Bounds on the sum of L1 influences

Yuval Filmus*

Institute for Advanced Study, Princeton, NJ

Hamed Hatami*

McGill University, Montreal, QC

April 15, 2014

Abstract

Let $f\colon \{-1,1\}^n \to [-1,1]$ have degree d as a multilinear polynomial. It is well-known that the total influence of f is at most d. Aaronson and Ambainis asked whether the total L_1 influence of f can also be bounded as a function of d. Bačkurs and Bavarian answered this question in the affirmative, providing a bound of $O(d^3)$ for general functions and $O(d^2)$ for homogeneous functions. We improve on their results by providing a bound of d^2 for general functions and $O(d \log d)$ for homogeneous functions.

1 Introduction

Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. The *influence* of the ith variable is

$$\mathrm{Inf}_i[f] = \Pr_{x \sim \{-1,1\}^n}[f(x) \neq f(x \oplus e_i)],$$

where $x \oplus e_i$ is obtained from x by flipping the ith coordinate. The *total influence* of the function is

$$\mathrm{Inf}[f] = \sum_{i=1}^n \mathrm{Inf}_i[f].$$

We define $\deg f$ as the degree of the unique multilinear polynomial representing f. It is well-known that $\mathrm{Inf}[f] \leq \deg f$, and much of the usefulness of influence in the study of Boolean functions rests on this property.

The notion of influence can be extended in several ways to real-valued functions $f\colon \{-1,1\}^n \to \mathbb{R}$. For each $p > 0$, one can define

$$\mathrm{Inf}_i^{(p)}[f] = \mathop{\mathbb{E}}_{x \sim \{-1,1\}^n}\left[\left|\frac{f(x) - f(x \oplus e_i)}{2}\right|^p\right], \quad \mathrm{Inf}^{(p)}[f] = \sum_{i=1}^n \mathrm{Inf}_i^{(p)}[f].$$

When f is Boolean, all these definitions agree with the original definition. It is well-known that $\mathrm{Inf}^{(2)}[f] \leq \deg f \cdot \|f\|_\infty$. While studying the query complexity of partial functions, Aaronson and Ambainis [AA11] asked whether $\mathrm{Inf}^{(1)}[f]$ can be bounded similarly. In other words, does every $f\colon \{-1,1\}^n \to [-1,1]$ of degree d satisfy $\mathrm{Inf}^{(1)}[f] = O(d^{O(1)})$?

Bačkurs and Bavarian [BB14] answered this in the affirmative, showing that $\mathrm{Inf}^{(1)}[f] = O(d^3)$. When f is homogeneous (that is, the unique multilinear polynomial representing f is homogeneous), they obtain an improved bound $\mathrm{Inf}^{(1)}[f] = O(d^2)$.

*Research conducted at the Simons Institute for the Theory of Computing during the 2013 fall semester on Real Analysis in Computer Science.

Our results Our main result is the bound $\mathrm{Inf}^{(1)}[f] \le d^2$. When f is homogeneous, we are able to show that $\mathrm{Inf}^{(1)}[f] = O(d \log d)$, and when f is both homogeneous and Boolean, we have a tight bound $\mathrm{Inf}^{(1)}[f] \le d$. When f is symmetric, we show that $\mathrm{Inf}^{(1)}[f] \le d + \tilde{O}(d/n)$. Following Bačkurs and Bavarian, we conjecture that the bound $\mathrm{Inf}^{(1)}[f] \le d$ holds for *all* functions $f\colon \{-1,1\}^n \to [-1,1]$ of degree d.

Paper organization Section 2 defines various notations used in the paper. Section 3 contains our upper bounds. Section 4 describes several functions for which the conjectured bound $\mathrm{Inf}^{(1)}[f] \le d$ is tight or almost tight. Section 5 contains several conjectures which would result in improvements to our main theorems. We believe that these conjectures are interesting in their own right.

Acknowledgements We thank Mohammad Bavarian, Guy Kindler, Elchanan Mossel and Krzysztof Oleszkiewicz for helpful discussions. This material is based upon work supported by the National Science Foundation under agreement No. DMS-1128155. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

2 Definitions

We use the notation $[n] = \{1, \dots, n\}$. The complement of a set $S \subseteq [n]$ will be denoted $\overline{S} = [n] \setminus S$. Probabilities or expectations over $\{-1,1\}^n$ are always with respect to the uniform probability measure. The point $(1, \dots, 1) \in \{-1,1\}^n$ will be denoted $\mathbf{1}$. A point $(x_1, \dots, x_n) \in \mathbb{R}^n$ will be abbreviated by $\mathbf{x}$.

