Electronic Colloquium on Computational Complexity, Report No. 19 (2014)

Inequalities and tail bounds for elementary symmetric polynomials

Amir Yehudayoff* Parikshit Gopalan†

Abstract

This paper studies the elementary symmetric polynomials $S_k(x)$ for $x \in \mathbb{R}^n$. We show that if $|S_k(x)|, |S_{k+1}(x)|$ are small for some $k > 0$ then $|S_\ell(x)|$ is also small for all $\ell > k$. We use this to prove probability tail bounds for the symmetric polynomials when the inputs are only t-wise independent, that may be useful in the context of derandomization. We also provide examples of t-wise independent distributions for which our bounds are essentially tight.

1 Introduction

The elementary symmetric polynomials are a basic family of functions that are stable under any permutation of the inputs. The k'th symmetric polynomial is defined as[1]

$$S_k(a) = \sum_{T \subseteq [n] : |T| = k} \prod_{i \in T} a_i$$

for all $a = (a_1, a_2, \ldots, a_n)$. They are defined over any field but we study them over the real numbers. They appear as coefficients of a univariate polynomial with roots $-a_1, \ldots, -a_n \in \mathbb{R}$. That is,

$$\prod_{i \in [n]} (\xi + a_i) = \sum_{k=0}^{n} \xi^k S_{n-k}(a).$$

This work focuses on their growth rates. Specifically, we study how local information on $S_k(a)$ for two consecutive values of k implies global information for all larger values

*Department of Mathematics, Technion–IIT. amir.yehudayoff@gmail.com. Horev fellow – supported by the Taub foundation. Research supported by ISF and BSF

†Microsoft Research – Silicon Valley. parik@microsoft.com.

[1]We omit the dependence on n from the notation. It is clear from the context.

ISSN 1433-8092

of k. Inequalities in these polynomials have been studied in mathematics, dating back to classical results of Newton and Maclaurin. For a survey of such inequalities, we refer the reader to [4].

An interesting property over the real numbers is that if $p(\xi)$ is a real univariate polynomial of degree n with n nonzero roots and $p'(0) = p''(0) = 0$ then $p \equiv 0$. This follows by simple properties of the symmetric polynomials over the real numbers: We may write

$$p(\xi) = \prod_{i \in [n]} (\xi b_i + 1) = \sum_{k=0}^{n} \xi^k S_k(b).$$

The condition on the derivatives of p is equivalent to $S_1(b) = S_2(b) = 0$, and the following fact completes the argument.

Fact A. *Over the real numbers, if $S_1(b) = S_2(b) = 0$ then $b = 0$.*

This does not hold over all fields, for example, the polynomial $p(\xi) = \xi^3 + 1$ is of degree three, has three nonzero complex roots and $p'(0) = p''(0) = 0$.

A robust version of Fact A was recently proved in [3]: For every $a \in \mathbb{R}^n$ and $k \in [n]$,

$$|S_k(a)| \leq \left(S_1^2(a) + 2|S_2(a)|\right)^{k/2}. \tag{1}$$

That is, if $S_1(a), S_2(a)$ are small in absolute value, then so is everything that follows. We provide an essentially optimal bound.

Theorem 1. *For every $a \in \mathbb{R}^n$ and $k \in [n]$,*

$$|S_k(a)| \leq \left(\frac{6e(S_1^2(a) + |S_2(a)|)^{1/2}}{k^{1/2}}\right)^k.$$

The parameters promised by Theorem 1 are tight up to an exponential in k which is often too small to matter (we do not attempt to optimise the constants). For example, if $a_i = (-1)^i$ for all $i \in [n]$ then $|S_1(a)| \leq 1$ and $|S_2(a)| \leq n + 1$ but $S_k(a)$ is roughly $(n/k)^{k/2}$.

The argument is quite general, and similar bounds may be obtained for functions that are recursively defined. Our proof is analytic and uses the method of Lagrange multipliers, and is very different from that of [3] which relied on the Newton-Girrard identities.

Stronger bounds are known when the inputs are nonnegative. When $a_i \geq 0$ for all $i \in [n]$, the classical Maclaurin inequalities imply that

$$S_k(a) \leq \left(\frac{e}{k}\right)^k (S_1(a))^k.$$

In contrast, when we do not assume non-negativity, one cannot hope for such bounds to hold under the assumption that $|S_1(a)|$ or any single $|S_k(a)|$ is small (cf. the alternating signs example above).

A more general statement than Fact A actually holds (see Appendix A for a proof).

