$

It is not difficult to express λ as the minimum between three possible eigenvalues of G. If $\mu = \min\{\mu_i : \mu_i > \mu_r\}$, then $\lambda = \min\{\mu d^{n-1}, \mu_r \mu_2 d^{n-2}, \mu_r^3 d^{n-3}\}$. In any case, since $\mu_2 < d$ and $\mu_r > -d$, there exists some $\eta = \eta(H) > 0$ (which depends only on the base graph H, and not on n), such that $\lambda = (\mu_r + \eta) d^{n-1}$. Substituting in (12) and recalling that $D = d^n$ and

$\lambda_m = \mu_r d^{n-1}$ we conclude that

$$\delta \leq \frac{(d-\mu_r)}{\eta} \frac{|I|}{m} \epsilon \, .$$

Now, for any function g let $(\widehat{g})$ be the vector of coefficients of its expansion according to the characters u_S described in section 2, and $(\bar{g})$ the vector of coefficients of its expansion according to the set v_S. Since each of these two sets is, up to a multiplicative constant, an orthonormal basis the transformation between these two expansions is norm-preserving up to a multiplicative constant:

$$m \|(\widehat{g})\|_2^2 = \|(\bar{g})\|_2^2 \, .$$

Furthermore, what is crucial to us, is that this transformation is norm-preserving levelwise: for any $0 \leq k \leq n$ the l_2 norm of $\widehat{g}$ on the k'th level is precisely $\sqrt{m}$ times the l_2 norm of $\bar{g}$ on the k'th level:

$$m \sum_{|S|=k} \left|\widehat{g}(S)\right|^2 = \sum_{|S|=k} \left|\bar{g}(S)\right|^2 .$$

This is true because for any such k

$$\operatorname{Span}\left(\{u_S : |S| = k\}\right) = \operatorname{Span}\left(\{v_S : |S| = k\}\right) ,$$

giving a simultaneous level-wise decomposition of the space of functions into orthogonal subspaces.

Returning to our function f we perform this change of basis, and express f as a linear combination of the characters u_S and conclude that

$$\sum_{|S|>1} \left|\widehat{f}(S)\right|^2 = \frac{\sum_{|S|>1} c_S^2}{m} = \frac{\delta m}{m} \leq \frac{(d-\mu_r)}{\eta} \frac{|I|}{m} \epsilon \, .$$

The desired result now follows from Lemma 2.4. □

5.1 Examples. There are various examples of graphs H which satisfy the assumptions of Theorems 1.4 and 1.5. These include complete graphs, line graphs of regular graphs which contain a perfect matching, Kneser graphs, some strongly regular graphs and appropriate classes of random regular graphs. Here are the details:

- Let $k \geq 3$ and let $G = (V, E)$ be a k-regular graph on n vertices which contains a perfect matching, i.e. a set of $n/2$ pairwise disjoint edges. The line graph $L(G)$ of G has a vertex for every edge of G and two vertices of $L(G)$ are adjacent if as edges of G they share a vertex. By definition, $L(G)$ is a $d = 2(k-1)$-regular graph on $r = nk/2$ vertices. Since an independent set in $L(G)$ corresponds to a matching in G

we have that $\alpha(L(G)) = n/2$. As noted by A. Hoffman, the smallest eigenvalue of the adjacency matrix $A_{L(G)}$ of $L(G)$ is easy to compute. Let $B = (b_{v,e})$ be the vertex-edge incidence matrix which is defined as $b_{v,e} = 1$ iff $v \in e$. Then $A_{L(G)} = B^T B - 2I$. Since $B^T B$ is positive semidefinite, all its eigenvalues are non-negative. Hence, the eigenvalues of $A_{L(G)}$ are ≥ -2. Moreover, we have that the rank of $B^T B$ satisfies $r(B^T B) = r(B) \leq |V(G)| < |E(G)|$. So, $B^T B$ has at least one 0 eigenvalue. This implies that the smallest eigenvalue of $A_{L(G)}$ is $\mu_r = -2$. Now it is easy to check that $L(G)$ satisfies the assumptions of Theorems 1.4 and 1.5. If, in addition, G is class 1-Vizing (that is, it is k-edge colorable), then $H = L(G)$ satisfies the assumption in the remark following Theorem 1.4, and hence any optimal coloring of H^n is determined by one of the factors.