Functions In this paper we consider functions $f\colon \{-1,1\}^n \to \mathbb{R}$. A function f is *Boolean* if f only attains the values ± 1. We think of a function $f\colon \{-1,1\}^n \to \mathbb{R}$ as having n input variables $x_1, \dots, x_n$ which are ± 1-valued. Every such function has a unique expansion as a multilinear polynomial over the variables $x_1, \dots, x_n$; this expansion is known as the *Fourier expansion* of f. Each set $S \subseteq [n]$ corresponds to a multilinear monomial $\chi_S = \prod_{i \in S} x_i$ known as a *Fourier character* or a *Walsh function*. The coefficient of χ_S in the expansion of f is known as the *Fourier coefficient* $\hat{f}(S)$.

The *degree* of f, denoted by $\deg f$, is the degree of its Fourier expansion. If all monomials appearing in the Fourier expansions of f have the same degree, then f is *homogeneous*. If $f(\mathbf{x})$ depends only on $x_1 + \cdots + x_n$ then f is *symmetric*.

Influence For $x \in \{-1,1\}^n$, we define $x \oplus e_i$ as the vector obtained from x by flipping the ith coordinate. For a function $f\colon \{-1,1\}^n \to \mathbb{R}$ and $i \in [n]$, we define

$$f_i(\mathbf{x}) = \frac{f(\mathbf{x}) - f(\mathbf{x} \oplus e_i)}{2} = \sum_{\substack{S \subseteq [n] \\ i \in S}} \hat{f}(S) \chi_S = x_i \frac{\partial f}{\partial x_i}(\mathbf{x}).$$

The ith *influence* of f is $\mathrm{Inf}_i[f] = \|f_i\|_1$ (in the introduction, we denoted this quantity by $\mathrm{Inf}_i^{(1)}[f]$, but for brevity we remove the superscript in the rest of the paper). The *total influence* of f is $\mathrm{Inf}[f] = \sum_{i=1}^n \mathrm{Inf}_i[f]$. Alternatively, if we define

$$\Delta(f)(\mathbf{x}) = \sum_{i=1}^n |f_i(\mathbf{x})|,$$

then $\mathrm{Inf}[f] = \|\Delta(f)\|_1$. When f is Boolean, $\Delta(f)(\mathbf{x})$ is the *sensitivity* of f at $\mathbf{x}$, which is the number of indices $i \in [n]$ such that $f(\mathbf{x} \oplus e_i) \neq f(\mathbf{x})$. The quantity $\mathrm{Inf}[f]$ is also known as the *average sensitivity* of f, and $\mathbb{S}(f) = \|\Delta(f)\|_\infty$ is also known as the *maximum sensitivity* of f.

Noise For $f\colon \{-1,1\}^n \to \mathbb{R}$ and $x \in \mathbb{R}$, the noise operator T_ρ takes the function f to the function $T_\rho f$ given by

$$T_\rho f = \sum_{S \subseteq [n]} \rho^{|S|} \hat{f}(S) \chi_S.$$

When $|\rho| \leq 1$, the noise operator has the following alternative interpretation. Fix a point $\mathbf{x} \in \{-1,1\}^n$. For each $i \in [n]$, independently let y_i be the unique ± 1-valued random variable such that $\mathbb{E}[x_i y_i] = \rho$. Then

$$T_\rho f(\mathbf{x}) = \mathop{\mathbb{E}}_{\mathbf{y}}[f(\mathbf{y})].$$

Chebyshev polynomials For each $d \geq 0$, the Chebyshev polynomial (of the first kind) T_d is the unique univariate polynomial such that $T_d(\cos\theta) = \cos(d\theta)$. The polynomial T_d has degree d, and is given by the recurrence $T_{d+1}(x) = 2xT_d(x) - T_{d-1}(x)$ with base cases $T_0(x) = 1$ and $T_1(x) = x$.

3 Upper bounds

In the rest of this section, we assume that $f\colon \{-1,1\}^n \to [-1,1]$ has degree d. We prove the following upper bounds on the total influence:

1. $\mathrm{Inf}[f] \leq d^2$.
2. If f is homogeneous then $\mathrm{Inf}[f] = O(d \log d)$.
3. If f is homogeneous and Boolean then $\mathrm{Inf}[f] \leq d$.
4. If f is symmetric then $\mathrm{Inf}[f] \leq d + O(\frac{d}{n} \log(dn))$.

The upper bound d^2 for general functions uses a Bernestein–Markov type inequality. The classical Bernestein–Markov theorem provides an upper bound on the derivative of a polynomial that is bounded in an interval.

Proposition 3.1 (Bernstein–Markov)**.** *Let $p\colon [-1,1] \to [-1,1]$ be a polynomial of degree d. For every $x \in [-1,1]$,*

$$p'(x) \leq \min\left(d^2, \frac{d}{\sqrt{1-x^2}}\right).$$

The generalization that we will use, due to Sarantopoulos [Sar91], extends Proposition 3.1 to Banach spaces. Using the classical Bernstein–Markov theorem instead results in the slightly weaker upper bound $2d^2$.

