One-sided linearity testing

Yuval Filmus et al.

September 28, 2019

Based on joint work with Noam Lifshitz (HUJI), Dor Minzer (IAS) and Elchanan Mossel (MIT).

1 Judgement aggregation

Social choice theory abounds in paradoxes. Arrow's theorem and the Gibbard–Satterthwaite theorem are well-known. Here is another example. Suppose that I survey a group of people about their opinion of various types of chocolates:

Count	Like white chocolate?	Like dark chocolate?	Like both?
60	Yes	No	No
50	No	Yes	No
40	Yes	Yes	Yes
Total Yes	100	90	40
Total No	50	60	110

Although a majority of people like white chocolate and a majority of people like dark chocolate, only a minority of people like both. This shows that the majority function is not an admissible judgement aggregation function in this setting.

Which functions $f\colon \{0,1\}^n \to \{0,1\}$ are admissible? Such functions have to satisfy

$$f(x) \land f(y) = f(x \land y)$$

for all $x, y \in \{0,1\}^n$. In contrast to the impossibility results mentioned above, in this case dictatorships are not the only examples. But the additional examples are not too helpful: a non-constant function f satisfies the equation above ("has AND as a polymorphism") iff it is a conjunction of some of the coordinates. The same situation holds if we allow different aggregation functions for different columns: if $f(x) \land g(y) = h(x \land y)$ for all $x, y \in \{0,1\}^n$ then necessarily $f = g = h$.

Ilan Nehama asked whether we can enlarge the class of admissible functions by allowing aggregation functions which are only *approximately* admissible, in the sense that $f(x) \land f(y) = f(x \land y)$ only holds for most inputs, say 99% with respect to the uniform distribution over x, y (while this is not the most realistic input distribution, it is the one most commonly considered in social choice theory).

Gil Kalai famously proved (using Fourier analysis) an approximate version of Arrow's theorem, which states that if an aggregation function approximately satisfies Arrow's axioms, then it is an approximate dictatorship. Falik and Friedgut extended this to the Gibbard–Satterthwaite theorem. Does the same hold in our case?

Ilan Nehama gave an affirmative answer to this question. He showed that if $f(x) \wedge f(y) = f(x \wedge y)$ holds with probability $1 - \epsilon$, then f is δ-close to an AND, where δ is polynomial in $n, 1/\epsilon$. (He also showed a similar result for the three-function version.) His result is a bit unsatisfying, since in the results of Kalai and Falik–Friedgut, the error δ is independent of n. Our main result rectifies this deficiency, albeit at the price of getting a worse dependence on ϵ.

2 Linearity testing

In the chocolate example above, the answer to the third question was the conjunction of the first two questions. If instead the third question was "do you like exactly one of the two types?", then instead of conjunction we would get exclusive or, and so an admissible aggregation function would have to satisfy

$$f(x) \oplus f(y) = f(x \oplus y).$$

Once again, dictatorships are not the only examples: every XOR satisfies this equation, and these are the only solutions. Moreover, Blum, Luby and Rubinfeld famously showed that if f passes this "test" with high probability, then it must be close to an XOR. Their proof used the technique of self-correction. Later Bellare, Coppersmith, Håstad, Kiwi and Sudan gave a completely different proof, using Fourier analysis.

The self-correction proof is based on the fact that XORs satisfy $f(x) = f(y) \oplus f(x \oplus y)$. If we choose y at random, then both y and $x \oplus y$ are individually random, and this allows us to self-correct an approximate XOR by taking its value at x to be the majority vote of $f(y) \oplus f(x \oplus y)$. This kind of technique seems not available when replacing XORs with ANDs.

The Fourier-based proof uses the happy coincidence that the XORs are essentially the same as the Fourier characters, and so XORs (or approximate XORs) have a particularly simple Fourier expansion. In contrast, ANDs are not orthogonal, and so it is not clear how to extend this kind of proof from XORs to ANDs.

3 One-sided noise formulation

Suppose that a function $f\colon \{0,1\}^n \to \{0,1\}$ satisfies $f(x \wedge y) = f(x) \wedge f(y) = f(x)f(y)$. If we take expectation over y, then we get

$$\mathop{\mathbb{E}}_{y \sim \mu_{1/2}}[f(x \wedge y)] = \mathop{\mathbb{E}}_{\mu_{1/2}}[f] \cdot f(x).$$

This suggests defining an operator $T_\downarrow$ which maps the function f to the left-hand side of this equation:

$$(T_\downarrow f)(x) = \mathop{\mathbb{E}}_{y\sim\mu_{1/2}}[f(x \wedge y)].$$

The equation above then reads

$$T_\downarrow f = \lambda f, \text{ where } \lambda = \mathop{\mathbb{E}}_{\mu_{1/2}}[f].$$

That is, if f is an admissible aggregation function, then it is an eigenvector of $T_\downarrow$.

