

# ON GRAPHS AND THE GOTSMAN-LINIAL CONJECTURE FOR $d = 2$

SPUR+ FINAL PAPER, SUMMER 2017

HYO WON KIM AND CHRISTOPHER MALDONADO  
MENTOR JAKE WELLENS  
PROJECT SUGGESTED BY JAKE WELLENS

ABSTRACT. Given a polynomial  $p(x) : \{-1, 1\}^n \rightarrow \mathbb{R}^\times$ , the associated polynomial threshold function (PTF) is a boolean function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  defined by  $f(x) = \text{sgn}(p(x))$ . A conjecture of Gotsman and Linial [2] posits that among all degree- $d$  PTFs, the one with the largest influence corresponds to the symmetric degree- $d$  polynomial which alternates sign at the  $d + 1$  values of  $\sum_{i=1}^n x_i$  closest to 0, having influence  $\Theta(d\sqrt{n})$ . We give an infinite class of counterexamples when  $d = 2$ , thereby disproving the conjecture as originally stated. However, as a theorem of Kane [3] shows, for constant  $d$ , and any degree- $d$  PTF  $f$  on  $n$  variables,  $\mathbf{I}[f] = O(\sqrt{n} \cdot \text{poly log}(n))$ , so at least the conjectured bound  $O(d\sqrt{n})$  is not too far off for small  $d$ . We examine the case  $d = 2$ , i.e. when  $f(x) = \text{sgn}(x^T A x + b^T x + c)$ , and using only elementary methods, we remove the  $\text{poly log}(n)$  from Kane's bound in a variety of special cases, based on graph properties of the matrix  $A$ , interpreted as a weighted adjacency matrix.

## 1. INTRODUCTION AND MAIN RESULTS

An useful way to measure the complexity of a boolean function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is through its *influence* or *average sensitivity*, defined as

$$\mathbf{I}[f] := \sum_{i=1}^n \Pr_{x \sim \{-1, 1\}^n} [f(x_1, \dots, x_i, \dots, x_n) \neq f(x_1, \dots, -x_i, \dots, x_n)].$$

Identifying boolean functions with subsets of the hypercube graph on  $\{-1, 1\}^n$ , influence counts (up to a factor of  $2^{-n+1}$ ) the number of *boundary edges* of a set. By way of Fourier analysis, influence can be related to many other complexity measures, such as noise sensitivity<sup>1</sup>, spectral concentration, decision tree and circuit complexity – see [5] for a largely self-contained summary of these relationships. To be imprecise, the general trend is that *low-influence functions are “easier to compute”* in many models of computation than are those with high influence. Thus, lower bounds on influence can provide circuit lower bounds (e.g.  $\text{PARITY} \notin \text{AC}^0$ ), while upper bounds on influence can provide algorithms for PAC learning classes of functions from random examples. This paper is concerned with upper bounding the influence of a class of boolean functions called *quadratic threshold functions*.

*Polynomial threshold functions* (PTFs), more generally, are boolean functions  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  which can be represented as the sign of a polynomial, i.e.

$$f(x) = \text{sgn}(p(x_1, \dots, x_n))$$

for some multilinear polynomial  $p$ .<sup>2</sup> When the degree of  $p$  is equal to  $d$ , we say  $f$  is a degree- $d$  PTF. Perhaps most common in applications are PTFs of degree  $d = 1$  or  $d = 2$ , known as *linear threshold functions* (LTFs) and *quadratic threshold functions* (QTFs), respectively. LTFs have been extensively studied in the literature dating back to the 1960’s, as such functions play a vital role in computational models of neural networks (see [4] for a classical introduction). In particular, it has long been known (as early as 1971, see [6]) that, among linear threshold functions, the majority function

$$\text{MAJ}_n(x_1, \dots, x_n) = \text{sgn}(x_1 + \dots + x_n)$$

has the largest total influence of  $n$ -variable LTFs. This influence can be computed explicitly and shown to be  $\Theta(\sqrt{n})$ . This fact has been shown ([1],[7]) to imply a noise-sensitivity bound of  $\text{NS}_\delta(f) \leq O(\sqrt{\delta})$  for any LTF  $f$ , which in turn implies  $f$  is  $\epsilon$ -concentrated up to degree  $O(1/\epsilon^2)$ , and therefore  $f$  can be PAC learned<sup>3</sup> from random examples in time  $n^{O(1/\epsilon^2)}$ .

Like their linear counterparts, quadratic threshold functions show up in a variety of real-world applications, such as Boltzmann machines (a type of stochastic neural network) and the Ising model (for describing the statistics of ferromagnetic spin configurations), although much less is known about QTFs than LTFs. We do not

<sup>1</sup>The *noise sensitivity*  $\text{NS}_\delta(f)$  of  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is defined as  $\Pr_{x \sim \{-1, 1\}^n} [f(x) \neq f(y)]$ , where each bit  $y_i$  of  $y$  is formed independently by flipping  $x_i$  with probability  $\delta$ .

<sup>2</sup>For obvious reasons, such functions are also called *threshold of parities circuits*. Note also that any boolean function on  $n$  variables is equal to the sign of its Fourier transform, and is therefore trivially a degree- $n$  PTF. Usually one is concerned with functions having *low-degree* PTF representations, often much lower than the degree of the function’s Fourier transform.

<sup>3</sup>This is not the first or the fastest algorithm for learning halfspaces. However, since  $\text{NS}_\delta(f_1 \wedge f_2) \leq \text{NS}_\delta(f_1) + \text{NS}_\delta(f_2)$ , this also yields the first algorithm for PAC learning *intersections* of  $O(1)$  halfspaces in polynomial time.

know which QTF has the largest influence, nor do we know tight bounds on influence, noise stability or spectral concentration.