Sarantopoulos's theorem concerns polynomials in general Banach spaces. Since in this paper we only need the finite dimensional case, to avoid introducing unnecessary terminology, we will state Sarantopoulos's theorem for the special case of finite dimensional Banach spaces. Recall that for a *finite dimensional* Banach space $E = (\mathbb{R}^n, \|\cdot\|)$, the Fréchet derivative of a differentiable function $f\colon E \to \mathbb{R}$ at a point $\mathbf{x}$ is the linear operator $Df(\mathbf{x})\colon E \to \mathbb{R}$ defined as

$$Df(\mathbf{x})\colon \mathbf{y} \mapsto \sum_{i=1}^n y_i \frac{\partial f}{\partial x_i}(\mathbf{x}).$$

Proposition 3.2 (Sarantopoulos [Sar91, Theorems 1 and 2])**.** *Let $E = (\mathbb{R}^n, \|\cdot\|)$ be a finite dimensional Banach space and $P\colon \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d satisfying $|P(\mathbf{x})| \leq 1$ for all $\|\mathbf{x}\| \leq 1$. Then $|DP(\mathbf{x})\mathbf{y}| \leq \min\left(d^2, \frac{d}{\sqrt{1-\|x\|^2}}\right)$ for all $\|\mathbf{x}\|, \|\mathbf{y}\| \leq 1$, where DP is the Fréchet derivative of P.*

Theorem 3.1. *Let $f\colon \{-1,1\}^n \to [-1,1]$ be a function of degree d. Then*

$$\mathrm{Inf}[f] \leq \|\Delta(f)\|_\infty \leq d^2.$$

Proof. Clearly $\mathrm{Inf}[f] = \|\Delta(f)\|_1 \le \|\Delta(f)\|_\infty$, and so it suffices to show that $|\Delta(f)(\mathbf{x})| \le d^2$ for all $\mathbf{x} \in \{-1,1\}^n$. Consider now $[-1,1]^n$ as the unit ball in the Banach space $(\mathbb{R}^n, \|\cdot\|_\infty)$. The Fréchet derivative of f at the point $\mathbf{x}$ is the linear operator

$$Df(\mathbf{x})\mathbf{y} = \sum_{i=1}^n y_i \frac{\partial f}{\partial x_i}(\mathbf{x}) = \sum_{i=1}^n y_i \frac{f_i(\mathbf{x})}{x_i}.$$

In particular, for every $\mathbf{x} \in \{-1,1\}^n$, there is some $\mathbf{y} \in \{-1,1\}^n$ such that

$$\Delta(f)(\mathbf{x}) = Df(\mathbf{x})\mathbf{y}.$$

Proposition 3.2 immediately implies that $|\Delta(f)(\mathbf{x})| \le d^2$ for all $\mathbf{x} \in \{-1,1\}^n$. □

The argument in fact gives a bound on $\|\Delta(f)\|_\infty$, and in this respect, it is tight. Indeed, consider the functions $f_n(x_1,\ldots,x_n) = T_d(\frac{x_1+\cdots+x_n}{n})$. At the point $\mathbf{1}$ we have

$$\lim_{n\to\infty} \Delta(f_n)(\mathbf{1}) = \lim_{n\to\infty} n \left| \frac{T_d(1) - T_d(1-\frac{2}{n})}{2} \right| = T_d'(1) = d^2.$$

The upper bound $O(d\log d)$ for homogeneous functions uses a result of Harris [Har97].

Proposition 3.3 (Harris [Har97])**.** *Let h be a real polynomial satisfying $h(\epsilon) \le (1+|\epsilon|)^d$ for all $\epsilon \in \mathbb{R}$. Then $h'(0) = O(d\log d)$.*

We comment that Révész and Sarantopoulos [RS03] show that the bound $O(d\log d)$ is optimal.

Theorem 3.2. *Let $f\colon \{-1,1\}^n \to [-1,1]$ be a homogeneous function of degree d. Then*

$$\mathrm{Inf}[f] \le \|\Delta(f)\|_\infty \le O(d\log d).$$

Proof. Since $\mathrm{Inf}[f] = \|\Delta(f)\|_1 \le \|\Delta(f)\|_\infty$, it suffices to show that $|\Delta(f)(\mathbf{1})| \le O(d\log d)$. Let S be the set of $i \in [n]$ such that $f_i(\mathbf{1}) \ge 0$. Then

$$|\Delta(f)(\mathbf{1})| = \sum_{i\in S} f_i(\mathbf{1}) - \sum_{i\in\overline{S}} f_i(\mathbf{1}).$$