Fact B. *Over the reals, if $S_k(a) = S_{k+1}(a) = 0$ for $k > 0$ then $S_\ell(a) = 0$ for all $\ell \geq k$.*

We prove a robust version of this fact as well: a twice-in-a-row bound on the increase of the symmetric functions implies a bound on what follows.

Theorem 2. *For every $a \in \mathbb{R}^n$, if $S_k(a) \neq 0$ and*

$$\left|\binom{k+1}{k}\frac{S_{k+1}(a)}{S_k(a)}\right| \leq C \quad \textit{and} \quad \left|\binom{k+2}{k}\frac{S_{k+2}(a)}{S_k(a)}\right| \leq C^2$$

then for every $1 \leq h \leq n-k$,

$$\left|\binom{k+h}{k}\frac{S_{k+h}(a)}{S_k(a)}\right| \leq \left(\frac{6eC}{h^{1/2}}\right)^h.$$

We now discuss applications of our bounds in the context of pseudorandomness.

1.1 Tail bounds under limited independence

Pseudorandomness studies the possibility of removing randomness from randomized algorithms while maintaining functionality. A central notion in this study is t-wise independence. The n random variables $X_1, \ldots, X_n$ are t-wise independent if every subset of k of them are independent. It turns out that one may produce t-wise independent distributions using much fewer bits than are required for producing a fully independent distribution [1, 2], and this is of course useful for derandomization.

The authors of [3] used this idea to construct pseudorandom generators for several families of tests including read-once DNF formulas and combinatorial rectangles. A key part of their proof was to show that the expected value of $\prod_{i\in[n]}(1+X_i)$ does not significantly change between the case the inputs are independent and the case the inputs are only t-wise independent, under the assumption that $\mathbb{E}[X_i] = 0$ for all $i \in [n]$ and $\sum_i \mathsf{Var}[X_i] \ll 1$.

A standard way to estimate $\prod_{i\in[n]}(1+X_i)$ is by taking a logarithm and using known concentration bounds for sums of independent random variables. This method gives an estimate of $\prod_{i\in[n]}(1+X_i)$ that is good up to some $1 \pm \epsilon$ power rather than an additive factor. Observing that

$$\prod_{i\in[n]}(1+X_i) = \sum_{\ell=0}^{n} S_\ell(X_1, \ldots, X_n),$$

we may obtain better approximations by understanding the symmetric polynomials. A key ingredient of [3] is indeed about controlling

$$\sum_{\ell=k}^{n} |S_\ell(X_1,\ldots,X_n)|,$$

assuming the distribution is $O(k)$-wise independent (but under stronger assumptions involving higher moments).

Let $X=(X_1,\ldots,X_n)$ be a vector of real valued random variables so that

$$\mathbb{E}[X_i]=0$$

for all $i\in[n]$, and denote

$$\sigma^2=\sum_{i\in[n]}\mathsf{Var}[X_i].$$

Let $\mathcal{U}$ denote the distribution where the coordinates of X are independent. It is easy to show (see Lemma 4) that

$$\mathbb{E}_{X\in\mathcal{U}}[|S_\ell(X)|]\le\frac{\sigma^\ell}{\sqrt{\ell!}}.$$

In particular, if $\sigma^2<1$ then $\mathbb{E}[|S_\ell|]$ decays exponentially with ℓ. For $t>0$ and $t\sigma\le 1/2$, we may also conclude (see Corollary 5)

$$\Pr_{X\in\mathcal{U}}\left[\sum_{\ell=k}^{n}|S_\ell(X)|\ge 2(t\sigma)^k\right]\le 2t^{-2k}. \tag{2}$$

Bounding $\mathbb{E}[|S_\ell|]$ for $\ell\le k$ for more general X only requires the distribution to be $(2k)$-wise independent. It can be shown (see Section 4) that this is not enough to get strong bounds on $\mathbb{E}[|S_\ell|]$ for $\ell>2k+1$. Nevertheless, we are able to show a tail bound which holds under limited independence, due to properties of the symmetric polynomials.