- Let t, k be two positive integers, $t > 2k$. The *Kneser graph* $K(t,k)$ is the graph whose vertices are all k-subsets of a t-element set, where two vertices are adjacent iff the corresponding sets are disjoint. The number of vertices of this graph is clearly $r = \binom{t}{k}$, and it is well known that its eigenvalues are all numbers of the form $(-1)^j \binom{t-k-j}{k-j}$, for $j \in \{0, 1, \dots, k\}$ (see, e.g. [Lo] for a proof.) In particular, the largest eigenvalue is the degree of regularity $d = \mu_1 = \binom{t-k}{k}$ and the smallest is $\mu_r = -\binom{t-k-1}{k-1}$. The independence number of $K(t,k)$ is the maximum possible cardinality of an intersecting family of k-subsets of a t-element set. This maximum is determined by the Erdős–Ko–Rado theorem [EKR], and is the number of all k-sets containing a single element, that is
$$\binom{t-1}{k-1} = \frac{\binom{t-k-1}{k-1}\binom{t}{k}}{\binom{t-k}{k} + \binom{t-k-1}{k-1}} = \frac{-\mu_r r}{d - \mu_r}\,.$$
Therefore, for each admissible t and k, $H = K(t,k)$ satisfies the assumptions of Theorems 1.4 and 1.5. In particular, $K(5,2)$, which is the Petersen graph, satisfies these assumptions.
- Let a be an even integer and let $q = p^a$ be a prime power which is congruent to 1 modulo 4 so that -1 is a square in the finite field $GF(q)$. Let P_q be the graph whose vertices are all elements of $GF(q)$ and two vertices are adjacent if and only if their difference is a quadratic residue in $GF(q)$. This graph is usually called the *Paley graph.* It is easy to see that P_q is $d = (q-1)/2$-regular. To compute the eigenvalues of P_q one can use the fact that it is strongly regular, i.e. the number of common neighbors of any two vertices of it has the same

value $-(q-5)/4$ if they are adjacent, and the same value $-(q-1)/4$ if they are not. This implies that the smallest eigenvalue of the adjacency matrix of P_q is $\mu_q = -(\sqrt{q}+1)/2$ (see, e.g. [KrS]). It is easy to see that all elements of the subfield $GF(\sqrt{q}) \subset GF(q)$ are quadratic residues in $GF(q)$. This implies that for every quadratic non-residue $\beta \in GF(q)$ all elements of any multiplicative coset $\beta GF(\sqrt{q})$ form an independent set of size $\sqrt{q}$. Hence for this graph we get

$$\alpha(P_q) = \sqrt{q} = \frac{(\sqrt{q}+1)/2}{(q-1)/2 + (\sqrt{q}+1)/2} q = \frac{-\mu_q}{d-\mu_q} q\,.$$

The graph P_q also satisfies the condition in the remark following Theorem 1.4. Indeed, it is easy to see that the sets $\{\alpha + \beta GF(\sqrt{q}) : \alpha \in GF(\sqrt{q})\}$ form a partition of of its set of vertices into $\sqrt{q}$ independent sets.

- Another construction of a strongly regular graph which satisfies the assumptions of Theorems 1.4 and 1.5 is due to Delsarte and Goethals and to Turyn (see, e.g. [KrS]). Let q be a prime power and let $V(H)$ be the elements of the two dimensional vector space over $GF(q)$. Thus H has $r = q^2$ vertices. Partition the $q+1$ lines through the origin of the space into two sets P and N, where $|P| = k$. Two vertices x and y of the graph H are adjacent if $x - y$ is parallel to a line in P. It is easy to check that H is $d = k(q-1)$-regular. Since all the points on any line from P form an independent set in H, we have $\alpha(H) \geq q$ and $\chi(H) \leq q$. It is known that the smallest eigenvalue of H is $\mu_r = -k$. Therefore, by Lemma 5.1, we obtain that

$$q \leq \alpha(H) \leq \frac{-\mu_r}{d-\mu_r} r = \frac{k}{k(q-1)+k} q^2 = q\,,$$

showing that H satisfies the assumptions of Theorems 1.4 and 1.5, as well as that of the remark between them.

- Let $k \geq 3$ be a fixed integer, let r be an integer divisible by k, and let $d < r(k-1)/k$ be another integer, divisible by $k-1$. Let H be a random k-colorable graph obtained as follows. The vertex set of H is the union of k pairwise disjoint sets $A_1, A_2, \ldots, A_k$, each of size r/k. The edges of H are the union of $\frac{d}{k-1}\binom{k}{2}$ random matchings M_{ijs}, where for $1 \leq i < j \leq k$, and $1 \leq s \leq d/(k-1)$, M_{ijs} is a random perfect matching between A_i and A_j. The resulting graph H is obviously d-regular, and it is not difficult to see that $-d/(k-1)$ is an eigenvalue of it with multiplicity at least $k-1$. In fact, each vector which is constant on the members of each A_i, where the sum of these k constants is 0, is an eigenvector of this eigenvalue. As shown in [AK], for any fixed k, and all $d > d_0(k)$ and $r > r_0(k, d)$, all the other eigenvalues of H besides d and $-d/(k-1)$ are in absolute value at most $O(\sqrt{d})$ with high probability. It follows that with high probability H satisfies the assumptions of Theorems 1.4 and 1.5, as well as the assumption in the remark between them. Another, similar class of graphs satisfying these assumptions can be constructed by replacing each bipartite graph between A_i and A_j in the construction above by an appropriate $d/(k-1)$-regular pseudo-random graph. We omit the detailed description.