Define the bivariate polynomial $g(x,y) = f(\overbrace{x,\ldots,x}^{S}, \overbrace{y,\ldots,y}^{\overline{S}})$. Since f is multilinear, its extension to the continuous cube $[-1,1]^n$ is also bounded in absolute value by 1. This, together with homogeneity of f, implies $|g(x,y)| \le \max(|x|,|y|)^d$. In particular, the function $h(\epsilon) = g(1+\epsilon, 1-\epsilon)$ is a polynomial satisfying $h(\epsilon) \le \max(|1+\epsilon|,|1-\epsilon|)^d = (1+|\epsilon|)^d$. Proposition 3.3 implies that $h'(0) = O(d\log d)$. We have

$$h(\epsilon) = \sum_{T\subseteq[n]} \hat{f}(T)(1+\epsilon)^{|T\cap S|}(1-\epsilon)^{|T\cap\overline{S}|},$$

and so

$$h'(0) = \sum_{T\subseteq[n]} \hat{f}(T)(|T\cap S| - |T\cap\overline{S}|).$$

On the other hand,

$$f_i(\mathbf{1}) = \sum_{\substack{T\subseteq[n]\\ i\in T}} \hat{f}(T)\chi_T(\mathbf{1}) = \sum_{\substack{T\subseteq[n]\\ i\in T}} \hat{f}(T),$$

and so

$$|\Delta(f)(\mathbf{1})| = \sum_{i\in S} f_i(\mathbf{1}) - \sum_{i\in\overline{S}} f_i(\mathbf{1}) = \sum_{T\subseteq[n]} \hat{f}(T)(|T\cap S| - |T\cap\overline{S}|).$$

Therefore $|\Delta(f)(\mathbf{1})| = h'(0) = O(d\log d)$. □

In Section 5 we discuss a variant of this argument which could result in better bounds.

When f is not only homogeneous but also Boolean, we can use a bootstrapping argument to prove an optimal bound of d.

Theorem 3.3. *Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be a homogeneous Boolean function of degree d. Then*

$$\mathrm{Inf}[f] \le \|\Delta(f)\|_\infty \le d.$$

Proof. We start by proving the theorem for odd d. For any homogeneous Boolean function $g\colon \{-1,1\}^m \to \{-1,1\}$ of odd degree D we can define a homogeneous Boolean function $g \circ f\colon \{-1,1\}^{mn} \to \{-1,1\}$ of odd degree Dd by

$$(g \circ f)(\mathbf{x}_1, \ldots, \mathbf{x}_m) = g(f(\mathbf{x}_1), \ldots, f(\mathbf{x}_m)).$$

We claim that $\mathbb{S}(g \circ f) = \mathbb{S}(g)\mathbb{S}(f)$ (recall that $\mathbb{S}(f) = \|\Delta(f)\|_\infty$). Indeed, $\mathbb{S}(g \circ f) \le \mathbb{S}(g)\mathbb{S}(f)$ follows directly from the interpretation of $\Delta(g \circ f)(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ as the number of indices whose flipping changes the value of $g \circ f$. For the other direction, let $\mathbf{x} \in \{-1,1\}^n$ be a point such that $\Delta(f)(\mathbf{x}) = \mathbb{S}(f)$. Since f is homogeneous and has odd degree, it is an odd function, and so $\Delta(f)(-\mathbf{x}) = \mathbb{S}(f)$ and $f(-\mathbf{x}) = -f(\mathbf{x})$. Therefore we can assume without loss of generality that $f(\mathbf{x}) = 1$. Let $\mathbf{y} \in \{-1,1\}^m$ be a point such that $\Delta(g)(\mathbf{y}) = \mathbb{S}(g)$. For $i \in [m]$, define $\mathbf{x}_i$ to be $\mathbf{x}$ if $y_i = 1$ and $-\mathbf{x}$ if $y_i = -1$. It is not hard to check that $\Delta(g \circ f)(\mathbf{x}_1, \ldots, \mathbf{x}_m) = \mathbb{S}(g)\mathbb{S}(f)$, showing that $\mathbb{S}(g \circ f) = \mathbb{S}(g)\mathbb{S}(f)$.

Let now $f^{\circ k}$ be obtained by composing f with itself k times. Since $\deg f^{\circ k} = d^k$, Theorem 3.2 shows that $\mathbb{S}(f)^k = \mathbb{S}(f^{\circ k}) = O(d^k k \log d)$. Taking kth roots and the limit $k \to \infty$, we deduce $\mathbb{S}(f) \le d$. This completes the proof of the theorem for odd d.

Suppose now that d is even. Define a homogeneous Boolean function $r\colon \{-1,1\}^{n+1} \to \{-1,1\}$ of degree $d+1$ by $r(\mathbf{x}, y) = yf(\mathbf{x})$. Clearly $\Delta(r)(\mathbf{x}, y) = \Delta(f)(\mathbf{x}) + 1$ for all $(\mathbf{x}, y) \in \{-1,1\}^{n+1}$, and so $\mathbb{S}(r) = \mathbb{S}(f) + 1$. Since $d+1$ is odd, we already know that $\mathbb{S}(r) \le d+1$, and so $\mathbb{S}(f) \le d$. □

This bound is attained, for example, by Fourier characters. Other cases of equality are discussed in Section 4.