Theorem 3. *Let $\mathcal{D}$ denote a distribution over $X=(X_1,\ldots,X_n)$ where $X_1,\ldots,X_n$ are $(2k+2)$-wise independent. Assume $\mathbb{E}[X_i]=0$ for all $i\in[n]$, and denote $\sigma=\sum_{i\in[n]}\mathsf{Var}[X_i]$. For $t>0$ and*[2] *$16et\sigma\le 1$,*

$$\Pr_{X\in\mathcal{D}}\left[\sum_{\ell=k}^{n}|S_\ell(X)|\ge 2(6et\sigma)^k\right]\le 2t^{-2k}, \tag{3}$$

[2] A weaker but more technical assumption on t,σ,k suffices, see Equation (16).

and if we denote $p(x_1, \ldots, x_n) = \prod_{i \in [n]} (1 + x_i)$ *then*[3]

$$\Pr_{X \in \mathcal{D}} \left[|p(X) - 1| \geq 2(6et\sigma)^k \right] \leq 2t^{-2k}.$$

Comparing Equation (3) to Equation (2) we see that it has a similar asymptotic behaviour. A key ingredient in our proof is to show that although we cannot upper bound the expectation of $|S_\ell|$ for large ℓ under k-wise independence alone, we can still show good tail bounds for it. Lemma 6 below shows that for $t > 0$ and for $\ell \geq k$, the bounds

$$|S_\ell(X)| \leq (6et\sigma)^\ell \left(\frac{k}{\ell} \right)^{\ell/2}$$

hold except with $\mathcal{D}$-probability $2t^{-2k}$.

In Section 4, we give an example of a $(2k+2)$-wise independent distribution where $\mathbb{E}[|S_\ell|]$ for $\ell \in \{2k+3, \ldots, n-2k-3\}$ is much larger than under the uniform distribution. This shows that one can only hope to show tail bounds for larger ℓ. The same example also shows that our tail bounds are close to tight.

To obtain their concentration bounds, the authors of [3] used Equation (1) which required them to start with a stronger assumption on $X_1, \ldots, X_n$: They require a bound on higher moments, i.e.,

$$\mathbb{E}[X_i^{2k}] \leq (2k)^{2k} \sigma_i^{2k}$$

for all $i \in [n]$. They additionally require $\sigma = o(1)$ as opposed to being bounded by some constant. Bounding the higher moments introduces technical difficulties and case analyses in their proofs. In contrast, bounding the second moments (as we require here) is immediate. Theorem 3 can be used to simplify their proofs.

Theorem 2 is proved by a simple reduction from the case of general k to the case of $k = 0$, and then applying Theorem 1. In place of Theorem 1, one could plug in the bound given by Equation (1). However, it seems that the resulting bound will not be strong enough to prove Theorem 3, and the asymptotic improvement given by Theorem 1 is crucial.

2 Inequalities for symmetric polynomials

Proof of Theorem 1. It will be convenient to use

$$E_2(a) = \sum_{i \in [n]} a_i^2.$$

[3] Observe $\mathbb{E}_{X \in \mathcal{U}}[p(X)] = 1$.

By Newton's identity, $E_2 = S_1^2 - 2S_2$ so for all $a \in \mathbb{R}^n$,

$$S_1^2(a) + E_2(a) \leq 2(S_1^2(a) + |S_2(a)|).$$

It therefore suffices to prove that for all $a \in \mathbb{R}^n$ and $k \in [n]$,

$$S_k^2(a) \leq \frac{(16e^2(S_1^2(a) + E_2(a)))^k}{k^k}.$$

We prove this by induction. For $k \in \{1, 2\}$, it indeed holds. Let $k > 2$. Our goal will be upper bounding the maximum of the (projectively defined) function

$$\phi_k(\xi) = \frac{S_k^2(\xi)}{(S_1^2(\xi) + E_2(\xi))^k}$$

under the constraint that $S_1(\xi)$ is fixed. Choose $a \neq 0$ to be a point that achieves the maximum of ϕ_k. We assume, without loss of generality, that $S_1(a)$ is non-negative (if $S_1(a) < 0$, consider $-a$ instead of a). There are two cases to consider:

The first case is that for all $i \in [n]$,

$$a_i \leq \frac{2k^{1/2}(S_1^2(a) + E_2(a))^{1/2}}{n}. \tag{4}$$

In this case we do not need the induction hypothesis and can in fact replace each a_i by its absolute value. Let $P \subseteq [n]$ be the set of $i \in [n]$ so that $a_i \geq 0$. Then by Equation (4),

$$\sum_{i \in P} |a_i| \leq 2k^{1/2}(S_1^2(a) + E_2(a))^{1/2}.$$

Note that

$$S_1(a) = \sum_{i \in P} |a_i| - \sum_{i \notin P} |a_i| \geq 0.$$

Hence

$$\sum_{i \notin P} |a_i| \leq \sum_{i \in P} |a_i| \leq 2k^{1/2}(S_1^2(a) + E_2(a))^{1/2}.$$

