Locating Linear Decision Lists within TC^0

Meena Mahajan

The Institute of Mathematical Sciences, HBNI, Chennai.

Workshop on Boolean Devices,
Simons Institute for the Theory of Computing.
10-14 Sep, 2018

Single tests: Linear Threshold Functions LTF, $\widehat{\text{LTF}}$

- $f : \{0,1\}^n \rightarrow \{0,1\}$ is an LTF if
  $\exists\ a_0, a_1, \ldots, a_n \in \mathbb{R}$, $\forall x \in \{0,1\}^n$, $f(x) = 1 \iff a_0 + \sum_i a_i x_i \geq 0$.

Single tests: Linear Threshold Functions LTF, $\widehat{\text{LTF}}$

- $f : \{0,1\}^n \to \{0,1\}$ is an LTF if
  $\exists\ a_0, a_1, \ldots, a_n \in \mathbb{R}, \forall x \in \{0,1\}^n, f(x) = 1 \iff a_0 + \sum_i a_i x_i \geq 0.$
- $\forall$LTF f, $\exists a_0, a_1, \ldots, a_n \in \mathbb{Z}$ describing f, with $|a_i| \leq 2^{O(n \log n)}$.
  [Muroga 1971]

Single tests: Linear Threshold Functions LTF, $\widehat{\text{LTF}}$

- $f : \{0,1\}^n \to \{0,1\}$ is an LTF if
  $\exists\ a_0, a_1, \ldots, a_n \in \mathbb{R}$, $\forall x \in \{0,1\}^n$, $f(x) = 1 \iff a_0 + \sum_i a_i x_i \geq 0$.
- $\forall$LTF f, $\exists a_0, a_1, \ldots, a_n \in \mathbb{Z}$ describing f, with $|a_i| \leq 2^{O(n \log n)}$.
  [Muroga 1971]
- $\widehat{\text{LTF}}$: those LTFs described by vectors $\tilde{a}$ with each $|a_i| \leq n^{O(1)}$.
  (i.e. Closure of MAJ under polynomial projection-reductions.)
- GreaterThan GT is an LTF. ($\text{GT}(x,y) = 1 \iff \sum_i 2^i (x_i - y_i) \geq 1$.)
  GT is not an $\widehat{\text{LTF}}$. (All rows of communication matrix of GT are distinct. So all $\langle a, x \rangle$ values must be distinct.)

Sequential Tests: Linear Decision Lists LDL

- A decision list DL of length ℓ computing $f : \{0,1\}^n \rightarrow \{0,1\}$: a sequence $(f_1, b_1), (f_2, b_2), \ldots, (f_{\ell-1}, b_{\ell-1}), (1, b_l)$ such that

$$
\begin{array}{llll}
f(x) = & \text{if} & f_1(x) & \text{then } b_1 \\
 & \text{elseif} & f_2(x) & \text{then } b_2 \\
\vdots & \vdots & \vdots & \\
 & \text{elseif} & f_{\ell-1}(x) & \text{then } b_{\ell-1} \\
 & \text{else} & & b_\ell .
\end{array}
$$

Sequential Tests: Linear Decision Lists LDL

- A decision list DL of length ℓ computing $f : \{0,1\}^n \to \{0,1\}$: a sequence $(f_1, b_1), (f_2, b_2), \ldots, (f_{\ell-1}, b_{\ell-1}), (1, b_l)$ such that

$$
\begin{array}{llll}
f(x) = & \text{if} & f_1(x) & \text{then } b_1 \\
 & \text{elseif} & f_2(x) & \text{then } b_2 \\
\vdots & \vdots & \vdots & \\
 & \text{elseif} & f_{\ell-1}(x) & \text{then } b_{\ell-1} \\
 & \text{else} & & b_\ell .
\end{array}
$$

- LDL: All f_i are LTFs.
- $\widehat{\text{LDL}}$: All f_i are $\widehat{\text{LTF}}$s.

  Perhaps better notation: DL(LTF) and DL($\widehat{\text{LTF}}$)?