The operator $T_\downarrow$ is a one-sided analog of the more familiar two-sided noise operator common in Fourier analysis. Whereas the two-sided noise operator flips each bit with some probability, the one-sided noise operator only changes coordinates from 1 to 0 (in this case, with probability 1/2).

It is not hard to check that $T_\downarrow$ has $n + 1$ eigenspaces, corresponding to the eigenvalues $1, \ldots, 2^{-n}$, the kth eigenspace spanned by the k-ANDs. Furthermore, any Boolean eigenvector must be an AND (rather than just a linear combination of ANDs of the same width). This suggests a possible avenue for understanding approximate solutions to $f(x \wedge y) = f(x) \wedge f(y)$: understand approximate eigenvectors of $T_\downarrow$.

Understanding the approximate eigenvectors of the usual two-sided noise operator T_ρ is relatively easy, since the eigenspaces are orthogonal: starting with

$$T_\rho f = \sum_d \rho^d f^{=d},$$

we immediately obtain that if $T_\rho f \approx \lambda f$ then

$$\sum_d |\lambda - \rho^d| \|f^{=d}\|^2 \approx 0,$$

and so $\lambda \approx \rho^d$ for some d and $f \approx f^{=d}$. It follows that f must be close to a homogeneous degree d Boolean function (not necessarily an XOR!). The same kind of argument unfortunately fails for the one-sided noise operator, since its eigenspaces are not orthogonal.

4 Phantom solutions

A worse problem is that stability actually fails: there are approximate eigenvectors of $T_\downarrow$ that are *not* close to eigenvectors! Here is the simplest example:

$$f(x) = \begin{cases} x_1 \vee x_2 & \text{if } |x| \geq n/3, \\ x_1 \oplus x_2 & \text{if } |x| < n/3. \end{cases}$$

If we sample $x \sim \mu_{1/2}$, then it is highly likely to be in the first case, so almost all x satisfy $f(x) = x_1 \vee x_2$. In contrast, if we sample $x, y \sim \mu_{1/2}$, then $x \wedge y$ is highly likely to be in the second case, and so almost all x satisfy

$$T_\downarrow f(x) \approx \mathop{\mathbb{E}}_{y_1,y_2}[(x_1 \wedge y_1) \oplus (x_2 \wedge y_2)].$$

If $x_1 = x_2 = 0$ then $T_\downarrow f(x) \approx 0$. In all other cases, the probability that $(x_1 \wedge y_1) \oplus (x_2 \wedge y_2) = 1$ is $1/2$, and so $T_\downarrow f(x) \approx 1/2$. In total,

$$T_\downarrow f \approx \frac{1}{2} f.$$

Here is an even more striking example. Let $\lambda \in (0,1)$ be arbitrary. We construct a function f by taking $f(x) = 1$ whenever $|x| \geq n/3$, and letting $f(x) = 1$ with probability λ for all other x. With high probability, the resulting function will satisfy $T_\downarrow f \approx \lambda f$.

In both examples, the approximate eigenvalue is quite far from the expectation: in the first example $\mathbb{E}[f] \approx 3/4$ whereas $\lambda \approx 1/2$, and in the second one $\mathbb{E}[f] \approx 1$ whereas λ could be arbitrary. Therefore, while these are "phantom" eigenvectors, they do not correspond to solutions of the original problem.

A second feature of these examples is that the function f is defined one way for inputs with "large" weight and another way for inputs with "small" weight. In fact, if we separate the two parts, then the approximate equation becomes an exact equation:

$$T_\downarrow(x_1 \oplus x_2) = \frac{1}{2}(x_1 \vee x_2).$$

Similarly, in the second example we have

$$T_\downarrow \lambda = \lambda \cdot 1,$$

although the function λ is no longer Boolean. All of this suggests considering the equation

$$T_\downarrow f = \lambda g,$$

where $f\colon \{0,1\}^n \to [0,1]$, $g\colon \{0,1\}^n \to \{0,1\}$, and $\lambda \in (0,1)$.

As further motivation, let us note that any such solution naturally corresponds to an approximate solution $T_\downarrow h \approx \lambda h$, given by

$$h(x) = \begin{cases} g(x) & \text{if } |x| \geq n/3, \\ 1 \text{ w.p. } f(x) & \text{if } |x| < n/3, \end{cases}$$

in the sense that with high probability, $T_\downarrow h$ is extremely close to λh.

5 Solving the fundamental equation

As a first step towards our eventual goal, we consider the equation

$$T_\downarrow f = \lambda g,$$

where $f\colon \{0,1\}^n \to [0,1]$ and $g\colon \{0,1\}^n \to \{0,1\}$. It turns out that the non-zero solutions are

$$g = \bigwedge_{i=1}^{m} \bigvee_{j \in A_i} x_j, \quad f = 2^m \lambda \bigwedge_{i=1}^{m} \bigoplus_{j \in A_i} x_j,$$

where $m \leq \log_2(1/\lambda)$ and $A_1, \ldots, A_m$ are disjoint.