Overall we have

$$\sum_{i \in [n]} |a_i| \leq 4k^{1/2}(S_1^2(a) + E_2(a))^{1/2}.$$

We then bound

$$\begin{aligned}|S_k(a_1,\ldots,a_n)| &\le S_k(|a_1|,\ldots,|a_n|)\\ &\le \left(\frac{e}{k}\right)^k \left(\sum_{i\in[n]} |a_i|\right)^k \quad \text{By the Maclaurin identities}\\ &\le \left(\frac{4e}{\sqrt{k}}\right)^k (S_1^2(a)+E_2(a))^{k/2}.\end{aligned}$$

The second case is that there exists $i_0 \in [n]$ so that

$$a_{i_0} > \frac{2k^{1/2}(S_1^2(a)+E_2(a))^{1/2}}{n}. \tag{5}$$

In this case we use induction and Lagrange multipliers. For simplicity of notation, for a function F on $\mathbb{R}^n$ denote

$$F(-i) = F(a_1,a_2,\ldots,a_{i-1},a_{i+1},\ldots,a_n)$$

for $i \in [n]$. So, for every $\delta \in \mathbb{R}^n$ so that $\sum_i \delta_i = 0$ we have $\phi_k(a+\delta) \le \phi_k(a)$. Hence[4], for all δ so that $\sum_i \delta_i = 0$,

$$\begin{aligned}\phi_k(a) &\ge \frac{S_k^2(a+\delta)}{(S_1^2(a+\delta)+E_2(a+\delta))^k}\\ &\ge \frac{(S_k(a)+\sum_i \delta_i S_{k-1}(-i)+O(\delta^2))^2}{(S_1^2(a)+E_2(a)+2\sum_i a_i\delta_i+O(\delta^2))^k}\\ &\ge \frac{S_k^2(a)+2S_k(a)\sum_i \delta_i S_{k-1}(-i)+O(\delta^2)}{(S_1^2(a)+E_2(a))^k+2k(S_1^2(a)+E_2(a))^{k-1}\sum_i a_i\delta_i+O(\delta^2)}.\end{aligned}$$

Hence, for all δ close enough to zero so that $\sum_i \delta_i = 0$,

$$\frac{S_k^2(a)}{(S_1^2(a)+E_2(a))^k} \ge \frac{S_k^2(a)+2S_k(a)\sum_i \delta_i S_{k-1}(-i)+O(\delta^2)}{(S_1^2(a)+E_2(a))^k+2k(S_1^2(a)+E_2(a))^{k-1}\sum_i a_i\delta_i+O(\delta^2)},$$

or

$$\sum_i \delta_i \left(a_i S_k(a)k - (S_1^2(a)+E_2(a))S_{k-1}(-i)\right) \ge 0.$$

[4]Here and below, $O(\delta^2)$ means of absolute value at most $C\cdot\|\delta\|_\infty$ for $C = C(n,k) \ge 0$.

For the above equality to hold, it must be that there is λ so that for all $i \in [n]$,

$$a_i S_k(a)k - (S_1^2(a) + E_2(a))S_{k-1}(-i) = \lambda.$$

Sum over i to get

$$\lambda n = S_1(a)S_k(a)k - (S_1^2(a) + E_2(a))(n - (k-1))S_{k-1}(a).$$

Thus, for all $i \in [n]$,

$$\begin{aligned} &a_i S_k(a)k - (S_1^2(a) + E_2(a))S_{k-1}(-i) \\ &\qquad = \frac{S_1(a)S_k(a)k - (S_1^2(a) + E_2(a))(n - (k-1))S_{k-1}(a)}{n}, \end{aligned}$$

or

$$\begin{aligned} &S_k(a)k\left(a_i - \frac{S_1(a)}{n}\right) \\ &\quad = (S_1^2(a) + E_2(a))(S_{k-1}(-i) - S_{k-1}(a)) + \frac{(k-1)(S_1^2(a) + E_2(a))S_{k-1}(a)}{n}. \end{aligned}$$

This specifically holds for i_0, so using (5) we have

$$\begin{aligned} &\left|S_k(a)k\frac{a_{i_0}}{2}\right| \\ &< \left|S_k(a)k\left(a_{i_0} - \frac{S_1(a)}{n}\right)\right| \\ &\le \left|(S_1^2(a) + E_2(a))a_{i_0}S_{k-2}(-i_0)\right| + \left|\frac{(k-1)(S_1^2(a) + E_2(a))S_{k-1}(a)}{n}\right|, \end{aligned}$$

or

$$\begin{aligned} &|S_k(a)| \\ &\le \left|\frac{2(S_1^2(a) + E_2(a))S_{k-2}(-i_0)}{k}\right| + \left|\frac{2(k-1)(S_1^2(a) + E_2(a))S_{k-1}(a)}{nka_{i_0}}\right| \\ &< \left|\frac{2(S_1^2(a) + E_2(a))S_{k-2}(-i_0)}{k}\right| + \left|\frac{(S_1^2(a) + E_2(a))^{1/2}S_{k-1}(a)}{k^{1/2}}\right|. \end{aligned}$$