What LDLs and $\widehat{\text{LDL}}$s can do

- GT not in $\widehat{\text{LTF}}$, but GT has short (poly-size) $\widehat{\text{LDL}}$.
  $(x_1 > y_1?, 1), (x_1 < y_1?, 0), \ldots, (x_n > y_n?, 1), (1, 0)$

What LDLs and $\widehat{\text{LDL}}$s can do

- GT not in $\widehat{\text{LTF}}$, but GT has short (poly-size) $\widehat{\text{LDL}}$.
  $(x_1 > y_1?, 1), (x_1 < y_1?, 0), \ldots, (x_n > y_n?, 1), (1, 0)$
- PARITY is not an LTF.
  (An LTF is either monotone or anti-monotone in each of its variables.)
  But PARITY has short $\widehat{\text{LDL}}$.

What LDLs and $\widehat{\text{LDL}}$s can do

- GT not in $\widehat{\text{LTF}}$, but GT has short (poly-size) $\widehat{\text{LDL}}$.
  $(x_1 > y_1?, 1), (x_1 < y_1?, 0), \ldots, (x_n > y_n?, 1), (1, 0)$
- PARITY is not an LTF.
  (An LTF is either monotone or anti-monotone in each of its variables.)
  But PARITY has short $\widehat{\text{LDL}}$.
- In fact, all symmetric functions have short $\widehat{\text{LDL}}$s.
  $f(x) = 1 \iff (\text{SUM} = \sum_i x_i) \in \cup_{j=1}^{k} [A_j, B_j]$.
  $\widehat{\text{LDL}} : (\text{SUM} < A_1?, 0), (\text{SUM} \leq B_1?, 1), \ldots, (\text{SUM} \leq B_k?, 1), (1, 0)$.

Parallel tests: Depth-2 threshold circuits

- Perform tests in parallel, combine results.
- All symmetric functions in MAJ $\circ$ MAJ.
  $f(x) = 1 \iff \text{SUM} = \sum_i x_i \in \cup_{j=1}^{k} [A_j, B_j]$.
  Parallel tests: $\text{SUM} \geq A_i$?, $\text{SUM} \leq B_i$?.
  Combination: Number of successful tests $\geq k+1$?

Parallel tests: Depth-2 threshold circuits

- Perform tests in parallel, combine results.
- All symmetric functions in MAJ $\circ$ MAJ.
  $f(x) = 1 \iff \text{SUM} = \sum_i x_i \in \cup_{j=1}^{k} [A_j, B_j]$.
  Parallel tests: $\text{SUM} \geq A_i$?, $\text{SUM} \leq B_i$?.
  Combination: Number of successful tests $\geq k+1$?

  [Goldmann,Håstad,Razborov 1992]
- LTF $\subseteq$ MAJ $\circ$ MAJ. ($\subsetneq$ because of PARITY.)
- MAJ $\circ$ LTF $=$ MAJ $\circ$ MAJ.
  (If top-weight small, bottom weights don't matter.)
- MAJ $\circ$ MAJ $\subsetneq$ LTF $\circ$ MAJ.
  (If bottom-weights small, top weights matter.)

Sequential Lists versus Parallel Tests

- LDL $\subseteq$ LTF $\circ$ LTF. [Turán,Vatan 1997]
  (To implement LDL $(f_1, b_1), \ldots, (f_{\ell-1}, b_{\ell-1}), (1, b_\ell)$;
  Bottom layer: all f_is. Top gate: $\sum_i (-1)^{b_i+1} 2^{\ell-i} [f_i] > 0$?)
- LTF $\circ$ LTF with top gate weights $\pm 2^i$ on ith edge equals LDL.
- $\widehat{\text{LDL}} \subseteq$ LTF $\circ$ MAJ.

Some known lower bounds for parallel tests

- OddMaxBit $\text{OMB}(z) = 1$ if last 1 is in an odd-numbered position.

Some known lower bounds for parallel tests

- OddMaxBit $\mathsf{OMB}(z) = 1$ if last 1 is in an odd-numbered position.

  $\mathsf{OMB} \circ \mathsf{AND}_2$ not in $\mathsf{MAJ} \circ \mathsf{MAJ}$.

  ($\mathsf{MAJ} \circ \mathsf{MAJ}$ circuits have inverse-polynomial discriminators.
  [Hajnal,Maas,Pudlák,Szegedy,Turán 1993].
  $\mathsf{OMB} \circ \mathsf{AND}_2$ has inverse exponential discrepancy.
  [Buhrman,Vereschagin,deWolf 2007].)