Noting that  $\text{MAJ}_n$  is the unique unbiased LTF which is *symmetric* in its  $n$  variables, one might expect something similar to hold for PTFs of any degree  $d$  — namely that, among all degree  $d$  PTFs  $f$  on  $n$  variables, the maximum value of  $\mathbf{I}[f]$  is obtained by a symmetric function. In that case, the maximizer would be of the form

$$f(x) = \text{sgn} \left( p_d \left( \sum_{i=1}^n x_i \right) \right)$$

for some univariate degree- $d$  polynomial  $p_d(t)$ . In this case,  $f$  would be sensitive to  $x$  in the  $i$ th coordinate if changing the sign of  $x_i$  shifts  $\sum_{i=1}^n x_i$  across a root of  $p_d(t)$ , and since  $\Pr_{x \sim \{-1,1\}^n}[\sum_{i=1}^n x_i = k]$  decreases as  $k$  moves away from 0, the  $p$  which maximizes  $\mathbf{I}[f]$  will be one which alternates sign at the  $d+1$  values of  $\sum_{i=1}^n x_i$  closest to 0. Explicitly, if  $n$  is odd, then one may take

$$p_d(t) := t(t-2)(t+2)(t-4)(t+4) \cdots (t + (-1)^{\lfloor d/2 \rfloor} \cdot 2 \lfloor d/2 \rfloor),$$

while for even  $n$ , one may take  $p_d(t)$  to be the same as above with  $t+1$  plugged in for  $t$ . Indeed, this reasoning led Gotsman and Linial to make the following conjecture:

**Conjecture.** (*Gotsman-Linial [2]*): *The degree- $d$  PTF on  $n$  variables with maximal influence is fully symmetric. Equivalently, for any PTF  $f$  on  $n$  variables with degree  $d \geq 2$ ,*<sup>4</sup>

$$\mathbf{I}[f] \leq \mathbf{I}[\text{sgn}(p_d(x_1 + \cdots + x_n))] = \begin{cases} 2^{-n+1} \sum_{k=0}^{d-1} \binom{n}{\lfloor (n-k)/2 \rfloor} (n - \lfloor (n-k)/2 \rfloor), & n \text{ even} \\ 2^{-n+1} \sum_{k=0}^{d-1} \binom{n}{\lceil (n-k)/2 \rceil} (n - \lceil (n-k)/2 \rceil), & n \text{ odd} \end{cases} \\ =: \mathbf{I}_{GL}(n, d).$$

We show in this paper that the Gotsman-Linial conjecture is not, in general, true. Using a linear-programming based algorithm described in Section 3, we were able to identify the following 3-variable quadratic polynomial, which produces counterexamples to the conjecture in the case of  $d = 2$  for all odd  $n \geq 5$ .

$$p(x, y, z) := 2x(1 - 7y + z) + 4y - 7y^2 + 4yz + 6z + 3z^2,$$

Indeed:

**Theorem 1.1.** *For odd  $n \geq 5$ , let*

$$f_n(x) := \text{sgn} \left( p \left( x_1, \sum_{i=2}^{n-2} x_i, x_{n-1} + x_n \right) \right)$$

*Then,*

$$\frac{\mathbf{I}[f_n]}{\mathbf{I}_{GL}(2, n)} = 1 + \frac{1}{8n} + O\left(\frac{1}{n^2}\right).$$

<sup>4</sup>Note that the conjecture is usually stated using only the form given here for even  $n$ , with no mention of the parity of  $n$ . It is shown in Appendix B to this paper that this is an error, perhaps propagating from the originally ambiguous floor/ceiling notation in [2].

It is shocking, in our opinion, that the “influence maximizer is symmetric” principle which holds for linear threshold functions (as well as for  $d = n$ ) does not persist for  $d = 2$ . However, for most applications, it would suffice to know only that the Gotsman-Linial bound is of the *correct asymptotic order* as  $n \rightarrow \infty$ . It’s not difficult to show that, for  $d < \sqrt{n}$ ,  $\mathbf{I}_{GL}(n, d)$  is of order  $\Theta(d\sqrt{n})$ , and hence a natural relaxation of the original Gotsman-Linial conjecture is the following:

**Conjecture** (Weak Gotsman-Linial). *For any degree- $d$  threshold function  $f$  on  $n$  variables,*

$$\mathbf{I}[f] \leq O(d \cdot \sqrt{n}).$$

Using arguments from [1] and [5], the weaker conjecture would yield all of the usual desired properties for low-degree PTFs:  $O(d\sqrt{\delta})$  noise sensitivity,  $O(d^2/\epsilon^2)$ -concentration, and PAC learnability in time  $n^{O(d^2/\epsilon^2)}$ . To the best of our knowledge, the best known upper bound on influence for general PTFs is the following theorem of Kane [3]:

**Theorem** (Kane, 2013). *There exists a constant  $M$  such that, for any degree- $d$  PTF  $f$  on  $n$  variables,*

$$\mathbf{I}[f] \leq \sqrt{n} \cdot (2^d \cdot \log n)^{M \cdot d \log d}.$$

For constant  $d$ , Kane’s result comes close to the conjectured  $O(\sqrt{n})$ , differing only by a polylogarithmic factor. However, it quickly becomes trivial when  $d$  is allowed to grow with  $n$  (for  $d = \Omega(\sqrt{\log n / \log \log n})$  the bound is already  $\Omega(n)$ ), and even for constant  $d$ , it still leaves open the question of whether a polylogarithmic factor is necessary: can an asymmetric QTF, for instance, really have influence which is  $\omega(1)$ -times bigger than all symmetric QTFs?

In this paper, we answer this question for a few important classes of QTFs. We define the notion of *support* for a QTF

$$f(x_1, \dots, x_n) = \text{sgn} \left( \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \sum_{i=1}^n b_i x + c \right)$$

as the graph<sup>5</sup>  $G = (V, E)$  on  $n$  vertices such that  $E = \{(i, j) : a_{ij} \neq 0\}$ . Almost all pre-existing work on bounding the influence of threshold functions follows a similar paradigm: estimates are first derived in the Gaussian setting for “regular” PTFs (i.e. those for which no single coordinate has too much influence) using anti-concentration results for low-degree polynomials, then transferred to the Bernoulli setting by way of an invariance principle, and extended to “irregular” PTFs using a regularity lemma. (This approach was used in [1] to obtain the first non-trivial upper bound of  $O\left(n^{1-\frac{1}{4d+2}}\right)$  for the influence of degree- $d$  PTFs, and subsequently improved by Kane in [3].) Our approach uses none of this machinery — in fact, our methods are elementary and completely independent of any information about the coefficients other than which ones are non-zero.<sup>6</sup> Using a simple *induced subgraph covering lemma*, proved in Section 2, we are able to obtain the following:

<sup>5</sup>More precisely, since a boolean function may have many representations as a QTF, we say  $f$  is supported on  $G = (V, E)$  if it has a representation with  $\{(i, j) : a_{ij} \neq 0\} \subseteq E$ .