To apply induction we need to bound $S_1^2(-i_0) + E_2(-i_0)$ from above. Since

$$\begin{aligned} S_1^2(a) + E_2(a) - S_1^2(-i_0) - E_2(-i_0) &= a_{i_0}^2 + 2a_{i_0}S_1(-i_0) + a_{i_0}^2 \\ &= 2a_{i_0}S_1(a) \ge 0. \end{aligned}$$

we have the bound

$$S_1^2(-i_0) + E_2(-i_0) \le S_1^2(a) + E_2(a).$$

Finally, by induction and the upper bound above,

$$\begin{aligned}
|S_k(a)| &\le \frac{2(S_1^2(a)+E_2(a))}{k}\frac{(16e^2(S_1^2(-i_0)+E_2(-i_0)))^{(k-2)/2}}{(k-2)^{(k-2)/2}} \\
&\quad + \frac{(S_1^2(a)+E_2(a))^{1/2}}{k^{1/2}}\frac{(16e^2(S_1^2(a)+E_2(a)))^{(k-1)/2}}{(k-1)^{(k-1)/2}} \\
&\le \frac{(16e^2(S_1^2(a)+E_2(a)))^{k/2}}{k^{k/2}}\left(\frac{2}{16e^2\left(1-\frac{2}{k}\right)^{(k-2)/2}} + \frac{1}{4e\left(1-\frac{1}{k}\right)^{(k-1)/2}}\right) \\
&< \frac{(16e^2(S_1^2(a)+E_2(a)))^{k/2}}{k^{k/2}}.
\end{aligned}$$

□

Proof of Theorem 2. The proof is by reduction to Theorem 1. Assume $a_1, \ldots, a_m$ are nonzero and $a_{m+1}, \ldots, a_n$ are zero. Denote $a' = (a_1, \ldots, a_m)$ and notice that for all[5] $k \in [n]$,

$$S_k(a) = S_k(a').$$

Write

$$p(\xi) = \prod_{i\in[m]}(\xi a_i + 1) = \sum_{k=0}^{m} \xi^k S_k(a).$$

Derive k times to get

$$\begin{aligned}
p^{(k)}(\xi) = S_k(a)k!\Bigg(&\binom{m}{k}\frac{S_m(a)}{S_k(a)}\xi^{m-k} + \binom{m-1}{k}\frac{S_{m-1}(a)}{S_k(a)}\xi^{m-k-1} + \ldots \\
&\ldots + \binom{k+1}{k}\frac{S_{k+1}(a)}{S_k(a)}\xi + 1\Bigg).
\end{aligned}$$

Since p has m real roots, $p^{(k)}$ has $m-k$ real roots. Since $p^{(k)}(0) \neq 0$, there is $b \in \mathbb{R}^{m-k}$ so that

$$p^{(k)}(\xi) = S_k(a)k! \prod_{i\in[m-k]}(\xi b_i + 1).$$

For all $h \in [m-k]$,

$$S_h(b) = \binom{k+h}{k}\frac{S_{k+h}(a)}{S_k(a)}.$$

[5] For $k > m$ we have $S_k(a) = 0$ so there is nothing to prove.

By assumption,

$$|S_1(b)| \leq C \quad \text{and} \quad |S_2(b)| \leq C^2.$$

Theorem 1 implies

$$|S_h(b)| = \left| \binom{k+h}{k} \frac{S_{k+h}(a)}{S_k(a)} \right| \leq \frac{(6eC)^h}{h^{h/2}}.$$

□

3 Tail bounds under limited independence

In this section we work with the following setup: Let $X = (X_1, \ldots, X_n)$ be a vector of real valued random variables so that $\mathbb{E}[X_i] = 0$ for all $i \in [n]$. Denote $\sigma_i^2 = \mathsf{Var}[X_i]$ and

$$\sigma^2 = \sum_{i \in [n]} \sigma_i^2.$$

The goal is proving a tail bound on the behaviour of the symmetric functions under limited independence.

We start by obtaining tail estimates, under full independence. Let $\mathcal{U}$ denote the distribution over $X = (X_1, \ldots, X_n)$ where $X_1, \ldots, X_n$ are independent.