Some known lower bounds for parallel tests

- OddMaxBit $\mathrm{OMB}(z) = 1$ if last 1 is in an odd-numbered position.
  $\mathrm{OMB} \circ \mathrm{AND}_2$ not in $\mathrm{MAJ} \circ \mathrm{MAJ}$.
  ($\mathrm{MAJ} \circ \mathrm{MAJ}$ circuits have inverse-polynomial discriminators.
  [Hajnal,Maas,Pudlák,Szegedy,Turán 1993].
  $\mathrm{OMB} \circ \mathrm{AND}_2$ has inverse exponential discrepancy.
  [Buhrman,Vereschagin,deWolf 2007].)

- $\mathrm{LTF} \circ \mathrm{MAJ} \subsetneq \mathrm{LTF} \circ \mathrm{LTF}$: [Chattopadhyay,Mande 2018].
  (If top-weight large, bottom weights do matter.)

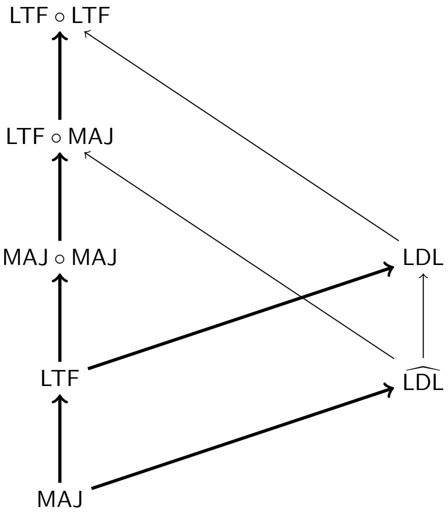
Known lower bounds for sequential tests

- $\mathsf{OMB} \circ \mathsf{AND}_2$ is in $\widehat{\mathsf{LDL}}$.
  So discrepancy does not give lower bounds for lists.

Known lower bounds for sequential tests

- $\text{OMB} \circ \text{AND}_2$ is in $\widehat{\text{LDL}}$.
  So discrepancy does not give lower bounds for lists.
- InnerProduct
  $\text{IP}_n(x, y) = (x_1 \wedge y_1) \oplus \ldots \oplus (x_n \wedge y_n) = \text{PARITY} \circ \text{AND}_2$.
  IP_n requires LDL length at least $2^{n/2}$. [Turán,Vatan 1997]
- IP_n requires exponential size in $\text{MAJ} \circ \text{MAJ}$ as well.
  [Hajnal,Maas,Pudlák,Szegedy,Turán 1993]

Polynomial size: Picture so far



A new result

Questions posed in [Turán,Vatan 1997] (restricted to poly-size):

- Are LDLs strictly weaker than LTF $\circ$ LTF?
- Are LDLs incomparable with MAJ $\circ$ MAJ?

A new result

Questions posed in [Turán,Vatan 1997] (restricted to poly-size):

- Are LDLs strictly weaker than $\mathrm{LTF} \circ \mathrm{LTF}$?
- Are LDLs incomparable with $\mathrm{MAJ} \circ \mathrm{MAJ}$?

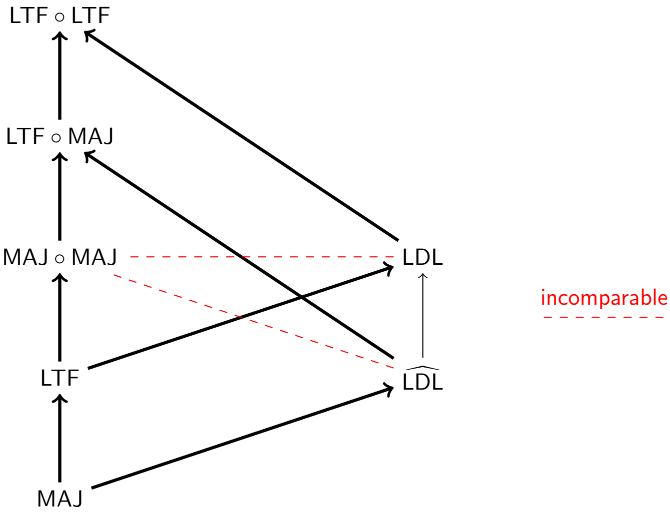
- We answer both affirmatively, with the function $\mathrm{MAJ} \circ \mathrm{XOR}$.