<sup>6</sup>Of course, one can always perturb coefficients slightly without changing a PTF’s truth table, so this is still in some sense an “open” condition.

**Theorem 1.2.** *If  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is a QTF supported on a graph  $G$  with fractional chromatic number<sup>7</sup>  $\chi_f(G)$ , then*

$$\mathbf{I}[f] \leq \sqrt{\chi_f(G)} \cdot \sqrt{n}.$$

From this theorem we can deduce that the Weak Gotsman-Linial conjecture holds for QTFs supported on  $k$ -partite graphs, planar graphs, graphs of bounded degree, to name a few. While such QTFs make up a “small fraction” of the space of all QTFs, they do encompass many of the functions that actually appear in practice. Restricted Boltzmann machines (RBMs), for example, commonly used in deep learning networks, are supported on bipartite graphs. The Ising model, too, is most often applied to graphs covered by this theorem: since particles are most often arranged in lattices, the graph of nearest-neighbor interactions is usually  $O(1)$  colorable (where the constant depends only on the lattice structure and not the number of vertices/particles).

Using a slight modification of an argument from [1], we are also able to prove a similar bound for graphs with few edges:

**Theorem 1.3.** *If  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is a QTF supported on a graph  $G = (V, E)$ , then*

$$\mathbf{I}[f] \leq \sqrt{n + \sqrt{2|E|n}}.$$

We also remark that our covering lemma gives a way of “throwing out” a number of problematic vertices. Indeed, using this observation, one can extend Theorems 1.2 and 1.3 to a slightly wider class of graphs: if  $f$  is supported on a graph  $G'$  which satisfies  $G' \setminus H = G$  for some induced subgraph  $H$ , then  $f$  obeys the same influence bound as if it were supported on  $G$ , with an extra additive term

$$\mathbf{I}[H] := \max_{g \in \text{QTFs supported on } H} \mathbf{I}[g].$$

In particular, combining this observation with Kane’s  $\sqrt{n} \cdot \text{poly log}(n)$  bound, we conclude that the weak Gotsman-Linial conjecture holds for QTFs supported on graphs which, after removing  $n/\Omega(\text{poly log}(n))$  vertices, have  $O(1)$  fractional chromatic number or  $O(n)$  edges. As an example, this applies to graphs with a small ( $O(n/\text{poly log}(n))$ ) vertex cover.

We prove our main theorems in Section 2. In Section 3, we discuss counterexamples to the Gotsman-Linial conjecture, and raise a few interesting questions regarding the effect of a QTF’s support on its influence.

## 2. INFLUENCE BOUNDS FOR QTFs SUPPORTED ON SPECIAL GRAPHS

### 2.1. Preliminaries.

**Definition 2.1.** *The influence of the  $i$ th coordinate on  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  is defined as*

$$\text{Inf}_i[f] := \Pr_{\mathbf{x} \sim \{-1, 1\}^n} [f(\mathbf{x}) \neq f(\mathbf{x}^{\oplus i})],$$

<sup>7</sup>This is a standard term from graph theory, and we give the precise definition in Section 2. For now, the unfamiliar reader may replace  $\chi_f(G)$  by the chromatic number  $\chi(G)$ , since  $\chi_f(G) \leq \chi(G)$ .

where  $x^{\oplus i} = (x_1, x_2, \dots, -x_i, x_{i+1}, \dots, x_n)$ . Moreover, the (total) influence is

$$\mathbf{I}[f] := \sum_{i=1}^n \text{Inf}_i[f].$$

**Definition 2.2.** We say that a QTF  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  of the form

$$f(x) = \text{sgn} \left( \sum_{i < j} a_{i,j} x_i x_j + \sum_{i=1}^n a_i x_i + a_0 \right)$$

is supported on the graph  $G = (V, E)$ , if  $\{(i, j) : a_{i,j} \neq 0\} \subset E$ . Moreover, let  $Q_G$  denote the set of QTFs supported on  $G$ .

**Definition 2.3.** Let  $G = (V, E)$  be a graph on  $n$  vertices. We define the maximal QTF influence of  $G$  to be

$$\mathbf{I}[G] := \max_{f \in Q_G} \mathbf{I}[f].$$

**Definition 2.4.** Let  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  be a real Boolean function. The  $i$ th discrete derivative operator maps  $f$  to the function  $D_i f : \{-1, 1\}^{n-1} \rightarrow \mathbb{R}$  defined by

$$D_i f(x) = \frac{f(x^{i \rightarrow 1}) - f(x^{i \rightarrow -1})}{2},$$

where  $x^{i \rightarrow a} = (x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$ .

For a fixed  $x$ , note that  $D_i f(x) = \pm 1$  whenever  $f(x) \neq f(x^{\oplus i})$ , so we obtain the useful fact:

**Fact 2.5.** For all boolean functions  $f$ ,

$$\text{Inf}_i[f] = \mathbb{E}_{x \sim \{-1, 1\}^n} [|D_i f(x)|]$$

The following lemma is a restatement of the fact that among LTFs, majority maximizes the total influence. For the sake of completion, this result is proved in Appendix A.

**Lemma 2.6.** For an edgeless graph  $G$  on  $n$  vertices, we have

$$\mathbf{I}[G] = \mathbf{I}[\text{MAJ}_n] = \sqrt{2/\pi} \sqrt{n} + O(1).$$

**Definition 2.7.** Let  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ , and  $J \subset [n]$ . Then  $f_{J|z} : \{-1, 1\}^{|J|} \rightarrow \mathbb{R}$  is the restriction of  $f$  to  $J$ , acquired by fixing the coordinates in  $[n] \setminus J$  to  $z$ .