Lemma 4. $\mathbb{E}_{X \in \mathcal{U}}[S_\ell^2(X)] \leq \frac{\sigma^{2\ell}}{\ell!}$.

Proof. Since the expectation of X_i is zero for all $i \in [n]$,

$$\begin{aligned} \mathbb{E}[S_\ell^2(X)] &= \sum_{T, T' \subset [n] : |T| = |T'| = \ell} \mathbb{E}\left[\prod_{t \in T} X_t \prod_{t' \in T'} X_{t'} \right] \\ &= \sum_{T \subset [n] : |T| = \ell} \mathbb{E}\left[\prod_{t \in T} X_t^2 \right] = \sum_{T \subset [n] : |T| = \ell} \prod_{t \in T} \sigma_t^2 \\ &\leq \frac{1}{\ell!} \left(\sum_{i \in [n]} \sigma_i^2 \right)^\ell = \frac{\sigma^{2\ell}}{\ell!}. \end{aligned}$$

□

Corollary 5. *For $t > 0$ and $\ell \in [n]$, by Markov's inequality,*

$$\Pr_{X \in \mathcal{U}} \left[|S_\ell(X)| \geq \left(\frac{e^{1/2} t \sigma}{\ell^{1/2}} \right)^\ell \geq \frac{(t\sigma)^\ell}{\sqrt{\ell!}} \right] \leq \frac{1}{t^{2\ell}}. \tag{6}$$

If $2e^{1/2}t\sigma \le k^{1/2}$ then by the union bound

$$\Pr_{X\in\mathcal{U}}\left[\sum_{\ell=k}^{n}|S_\ell(X)| \ge 2\left(\frac{e^{1/2}t\sigma}{k^{1/2}}\right)^k\right] \le \frac{1}{t^{2k}-t^{2(k-1)}}. \tag{7}$$

We now consider limited independence.

Lemma 6. *Let $\mathcal{D}$ denote a distribution over $X=(X_1,\ldots,X_n)$ where $X_1,\ldots,X_n$ are $(2k+2)$-wise independent. Let $t\ge 1$. Except with $\mathcal{D}$-probability at most $2t^{-2k}$, the following bounds hold for all $\ell\in\{k,\ldots,n\}$:*

$$|S_\ell(X)| \le (6et\sigma)^\ell\left(\frac{k}{\ell}\right)^{\ell/2}. \tag{8}$$

Proof. In the following the underlying probability distribution over X is $\mathcal{D}$. By Lemma 4, for $i\in\{k,k+1\}$,

$$\mathbb{E}[S_i^2(X)] \le \frac{\sigma^{2i}}{i!}.$$

Hence by Markov's inequality,

$$\Pr\left[|S_i(X)| \ge \frac{(t\sigma)^i}{\sqrt{i!}}\right] \le t^{-2i}.$$

From now on, condition on the event that

$$|S_k(X)| \le \frac{(t\sigma)^k}{\sqrt{k!}} \text{ and } |S_{k+1}(X)| \le \frac{(t\sigma)^{k+1}}{\sqrt{(k+1)!}}, \tag{9}$$

which occurs with probability at least $1-2t^{-2k}$. Fix $x=(x_1,\ldots,x_n)$ such that Equation (9) holds.

We claim that there must exist $k_0\in\{0,\ldots,k-1\}$ for which the following bounds hold:

$$|S_{k_0}(x)| \ge \frac{(t\sigma)^{k_0}}{\sqrt{k_0!}}, \tag{10}$$

$$|S_{k_0+1}(x)| \le \frac{(t\sigma)^{k_0+1}}{\sqrt{(k_0+1)!}}, \tag{11}$$

$$|S_{k_0+2}(x)| \le \frac{(t\sigma)^{k_0+2}}{\sqrt{(k_0+2)!}}. \tag{12}$$

To see this, mark point $j \in \{0, \ldots, k+1\}$ as *high* if

$$|S_j(x)| \geq \frac{(t\sigma)^j}{\sqrt{j!}}$$

and *low* if

$$|S_j(x)| \leq \frac{(t\sigma)^j}{\sqrt{j!}}.$$

A point is marked both high and low if equality holds. Observe that 0 is marked high (and low) since $S_0(x) = 1$ and k and $k+1$ are marked low by Equation (9). This implies the existence of a triple k_0, k_0+1, k_0+2 where the first point is high and the next two are low.

