Advances in Hardness Magnification

Talk at Simons Institute – 9/December/2019

Igor Carboni Oliveira

1. Context for this talk. A lot has happened in hardness magnification since last year. The goal of this talk is to put these advances into context, and to provide more details about a couple of recent papers in this area. This presentation will be complemented by Ryan's talk on Thursday, which will focus on a different set of results.

Natural Proofs is a theory that classifies boolean devices into "weak" or "strong". Strong circuit classes can support pseudorandom functions families. For example, AC^0, $\mathsf{AC}^0[p]$, and $\mathsf{MAJ} \circ \mathsf{MAJ}$ are weak, while TC^0 and NC^1 are believed to be strong. *This theory focuses on computational models and lower bound techniques.*

While successful, natural proofs is in some sense an *incomplete theory*, specially in light of recent advances in hardness magnification. Its main limitation is perhaps to be oblivious to *computational problems*, and individual problems with special properties (such as random self reducibility, downward self reducibility, etc) have important applications in algorithms and complexity. For instance, while strong lower bounds are known against $\mathsf{AC}^0[2]$, which is a weak circuit class, it seems hard to show that $\mathsf{InnerProduct} \notin \mathsf{AC}^0 \circ \mathsf{XOR}$. Somehow the computational problem in hand can make a big difference, and the theory of natural proofs doesn't address this situation.

Hardness magnification (see [OS18] and references therein) shows that, for many natural computational problems, a weak lower bound can be "magnified" into a strong one.[1]

Example 1: Detecting huge cliques [CHOPRS'20, adapting OS'18].
If $(N - (\log N)^{50})\text{-}\mathsf{Clique} \notin \mathsf{AC}^0[N^{2+\Omega(1)}]$, then $\mathsf{NP} \not\subseteq \mathsf{NC}^1$.

The hope is that by focusing on specific problems, we might be able to sidestep the natural proofs barrier.

Example 2: Estimating complexity [OS'18].
If $\exists \gamma > 0$ such that $\forall \varepsilon > 0$ $\mathsf{MKtP}[N^\varepsilon, N^\varepsilon + 10 \cdot \log N] \notin \mathsf{Circuit}[N^{1+\gamma}]$, then $\mathsf{EXP} \not\subseteq \mathsf{Circuit}[\mathsf{poly}]$.

This result and its proof have been influential in subsequent developments. In particular, the following questions have been addressed:

1. It could be easier to prove a lower bound by exploring instances very close to the threshold N^ε. Is there a magnification theorem without a gap, i.e., for the total problem $\mathsf{MKtP}[N^\varepsilon]$? Yes [MMW'19].

2. We don't know lower bounds against $\mathsf{Circuit}[N^{1+\gamma}]$. Is there a fine-grained version of this magnification theorem for small formulas and other boolean devices? Yes [OPS'19].

[1] Hardness magnification can be seen as a *dual* theory to results of the form: *slightly faster algorithms* (such as SAT and learning) imply major lower bounds.

3. Is there something especial about meta-complexity problems such as MKtP?

– No [CMMW'19, CJW'19]: Any *sparse enough explicit language* will do.

– Yes [CHOPRS'20]: Weak lower bounds against *variants* such as $\mathsf{MCSP}[(\log N)^C, N/(\log N)^C]$ *provably* show that there exist no natural property against $\mathsf{Circuit}[\mathsf{poly}]$. Moreover, using a connection between natural properties and learning algorithms, this implies that *weak worst-case lower bounds for this problem also rule out sub-exponential time learning algorithms.*

2. Hardness Magnification Frontiers. Many magnification theorems have been discovered, and a concise summary of them can be found in [CHOPRS'20, Appendix]. In order to better understand these results, we introduce the (informal) notion of a *Hardness Magnification Frontier* [CHOPRS'20].

HM Frontier: There is a natural problem Q and a computational model $\mathcal{C}$ such that:

1. (*Magnification*) $Q \notin \mathcal{C}$ implies $\mathsf{NP} \not\subseteq \mathsf{NC}^1$ or a similar breakthrough.
2. (*Evidence of Hardness*) $Q \notin \mathcal{C}$ under a standard conjecture.
3. (*Lower Bound against* $\mathcal{C}$) $L \notin \mathcal{C}$, where L is a simple function like Parity.
4. (*Lower Bound for* Q) $Q \notin \mathcal{C}^-$, where $\mathcal{C}^-$ is slightly weaker than $\mathcal{C}$.