  Easy to see $\mathrm{MAJ} \circ \mathrm{XOR}$ is in $\mathrm{MAJ} \circ \mathrm{MAJ}$.

  (Parallel tests: $x_i + y_i \leq 1$?, $x_i + y_i \geq 1$?, Combination: Number of successful tests $\geq 3n/2$?)

  We show $\mathrm{MAJ} \circ \mathrm{XOR}$ has no short LDL.

Polynomial size: Updated Picture



MAJ $\circ$ XOR hard for LDL: Proof Idea

- Short decision list implies large monochromatic squares.
- Upper bound on size of monochromatic squares of MAJ $\circ$ XOR.

Proof Step 1

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$.
Monochromatic square: $A, B \subseteq \{0,1\}^n$, $|A| = |B|$, $A \times B \subseteq f^{-1}(b)$.
(Square size $= |A| = |B|$)

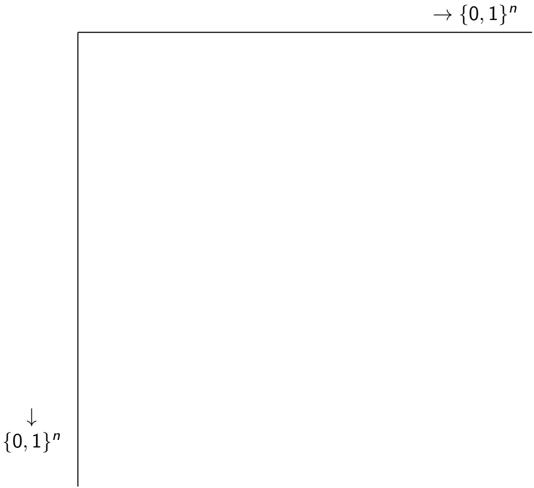
Theorem (extracted from [Turán,Vatan 1997])

If $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ has no monochromatic square of size $t+1$, then any LDL for f must have size at least $2^n/t$.

Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $\quad f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.



Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.

$\langle B_1, y\rangle$
decreasing

$\to \{0,1\}^n$

$\langle A_1, x\rangle$
decreasing

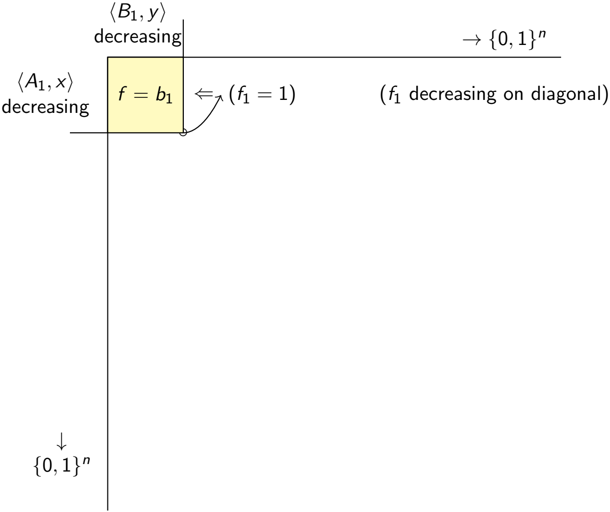
(f_1 decreasing on diagonal)

$\downarrow$
$\{0,1\}^n$

Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

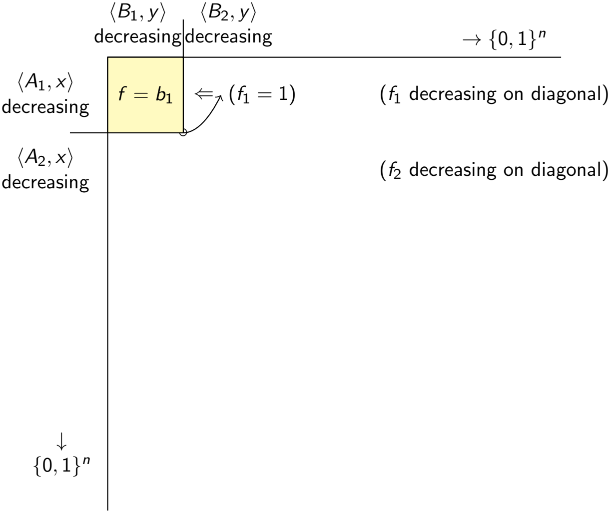
LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $\quad f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.



Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

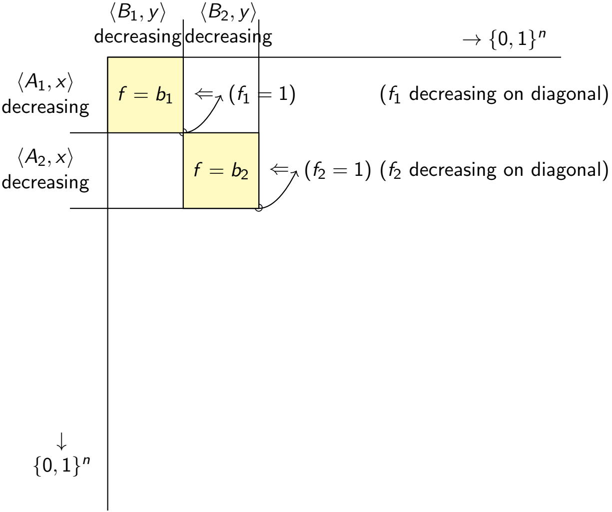
LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.



Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

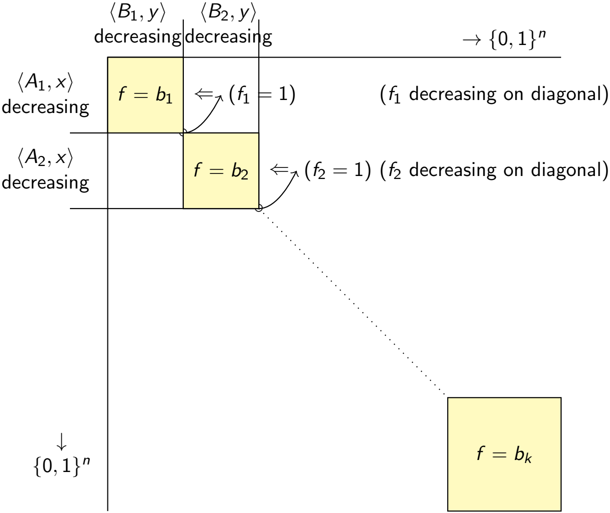
LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $\quad f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.



Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

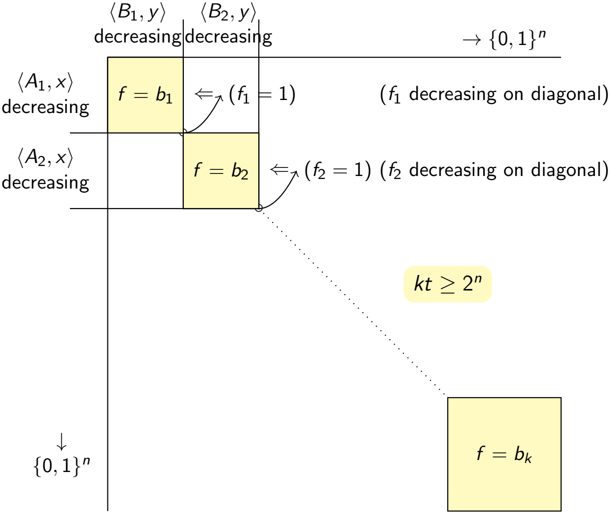
LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $\quad f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.



Proof Step 1 (cont'd)

$f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, largest monochromatic square size t.

LDL : $(f_1, b_1), (f_2, b_2), \ldots, (f_k, b_k)$; $\quad f_i : \langle A_i, x\rangle + \langle B_i, y\rangle \geq T_i$?.



Proof Step 2

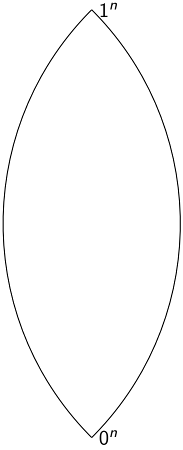
$\text{MAJ} \circ \text{XOR}(x, y) = 1$ iff Hamming distance $d_H(x, y) \geq n/2$.

