Beyond natural proofs: hardness magnification and locality

Beyond Natural Proofs: Hardness Magnification and Locality

LIJIE CHEN, Massachusetts Institute of Technology, USA
SHUICHI HIRAHARA, National Institute of Informatics, Japan
IGOR CARBONI OLIVEIRA, University of Warwick, UK
JÁN PICH, University of Oxford, UK
NINAD RAJGOPAL, University of Warwick, UK
RAHUL SANTHANAM, University of Oxford, UK

Hardness magnification reduces major complexity separations (such as EXP $\not\subseteq$ NC^1) to proving lower bounds for some natural problem Q against weak circuit models. Several recent works [11, 13, 14, 40, 42, 43, 46] have established results of this form. In the most intriguing cases, the required lower bound is known for problems that appear to be significantly easier than Q, while Q itself is susceptible to lower bounds, but these are not yet sufficient for magnification.

In this work, we provide more examples of this phenomenon and investigate the prospects of proving new lower bounds using this approach. In particular, we consider the following essential questions associated with the hardness magnification program:

— *Does hardness magnification avoid the natural proofs barrier of Razborov and Rudich* [51]*?*
— *Can we adapt known lower-bound techniques to establish the desired lower bound for Q?*

We establish that some instantiations of hardness magnification overcome the natural proofs barrier in the following sense: slightly superlinear-size circuit lower bounds for certain versions of the minimum circuit-size problem imply the non-existence of natural proofs. As the non-existence of natural proofs implies the non-existence of efficient learning algorithms, we show that certain magnification theorems not only imply strong worst-case circuit lower bounds but also rule out the existence of efficient learning algorithms.

Hardness magnification might sidestep natural proofs, but we identify a source of difficulty when trying to adapt existing lower-bound techniques to prove strong lower bounds via magnification. This is captured

The reader can refer to Reference [47] for an alternative exposition of the results in this article and additional observations.

Lijie Chen is supported by Grants No. NSF CCF-1741615 and NSF CCF-2127597, a Google Faculty Research Award, and an IBM Fellowship. Igor C. Oliveira received support from the Royal Society University Research Fellowship Grant No. URF\R1\191059. Ján Pich was supported by the Royal Society University Research Fellowship Grant No. URF\R1\211106, and in part by Grant No. 19-05497S of GA ČR. Rahul Santhanam was partially funded by the EPSRC New Horizons Grant No. EP/V048201/1: "Structure versus Randomness in Algorithms and Computation." This work was supported in part by the European Research Council under the European Union's Seventh Framework Programme (Grant No. FP7/2007-2014)/ERC Grant Agreement No. 615075. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodovska-Curie Grant Agreement No. 890220. Shuichi Hirahara is supported by ACT-I, JST. Ninad Rajgopal was supported in part by Tom Gur's UKRI Future Leaders Fellowship Grant number: MR/S031545/1.

Authors' addresses: L. Chen, Massachusetts Institute of Technology, 32 Vassar St, Cambridge, MA 02139, USA; email: lijieche@mit.edu; S. Hirahara, Department of Computer Science, University of Warwick, CV4 7AL, UK; email: s_hirahara@nii.ac.jp; I. C. Oliveira and N. Rajgopal, University of Warwick, UK; emails: {igor.oliveira, ninad.rajgopal}@warwick.ac.uk; J. Pich and R. Santhanam, Department of Computer Science, University of Oxford, Parks road, Oxford, OX1 3QD, UK; emails: {jan.pich, rahul.santhanam}@cs.ox.ac.uk.

0004-5411/2022/08-ART25 $15.00
https://doi.org/10.1145/3538391

by a *locality barrier*: existing magnification theorems *unconditionally* show that the problems Q considered above admit highly efficient circuits extended with small fan-in oracle gates, while lower-bound techniques against weak circuit models quite often easily extend to circuits containing such oracles. This explains why direct adaptations of certain lower bounds are unlikely to yield strong complexity separations via hardness magnification.

CCS Concepts: • **Theory of computation**;

Additional Key Words and Phrases: Circuit complexity, natural proofs, hardness magnification

ACM Reference format:
Lijie Chen, Shuichi Hirahara, Igor Carboni Oliveira, Ján Pich, Ninad Rajgopal, and Rahul Santhanam. 2022. Beyond Natural Proofs: Hardness Magnification and Locality. *J. ACM* 69, 4, Article 25 (August 2022), 49 pages.
https://doi.org/10.1145/3538391

1 INTRODUCTION

Proving circuit-size lower bounds for explicit Boolean functions is a central problem in Complexity Theory. Unfortunately, it is also notoriously hard, and arguments ruling out a wide range of approaches have been discovered. The most prominent of them is the *natural proofs barrier* of Razborov and Rudich [51].

A candidate approach for overcoming this barrier was investigated recently by Oliveira and Santhanam [46]. *Hardness Magnification* identifies situations where strong circuit lower bounds for explicit Boolean functions (e.g., $\mathsf{NP} \not\subseteq \mathsf{P/poly}$) follow from much weaker (e.g., slightly superlinear) lower bounds for specific natural problems. As discussed in Reference [46], in some cases the lower bounds required for magnification are already known for explicit problems, but not yet for the problem for which the magnification theorem holds. This approach to lower bounds has attracted the interest of several researchers, and a number of recent works have proved magnification results [11, 13, 14, 40, 42, 43] (see also References [3, 39, 41, 55] for related previous work). We provide a concise review of existing results in Appendix A.1.