We close this section by showing an upper bound of $d + O(\log d)$ for symmetric functions. We use the classical Bernstein-Markov theorem on real polynomials.

Theorem 3.4. *Let $f\colon \{-1,1\}^n \to [-1,1]$ be a symmetric function of degree d. Then*

$$\mathrm{Inf}[f] \le d + O\left(\frac{d \log(dn)}{n}\right).$$

Proof. Theorem 3.1 allows us to assume that $d \ge 2$. Since f is symmetric, we can write $f(\mathbf{x}) = p(\frac{x_1 + \cdots + x_n}{n})$ for some polynomial $p\colon [-1,1]^n \to [-1,1]$ of degree d. We can calculate explicitly

$$\mathrm{Inf}[f] = n\,\mathrm{Inf}_n[f] = n \mathop{\mathbb{E}}_{\mathbf{x} \in \{-1,1\}^{n-1}} \left[\left| \frac{p(\frac{S+1}{n}) - p(\frac{S-1}{n})}{2} \right| \right], \text{ where } S = x_1 + \cdots + x_{n-1}.$$

The mean value theorem shows that for some $\theta_S \in [-1,1]$,

$$\frac{1}{2}|p(\tfrac{S+1}{n}) - p(\tfrac{S-1}{n})| = \frac{1}{n}|p'(\tfrac{S+\theta_S}{n})| \le \frac{1}{n} \min\left(d^2, \frac{d}{\sqrt{1 - (|S|+1)^2/n^2}}\right),$$

using Proposition 3.1. Let $T = \sqrt{n \log(dn)}$. Then

$$\begin{aligned}
\operatorname{Inf}[f] &= n \mathop{\mathbb{E}}_{\mathbf{x} \in \{-1,1\}^{n-1}} \left[\left| \frac{p(\frac{S+1}{n}) - p(\frac{S-1}{n})}{2} \right| \right] \\
&\leq \frac{d}{\sqrt{1 - (T+1)^2/n^2}} + d^2 \Pr[|S| > T] \\
&\leq \frac{d}{\sqrt{1 - O(\log(dn)/n)}} + 2d^2 e^{-2T^2/n} \\
&\leq d + O\left(\frac{d \log(dn)}{n}\right) + \frac{2}{n^2} = d + O\left(\frac{d \log(dn)}{n}\right),
\end{aligned}$$

using Hoeffding's bound in the second inequality. □

4 Tight examples

Following Bačkurs and Bavarian [BB14], we conjecture that the total influence of a function $f\colon \{-1,1\}^n \to [-1,1]$ of degree d is at most d. In this section we discuss several examples of functions f which achieve or almost achieve this bound.

Boolean homogeneous functions attaining the bound When f is Boolean and homogeneous, Theorem 3.3 shows that a function with total influence d must have sensitivity d at all points (recall that the sensitivity of f at a point $\mathbf{x}$ is the number of indices $i \in [n]$ such that flipping the ith input changes the value of the function). The quintessential example of such a function is a Fourier character of degree d, that is χ_S for some set $S \subseteq [n]$ of cardinality $|S| = d$. Another example (with $d = 2$) is the function

$$f_4(x,y,z,w) = \frac{x(z+w) + y(z-w)}{2}.$$

For an arbitrary degree $d \geq 2$, the function $f_4(x_1,x_2,x_3,x_4)x_5 \cdots x_{d+2}$ has total influence d. This shows that even when f is Boolean and homogeneous, characters are not the unique functions having total influence d.

Non-Boolean functions attaining the bound The following two quadratic functions satisfy $\|f\|_\infty = 1$ and $\Delta f \equiv 2$, and in particular have total influence 2:

$$\begin{aligned}
s(x,y,z,w) &= \frac{xy - zw}{2} + \frac{\sqrt{2}-1}{8}(xz + yw), \\
t(x,y,z,w) &= \frac{xy - zw}{2} + \frac{\sqrt{2}-1}{16}(x+y)(z+w).
\end{aligned}$$

Symmetric functions almost attaining the bound Let $P\colon [-1,1] \to [-1,1]$ be a polynomial of degree d, and consider the corresponding symmetric function $f(\mathbf{x}) = P(\frac{x_1 + \cdots + x_n}{n})$. For large n we have

$$\operatorname{Inf}[f] = n \operatorname{Inf}_n[f] \approx n \left| \frac{P(\frac{1}{n}) - P(-\frac{1}{n})}{2} \right| \approx |P'(0)|.$$

The Bernstein–Markov theorem (Proposition 3.1) shows that $P'(0) \leq d$. When d is odd, setting P to the Chebyshev polynomial T_d we have $P'(0) = d$, and as $n \to \infty$, the estimates above can be made precise to show that $\operatorname{Inf}[f] \to d$. One could wonder whether these functions provide a counter-example to the conjecture that $\operatorname{Inf}[f] \leq \deg f$. However, it is not difficult to see that in a deleted neighborhood of 0, $T_d'(0) < d$, and so for large n the estimates show that $\operatorname{Inf}[f] < d$, that is, the limit is approached from below. Numerical experiments suggest that $\operatorname{Inf}[f] < d$ holds also for small n.