It is not too hard to check that indeed $T_\downarrow f = \lambda g$. The proof of the other direction is more complicated. We will explain the first step in some detail, and then outline the rest of the argument.

Let us consider a pair of inputs $z \leq x$ (this means that $z_i \leq x_i$ for all i). If $g(x) = 0$ then $T_\downarrow f(x) = \lambda g(x) = 0$. Since $T_\downarrow f(x)$ is the average of $f(y)$ for all $y \leq x$ and f is non-negative, necessarily $f(y) = 0$ for all $y \leq x$. In particular, $f(y) = 0$ for all $y \leq z$, and so $g(z) = 0$ as well. In other words, we have shown that g is monotone.

The second step is characterizing the possible options for g, using only the non-negativity of f. An elementary (but not immediate) argument shows that g must be an "AND-OR", that is, a function which is the conjunction of monotone disjunctions on disjoint sets of variables. Since $T_\downarrow$ is injective, we can infer that f is the corresponding "AND-XOR", up to multiplying by $2^m\lambda$ (where m is the number of ORs). Since $f \leq 1$, we can bound m by $\log_2(1/\lambda)$.

More on the second step Since g is monotone, we can consider its minterms. A simple calculation shows that if x is a minterm of g, then $f(x) = 2^{|x|}\lambda$, where $|x|$ is the Hamming weight of x.

The crucial observation, driving the entire proof, is that if $g(x) = 1$ and z is "disjoint" from x (meaning that $x_i \wedge z_i = 0$ for all i) then

$$\sum_{y \leq x} f(y \vee z) = 2^{|x|}\lambda,$$

where $y \vee z$ is the bitwise OR of y, z; the proof uses Möbius inversion. This immediately implies that all minterms of g have the same size m.

A simple induction shows that $f(x) \in \{0, 2^m\lambda\}$ for all x, allowing us to rephrase the crucial observation: if $g(x) = 1$ and z is disjoint from x, then there is exactly one $y \leq x$ such that $f(y \vee z) \neq 0$.

At this point we can unravel the structure of g. For each minterm x of g, we define a coloring $\chi_x \colon [n] \to \{0, \ldots, m\}$ as follows. Let $i_1, \ldots, i_m$ be the indices of the ones of x. We define $\chi_x(i_j) = j$. If $i \neq i_1, \ldots, i_m$, then the crucial observation implies that $f(y \vee 1_i) \neq 0$ for a single $y \leq x$, which can either be x itself, in which case we set $\chi_x(i) = 0$, or $x|_{i_j=0}$, in which case we set $\chi_x(i) = j$.

It turns out that (up to permutation) the functions χ_x coincide for all minterms, so we let $\chi = \chi_x$ for an arbitrary minterm x. A short argument shows that $g = \bigwedge_{j=1}^m \bigvee_{i \in \chi^{-1}(j)} x_i$. It easily follows that f is of the required form, since $T_\downarrow g = \lambda f$, and $T_\downarrow$ is invertible.

6 Deducing stability

Suppose now that we only know that $T_\downarrow f \approx \lambda g$. Carefully modifying the argument characterizing the exact solutions to handle approximate solutions, we can deduce that f, g are close to an AND-XOR,AND-OR pair, but the parameters are quite bad: we take an exponential hit in the number of coordinates n. A different argument is needed in order to eliminate the dependence on n. This argument will eventually involve an application of the "naive argument" just mentioned, for a *constant* value of n.

The main observation driving our approach is that $T_\downarrow$ is a "low-pass" operator, in the sense that the "high-degree" parts of g must be small. Our model here is the corresponding property of the two-sided noise operator: since $T_\rho f = \sum_d \rho^d f^{=d}$, Parseval's identity shows that bounded f satisfy $\|(T_\rho f)^{\geq d}\|^2 \leq \rho^d$. A similar, but more subtle, property holds for $T_\downarrow$.

If we are interested in the properties of $T_\rho f$ around the middle slice, then the relevant values of f are also around the middle slice. In contrast, if we are interested in the properties of $T_\downarrow f$ around the middle slice, then the relevant values of f are around the *quarter* slice. This suggests that the (usual) Fourier expansion of $T_\downarrow f$ might depend on the biased Fourier expansion of f with respect to $\mu_{1/4}$. Indeed, a simple (but eye-opening) calculation shows that

$$\widehat{T_\downarrow f}(S) = (1/\sqrt{3})^{|S|} \hat{f}(S),$$

where the Fourier coefficient on the left is with respect to $\mu_{1/2}$, and the one on the right is with respect to $\mu_{1/4}$. Since f is bounded in our case, this implies that the high-degree parts of $T_\downarrow f$, and so of g, are exponentially small.