Let $\gamma > 0$ be the smallest number so that the following inequalities hold:

$$|S_{k_0+1}(x)| \leq |S_{k_0}(x)| \frac{\gamma}{\sqrt{k_0+1}}, \tag{13}$$

$$|S_{k_0+2}(x)| \leq |S_{k_0}(x)| \frac{\gamma^2}{\sqrt{(k_0+1)(k_0+2)}}. \tag{14}$$

By definition, one of Equations (13) and (14) holds with equality so

$$|S_{k_0}(x)| = \max\left\{ \frac{|S_{k_0+1}(x)|\sqrt{k_0+1}}{\gamma}, \frac{|S_{k_0+2}(x)|\sqrt{(k_0+1)(k_0+2)}}{\gamma^2} \right\}.$$

Observe further that $\gamma \leq t\sigma$ by Equations (10), (11) and (12). Combining this with the bounds in Equations (11) and (12)

$$|S_{k_0}(x)| \leq \max\left\{ \frac{(t\sigma)^{k_0+1}}{\gamma\sqrt{k_0!}}, \frac{(t\sigma)^{k_0+2}}{\gamma^2\sqrt{k_0!}} \right\} = \frac{(t\sigma)^{k_0+2}}{\gamma^2\sqrt{k_0!}}. \tag{15}$$

Equations (13) and (14) let us apply Theorem 2 with $C = \gamma\sqrt{k_0+1}$ and $h \geq 3$ to get

$$\left| \frac{S_{k_0+h}(x)}{S_{k_0}(x)} \right| \leq (6e\gamma)^h \frac{(k_0+1)^{h/2}}{h^{h/2}\binom{k_0+h}{k_0}}.$$

Bounding $|S_{k_0}|$ by Equation (15), we get

$$|S_{k_0+h}(x)| \leq (6e\gamma)^h \frac{(k_0+1)^{h/2}}{h^{h/2}\binom{k_0+h}{k_0}} \frac{(t\sigma)^{k_0+2}}{\gamma^2\sqrt{k_0!}} \leq (6et\sigma)^{k_0+h} \frac{(k_0+1)^{h/2}}{h^{h/2}\sqrt{k_0!}\binom{k_0+h}{h}}.$$

Since

$$\binom{k_0+h}{h} \geq \max\left\{ \left(\frac{k_0+h}{k_0}\right)^{k_0}, \left(\frac{k_0+h}{h}\right)^{h} \right\} \geq \frac{(k_0+h)^{(k_0+h)/2}}{k_0^{k_0/2} h^{h/2}},$$

we have

$$\frac{(k_0+1)^{h/2}}{h^{h/2}\sqrt{k_0!}\binom{k_0+h}{h}} \leq \left(\frac{k_0+1}{h}\right)^{h/2} \frac{k_0^{k_0/2} h^{h/2}}{(k_0+h)^{(k_0+h)/2}} \leq \left(\frac{k_0+1}{k_0+h}\right)^{(k_0+h)/2}.$$

Therefore, denoting $\ell = k_0 + h$, since $k_0 + 1 \leq k$,

$$|S_\ell(x)| \leq (6et\sigma)^\ell \left(\frac{k}{\ell}\right)^{\ell/2}.$$

□

Proof of Theorem 3. As in Lemma 6, fix $x = (x_1, \ldots, x_n)$ such that Equation (9) holds (the random vector X has this property with $\mathcal{D}$-probability at least $1 - 2t^{-2k}$). By the proof of lemma, since by assumption $6et\sigma < 1/2$,

$$\sum_{\ell=k}^{n} |S_\ell(x)| \leq \frac{(t\sigma)^k}{k!} + \frac{(t\sigma)^{k+1}}{\sqrt{(k+1)!}} + \sum_{\ell=k+2}^{n} (6et\sigma)^\ell \left(\frac{k}{\ell}\right)^{\ell/2} \leq 2(6et\sigma)^k. \tag{16}$$

□

4 On the tightness of the tail bounds

We conclude by showing that $(2k+2)$-wise independence is insufficient to fool $|S_\ell|$ for $\ell > 2k+2$ in expectation. We use a modification of a simple proof due to Noga Alon of the $\Omega(n^{k/2})$ lower bound on the support size of a k-wise independent distribution on $\{-1,1\}^n$, which was communicated to us by Raghu Meka.