5 Conjectures

In this section we discuss two directions for improving our results. The first direction aims at improving Theorem 3.1 to a bound of $O(d^{3/2})$ on the total influence. The second direction aims at improving Theorem 3.2 to a bound of $O(d)$ on the total influence of homogeneous functions.

5.1 General functions

We start by proving an $O(d^{3/2})$ bound on the total influence of homogeneous functions. While Theorem 3.2 provides a better upper bound of $O(d \log d)$, this new method could potentially extend to general functions. The proof uses Sarantopoulos's extension (Proposition 3.2) of the Markov–Bernstein theorem. (We could also use the classical Bernstein's theorem.)

Theorem 5.1. *Let $f\colon \{-1,1\}^n \to [-1,1]$ be a homogeneous function of degree d. Then*

$$\mathrm{Inf}[f] \le \|\Delta(f)\|_\infty \le O(d^{3/2}).$$

Proof. Let $\alpha = 1 - 1/d$, and define $g = T_\alpha f$. Note that $g(\mathbf{x}) = f(\alpha\mathbf{x})$ and similarly $\Delta(g)(\mathbf{x}) = \Delta(f)(\alpha\mathbf{x})$. Since $|\alpha| \le 1$, the interpretation of $T_\alpha f$ as an averaging operator shows that $\|g\|_\infty \le \|f\|_\infty \le 1$. As in the proof of Theorem 3.1, Proposition 3.2 shows that for all $\mathbf{x} \in \{-1,1\}^n$,

$$\Delta(g)(\mathbf{x}) = \Delta(f)(\alpha\mathbf{x}) \le \frac{d}{\sqrt{1-\alpha^2}}.$$

Since g is homogeneous, $\Delta(g)(\mathbf{x}) = \alpha^d \Delta(f)(\mathbf{x})$, and so

$$\Delta(f)(\mathbf{x}) \le \alpha^{-d} \frac{d}{\sqrt{1-\alpha^2}} = O(d^{3/2}). \qquad \square$$

When f is not homogeneous, we can try to fix the argument as follows.

Lemma 5.1. *Let $f\colon \{-1,1\}^n \to [-1,1]$ be a function of degree d. For all $\alpha \in [-1,1]$ we have*

$$\mathrm{Inf}[f] \le \max_{i \in [n]} \frac{\|f_i\|_1}{\|T_\alpha f_i\|_1} \frac{d}{\sqrt{1-\alpha^2}}.$$

Proof. Fix α, and let $g = T_\alpha f$. As in Theorem 5.1,

$$\|\Delta(g)\|_\infty \le \frac{d}{\sqrt{1-\alpha^2}}.$$

On the other hand, as $g_i = T_\alpha f_i$,

$$\mathrm{Inf}[f] = \sum_{i=1}^n \|f_i\|_1 = \sum_{i=1}^n \frac{\|f_i\|_1}{\|T_\alpha f_i\|_1} \|g_i\|_1 \le \left(\max_{i\in[n]} \frac{\|f_i\|_1}{\|T_\alpha f_i\|_1}\right) \mathrm{Inf}[g] \le \left(\max_{i\in[n]} \frac{\|f_i\|_1}{\|T_\alpha f_i\|_1}\right) \frac{d}{\sqrt{1-\alpha^2}}. \qquad \square$$

This prompts the following definition.

Definition 5.1. Let $d \ge 1$ and $\alpha \in [-1,1]$. Define

$$C_{d,\alpha} = \sup_f \frac{\|f\|_1}{\|T_\alpha f\|_1},$$

where the supremum ranges over all n and all functions $f\colon \{-1,1\}^n \to \mathbb{R}$ of degree at most d.

We can restate the conclusion in Lemma 5.1 as follows:

$$\mathrm{Inf}[f] \le C_{d,\alpha} \frac{d}{\sqrt{1-\alpha^2}}.$$

In particular, if $C_{d,1-1/d} = O(1)$ then $\mathrm{Inf}[f] = O(d^{3/2})$.

The best bound on $C_{d,\alpha}$ we can prove is the following.

Lemma 5.2. *For all functions $f\colon \{-1,1\}^n \to \mathbb{R}$ of degree d and all $\alpha \in (0,1]$,*

$$\|T_\alpha f\|_1 \ge \alpha^{\min(d^2,n)} \|f\|_1.$$

In particular, $C_{d,\alpha} \le \alpha^{-d^2}$.