A frontier gives hope that the required lower bound in Item 1 is true (thanks to Item 2), and that it might be within the reach of known techniques (thanks to Items 3 and 4, which provide evidence that we can analyse both the circuit model and the problem).

We note that not every hardness magnification theorem achieves a frontier in the sense above. For instance, [CT'19] shows that for certain NC^1-complete problems, lower bounds against depth-d TC^0 circuits of size $N^{1+c^{-d}}$ would imply $\mathsf{NC}^1 \not\subseteq \mathsf{TC}^0$. However, such lower bounds are *not* known for an explicit function for the constant $c > 1$ provided by their result, let alone for these particular NC^1-complete problems. In other words, Item 3 of HM Frontier is not satisfied.

For this talk, the following HM Frontier is relevant. Let $N = 2^n$ be the input length for MCSP and for the other problems appearing next.

(A) HM Frontier for $Q = \mathsf{MCSP}$ and $\mathcal{C} = \mathsf{Formula\text{-}XOR}$:

A1. If $\mathsf{MCSP}[2^{o(n)}] \notin \mathsf{Formula}[N^{1.01}]\text{-}\mathsf{XOR}$ then $\mathsf{NP} \not\subseteq \mathsf{Formula}[N^k]$ for all k [CJM'19].
A2. $\mathsf{MCSP}[2^{o(n)}] \notin \mathsf{Circuit}[N^{\omega(1)}]$ under standard cryptographic assumptions [RR'97].
A3. $\mathsf{InnerProduct}_N \notin \mathsf{Formula}[N^{1.99}]\text{-}\mathsf{XOR}$ [Tal'17].
A4. There exists a constant $c > 1$ such that $\mathsf{MCSP}[n^c, 2^{n/c}] \notin \mathsf{Formula}[N^{1.99}]$ [OPS'19].

As a consequence of Items A1 and A4, it seems that we are a layer of XOR gates away from a breakthrough in complexity theory![2] During the rest of the talk, we will dive into the $\mathsf{Formula} \circ \mathsf{XOR}$ model, and discuss a challenge in implementing the hardness magnification program.

3. The Formula-XOR model and a new lower bound for MCSP. This extension of de Morgan formulas is interesting for other reasons beyond hardness magnification. A long-standing problem in circuit complexity is to prove better than N^3 formula size lower bounds. Observe that $\mathsf{Formula}[N^{1+\varepsilon}] \circ \mathsf{XOR} \subseteq \mathsf{Formula}[N^{3+\varepsilon}]$, so gaining a better understand of the former model seems necessary. On the technical side, parity functions available at the leaves create significant difficulties for the technique of random restrictions. Moreover, the well-known $\mathsf{Andreev}_N$ function, which requires formulas of nearly cubic size, can be computed by almost linear size $\mathsf{Formula} \circ \mathsf{XOR}$ devices. Until very recently, the only known result for this model was the lower bound for the inner product function mentioned above.

In [KKLMO'20$^+$], we provide learning algorithms, satisfiability algorithms, PRGs, and new lower bounds for $\mathsf{Formula}[N^{1.99}] \circ \mathsf{XOR}$ and its extensions. In this talk, we will discuss a new lower bound for MCSP that is connected to the HM Frontier A described above.

Theorem 1. *There exists a function $\lambda(n) = 2^{\Theta(n)}$ such that $\mathsf{MCSP}[\lambda(n)] \notin \mathsf{Formula}[N^{1.99}] \circ \mathsf{XOR}$.*

Proof Sketch. We construct a PRG $G\colon \{0,1\}^\ell \to \{0,1\}^N$ that ε-fools $\mathsf{Formula}[s] \circ \mathsf{XOR}$, where:

- $\ell = O(\sqrt{s} \cdot \log s \cdot \log(1/\varepsilon) + \log N)$, and
- every string in the output of G has circuit complexity $T = O(\ell \cdot \mathsf{poly}(\log N))$.