At this point we appeal to the junta theorem of Bourgain (reproved by Kirshner, Kindler and O'Donnell with better parameters), which states that a Boolean function whose Fourier mass drops fast enough must be close to a junta. Applied with the proper parameters, Bourgain's theorem implies that g is close to some (Boolean) junta G.

At this point it is tempting to observe that $T_\downarrow F \approx \lambda G$ for a proper averaging F of f, and apply the naive argument. Unfortunately, the size of the junta and the quality of the approximation in Bourgain's result are tied together in such a way that we cannot really conclude anything using the naive argument (this is because of the exponential dependence on the size of the junta).

Consider, however, the functions g_z obtained by fixing the inputs outside the junta to some values z, and the corresponding functions f_z, defined so that on average $T_\downarrow f_z \approx \lambda g_z$, with closeness comparable to the original closeness parameter (this is the essential difference between this argument and the failed attempt, in which the error is a function of the size of the junta). If we choose $z, w \sim \mu_{1/2}$ then on average $g_z \approx G \approx g_w$, and so we can find a value of z such that both $T_\downarrow f_z \approx \lambda g_z$, and on average $g_w \approx g_z$. The former property implies that g_z is close to an AND-OR, and the latter property implies that $g \approx g_z$. In total, we have shown that g is close to an AND-OR.

A further argument (which involves "trimming" the g_z by removing large clauses) shows that f is close to the corresponding AND-XOR, in the sense that if we average f over all variables outside the AND-OR (using the $\mu_{1/4}$ measure!), then we get a function which is close to the appropriate constant multiple of the corresponding AND-XOR. (We cannot expect closeness without averaging, as the second phantom example demonstrates.)

Finally, let us go back to the original motivation. Consider a function f satisfying $f(x \wedge y) = f(x) \wedge f(y)$ on most inputs x, y. Then $T_\downarrow f \approx \lambda f$, for $\lambda = \mathbb{E}[f]$. If λ is small then $f \approx 0$. Otherwise, we can apply the preceding argument (which has a dependence on λ) to conclude that f is close to an AND-OR of width at most $\log_2(1/\lambda)$ (plus a bit). Since $\mathbb{E}[f] = \lambda$, this is only possible if f is close to an AND.

7 Extensions

Our argument works more or less the same if closeness is with respect to μ_p rather that with respect to $\mu_{1/2}$, for an arbitrary constant p. More interesting is what happens when the one-sided noise is larger or smaller.

If the noise is larger, then the only solutions to $T_\downarrow f = \lambda g$ are conjunctions, and we can also prove a robust version of this statement.

In contrast, if the noise is smaller, then new solutions appear; in fact, any monotone g is a solution for small enough noise. We can still characterize the solutions to $T_\downarrow f \approx \lambda g$: in all of them, g must be close to a junta which is part of an exact solution, the size of the junta depending on both λ and the amount of noise. (Curiously, we cannot say quite the same about f.)

8 Monomial testing

Parnas, Ron and Samorodnitsky (*Testing basic Boolean formulae*) considered property testing for monomials (which are conjunctions of literals). In particular, they considered the following natural test for membership in $\{x_1, \ldots, x_n\}$, which they were unable to analyze:

1. Sample a few values of f, and check that their average is close to 1/2.

2. Sample a few pairs x, y, and check that $f(x \wedge y) = f(x) \wedge f(y)$.

Our analysis implies that this test is sound.

(In their paper, they suggest adding linearity testing as a third step: sample a few pairs x, y, and check that $f(x \oplus y) = f(x) \oplus f(y)$. Since linearity testing is known to be sound, a function passing the test must be an XOR. For XORs one can explicitly calculate the success probability of the test $f(x \wedge y) = f(x) \wedge f(y)$, and deduce that a function passing the test must be of the form x_i.)

9 Open questions

Our work leaves many questions open. The most interesting one is to improve the dependence between the failure probability of the test and the closeness of f to an AND. We conjecture that the correct dependence is polynomial, but our arguments only show an exponential (or worse) dependence.

More generally, we can consider judgement aggregation in other scenarios. Consider a general situation in which there are m issues, and several "corollaries" which depend only on these m issues (this is called the *truth-functional* setting). Dokow and Holzman showed that non-dictatorship solutions arise only in three scenarios: all the corollaries are conjunctions (or their negations); all of them are disjunctions (or their negations); all of them are affine. Can we prove a robust version of this result?

Even more generally, one can consider an arbitrary set of allowed values. The valid judgement aggregation functions depend on universal-algebraic properties of the set of allowed values. Are these results robust?