6 Concluding Remarks

– The weak product considered here is similar to the AND product considered in the definition of the Shannon capacity of graphs. The n-th *AND power* of an undirected graph $H = (V, E)$ is the graph denoted by $H^{\wedge n}$ whose vertex set is V^n in which distinct vertices $(x_1 \dots x_n)$ and $(x'_1 \dots x'_n)$ are connected if $\{x_i, x'_i\} \in E$ for all $i \in \{1, 2, \dots, n\}$ such that $x_i \neq x'_i$. The *Shannon capacity* of H is the limit of $(\alpha(H^{\wedge n})\ )^{1/n}$, as n tends to infinity. Although the definition here is very similar to that of a weak product, the investigation of the independence numbers of these powers is far more complicated than that of the independence numbers in weak powers. In fact, even the Shannon capacity of a cycle of length 7 is not known. See [Lo], [A2] and their references for more details.

– The n-th (*sparse*) *power* of an undirected graph $H = (V, E)$, denoted here by H^{*n}, is the graph whose vertex set is V^n in which distinct vertices $(x_1 \dots x_n)$ and $(x'_1 \dots x'_n)$ are connected iff there exists a single index i such that $x_j = x'_j$ for all $j \neq i$ and x_i and x'_i are connected in H. There are several papers dealing with the asymptotic behaviour of the independence number of H^{*n}, for a fixed graph H. See, e.g. [HHP] and its references. In particular, it is known that for every fixed graph H, the limit $\lim_{n \mapsto \infty} \alpha(H^{*n})/|V|^n$ exists, and is at least $1/\chi(H)$ and at most the reciprocal of the fractional chromatic number of H. Moreover, for every Cayley graph of an Abelian group this limit is precisely the reciprocal of the fractional chromatic number of G.

– Most of the results proved here can be extended to independent sets in weak products of different regular factors H_i, and there is no need to assume all the factors are identical. We omit the detailed statements of the corresponding results.

– The assumptions in Theorems 1.4 and 1.5 that the base graph H is connected and non-bipartite are essential. Indeed, if H is either disconnected or bipartite, then H^n has many connected components, and as in each of them one can choose a maximum independent set separately, no single coordinate can determine the membership in a maximum independent set of H^n. Moreover, if H is bipartite then so is H^n, and as it is regular, its independence number is precisely half the number of vertices. There are many maximum size independent sets. In fact, if J is a maximum size independent set in H, $|J| = |H|/2$, and $A \subset \{1, 2, \dots, n\}$ is any set of odd size, then the set

$$I = \big\{(x_1, \dots, x_n) : |\{i : i \in A \ \text{ and } \ x_i \in J\}| > |A|/2\}$$

is a maximum size independent set in H^n. Therefore, there are no uniqueness or stability results in this case. In case H is disconnected, one can consider each component separately, and use the assertions of Theorems 1.4 and 1.5 for products of possibly distinct factors to obtain some information about the structure of maximum and near maximum independent sets in H^n.

– The proof of Theorem 1.5 works even if the set I considered is not independent, but contains a small number of edges. Indeed, the assertion of Lemma 5.1, part (ii), holds, with an extra constant factor, even if the induced subgraph of G on I contains some $\delta m(\lambda - \lambda_m)/4$ edges. This can be used to obtain results about colorings of H^n in which the number of monochromatic edges is relatively small.

– The proof of Theorem 1.5 is based on the fact that a change of basis enables us to prove several results from discrete Harmonic Analysis for general tensor products of an arbitrary fixed orthogonal basis, and not only for characters. We believe that this simple fact is likely to lead to additional interesting consequences.

Note added in proof. In a recent paper [LT], Larose and Tardif study the structure of maximum independent sets in graph powers. They prove Proposition 1.3, and describe an example of vertex transitive graphs H for which the assertion of Theorem 1.4 fails.

Acknowledgments. We wish to thank David Galvin, Johan Håstad, Jeff Kahn, Guy Kindler, Assaf Naor, Jarik Nešetřil and Prasad Tetali for useful remarks, and Claude Tardif for informing us about [LT].

References

[A1] N. Alon, Spectral techniques in graph algorithms, Springer Lecture Notes in Computer Science 1380 (C.L. Lucchesi, A.V. Moura, eds.) (1998), 206–215.

[A2] N. Alon, Graph powers, in "Contemporary Combinatorics (B. Bollobás, ed.), Bolyai Society Mathematical Studies, Springer (2002), 11–28.