Since $\mathsf{MCSP}[\lambda]$ for $T < \lambda < o(N/\log N)$ is a distinguisher for G, $\mathsf{MCSP} \notin \mathsf{Formula}[N^{1.99}] \circ \mathsf{XOR}$.

Our PRG Framework. Let $\mathcal{G}$ be a class of Boolean functions. We show how to get PRGs for $\mathsf{Formula} \circ \mathcal{G}$ from *small-error* PRGs for $\mathsf{XOR} \circ \mathcal{G}$. For convenience, we switch notation to $\{-1,+1\}$.

Lemma 2 (PRG Lemma)**.** *If a distribution $\mathcal{D}$ supported over $\{-1,+1\}^N$ δ-fools the XOR of $\sqrt{s} \cdot \log(1/\varepsilon)$ functions in $\mathcal{G}$, where $\delta = \exp(-\sqrt{s} \cdot \log(s) \cdot \log(1/\varepsilon))$, then $\mathcal{D}$ also ε-fools $\mathsf{Formula}[s] \circ \mathcal{G}$.*

In our application, $\mathcal{G} = \mathsf{XOR}$, and so $\mathsf{XOR} \circ \mathcal{G} = \mathsf{XOR}$! There are well-known PRGs (small-bias generators) that δ-fool the class of parity functions on N inputs: The seed length is $O(\log(1/\delta) + \log N)$, and the circuit complexity of each output string is not much larger. So to complete the proof of Theorem 1 it remains to prove the PRG Lemma.

Proof of the PRG Lemma. We will use the following powerful result, which was established in a sequence of works using techniques from quantum computing:

Pointwise Approximation by Polynomials. For every function $F(y_1, \ldots, y_m)$ in $\mathsf{Formula}[s]$ and $\varepsilon > 0$, there is a polynomial $p(y_1, \ldots, y_m) \in \mathbb{R}[y_1, \ldots, y_m]$ of degree $d \leq \sqrt{s} \cdot \log(1/\varepsilon)$ such that

$$\big|F(a) - p(a)\big| \;\leq\; \varepsilon/3 \quad \text{for every } a \in \{-1,+1\}^m.$$

[2] For the specialist, we also highlight that HM Frontier A.4 indicates that proving worst-case lower bounds for MCSP can be significantly easier than constructing PRGs: note the parameter n^c in the lower bound for $\mathsf{MCSP}[n^c]$, and compare it to the seed length/output circuit complexity of the best known PRGs for formulas on N variables.

Given this result, we proceed as follows. Let $C = F(g_1, g_2 \ldots, g_s)$ be a function in $\mathsf{FORMULA}[s] \circ \mathcal{G}$, where F is a formula, and $g_1, g_2, \ldots, g_s$ are functions from the class $\mathcal{G}$. Let $\mathcal{U}$ be the uniform distribution over $\{-1,1\}^N$. We need to show

$$\mathbf{E}[C(\mathcal{D})] \overset{\varepsilon}{\approx} \mathbf{E}[C(\mathcal{U})].$$

Let p be a $(\varepsilon/3)$-approximating polynomial for F of degree $d = \sqrt{s} \cdot \log(1/\varepsilon)$, as described above. Let us replace F, the formula part of C, with p, and let

$$\widetilde{C} = p(g_1, g_2 \ldots, g_s).$$

Since $\tilde{C}$ approximates C on every input, we have $\mathbf{E}[\widetilde{C}(\mathcal{U})] \overset{\varepsilon/3}{\approx} \mathbf{E}[C(\mathcal{U})]$ and $\mathbf{E}[\widetilde{C}(\mathcal{D})] \overset{\varepsilon/3}{\approx} \mathbf{E}[C(\mathcal{D})]$. Thus to prove the PRG Lemma it suffices to show $\mathbf{E}[\widetilde{C}(\mathcal{D})] \overset{\varepsilon/3}{\approx} \mathbf{E}[\widetilde{C}(\mathcal{U})]$. We have

$$\begin{aligned}
\mathop{\mathbf{E}}_{x \sim \mathcal{D}}[\widetilde{C}(x)] &= \mathop{\mathbf{E}}_{x \sim D}\left[\sum_{\substack{S \subseteq [s]: \\ |S| \le d}} \hat{p}(S) \cdot \prod_{i \in S} g_i(x)\right] \\
&= \sum_{\substack{S \subseteq [s]: \\ |S| \le d}} \hat{p}(S) \cdot \mathop{\mathbf{E}}_{x \sim \mathcal{D}}\left[\prod_{i \in S} g_i(x)\right].
\end{aligned}$$