The following simple lemma provides the basis for many of the results used in the proofs of our main theorems.

**Lemma 2.8.** Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be a Boolean function and  $J \subset [n]$ . Then, for  $i \in J$ ,

$$\text{Inf}_i[f] = \mathbb{E}_{z \in \{-1, 1\}^{|J^c|}} [\text{Inf}_i[f_{J|z}]].$$

*Proof.* We have that

$$\begin{aligned} \text{Inf}_i[f] &= \mathbb{E}_{x \in \{-1, 1\}^n} [D_i f^2] \\ &= \sum_{z \in \{-1, 1\}^{|J^c|}} \frac{1}{2^{|J^c|}} \cdot \mathbb{E}_{x \in \{-1, 1\}^{|J|}} [D_i f_{J|z}^2] \\ &= \mathbb{E}_{z \in \{-1, 1\}^{|J^c|}} [\text{Inf}_i[f_{J|z}]]. \end{aligned}$$

□

A boolean function's influence can also be defined in terms of its Fourier coefficients. A brief overview of Fourier analysis is included in Appendix A for completion.

## 2.2. Proof of Main Theorems.

**Lemma 2.9** (Covering Lemma). *Let  $G = (V, E)$  be a graph that admits a covering by induced graphs  $G_1, \dots, G_k$ . Then,*

$$\mathbf{I}[G] \leq \mathbf{I}[G_1] + \dots + \mathbf{I}[G_k].$$

*Proof.* Without loss of generality, assume that the graphs  $G_i$  are disjoint. Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be a QTF supported on  $G$ . By Lemma 2.8,

$$\sum_{j \in G_i} \text{Inf}_j[f] = \mathbb{E}_{z \sim \{-1, 1\}^{|G_i^c|}} [\mathbf{I}[f_{G_i|z}]].$$

Since  $f_{G_i|z}$  is supported on  $G_i$  for any  $z \in \{-1, 1\}^{|G_i^c|}$ ,

$$\mathbb{E}_z [\mathbf{I}[f_{G_i|z}]] \leq \mathbf{I}[G_i],$$

and so

$$\mathbf{I}[f] = \sum_{i=1}^k \sum_{j \in G_i} \text{Inf}_j[f] \leq \mathbf{I}[G_1] + \dots + \mathbf{I}[G_k].$$

□

**Corollary 2.10.** *For any  $f \in Q_G$  for  $n$ -vertex graphs  $G$  with chromatic number  $\chi(G)$ ,*

$$\mathbf{I}[f] \leq \sqrt{\chi(G)} \sqrt{n}$$

*in fact, we can also show*

$$\mathbf{I}[G] \leq \sqrt{2/\pi} \sqrt{\chi(G)} \sqrt{n} + O(\chi(G))$$

*Proof.* Let  $G_1, \dots, G_{\chi(G)}$  be the monochromatic sets formed when  $\chi(G)$ -coloring  $G$ . Each  $G_i$  is an independent set, and thus  $f$  is an LTF for all  $f \in G_i$ . Therefore,  $\mathbf{I}[G_i] = \mathbf{I}[\text{MAJ}_{|G_i|}]$ . Then, from the covering lemma, we find

$$\mathbf{I}[G] \leq \sum_{i=1}^{\chi(G)} \mathbf{I}[\text{MAJ}_{|G_i|}] \leq \sqrt{2/\pi} \sum_{i=1}^{\chi(G)} \sqrt{|G_i|} + O(\chi(G))$$

As  $\sum_i |G_i| = n$ , by Cauchy-Schwarz, we can show

$$\sum_{i=1}^{\chi(G)} \sqrt{|G_i|} \leq \sqrt{\chi(G)} \sqrt{n}.$$

This gives us

$$\mathbf{I}[G] \leq \sqrt{2/\pi} \sqrt{\chi(G)} \sqrt{n} + O(\chi(G))$$

proving the corollary. □

**Remark.** *This corollary, combined with the facts that planar graphs are 4-colorable and that a graph with maximum degree  $\Delta$  has  $\chi(G) \leq \Delta + 1$ , we can prove that  $\mathbf{I}[G] \leq O(\sqrt{n})$  for any planar graphs or graphs with bounded degree.*

We can improve the Covering Lemma by allowing a non-uniform probability distribution on the elements of the cover.

**Lemma 2.11** (Randomized Covering Lemma). *Let  $G = (V, E)$  be a graph with induced cover  $C = \{G_1, \dots, G_k\}$  and let  $P$  be a probability distribution on covering  $C$  such that  $\Pr_{G_i \sim P}[v \in G_i] \geq \epsilon, \forall v \in V$ . Then,*

$$\mathbf{I}[G] \leq \frac{1}{\epsilon} \cdot \mathbb{E}_{G_i \sim P}[\mathbf{I}[G_i]].$$

The proof follows the same approach as that of the Covering Lemma. This improved lemma allows us to change the chromatic number bound into one involving the fractional chromatic number of  $G$ .

**Definition 2.12.** *Let  $G = (V, E)$  be a graph and let  $\mathcal{I}$  denote the collection of independent sets in  $G$ . The fractional chromatic number  $\chi_f(G)$  is the optimal value of the linear program*

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{I}} x_S \\ \text{s.t.} \quad & \sum_{S \ni v} x_S \geq 1, \forall v \in V \end{aligned}$$

**Proof of Theorem 1.2.** If we scale the optimal solution of the LP by  $\chi_f(G)^{-1}$ , we obtain a probability distribution  $P$  on the independent sets of  $G$  such that  $\forall v \in V$ ,

$$p_v := \Pr_{S \sim P}[v \in S] \geq \chi_f(G)^{-1}.$$

By the Randomized Covering Lemma,

$$\mathbf{I}[f] \leq \chi_f(G) \cdot \mathbb{E}_{S \sim P}[\mathbf{I}[S]].$$

Since  $S$  is an independent set of  $G$ , if  $f \in Q_S$ , then  $f_{S|z}$  is an LTF for any  $z \in \{-1, 1\}^{|S^c|}$  and so

$$\mathbf{I}[f] \leq \chi_f(G) \cdot \mathbb{E}_{S \sim P}[\sqrt{|S|}] \leq \chi_f(G) \cdot \sqrt{\mathbb{E}_{S \sim P}[|S|]}.$$

