**Acknowledgement of priority**

Results stronger than those contained in this paper, with similar methods, have been obtained by Javier Cilleruelo [4] before my paper was written. My paper will not be published. Please do not cite it.

Mihalis Kolountzakis

# COEFFICIENTS OF SQUARES OF NEWMAN POLYNOMIALS

MIHAIL N. KOLOUNTZAKIS

Abstract. We show that there are polynomials $p_N$ of arbitrarily large degree $N$, with coefficients equal to 0 or 1 (Newman polynomials), such that

$$\liminf_{N\to\infty} N\|p_N^2\|_\infty / p_N^2(1) < 1,$$

where $\|q\|_\infty$ denotes the maximum coefficient of the polynomial $q$ and which, at the same time, are sparse: $p_N(1)/N \to 0$. This disproves a conjecture of Yu [8]. We build on some previous results of Berenhaut and Saidak [2] and Dubickas [5] whose examples lacked the sparsity. This sparsity we create from these examples by randomization.

A *Newman polynomial* is a polynomial whose coefficients are 0 or 1. This is a very natural object and this terminology is naturally not universal across mathematics. Many problems can be expressed using Newman polynomials and many of them turn out to be quite hard: it is often non-trivial to encode this 0-1 condition using the data of a specific problem. For instance (drawing from the author's experience) many questions that concern problems of tiling the integers by translations of finite sets can be expressed using divisibility and factorization properties of Newman polynomials (see for instance [6]). Similarly such properties of Newman polynomials play a major role in questions of phase retrieval [7] (how to recover the phase of the Fourier transform of an indicator function of a finite set of integers if one knows only the modulus of the Fourier transform). Several extremal problems concerning Newman polynomials are also of interest (see, for instance, the references in [2]).

In this note we disprove a conjecture of Yu [8] which concerns the size of the coefficients of squares of Newman polynomials. For a polynomial $p(x) = \sum_{j=0}^d p_j x^j$, with $p_d \neq 0$, we denote by $\|p\|_\infty$ the size of the maximal coefficient in absolute value and by $\|p\|_1 = \sum_{j=0}^d |p_j|$. We also write $\deg p = d$.

Write

$$R(p) = \frac{\|p^2\|_\infty}{\|p\|_1^2}.$$

*Date*: June 2008; revised November 2008.

Supported by research grant No 2569 from the Univ. of Crete.

For any polynomial $p$ with nonnegative coefficients we have $\|p\|_1^2 = \|p^2\|_1$ and, observing that the degree of $p^2$ is $2\deg p$, we obtain easily

$$R(p) \ge \frac{1}{2\deg p + 1}. \tag{1}$$

Yu [8] conjectured that if $p_n$ is a sequence of Newman polynomials with

$$\|p_n\|_1 = o(\deg p_n) \tag{2}$$

then

$$\liminf_n R(p_n) \deg p_n \ge 1. \tag{3}$$

By (1) the trivial right hand side in (3) would be 1/2.

That (3) fails if we ommit the sparsity condition (2) was shown by Berenhaut and Saidak [2] and by Dubickas [5]. They exhibited sequences of polynomials $p_n$ with $\liminf_n R(p_n)\deg p_n < 1$ (the liminf was 8/9 in the case of [2] and 5/6 in the case of [5]), but with $\|p_n\|_1 \ge c \deg p_n$ for some positive constant $c$.

Our purpose here is to show that the conjecture of Yu mentioned above fails. We will show that (3) fails even with the condition (2). For this we will use a sequence of "dense" polynomials $p_n$ which satisfies $\liminf_n R(p_n)\deg p_n = \rho < 1$ (such as any of those constructed in [2] or [5]) and will construct, for any $\rho < \rho' < 1$, another sequence of Newman polynomials $q_n$, with $\|q_n\|_1 = o(\deg q_n)$, which satisfies $\liminf_n R(q_n)\deg q_n = \rho' < 1$.

The remainder of this note is devoted to the proof of the following result.

**Theorem 1.** *Suppose there exists a sequence of Newman polynomials $p_n$, with degrees tending to infinity, and positive constants $c_0$, $\rho$ such that*

$$\|p_n\|_1 \ge c_0 \deg p_n \quad \text{and} \quad R(p_n) \le \rho \frac{1}{\deg p_n}. \tag{4}$$

*Then for every $\rho' > \rho$ there exists an infinite sequence of Newman polynomials $q_n$, with degrees tending to infinity, such that*

$$\|q_n\|_1 = o(\deg q_n) \quad \text{and} \quad R(q_n) \le \rho' \frac{1}{\deg q_n}. \tag{5}$$

*Proof.* The idea of the proof is to construct the polynomials $q_n$ from the $p_n$ by keeping a random subset of the monomials in $p_n$. This will achieve the sparsity condition (2) if we keep the monomials with small probability. At the same time we are able to control the size of the coefficients of $q_n^2$ using standard tail estimates.