Theorem

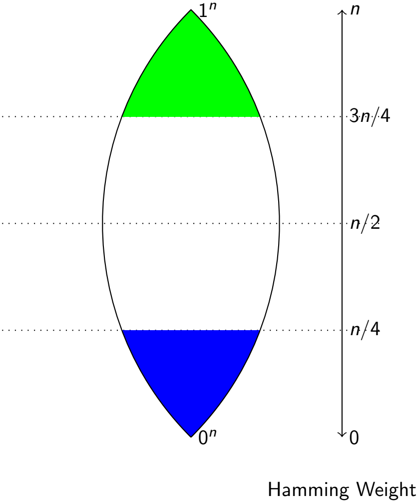
The largest monochromatic square in $\text{MAJ} \circ \text{XOR}$ *has size* $\sum_{i=0}^{\lfloor n/4 \rfloor} \binom{n}{i}$.

This size is at most $2^{0.2n}$, enough for an exp lower bound for LDL.

Proof Step 2 (cont'd)



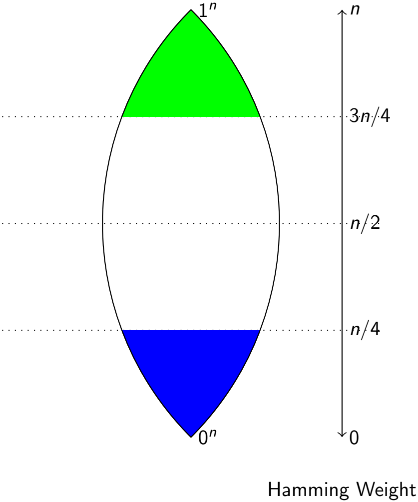
Proof Step 2 (cont'd)



A large 1-chromatic square

$$A = \{x \in \{0,1\}^n \mid HW(x) \leq n/4\}$$
$$B = \{x \in \{0,1\}^n \mid HW(x) \geq \lceil 3n/4 \rceil\}$$

Proof Step 2 (cont'd)



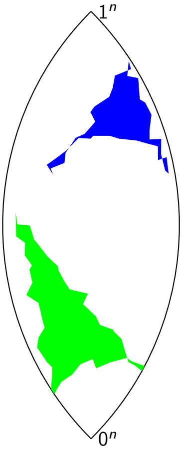
A large 1-chromatic square

$$A = \{x \in \{0,1\}^n \mid HW(x) \leq n/4\}$$
$$B = \{x \in \{0,1\}^n \mid HW(x) \geq \lceil 3n/4 \rceil\}$$

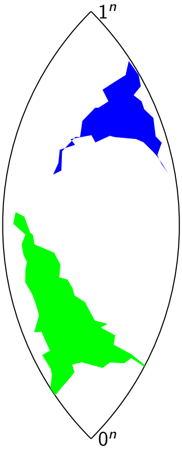
An isoperimetric inequality due to Harper guarantees no larger monochromatic square.

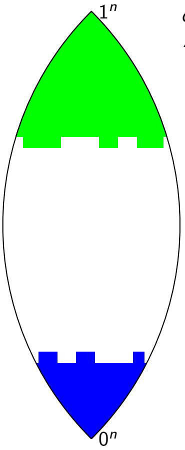
Harper's Theorem

$\forall A, B \subseteq \{0,1\}^n$



$\forall A, B \subseteq \{0,1\}^n$





$\exists A', B' \subseteq \{0,1\}^n$:
$|A'| = |A|$;
$|B'| = |B|'$
$d_H(A', B') \geq d_H(A, B)$;
A',B', Hamming Balls
centred on 0^n, 1^n.

Another separating example

- $\mathrm{OR} \circ \mathrm{EQ}$ has no large monochromatic squares.
  [Impaliazzo,Williams 2010]
  So no short LDL.
- $\mathrm{OR} \circ \mathrm{EQ}$ is in $\mathrm{MAJ} \circ \mathrm{LTF}$.
  (A more general result: $\mathrm{MAJ} \circ \mathrm{EQ} \subseteq \mathrm{MAJ} \circ \mathrm{LTF}$. [Hansen,Podolskii 2010])
  $\mathrm{MAJ} \circ \mathrm{LTF} = \mathrm{MAJ} \circ \mathrm{MAJ}$. [GHR 1992]

Spectral classes

Fourier representation of f as $\sum_{S \subseteq [n]} \widehat{f_S} \; \chi_S$.