Moreover, let  $q_v$  be the value in  $[0, 1]$  such that  $p_v \cdot (1 - q_v) = \chi_f(G)^{-1}$ . We construct a new probability distribution  $\tilde{P}$  by, for every independent set  $S$ , removing each element in  $S$  with probability  $q_v$ . Using distribution  $\tilde{P}$ ,

$$\Pr_{S \sim \tilde{P}}[v \in S] = p_v \cdot (1 - q_v) = \chi_f(G)^{-1},$$

and so

$$\mathbb{E}_{S \sim \tilde{P}}[|S|] = \sum_{v \in V} \Pr_{S \sim \tilde{P}}[v \in S] = n \cdot \chi_f(G)^{-1},$$

as desired.  $\square$

To prove Theorem 1.3, we'll borrow some simple ideas from [1], which were used therein to prove an influence bound of  $O(n^{1-\frac{1}{2^d}})$  for degree- $d$  PTFs.



**Proof of Theorem 1.3.** Let  $f(x) = \text{sgn}(p(x))$ , where the quadratic polynomial  $p(x) \neq 0$  for all  $x \in \{-1, 1\}^n$ . For some fixed  $x$ , if  $D_i f(x) > 0$ , then  $p(x^{i \rightarrow 1}) > p(x^{i \rightarrow -1})$  and so  $D_i f(x) = \text{sgn}(D_i p(x))$  whenever the left side is nonzero. Hence,

$$\text{Inf}_i[f] = \mathbb{E}_{x \sim \{-1, 1\}^n} [|D_i f|] = \mathbb{E}_x [\text{sgn}(D_i p) \cdot D_i f].$$

Moreover, using the fact that  $f = x_i \cdot D_i f + E_i$ ,

$$\begin{aligned} \mathbb{E}_x [x_i \cdot \text{sgn}(D_i p(x)) \cdot f] &= \mathbb{E}_x [\text{sgn}(D_i p(x)) \cdot D_i f] + \mathbb{E}_x [x_i \cdot \text{sgn}(D_i p(x)) \cdot E_i] \\ &= \mathbb{E}_x [\text{sgn}(D_i p(x)) \cdot D_i f] + \mathbb{E}_x [x_i] \mathbb{E}_x [\text{sgn}(D_i p(x)) \cdot E_i], \end{aligned}$$

where the right term is zero as  $E_i$  and  $\text{sgn}(D_i p(x))$  are independent from  $x_i$ . Thus,

$$\text{Inf}_i[f] = \mathbb{E}_x [x_i \cdot \text{sgn}(D_i p(x)) \cdot f].$$

Using Cauchy-Schwarz,

$$\begin{aligned} \mathbf{I}[f] &= \sum_{i \in [n]} \mathbb{E}_x [x_i \cdot \text{sgn}(D_i p) \cdot f] \leq \mathbb{E}_x \left[ \left| \sum_{i \in [n]} x_i \cdot \text{sgn}(D_i p) \right| \right] \\ &\leq \sqrt{n + \sum_{i \neq j} \mathbb{E}_x [x_i x_j \cdot \text{sgn}(D_i p) \cdot \text{sgn}(D_j p)]}. \end{aligned}$$

For ease of notation, let  $f_i(x) := \text{sgn}(D_i p(x))$ . Since  $f_i$  does not depend on  $x_i$ ,

$$\begin{aligned} \mathbb{E}_{x \sim \{-1, 1\}^n} [x_i x_j f_i f_j] &= \mathbb{E}_x \left[ x_j f_i \left( \frac{f_j(x^{i \rightarrow 1}) - f_j(x^{i \rightarrow -1})}{2} \right) \right] \\ &= \mathbb{E}_x [x_j f_i \cdot D_i f_j]. \end{aligned}$$

Similarly, since  $D_j f_i$  does not depend on  $x_j$ , we can repeat the same process with  $f_i$  to obtain

$$\mathbb{E}_x [x_i x_j f_i f_j] = \mathbb{E}_x [D_j f_i \cdot D_i f_j].$$

Moreover, using Cauchy-Schwarz,

$$\begin{aligned} \mathbb{E}_x [D_j f_i \cdot D_i f_j] &\leq \sqrt{\mathbb{E}_x [D_j f_i^2] \mathbb{E}_x [D_i f_j^2]} = \sqrt{\text{Inf}_j[f_i] \cdot \text{Inf}_i[f_j]} \\ &\leq \frac{\text{Inf}_j[f_i] + \text{Inf}_i[f_j]}{2}, \end{aligned}$$

where the last step follows by the AM-GM inequality. Therefore,

$$\mathbf{I}[f] \leq \sqrt{n + \sum_{i \in [n]} \mathbf{I}[\text{sgn}(D_i p)]}.$$

Since  $p(x) = \sum_{i < j} a_{i,j} x_i x_j + \sum_i b_i x_i + c$ , we have that, for a fixed  $i$ ,

$$\text{sgn}(D_i p(x)) = \text{sgn} \left( \sum_{j | (i,j) \in E} a_{i,j} x_j + b_i \right),$$

so  $\text{sgn}(D_i p(x))$  is an LTF on the variables adjacent to  $x_i$  in  $G$  and therefore

$$\mathbf{I}[\text{sgn}(D_i p(x))] \leq \sqrt{\deg(i)}.$$

Hence, by Cauchy-Schwarz,

$$\mathbf{I}[f] \leq \sqrt{n + \sum_{i \in [n]} \deg(i)} \leq \sqrt{n + \sqrt{2|E|n}}.$$

□

### 3. COUNTEREXAMPLES TO THE GOTSMAN-LINIAL CONJECTURE

**3.1. The first counterexample.** For  $n \leq 4$ , the number of boolean functions on  $n$  variables is  $2^{2^n} \leq 2^{16}$ , so an exhaustive search is tractable. In those cases, our exhaustive search verified that the Gotsman-Linial conjecture is true for  $d = 2, n \leq 4$ .