[AK] N. Alon, N. Kahale, A spectral technique for coloring random 3-colorable graphs, Proc. of the 26^{th} ACM STOC, ACM Press (1994), 346–355; also; SIAM J. Comput. 26 (1997), 1733–1748.

[AKRS] N. Alon, G. Kalai, M. Ricklin, L. Stockmeyer, Lower bounds on the competitive ratio for mobile user tracking and distributed job-scheduling, Proc. 33^{rd} IEEE FOCS, Pittsburgh, IEEE (1992), 334–343; also: Theoretical Computer Science 130 (1994), 175–201.

[ALM] N. Alon, N. Linial, R. Meshulam, Additive bases of vector spaces over prime fields, J. Combinatorial Theory, Ser. A 57 (1991), 203–210.

[BS] R.C. Baker, W.M. Schmidt, Diophantine problems in variables restricted to the values 0 and 1, J. Number Theory 12 (1980), 460–486.

[Be] W. Beckner, Inequalities in Fourier analysis, Annals of Mathematics, 102 (1975), 159–182.

[BenKS] I. Benjamini, G. Kalai, O. Schramm, Noise sensitivity of boolean functions and applications to percolation, Inst. Hautes Etudes Sci. Publ. Math. 90 (1999), 5–43 (2001).

[Bo] A. Bonami, Etude des coefficients Fourier des fonctiones de $L^p(G)$, Ann. Inst. Fourier (Grenoble) 20:2 (1970), 335–402.

[Bou1] J. Bourgain, Walsh subspaces of ℓ_p product spaces, Seminaire D'Analyse Fonctionnelle, Ecole Polytechnique, Centre De Mathematiques 1979-80 (1980), 4.1–4.9.

[Bou2] J. Bourgain, On the distribution of the Fourier spectrum of Boolean functions, Israel Journal of Math., to appear.

[BrNR] J.I. Brown, R.J. Nowakowski, D. Rall, The ultimate categorical independence ratio of a graph, SIAM J. Discrete Math. 9 (1996), 290–300.

[DFK] I. Dinur, E. Friedgut, G. Kindler, Bounded functions with decaying Fourier expansion are juntas, in preparation.

[DS] I. Dinur, S. Safra, On the importance of being biased, Proc. of the 34^{th} ACM STOC, ACM Press (2002), 33–42.

[EKR] P. Erdős, C. Ko, R. Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford (2) 12 (1961), 313–320.

[F] E. Friedgut, Boolean functions with low average sensitivity depend on few coordinates, Combinatorica 18:1 (1998), 27–36.

[FKN] E. Friedgut, G. Kalai, A. Naor, Boolean functions whose Fourier transform is concentrated on the first two levels, Adv. in Appl. Math., to appear.

[GL] D. Greenwell, L. Lovász, Applications of product colorings, Acta Math. Acad. Sci. Hungar. 25:3-4 (1974), 335–340.

[HHP] G. Hahn, P. Hell, S. Poljak, On the ultimate independence ratio of a graph, Europ. J. Combinatorics 16 (1995), 253–261.

[Ha] J. Håstad, private communication.

[KKL] J. Kahn, G. Kalai, N. Linial, The influence of variables on Boolean functions, Proc. 29-th Ann. Symp. on Foundations of Comp. Sci., Computer Society Press (1988), 68–80.

[KiS] G. Kindler, M. Safra, Noise insensitive Boolean functions are juntas, submitted.

[KrS] M. Krivelevich, B. Sudakov, Pseudo-random graphs, to appear.

[LT] B. Larose, C. Tardif, Projectivity and independent sets in powers of graphs, preprint (2003).

[LiMN] N. Linial Y. Mansour, N. Nisan, Constant-depth circuits, Fourier transform and learnability, Jour. Assoc. Comput. Mach. 40 (1993), 607–620.

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

[M] R. Meshulam, On subsets of finite abelian groups with no 3-term arithmetic progressions, Journal of Combinatorial Theory Ser. A. 71 (1995), 168–172.

[Mü] V. Müller, On colorings of graphs without short cycles, Discrete Math. 26 (1979), 165–179.

[T] M. Talagrand, On boundaries and influences, Combinatorica 17 (1997), 275–285.

Noga Alon, Schools of Mathematics and Computer Science, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel noga@math.tau.ac.il

Irit Dinur, NEC Research, Princeton, NJ 08540, USA iritd@nec-labs.com

Ehud Friedgut, Institute of Mathematics, Hebrew University, Jerusalem, Israel ehudf@math.huji.ac.il

Benny Sudakov, Department of Mathematics, Princeton University, Princeton, NJ 08540, USA and Institute for Advanced Study, Princeton, NJ 08540, USA bsudakov@math.princeton.edu

Submitted: May 2003
Revision: October 2003