- f is in PL_1 if $\sum_{S \subseteq [n]} |\widehat{f_S}| \leq \mathrm{poly}(n)$.
- f is in PL_∞ if $\max_{S \subseteq [n]} |\widehat{f_S}| \geq \frac{1}{\mathrm{poly}(n)}$.

Spectral classes

Fourier representation of f as $\sum_{S \subseteq [n]} \widehat{f_S}\ \chi_S$.

- f is in PL_1 if $\sum_{S \subseteq [n]} |\widehat{f_S}| \leq \mathrm{poly}(n)$.
- f is in PL_∞ if $\max_{S \subseteq [n]} |\widehat{f_S}| \geq \frac{1}{\mathrm{poly}(n)}$.
- The CompleteQuadratic function, a symmetric function, is not in PL_∞. [Bruck 1990].
- Known: $\mathrm{PL}_1 \subsetneq \mathrm{MAJ} \circ \mathrm{PARITY} \subsetneq \mathrm{LTF} \circ \mathrm{PARITY} \subsetneq \mathrm{PL}_\infty$. Also, many known relationships with depth-2 threshold circuits. [Bruck 1990], [Bruck,Smolensky 1990].

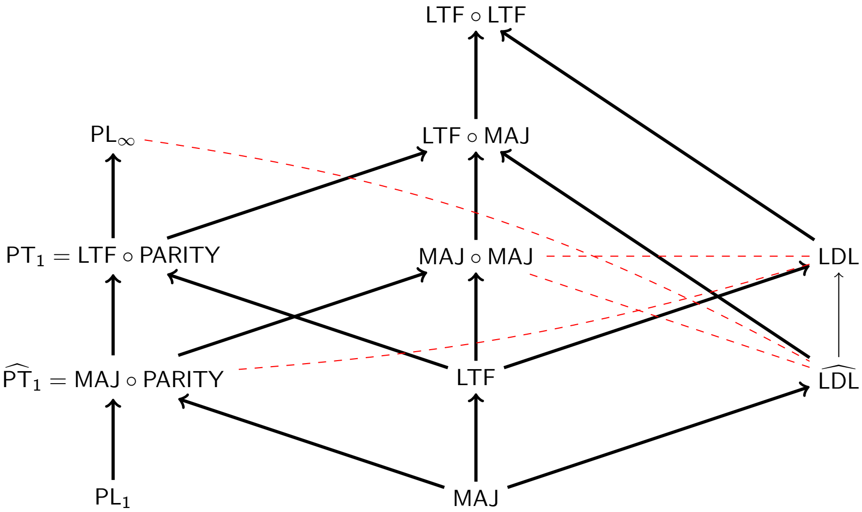
  ($\mathrm{MAJ} \circ \mathrm{PARITY} = \widehat{\mathrm{PT}}_1$ and $\mathrm{LTF} \circ \mathrm{PARITY} = \mathrm{PT}_1$: threshold tests on sparse polynomial with (poly-bounded) integer coefficients.)

Spectral classes

Fourier representation of f as $\sum_{S \subseteq [n]} \widehat{f_S}\ \chi_S$.

- f is in PL_1 if $\sum_{S \subseteq [n]} |\widehat{f_S}| \leq \mathrm{poly}(n)$.
- f is in PL_∞ if $\max_{S \subseteq [n]} |\widehat{f_S}| \geq \frac{1}{\mathrm{poly}(n)}$.
- The CompleteQuadratic function, a symmetric function, is not in PL_∞. [Bruck 1990].
- Known: $\mathrm{PL}_1 \subsetneq \mathrm{MAJ} \circ \mathrm{PARITY} \subsetneq \mathrm{LTF} \circ \mathrm{PARITY} \subsetneq \mathrm{PL}_\infty$. Also, many known relationships with depth-2 threshold circuits. [Bruck 1990], [Bruck,Smolensky 1990].
  ($\mathrm{MAJ} \circ \mathrm{PARITY} = \widehat{\mathrm{PT}}_1$ and $\mathrm{LTF} \circ \mathrm{PARITY} = \mathrm{PT}_1$: threshold tests on sparse polynomial with (poly-bounded) integer coefficients.)
- Easy: $\mathrm{MAJ} \circ \mathrm{XOR}$ is in $\mathrm{MAJ} \circ \mathrm{PARITY}$.