For  $n = 5$ , naively searching through every truth table becomes time consuming, so we sped up the process in two ways. First, owing to a mixture of intuition and optimism, we decided to search only through truth tables of functions symmetric in the last two coordinates – this reduces the size of the search space from  $2^{32}$  to  $2^{24}$ . Second, we screened for high influence first ( $> \mathbf{I}_{GL}(5, 2) = 3.125$ ) before testing for QTF-ness. Since an average function has influence

$$\mathbb{E}_f[\mathbf{I}[f]] = 2.5,$$

standard concentration bounds imply that most functions will fail to pass this screening.

To test whether a boolean function  $f$  can be represented as a QTF, we simply test the feasibility of the following LP:

$$\begin{aligned} &\text{find } q \in \mathbb{R}^{16} \\ &\text{such that } f_x(Tq)_x \geq 1, \forall x \in \{-1, 1\}^5 \end{aligned}$$

where the linear map  $T : \mathbb{R}^{16} \rightarrow \mathbb{R}^{32}$  is the evaluation map on the space of multilinear quadratic functions on  $\{-1, 1\}^5$ . If this LP is feasible, then any solution yields the coefficients of a quadratic polynomial  $q$  such that at all of the inputs  $x \in \{-1, 1\}^5$ ,  $\text{sgn}(q(x)) = f(x)$ . This method gave us the counterexample

$$\begin{aligned} f(x) = \text{sgn}(-7x_1x_2 - 7x_1x_3 + x_1x_4 + x_1x_5 - 7x_2x_3 + 2x_2x_4 + 2x_2x_5 + 2x_3x_4 \\ + 2x_3x_5 + 3x_4x_5 + x_1 + 2x_2 + 2x_3 + 3x_4 + 3x_5 - 4) \end{aligned}$$

with influence  $3.1875 > \mathbf{I}_{GL}(5, 2)$ .

**3.2. Generalizing to all odd  $n \geq 5$ .** Observe that the counterexample found in the previous section can be written as

$$\text{sgn}(p(x_1, x_2 + x_3, x_4 + x_5))$$

for

$$p(x, y, z) := 2x(1 - 7y + z) + 4y - 7y^2 + 4yz + 6z + 3z^2.$$

On 7 variables, the function

$$\text{sgn}(p(x_1, x_2 + x_3 + x_4 + x_5, x_6 + x_7))$$

has influence  $249/64 > \mathbf{I}_{GL}(7, 2) = 245/64$ . In fact, one can prove that the functions obtained in this way:

$$f_n(x) := \text{sgn} \left( p \left( x_1, \sum_{i=2}^{n-2} x_i, x_{n-1} + x_n \right) \right)$$

always have higher influence than  $\mathbf{I}_{GL}(n, 2)$  for odd  $n \geq 5$ .

**Proof of Theorem 1.1.** Let  $n \geq 9$  be odd and set  $m = n - 3$ . By considering the 6 possible functions obtained from  $f_n$  by restricting the values of  $x_1$  and  $x_{n-1} + x_n$ , we can find the influence of the remaining  $n - 3$  coordinates directly by counting, and then appeal to Lemma 2.8 to conclude that

$$\begin{aligned} \sum_{i=2}^{n-2} \text{Inf}_i[f_n] &= 2^{-m} \left[ \binom{m}{m/2} \cdot \frac{11m}{16} + \binom{m}{m/2-1} \left( \frac{15m}{16} + \frac{7}{8} \right) \right. \\ &\quad \left. + \binom{m}{m/2-2} \left( \frac{5m}{16} + \frac{3}{4} \right) + \binom{m}{m/2-3} \left( \frac{m+6}{15} \right) \right] \end{aligned}$$

The influences of  $x_1$ ,  $x_{n-1}$  and  $x_n$  can also be obtained directly:

$$\text{Inf}_1[f_n] = 2^{-m} \left[ \binom{m}{m/2} \cdot \frac{3}{4} + \binom{m}{m/2-1} \cdot \frac{5}{4} + \binom{m}{m/2-2} \cdot \frac{1}{4} \right]$$

$$\text{Inf}_{n-1}[f_n] = \text{Inf}_n[f_n] = 2^{-m} \left[ \binom{m}{m/2} \cdot \frac{3}{4} + \binom{m}{m/2-1} \cdot \frac{3}{4} + \binom{m}{m/2-2} \cdot \frac{1}{4} \right]$$

Adding these expressions together yields

$$\begin{aligned} \mathbf{I}[f_n] &= 2^{-m} \left[ \binom{m}{m/2} \left( \frac{11m}{16} + \frac{9}{4} \right) + \binom{m}{m/2-1} \left( \frac{15m}{16} + \frac{29}{8} \right) \right. \\ &\quad \left. + \binom{m}{m/2-2} \left( \frac{5m}{16} + \frac{3}{2} \right) + \binom{m}{m/2-3} \left( \frac{m}{16} + \frac{3}{8} \right) \right] \end{aligned}$$

while by Lemma B.1,

$$\mathbf{I}_{GL}(2, n) = 2^{-m} \binom{m+3}{m/2+1} \left( \frac{m+3}{4} \right).$$

Computing the ratio, one finds

$$\frac{\mathbf{I}[f_n]}{\mathbf{I}_{GL}(2, n)} = 1 + \frac{7}{32n} - \frac{3}{32(n-2)} + \frac{3}{16n^2} = 1 + \frac{1}{8n} + O\left(\frac{1}{n^2}\right).$$

□

Thus,  $\{f_n\}_{n \geq 5, \text{ odd}}$  provides an infinite family of counterexamples to the Gotsman-Linial conjecture for  $d = 2$ , beating the supposed upper bound by an additive  $\Omega(1/\sqrt{n})$ .

**3.3. Discussion.** We note that all of the counterexamples discovered have a complete support. This raises a natural question: can a graph with incomplete support still have influence greater than  $\mathbf{I}_{GL}$ ?