Proof. We start by showing that $\|T_\alpha f\|_1 \ge \alpha^n \|f\|_1$. Note first that for all $i \in [n]$, we have $\|f_i\|_1 \le \|f\|_1$. This follows from

$$|f_i(x)| + |f_i(x \oplus e_i)| = 2\left|\frac{f(x) - f(x \oplus e_i)}{2}\right| \le |f(x)| + |f(x \oplus e_i)|.$$

Now, it is well-known that

$$\frac{dT_{e^{-\epsilon}} f(\mathbf{x})}{d\epsilon} = -Lf(\mathbf{x}),$$

where the Laplacian Lf is given by $Lf = f_1 + \cdots + f_n$. Therefore

$$\frac{d\|T_{e^{-\epsilon}} f\|_1}{d\epsilon} \ge -\|Lf\|_1 \ge -\sum_{i=1}^n \|f_i\|_1 \ge -n\|f\|_1.$$

Let $\phi(\delta) = \|T_{e^{-\delta}}\|_1$. Applying the inequality above to $T_{e^{-\delta}} f$ shows that $\phi'(\delta) \ge -n\phi(\delta)$ and so $(\log \phi(\delta))' \ge -n$. Integrating, we obtain $\phi(\delta)/\phi(0) \ge e^{-\delta n}$. Taking $\delta = -\log \alpha$, we deduce

$$\|T_\alpha f\|_1 = \phi(\delta) \ge e^{-\delta n} \phi(0) = \alpha^n \|f\|_1.$$

We proceed with the proof that $\|T_\alpha f\|_1 \ge \alpha^{d^2} \|f\|_1$ (we thank K. Oleszkiewicz for help with this proof). Let X_d denote the vector space of all real-valued polynomials P of degree at most d satisfying $|P(x)| \le 1$ for all $|x| \le 1$. Standard results in functional analysis (see for example Rivlin's book [Riv75]) show that every linear functional $\Phi\colon X_d \to \mathbb{R}$ can be represented as

$$\Phi\colon P \mapsto \sum_i c_i P(\alpha_i)$$

for some points $\alpha_i \in [-1,1]$, in such a way that the maximum of ΦP over X_d is $\sum_i |c_i|$. Applying this result to the functional $\Phi f\colon f \mapsto f'(1)$, we obtain coefficients c_i, α_i satisfying $\sum_i |c_i| = d^2$ (according to Markov's inequality). Since Φ maps x^k to k for all $k \le d$, for such k we have

$$\sum_i c_i \alpha_i^k = k.$$

Since f has degree d, this implies that

$$\sum_i c_i T_{\alpha_i} f = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S \sum_i c_i \alpha_i^{|S|} = \sum_{S \subseteq [n]} |S| \hat{f}(S) \chi_S = Lf.$$

The interpretation of T_β as an average shows that $\|T_\beta f\|_1 \le \|f\|_1 \le 1$ for every $\beta \in [-1,1]$. In particular,

$$\|Lf\|_1 \le \sum_i \|c_i T_{\alpha_i} f\|_1 \le \sum_i |c_i| = d^2.$$

As in the preceding half of the proof, this implies that $\|T_\alpha f\|_1 \ge \alpha^{d^2} \|f\|_1$. ☐

Unfortunately, plugging this bound in Lemma 5.1 does not result in any improvement over Theorem 3.1.

5.2 Homogeneous functions

Theorem 3.2 shows that the total influence of a homogeneous function of degree d is at most $O(d \log d)$. The argument relies on a result of Harris [Har97] showing that a real polynomial satisfying $h(\epsilon) \leq (1 + |\epsilon|)^d$ for all $\epsilon \in \mathbb{R}$ also satisfies $h'(0) = O(d \log d)$; Révész and Sarantopoulos [RS03] show that the bound on $h'(0)$ is tight. Recall that the function $h(\epsilon)$ figures in the proof in the following way. For a certain set $S \subseteq [n]$, we define $g(x, y) = f(\overbrace{x, \ldots, x}^{S}, \overbrace{y, \ldots, y}^{\overline{S}})$. Since f is multilinear and homogeneous, $|g(x, y)| \leq \max(|x|, |y|)^d$, and so the function $h(\epsilon) = g(1 + \epsilon, 1 - \epsilon)$ satisfies $h(\epsilon) \leq (1 + |\epsilon|)^d$.

Révész and Sarantopoulos comment that every real polynomial h satisfying $h(\epsilon) \leq (1 + |\epsilon|)^d$ can be lifted to a bivariate homogeneous polynomial $g(x, y)$ given by $g(x, y) = y^d h(x/y)$. This polynomial satisfies $|g(x, y)| \leq |y|^d (1 + |x|/|y|)^d = (|x| + |y|)^d$, and so the bound on the derived $h'(0)$ can be achieved by some function $g(x, y)$ satisfying $|g(x, y)| \leq (|x| + |y|)^d$. In our case, we have the stronger guarantee $|g(x, y)| \leq \max(|x|, |y|)^d$. We can modify the proof to reflect this stronger guarantee.