For this section, let $X_1, \ldots, X_n$ be so that each X_i is uniform over $\{-1,1\}$. Thus $\sigma^2 = \sum_i \mathsf{Var}[X_i] = n$. By Lemma 4, we have

$$\mathbb{E}_{X \in \mathcal{U}}[|S_\ell(X)|] \leq \left(\mathbb{E}_{X \in \mathcal{U}}[S_\ell^2(X)]\right)^{1/2} \leq \frac{n^{\ell/2}}{\sqrt{\ell!}}. \tag{17}$$

In contrast we have the following:

Lemma 7. *There is a $(2k+2)$-wise independent distribution on $X = (X_1, X_2, \ldots, X_n)$*

in $\{-1,1\}^n$ *such that for every* $\ell \in [n]$,

$$\Pr_{X \in \mathcal{D}}\left[|S_\ell(X)| \geq \binom{n}{\ell}\right] \geq \frac{1}{3n^{k+1}}.$$

Specifically,

$$\mathbb{E}_{X \in \mathcal{D}}[|S_\ell(X)|] \geq \frac{\binom{n}{\ell}}{3n^{k+1}}. \tag{18}$$

Proof. Let $\mathcal{D}$ be a $(2k+2)$-wise independent distribution on $\{-1,1\}^n$ that is uniform over a set D of size $2(n+1)^{k+1} \leq 3n^{k+1}$. Such distributions are known to exist [1]. Further, by translating the support by some fixed vector if needed, we may assume that $(1,1,\ldots,1) \in D$. It is easy to see that every such translate also induces a $(2k+2)$-wise independent distribution. The claim holds since $S_\ell(1,\ldots,1) = \binom{n}{\ell}$. □

When e.g. $k = O(\log n)$, which is often the case of interest, for $2k+3 \leq \ell \leq n-(2k+3)$, the RHS of (18) is much larger than the bound guaranteed by Equation (17). The tail bound provided by Lemma 6 can not therefore be extended to a satisfactory bound on the expectation. Furthermore, applying Lemma 6 with

$$t = \frac{1}{6e}\sqrt{\frac{n}{\ell k}}$$

implies that for any $(2k+2)$-wise independent distribution,

$$\Pr\left[|S_\ell(X)| \geq \binom{n}{\ell}\right] \leq \Pr\left[|S_\ell(X)| \geq (6et\sqrt{n})^\ell \left(\frac{k}{\ell}\right)^{\ell/2}\right] \leq 2\left(\frac{36e^2k\ell}{n}\right)^k.$$

When $k\ell = o(n)$, this is at most $O(n^{-k+o(1)})$. Comparing this to the bound given in Lemma 7, we see that the bound provided by Lemma 6 is nearly tight.

Acknowledgements

We thank Nati Linial, Raghu Meka, Yuval Peres, Dan Spielman and Avi Wigderson for helpful discussions.

References

[1] Noga Alon, Laszlo Babai and Alon Itai. *A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem.* J. Algorithms 7(4), pages

567–583, 1986.

[2] Benny Chor and Oded Goldreich. *On the power of two-point based sampling.* J. Complexity 5(1), pages 96–86, 1989.

[3] Parikshit Gopalan, Raghu Meka, Omer Reingold, Luca Trevisan and Salil P. Vadhan. *Better Pseudorandom Generators from Milder Pseudorandom restrictions.* FOCS, pages 120–129, 2012.

[4] J. Michael Steele. *The Cauchy-Schwarz Master Class.* Cambridge University Press, 2004.

[5] G Polya and G. Szego. *Problems and Theorems in Anaysis II.* Springer Classics in Mathematics, 1976.

A Proof of Fact B

For a univariate polynomial $p(\xi)$ and a root $y \in \mathbb{R}$ of p, denote by $\mathsf{mult}(p, y)$ the multiplicity of the root y in p. We use the following property of polynomials $p(\xi)$ with real roots [5], which can be proved using the interlacing of the zeroes of $p(\xi)$ and $p'(\xi)$: If $\mathsf{mult}(p', y) \geq 2$ then $\mathsf{mult}(p, y) \geq \mathsf{mult}(p', y) + 1$.

Proof of Fact B. Let

$$p(\xi) = \prod_{i \in [n]} (\xi + b_i) = \sum_{k=0}^{n} \xi^k S_{n-k}(b).$$

Consider $p^{(n-k-1)}(\xi)$ which is the $(n-k-1)^{th}$ derivative of $p(\xi)$. Since $S_k(b) = S_{k+1}(b) = 0$ for $k > 0$, it follows that ξ^2 divides $p^{(n-k-1)}(\xi)$ and hence $\mathsf{mult}(p^{(n-k-1)}, 0) \geq 2$. Applying the above fact $n - k - 1$ times, we get $\mathsf{mult}(p, 0) \geq n - k + 1$ so $S_n(b) = \ldots = S_k(b) = 0$. □

ECCC ISSN 1433-8092

http://eccc.hpi-web.de