Let  $H$  and  $G$  be graphs, with  $H \subseteq G$ . Then, as we can change each coefficient of a PTF by a small amount without altering its truth table, we of course have  $\mathbf{I}[h] \leq \mathbf{I}[g]$ . However, it is not obvious when equality should hold. Suppose  $G_0 \subseteq G_1 \subseteq \dots \subseteq G_m = G$  is a chain of subgraphs of  $G$ , starting from the empty graph  $G_0$  and adding one edge at a time. If  $G = K_n$ , then since  $\mathbf{I}[G] - \mathbf{I}[G_0] = \tilde{\Theta}(\sqrt{n})$ , the average gap  $\mathbf{I}[G_{i+1}] - \mathbf{I}[G_i]$  is  $\tilde{\Theta}(n^{-3/2})$ , where the  $\tilde{\Theta}$  hides the possible polylogarithmic factors from Kane's upper bound. For a random sequence of edges, are the gaps likely to be evenly distributed, or does removing some small, special subset of edges account for most of the drop in influence? More generally, for non-isomorphic graphs  $G$  and  $G'$ , what can be said about  $\mathbf{I}[G]$  versus  $\mathbf{I}[G']$ ?

In order to gain some intuition, we decided to explicitly find the maximum influence attained by a function with each support. As there are only  $2^{2^4}$  boolean functions on four variables, we used the method outlined in the previous subsection to find these values. The results are shown in Table 1.

We can see that the maximal influence of 3.0 was only attained by a function with full support. As this also appeared to be the case in our search for a counterexample in the case of odd  $n$  for  $d = 2$ , it seems reasonable to conjecture that for each  $n$  and  $d$ , the influence maximizer will have full support. As a stronger conjecture, it seems feasible that any function violating Gotsman-Linial's influence bound must have full support.








$G$	$\mathbf{I}[G]$	$G$	$\mathbf{I}[G]$	$G$	$\mathbf{I}[G]$	$G$	$\mathbf{I}[G]$
	2		2		2.5		3
	2		2.5		2.5		

TABLE 1. Maximum influences of QTFs supported by graphs on 4 vertices

#### 4. ACKNOWLEDGEMENTS

We would like to thank our mentor Jake Wellens for proposing the project, and for showing us the wonders and responsibilities of producing research. We would also like to thank the organizers of SPUR+ program at MIT for making this research opportunity possible. In particular, we would like to thank Cris Negron for his support and advice.

#### REFERENCES

- [1] I. Diakonikolas, P. Harsha, A. Klivans, R. Meka, P Raghavendra, R. Servedio, & L. Y. Tan, *Bounding the average sensitivity and noise sensitivity of polynomial threshold functions*. STOC'10 - Proceedings of the 2010 ACM International Symposium on Theory of Computing (2010) 533-542.
- [2] C. Gotsman, N. Linial, *Spectral properties of threshold functions*, *Combinatorica* 14 (1994), 35 - 50.

- [3] D. M. Kane, *The correct exponent for the Gotsman-Linial conjecture*, Computational Complexity 23 (2014), 151-175.
- [4] M. Minsky and S. Papert, *Perceptrons*. MIT Press, Cambridge, MA., (1969).
- [5] R. O'Donnell, *Analysis of boolean functions*. <http://www.cs.cmu.edu/odonnell/boolean-analysis>, (2013).
- [6] P. O'Neil, *Hyperplane cuts of an  $n$ -cube*, Discrete Maths. 1 (1971), 193 -195.
- [7] Peres, Yuval. *Noise stability of weighted majority.*, arXiv preprint math/0412377 (2004).

## APPENDIX A. FOURIER ANALYSIS

**Proposition A.1** (Fourier Expansion). *Let  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  be a real Boolean function. Then, there exists a unique representation as a multilinear polynomial*

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \cdot x_S,$$

where  $x_S = \prod_{i \in S} x_i$ .

*Proof.* Note that we can interpret the set of functions as a vector space in  $\mathbb{R}^{2^n}$  as we can add functions point-wise and multiply by a scalar. By defining the inner-product,

$$\langle f, g \rangle = \mathbb{E}_{x \in \{-1, 1\}^n} [f(x) \cdot g(x)],$$

the characters  $x_S$  form an orthonormal basis. □

This representation is particularly useful as it allows us to “read off” properties of the function. In particular, the Fourier expansion allows us to show that the discrete derivative operator behaves like in the continuous case.

**Proposition A.2.** *Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be a Boolean function with Fourier expansion  $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \cdot x_S$ . Then,*

$$D_i f(x) = \sum_{S \ni i} \hat{f}(S) \cdot x_{S \setminus \{i\}}.$$

*Proof.* Note that the derivative operator is linear, so it suffices to look at the characters  $x_S$ . In this case,

$$D_i x_S = \begin{cases} x_{S \setminus \{i\}}, & \text{if } i \in S \\ 0, & \text{otherwise} \end{cases}$$

so the statement follows by Proposition A.1. □

This result allows us to obtain an expression for the influence of a Boolean function in terms of Fourier coefficients.

**Proposition A.3.** *Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be a Boolean function with Fourier expansion  $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \cdot x_S$ . Then,*

$$\text{Inf}_i[f] = \sum_{S \ni i} \hat{f}(S)^2.$$

*Proof.* By Definition 2.4, note that

$$D_i f(x) = \begin{cases} \pm 1, & \text{if } x_i \text{ is an influential coordinate} \\ 0, & \text{otherwise} \end{cases}$$

so  $(D_i f(x))^2$  is an indicator for the  $i$ th influence. Hence,

$$\text{Inf}_i[f] = \mathbb{E}_x[(D_i f(x))^2] = \sum_{S \ni i} \hat{f}(S)^2.$$

□

Similar to the discrete derivative operator, we can define the  $i$ th expectation operator where we take the average of a function at fixed values of  $x_i$ .