Definition 5.2. Let K_d be the supremum of $|h'(1)|$ over all polynomials satisfying $h(\epsilon) \leq \max(1, |\epsilon|^d)$ for all $\epsilon \in \mathbb{R}$.

Lemma 5.3. *Let $f \colon \{-1, 1\}^n \to [-1, 1]$ be a homogeneous function of degree d. Then*

$$\operatorname{Inf}[f] \leq \|\Delta(f)\|_\infty \leq 2K_d.$$

Proof. Since $\operatorname{Inf}[f] = \|\Delta(f)\|_1 \leq \|\Delta(f)\|_\infty$, it is enough to show that $|\Delta(f)(\mathbf{1})| \leq K_d$. Let S be the set of $i \in [n]$ such that $f_i(\mathbf{1}) \geq 0$. Then

$$|\Delta(f)(\mathbf{1})| = \sum_{i \in S} f_i(\mathbf{1}) - \sum_{i \in \overline{S}} f_i(\mathbf{1}).$$

Define the bivariate polynomial $g(x, y) = f(\overbrace{x, \ldots, x}^{S}, \overbrace{y, \ldots, y}^{\overline{S}})$. Since f is multilinear and homogeneous, $|g(x, y)| \leq \max(|x|, |y|)^d$. In particular, the function $h(\epsilon) = g(\epsilon, 1)$ is a polynomial satisfying $h(\epsilon) \leq \max(1, |\epsilon|)^d$. By definition, $|h'(1)| \leq K_d$. We have

$$h(\epsilon) = \sum_{T \subseteq [n]} \hat{f}(T) \epsilon^{|T \cap S|},$$

and so

$$h'(1) = \sum_{T \subseteq [n]} \hat{f}(T) |T \cap S|.$$

We conclude that

$$\left| \sum_{T \subseteq [n]} \hat{f}(T) |T \cap S| \right| \leq K_d.$$

Similarly,

$$\left| \sum_{T \subseteq [n]} \hat{f}(T) |T \cap \overline{S}| \right| \leq K_d.$$

As in the proof of Theorem 3.2,

$$|\Delta(f)(\mathbf{1})| = \sum_{T \subseteq [n]} \hat{f}(T) (|T \cap S| - |T \cap \overline{S}|) \leq 2K_d. \qquad \square$$

Harris [Har98a, Har98b] develops systematically a method aimed toward computing constants like K_d using Lagrange interpolation. Révész and Sarantopoulos [RS03] present a different framework which employs potential theory. We believe that these methods can be used to estimate K_d asymptotically. We conjecture that $K_d = \Theta(d)$, leading to a proof that $\mathrm{Inf}[f] \leq O(d)$ for homogeneous functions.

We have computed $K_1 = 1$ and $K_2 = 1 + \sqrt{2}$. The bound for $d = 1$ is attained for $h(x) = \pm x$. For $d = 2$, it is attained for

$$h(x) = \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)(x^2 - 1) + \frac{1}{\sqrt{2}}x.$$

The upper bound $K_1 \leq 1$ is trivial. The upper bound $K_2 \leq 1 + \sqrt{2}$ follows by Lagrange interpolation (following Harris) with the points $1 \pm \sqrt{2}$ (a priori, the method requires three points, but the third point cancels out in the calculation).

References

[AA11] Scott Aaronson and Andris Ambainis. The need for structure in quantum speedups. In *Innovations in Computer Science (ICS)*, pages 338–352, 2011.

[BB14] Artūrs Bačkurs and Mohammad Bavarian. On the sum of L1 influences. In *Conference on Computational Complexity (CCC)*, 2014.

[Har97] Lawrence A. Harris. A Bernstein–Markov theorem for normed spaces. *J. Math. Anal. Appl.*, 208:476–486, 1997.

[Har98a] Lawrence A. Harris. Coefficients of polynomials of restricted growth on the real line. *J. Approx. Theory*, 93:293–312, 1998.

[Har98b] Lawrence A. Harris. Optimal oscillation points for polynomials of restricted growth on the real line. In T. M. Rassias, editor, *Approximation Theory and Applications*, pages 85–106. Hadronic Press, 1998.

[Riv75] Theodore J. Rivlin. *The Chebyshev polynomials*. John Wiley and Sons, 1975.

[RS03] Szilárd Gy. Révész and Yannis Sarantopoulos. On Markov constants of homogeneous polynomials over real normed spaces. *East J. Approx.*, 9(3):277–304, 2003.

[Sar91] Yannis Sarantopoulos. Bounds on the derivatives of polynomials on Banach spaces. *Math. Proc. Camb. Phil. Soc.*, 307:307–312, 1991.