Write $N = \deg p_n$, assume $N$ is large, and notice that our assumptions on $p_n$ and (1) imply that

$$\left\|p_n^2\right\|_\infty \ge \frac{c_0^2 N^2}{2N+1} \ge \frac{c_0^2}{3} N. \tag{6}$$

Let $\alpha = \alpha(N) = N^{-1/10}$ and define the random polynomial

$$q_n(x) = \sum_{j=0}^{N} q_j x^j$$

by taking $q_j = \epsilon_j p_j$ with independent $\epsilon_j \in \{0,1\}$ being equal to 1 with probability $\alpha$ and 0 with probability $1-\alpha$. Write $(p_n^2)_j$ and $(q_n^2)_j$ for the coefficients of $x^j$ respectively in the polynomials $p_n^2$ and $q_n^2$.

It follows that $q_n$ is a Newman polynomial of degree at most $N$ and we have

$$\|q_n\|_1 = \sum_{j=0}^{N} \epsilon_j (p_n)_j, \quad (q_n^2)_k = \sum_{j=0}^{k} \epsilon_j \epsilon_{k-j} (p_n)_j (p_n)_{k-j} \quad (k = 0, 1, \ldots, 2N).$$

It follows immediately that

$$\mathbf{E}\|q_n\|_1 = \alpha \|p_n\|_1$$

and, if $k$ is odd, we also have

$$\mathbf{E}(q_n^2)_k = \alpha^2 (p_n^2)_k$$

(the reason for restricting $k$ to be odd is that then the products $\epsilon_j \epsilon_{k-j}$ that appear are products of independent variables).

If $k$ is even we have

$$\begin{aligned} \mathbf{E}(q_n^2)_k &= \alpha^2 (p_n^2)_k + \alpha(1-\alpha)(p_n)_{k/2}^2 \\ &= \alpha^2 (p_n^2)_k + \theta_k, \quad (0 \le \theta_k < 1). \end{aligned}$$

The random variables $\|q_n\|_1$ and $(q_n^2)_k$ are both sums of indicator (0-1 valued) random variables. In the case of $\|q_n\|_1$ these random variables are independent while $(q_n^2)_k$ can be written, depending on whether $k$ is odd or even, as follows.

When $k$ is odd we have

$$\begin{aligned} (q_n^2)_k &= \sum_{j=0}^{\lfloor k/2 \rfloor} \epsilon_j \epsilon_{k-j} (p_n)_j (p_n)_{k-j} + \sum_{j=\lfloor k/2 \rfloor + 1}^{k} \epsilon_j \epsilon_{k-j} (p_n)_j (p_n)_{k-j} \\ &=: X_{n,k,1} + X_{n,k,2}, \end{aligned} \tag{7}$$

while for even $k$ we have

$$\begin{aligned} (q_n^2)_k &= \sum_{j=0}^{k/2-1} \epsilon_j \epsilon_{k-j} (p_n)_j (p_n)_{k-j} + \sum_{j=k/2+1}^{k} \epsilon_j \epsilon_{k-j} (p_n)_j (p_n)_{k-j} + \epsilon_{k/2} (p_n)_{k/2}^2 \\ &=: Y_{n,k,1} + Y_{n,k,2} + Y_{n,k,3}. \end{aligned} \tag{8}$$

The random variables $X_{n,k,1}, X_{n,k,2}, Y_{n,k,1}, Y_{n,k,2}, Y_{n,k,3}$ defined above are all sums of *independent* indicator random variables. For such random variables we can control the probability of their deviation from their mean using the following well known result, which we are going to use with $\epsilon$ being a constant that depends only on $\rho$ and $\rho'$.

**Theorem A** (Chernoff [3], [1, Corollary A.1.14]) *If $X = X_1 + \cdots + X_k$, and the $X_j$ are independent indicator random variables (that is $X_j \in \{0, 1\}$), then for all $\epsilon > 0$*

$$\mathbf{Pr}\left[|X - \mathbf{E}X| > \epsilon \mathbf{E}X\right] \le 2e^{-c_\epsilon \mathbf{E}X},$$

*where $c_\epsilon > 0$ is a function of $\epsilon$ alone*

$$c_\epsilon = \min\{-\log(e^\epsilon (1+\epsilon)^{-(1+\epsilon)}), \epsilon^2/2\}.$$

Fixing $\epsilon > 0$, our purpose is to avoid the following "bad events":

$$\begin{aligned} E &= \{\|q_n\|_1 < (1-\epsilon)\mathbf{E}\|q_n\|_1\} \\ &= \{\|q_n\|_1 < (1-\epsilon)\alpha\|p_n\|_1\}, \end{aligned} \tag{9}$$

$$E_k = \left\{(q_n^2)_k > (1+\epsilon)\alpha^2 \left\|p_n^2\right\|_\infty\right\}, \quad k = 0, \ldots, 2N, \tag{10}$$

and

$$D = \left\{ \deg q_n \le \frac{c_0}{2} \deg p_n \right\}. \tag{11}$$

If none of these events holds then

$$\begin{aligned} R(q_n) \deg q_n &= \frac{\|q_n^2\|_\infty}{\|q_n\|_1^2} \deg q_n \\ &\le \frac{1+\epsilon}{(1-\epsilon)^2} \liminf_n R(p_n) \deg p_n \\ &\le \frac{1+\epsilon}{(1-\epsilon)^2} \rho, \end{aligned}$$

which can be made less than $\rho'$ for appropriately small $\epsilon$. The failure of $E$ and $D$ guarantees the sparseness of $q_n$ since

$$\|q_n\|_1 \le (1-\epsilon) N^{9/10} \quad \text{and} \quad \deg q_n > \frac{c_0}{2} N.$$

Therefore it remains to estimate from above the probability that none of the bad events $E$, $E_k$ holds.