**Definition A.4** (Expectation Operator). *Let  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  be a real Boolean function. Then, the  $i$ th expectation operator maps  $f$  to the function  $E_i f : \{-1, 1\}^{n-1} \rightarrow \{-1, 1\}$  defined by*

$$E_i f(x) = \mathbb{E}_{x_i \sim \{-1, 1\}} [f(x_1, \dots, x_i, \dots, x_n)] = \frac{f(x^{i \rightarrow 1}) + f(x^{i \rightarrow -1})}{2}.$$

**Proposition A.5.** *Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be a Boolean function with Fourier expansion  $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \cdot x_S$ . Then,*

$$E_i f(x) = \sum_{S | i \notin S} \hat{f}(S) \cdot x_S.$$

The proof follows the same structure as Proposition 2.4. This result allows us to rewrite a function on  $n$  variables in terms of those in at most  $n - 1$ .

**Proposition A.6.** *Let  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  be a real Boolean function. Then,*

$$f(x) = x_i \cdot D_i f(x) + E_i f(x).$$

**Theorem A.7** (Total influence of LTF's). *Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be an LTF. Then,  $\mathbf{I}[f] \leq \sqrt{n}$ .*

*Proof.* (Our proof follows [5].) Let  $f(x) = \text{sgn}(a_0 + \sum_{i \in [n]} a_i x_i)$  with real coefficients. Note that for a fixed  $i$ ,  $f(x^{i \rightarrow 1}) \geq f(x^{i \rightarrow -1})$  when  $a_i > 0$  and  $f(x^{i \rightarrow 1}) \leq f(x^{i \rightarrow -1})$  when  $a_i < 0$ . Hence, the discrete derivative  $\pm D_i f(x)$  will be a 0-1 indicator for the  $i$ th influence, where the sign depends on which case applies. Hence,

$$\text{Inf}_i[f] = \mathbb{E}_{x \sim \{-1, 1\}^n} [\pm D_i f(x)] = |\hat{f}(i)|$$

so by Cauchy-Schwarz,

$$\mathbf{I}[f] = \sum_{i \in [n]} |\hat{f}(i)| \leq \sqrt{n \cdot \sum_{i \in [n]} \hat{f}(i)^2} = \sqrt{n},$$

where we use the fact that

$$1 = \mathbb{E}_{x \sim \{-1, 1\}^n} [f(x)^2] = \sum_{i \in [n]} \hat{f}(i)^2.$$

□

**Lemma A.8.** *For the edgeless graph  $G = (V, \emptyset)$  on  $n$  vertices,*

$$\mathbf{I}[G] = \mathbf{I}[\text{MAJ}_n] = (\sqrt{2/\pi} + o_n(1))\sqrt{n}$$

*Proof.* The set  $Q_G$  is exactly the set of LTFs on  $n$  variables. If  $f$  is an LTF on  $n$  variables, then following the approach of the previous lemma,

$$\mathbf{I}[f] = \sum_{i \in [n]} |\hat{f}(i)| = \mathbb{E}_x[(x_1 + \cdots + x_n)f(x)] \leq \mathbb{E}_x[|x_1 + \cdots + x_n|]$$

with equality if and only if  $f(x) = \text{sgn}(x_1 + \cdots + x_n) = \text{MAJ}_n(x)$ . Thus,  $\mathbf{I}[G] = \mathbf{I}[\text{MAJ}_n]$ . To get the asymptotic formula, we apply the central limit theorem to conclude

$$\frac{x_1 + \cdots + x_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

and hence

$$\frac{1}{\sqrt{n}} \mathbb{E}[|x_1 + \cdots + x_n|] \rightarrow \mathbb{E}[|\mathcal{N}(0, 1)|] = \sqrt{2/\pi}.$$

□

## APPENDIX B. $\mathbf{I}_{GL}$ COMPUTATION

We show the derivation in the case of  $d = 2$  for the sake of simplicity, but the argument generalizes to all  $d$ .

**Lemma B.1.** *Let  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$  be a symmetric QTF. Then,*

$$\mathbf{I}[f] \leq \begin{cases} n \cdot 2^{-n+1} \binom{n}{\frac{n-1}{2}}, & n \text{ is odd} \\ n \cdot 2^{-n} \left[ 2 \binom{n}{\frac{n-2}{2}} + \binom{n}{\frac{n}{2}} \right], & n \text{ is even} \end{cases}.$$

*Proof.* Let  $f$  be a symmetric QTF with the form  $f(x) = \text{sgn}(p(t))$ , where  $p$  is a polynomial on the sum of variables  $t := \sum_{i \in [n]} x_i$ . Ideally, we would like the roots of  $p$  to be between the possible values of  $t$  so that changing the value of a coordinate corresponds to a change in the sign of  $p$ . Moreover, since the number of inputs  $x$  such that  $t(x) = k$  is  $\binom{n}{\frac{n-k}{2}}$  (a decreasing function of  $k$ ), we would like the roots to be close to 0. Hence,

- (1) When  $n$  is odd, we choose the roots of  $p$  to be at  $t = 0, 2$ . Thus, by the well-known binomial identity  $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ ,

$$\text{Inf}_1[f] = 2^{-n} \left[ \binom{n}{\frac{n-1}{2}} + \binom{n-1}{\frac{n-3}{2}} + \binom{n-1}{\frac{n-1}{2}} \right] = 2^{-n+1} \binom{n}{\frac{n-1}{2}}$$

where  $x_1$  is restricted to be -1 at the boundary value  $t = -1$  and, similarly, 1 at  $t = 3$ .

- (2) When  $n$  is even, we choose the roots of  $p$  to be at  $t = -1, 1$ . Similarly,

$$\text{Inf}_1[f] = 2^{-n} \left[ \binom{n}{\frac{n}{2}} + 2 \binom{n-1}{\frac{n-2}{2}} \right],$$

where, as before,  $x_1 = -1$  when  $t = -2$  and  $x_1 = 1$  when  $t = 2$ . Since  $f$  is symmetric,

$$\mathbf{I}[f] = n \cdot \text{Inf}_1[f].$$

□

Note that, for  $d = 2$  and  $n = 5$ , we obtain the value 3.125. We take care to calculate this value as there is ambiguity in the original conjecture that results in a higher value of 3.4375, which was verified through exhaustive search to be unattainable by symmetric QTFs.