For each $S \subseteq [s]$ as above, $\prod_{i \in S} g_i(x)$ computes the XOR of at most d functions from $\mathcal{G}$. Since $\mathcal{D}$ δ-fools this class,

$$\begin{aligned}
&= \sum_{\substack{S \subseteq [s]: \\ |S| \le d}} \hat{p}(S) \cdot \left(\mathop{\mathbf{E}}_{x \sim \mathcal{U}}\left[\prod_{i \in S} g_i(x)\right] + \delta_S\right) \qquad \text{(where } |\delta_S| \le \delta\text{)} \\
&= \mathop{\mathbf{E}}_{x \sim \mathcal{U}}[\widetilde{C}(x)] + \sum_{\substack{S \subseteq [s]: \\ |S| \le d}} \hat{p}(S) \cdot \delta_S.
\end{aligned}$$

We claim that the second term above is $\le \varepsilon/3$. Since $p(z) \in [1-\varepsilon/3, 1+\varepsilon/3]$ for every $z \in \{-1,1\}^s$,

$$|\hat{p}(S)| = \left|\mathop{\mathbf{E}}_{z \sim \{-1,1\}^s}\left[p(z) \cdot \prod_{i \in S} z_i\right]\right| \le 1 + \varepsilon/3 = O(1)$$

for every set S. The result then follows from our choice of δ and the upper bound on d. □

Since the function $\lambda(n)$ appearing above is $2^{\Omega(n)}$, we cannot apply HM Frontier A.1.

– *Perhaps there is a way to prove the required lower bound using a variant of this argument?*

– *Maybe we can proceed indirectly, and reduce* $\mathsf{InnerProduct}_N$ *to* $\mathsf{MCSP}[2^{o(n)}]$*?*

We will now discuss a significant difficulty when trying to establish the lower bound in HM Frontier A. The same challenge appears in many other hardness magnification contexts.

4. The Locality Barrier. First, we make the following observation. In the proof above, if we consider instead the model $\mathsf{Formula}[s] \circ \mathcal{O}_{N^{\alpha}} \circ \mathsf{XOR}$, i.e, we add a layer of oracle gates $\mathcal{O}$ of fan-in $\leq N^{\alpha}$, where α is *small*, the argument would still go through. Why? Because we approximated the top formula by a bounded-degree polynomial, and we could equally well simply compose this polynomial with *small-degree* exact polynomials for each oracle gate. In particular,

Proposition 3. $\mathsf{MCSP}[2^{(1-o(1))n}] \notin \mathsf{Formula}[N^{1.97}] \circ \mathcal{O}_{N^{0.01}} \circ \mathsf{XOR}$.

In general, a large number of lower bound techniques easily extend to boolean devices containing certain restricted configurations of oracle gates of bounded fan-in. Further examples include AC^0 and the method of random restrictions, and the Razborov-Smolensky polynomial approximation method for $\mathsf{AC}^0[\oplus]$. This means that such techniques are powerful, but also *insensitive to the presence of "local" computations.*

The *locality barrier* [CHOPRS'20] originates from the observation that, while a magnification theorem is an implication (weak to strong lower bound), most if not all existing theorems can be turned into an *unconditional* construction of circuits with oracles of small fan-in. For example,

Proposition 4. *The following results hold:*

- $(N-(\log N)^{50})\text{-}\mathsf{Clique} \in \mathcal{O}_{\mathsf{poly}(\log N)} \circ \mathsf{AC}^0[N^{2+o(1)}]$.
- $\mathsf{MCSP}[2^{o(n)}] \in \mathsf{Formula}[N^{1+o(1)}] \circ \mathcal{O}_{N^{o(1)}} \circ \mathsf{XOR}$.

Such results are often immediate consequences of the proof of hardness magnification theorems, which tend to reduce the original problem (such as MCSP) to a problem on a much smaller instance.

Therefore, for a lower bound technique to succeed for a problem such as $\mathsf{MCSP}[2^{o(n)}]$, it must be *non-local*, i.e., it cannot easily extend to small fan-in oracle gates.