Let us start with $\|q_n\|_1$. This is a sum of independent indicator random variables with mean $\mathbf{E}\|q_n\|_1 = \alpha \|p_n\|_1 \ge c_0 \alpha N = c_0 N^{9/10}$, hence Theorem A implies

$$\begin{aligned} \mathbf{Pr}\left[E\right] &\le \mathbf{Pr}\left[ \left| \|q_n\|_1 - \alpha \|p_n\|_1 \right| > \epsilon \alpha \|p_n\|_1 \right] \\ &\le 2 \exp(-c_\epsilon c_0 N^{0.9}), \end{aligned} \tag{12}$$

which tends to 0 with $N \to \infty$. One proves similarly that $\mathbf{Pr}\left[D\right] \to 0$.

The summands contributing to the random variables $(q_n^2)_k$ are indicator random variables but there are dependencies. That's why we need to use the breakups (7) and (8) above. The $X$ and $Y$ random variables defined there are sums of independent indicator random variables and we can apply Theorem A to them in order to control the probabilities of their deviations from their mean.

Let us deal with the case of odd $k$. The case of even $k$ is treated similarly. We separate the random variables $X_{n,k,1}, X_{n,k,2}$, $k = 0, 2, \ldots, 2N$, into two groups. In the first group we put those variables whose mean is at most $N^{1/10}$ and in the second group we put the remaining variables. For a given odd $k$ we have the following three cases: (a) both $X_{n,k,1}$ and $X_{n,k,2}$ are in the first group, (b) only one of them is, and (c) none is. Notice that the mean of an $X$ variable is equal to $\alpha^2 = N^{-2/10}$ times the maximum value that variable can take, which corresponds to the case of all relevant $\epsilon_j$ being equal to 1.

If (a) is the case then the $X$ variables are always at most $N^{3/10}$. From (6) it follows that $E_k$ cannot hold.

If (b) is the case then, assuming, without loss of generality, that the variable $X_{n,k,1}$ is in the first group, we get that always $X_{n,k,1} \le N^{3/10}$ as before. Therefore, for $E_k$ to hold it must be the case that

$$X_{n,k,2} \ge (1 + \frac{\epsilon}{2}) \alpha^2 \|p_n^2\|_\infty \ge (1 + \frac{\epsilon}{2}) \alpha^2 (p_n^2)_k \ge (1 + \frac{\epsilon}{2}) \mathbf{E} X_{n,k,2},$$

and from Theorem A we obtain

$$\mathbf{Pr}\left[E_k\right] \le \mathbf{Pr}\left[ X_{n,k,2} > (1 + \frac{\epsilon}{2}) \mathbf{E} X_{n,k,2} \right] \le 2 \exp(-c_{\epsilon/2} N^{1/10}).$$

Since the number of relevant $k$ is $O(N)$ this implies that the total probability of the $E_k$ for odd $k$ falling in case (b) tends to 0.

The case (c) is treated similarly to case (b).

We have proved that $\mathbf{Pr}\left[\cup_k E_k\right] \to 0$ as $n \to \infty$ by splitting the events into three groups and the proof of the Theorem is complete as this implies that there is a choice of the numbers $\epsilon_j \in \{0,1\}$ such that the events $E$, $E_k$ do not hold and this implies that the polynomial $q_n$ has the desired properties. □

## References


[1] N. Alon and J. Spencer, *The probabilistic method*, Second edition. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, 2000

[2] K.S. Berenhaut and F. Saidak, A note on the maximal coefficients of squares of Newman polynomials, J. Number Th. **125** (2007), 2, 285–288.

[3] H. Chernoff, A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Stat. **23** (1952), 493-509.

[4] J. Cilleruelo, Maximal coefficients of squares of Newman polynomials, `preprint`.

[5] A. Dubickas, Heights of powers of Newman and Littlewood polynomials, Acta Arith. **128** (2007), 2, 167–176.

[6] M.N. Kolountzakis, Translational tilings of the integers with long periods, Electr. J. Combinatorics **10** (2003), 1, R22.

[7] P. Lemke, S.S. Skiena, W.D. Smith, Reconstructing sets from interpoint distances, DIMACS Technical Report 2002-37.

[8] G. Yu, An upper bound on $B_2[g]$ sets, J. Number Th. **122** (2007), 211220.



Department of Mathematics, University of Crete, Knossos Ave., GR-714 09, Iraklio, Greece

*E-mail address*: `kolount@gmail.com`