Polynomial size: Updated Picture



Why we got interested in Decision Lists ...

- Cutting Planes CP: a proof system for certifying unsatisfiability. QCP: an augmented proof system for certifying falsity of a Quantified Boolean Formula QBF. [Beyersdorff,Chew,M,Shukla 2016]
- QBF false $\iff$ universal player has a winning strategy in the evaluation game.
- Short QCP proof implies short LDL computing winning strategy.

Transferring LDL lower bounds to QCP

- Find a function f hard for LDL.

Transferring LDL lower bounds to QCP

- Find a function f hard for LDL.
- Design a false QBF where any winning strategy of the universal player involves computing f.

Transferring LDL lower bounds to QCP

- Find a function f hard for LDL.
- Design a false QBF where any winning strategy of the universal player involves computing f.
- This is easy if f has a small circuit C. Define QBF $Q_{f,C}$:

$$\exists x_1 x_2 \ldots x_n \forall w \exists z_1 z_2 \ldots z_m \left[\begin{array}{c} (w \neq z_m) \\ (z_i = \text{value of } i\text{th gate of } C(x)) : \quad i \in [m] \end{array} \right]$$

(z_i clauses enforce $z_m = f(x)$.)

Only winning strategy: $w = f(x)$.

Another question about QCP

- CP^*, QCP^*: restricting CP and QCP to poly-bounded coefficients.
- CP^* weaker than CP???
- Is QCP^* weaker than QCP, not purely propositionally?
  (ie even given a SAT oracle? as formalised in
  [Beyersdorff,Hinde,Pich 2017] [Chen 2016])

Another question about QCP

- CP^*, QCP^*: restricting CP and QCP to poly-bounded coefficients.
- CP^* weaker than CP???
- Is QCP^* weaker than QCP, not purely propositionally?
  (ie even given a SAT oracle? as formalised in
  [Beyersdorff,Hinde,Pich 2017] [Chen 2016])
- $\widehat{\text{LDL}}$ lower bounds can be transferred to QCP^*.
  This isn't enough, but it's a start.

Another question about QCP

- CP^*, QCP^*: restricting CP and QCP to poly-bounded coefficients.
- CP^* weaker than CP???
- Is QCP^* weaker than QCP, not purely propositionally?
  (ie even given a SAT oracle? as formalised in
  [Beyersdorff,Hinde,Pich 2017] [Chen 2016])
- $\widehat{\text{LDL}}$ lower bounds can be transferred to QCP^*.
  This isn't enough, but it's a start.
- Why not enough?
  Short QCP proof implies short LDL computing winning strategy.
  Short QCP^* proof implies short $\widehat{\text{LDL}}$ computing winning strategy.

  But converse not true; An encoding of clique-coclique has a trivial winning strategy, no short QCP proof.

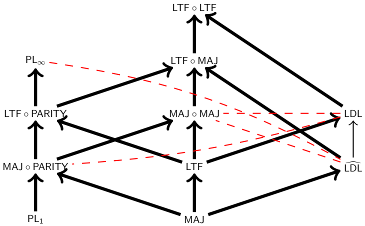
Summary

New Incomparabilities

The classes $\widehat{\text{LDL}}$ and LDL (poly-size) are incomparable with each of $\text{MAJ} \circ \text{MAJ}$, $\text{MAJ} \circ \text{PARITY}$, $\text{LTF} \circ \text{PARITY}$, PL_∞.

Questions

- Are $\widehat{\text{LDL}}$s weaker than LDLs?
- Are $\widehat{\text{LDL}}$s incomparable with LTFs?
- Are LDLs incomparable with $\text{LTF} \circ \text{MAJ}$?
- Is PL_1 in LDL? in $\widehat{\text{LDL}}$?



- Joint work with

Arkadev Chattopadhyay	TIFR, Mumbai
Nikhil Mande	TIFR, Mumbai
Nitin Saurabh	MPII, Saarbrücken

- Thanks to Rahul Santhanam, for inciting us to look for LDL lower bounds (to get unconditional QCP lower bounds).
- Thanks to Jaikumar Radhakrishnan for referring us to Harper's theorem.
- **Thank you for your attention!**