The locality barrier seems significant, and it indicates that *existing techniques are not refined enough to establish weak lower bounds for certain problems of interest.* Furthermore, it even rules out lower bounds via reductions: a problem such as $\mathsf{InnerProduct}_N$, which cannot be computed with small fan-in oracles, *provably* cannot be reduced to one that admits such computations.[3]

One is perhaps tempted to revisit the following question.

Q1. *Are the weak lower bounds required by hardness magnification true?*

Q2. *If so, can we overcome the locality barrier?*

5. Concluding remarks.

Beyond the locality barrier. In [O'19], a randomized analogue of Kt complexity has been introduced: "rKt".[4] It leads to the computational problem $\mathsf{MrKtP}[N^{\varepsilon}, N^{1-\varepsilon}]$. This problem admits hardness magnification theorems that are similar to the ones for MKtP. In particular, it provides an approach to showing that $\mathsf{Promise\text{-}BPE} \nsubseteq \mathsf{NC}^1$ from weak lower bounds against $\mathsf{Formula} \circ \mathsf{XOR}$. We note that

[3] A useful reduction for hardness magnification would need to meet certain complexity constraints, and under them, we arrive exactly at a class of reductions producing circuits that are ruled out by the locality barrier!

[4] See the slides on my webpage for a quick introduction.

the locality barrier also applies to MrKtP, in the sense that the problem becomes easy in the presence of bounded fan-in oracles: $\mathsf{MrKtP}[N^{o(1)}, N^{1-o(1)}] \in \mathsf{Promise\text{-}BPTIME}^{\mathcal{O}}[N^{1+o(1)}]$, where $\mathcal{O}$ has fan-in $N^{o(1)}$.

Crucially, for MrKtP we are able to address questions Q1 and Q2 above, but in the *uniform* world. More precisely, [O'19] unconditionally proved that $\mathsf{MrKtP}[N^{o(1)}, N^{1-o(1)}] \notin \mathsf{Promise\text{-}BPP}$. The proof employs and indirect diagonalization argument via pseudorandomness, and does not extend to computations with small fan-in oracles.[5] This shows that there are lower bound techniques and computational problems that, together, can be sensitive to such oracles. An additional example appears in [CHOPRS'20].

Looking ahead. At this point it is unclear if hardness magnification can lead to lower bounds. Nevertheless, these ideas are allowing us to obtain a better understanding of the difficulty of proving very weak results in complexity theory. Along the way, we are also establishing new connections (see e.g. [CHOPRS'20]). In particular, hardness magnification forces us to revisit combinatorial lower bound methods and the theory of natural proofs, and to understand why some problems behave differently than others. Even if we are unable to apply it to get lower bounds, it is possible that magnification will lead us to a more robust theory on the hardness of proving lower bounds.

Remark. I haven't really explained in this talk the techniques employed in the proof of hardness magnification theorems. In this direction, Ryan's talk will provide more details about a few exciting magnification theorems (e.g. [CJW'19, CJW'20$^+$]) discovered over the last few months.

Appendix. A list of recent works on hardness magnification:

[**OS'18**] Hardness magnification for natural problems (FOCS'18).

[**OPS'19**] Hardness magnification near state-of-the-art lower bounds (CCC'19).

[**CMMW'19**] Relations and equivalences between circuit lower bounds and Karp-Lipton theorems (CCC'19).

[**MMW'19**] Weak lower bounds on resource-bounded compression imply strong separations of complexity classes (STOC'19).

[**CT'19**] Bootstraping results for threshold circuits just beyond known lower bounds (STOC'19).

[**O'19**] Randomness and intractability in Kolmogorov complexity (ICALP'19).

[**CJW'19**] Hardness magnification for all sparse languages (FOCS'19).

[**CHOPRS'20**] Beyond natural proofs: hardness magnification and locality (ITCS'20).

[**KKLMO'20$^+$**] Algorithms and lower bounds for de Morgan formulas of low-communication leaf gates (Preprint).

[**CJW'20$^+$**] Sharp threshold results in computational complexity (Preprint).

[5] Roughly speaking, when diagonalizing in the presence of oracles, the easiness of MrKtP does not give us much mileage, since this problem refers to *standard computations*!