In this work, we are interested in understanding the prospects of proving new lower bounds using hardness magnification, including potential barriers.

1.1 Hardness Magnification Frontiers

While hardness magnification is a broad phenomenon, its most promising instantiations seem to occur in the setting of circuit classes such as NC^1. The potential of hardness magnification stems from establishing the following scenario.

> **HM Frontier:** There is a natural problem Q and a computational model C such that:
>
> 1. (*Magnification*) $Q \notin C$ implies $\mathsf{NP} \not\subseteq \mathsf{NC}^1$ or a similar breakthrough.
> 2. (*Evidence of Hardness*) $Q \notin C$ under a standard conjecture.
> 3. (*Lower Bound against C*) $L \notin C$, where L is a simple function like PARITY.
> 4. (*Lower Bound for Q*) $Q \notin C^-$, where C^- is slightly weaker than C.

A frontier of this form provides 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 the circuit model and the problem). Scenarios similar to HM Frontier were identified already by Oliveira and Santhanam [46], but they did not obtain

Item 4 (a non-trivial lower bound for the same problem that appears in the magnification theorem). Subsequent works have addressed this issue and achieved HM frontiers (with Item 4). A striking example, similar to HM frontier B presented below, appeared in Reference [43] (see also Reference [11]). Despite the number of works in this area, we note that the HM frontier is achieved only by some magnification theorems (Item 3 is often unknown; e.g., in the case of results in References [3, 14]).

To make our subsequent discussion more concrete, we provide five examples of HM frontiers. Some of these results are new or require an extension of previous work, and the relevant statements will be explained in more detail in Section 3. The list of frontiers is not meant to be exhaustive, but we have tried to cover different computational models.

(A) HM Frontier for $\mathsf{MKtP}[n^c, 2n^c]$ **and** AC^0**-XOR:**

A1. If $\mathsf{MKtP}[n^c, 2n^c] \notin \mathsf{AC}^0\text{-XOR}[N^{1.01}]$ for large $c > 1$, then $\mathsf{EXP} \not\subseteq \mathsf{NC}^1$ (Section 3.1).
A2. $\mathsf{MKtP}[n^c, 2n^c] \notin \mathsf{P/poly}$ for large enough c under exponentially secure PRFs [51].
A3. Majority $\notin \mathsf{AC}^0\text{-XOR}[2^{N^{o(1)}}]$ (immediate from Reference [49, 54]).
A4. $\mathsf{MKtP}[n^c, 2n^c] \notin \mathsf{AC}^0$ for any sufficiently large constant c (Section 3.1).

A. $\mathsf{MKtP}[s, t]$ refers to the promise problem of determining if an N-bit input has Levin Kolmogorov complexity at most s versus at least t (cf. Reference [43]). Here $N = 2^n$. The AC^0-XOR model is the extension of AC^0 where gates at the bottom layer of the circuit can compute arbitrary parity functions. $\mathsf{AC}^0\text{-XOR}[s]$ denotes AC^0-XOR circuits of size s where the size is measured as the number of gates. This circuit class has received some attention in recent years (cf. Reference [16]), and a few basic questions about AC^0 circuits with parity gates (such as constructing PRGs of seed length $o(n)$ and learnability using random examples) remain open for AC^0-XOR as well.

(B) HM Frontier for $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ **and** Formula-XOR**:**

B1. $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \notin \text{Formula-XOR}[N^{1.01}]$ implies $\mathsf{NQP} \not\subseteq \mathsf{NC}^1$ (Section 3.2).
B2. $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \notin \mathsf{P/poly}$ under standard cryptographic assumptions [51].
B3. InnerProduct $\notin \text{Formula-XOR}[N^{1.99}]$ (immediate consequence of Reference [57]).
B4. $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \notin \text{Formula}[N^{1.99}]$ ([26]; see also Reference [43]).

B. Here, NQP is nondeterministic quasi-polynomial time, InnerProduct is the Boolean function defined as $\text{InnerProduct}(x, y) = \sum_i x_i \cdot y_i \pmod 2$, where $x, y \in \{0,1\}^N$, Formula-XOR$[s]$ refers to the class of Boolean formulas over the De Morgan basis with at most s leaves, where each leaf is an XOR of arbitrary arity over the inputs,[1] and $\mathsf{MCSP}[s, t]$ denotes a promise **minimum circuit-size problem (MCSP)** over $N = 2^n$ input bits with YES inputs being truth tables of Boolean functions on n inputs that are computable by circuits of size s, and NO instances being truth tables of Boolean functions that are hard for circuits of size t.

[1]Note that $\text{Formula-XOR}[N^{1.01}] \subseteq \text{Formula}[N^{3.01}]$. A better understanding of the former class is therefore necessary before we can understand the power and limitations of super-cubic formulas, which is a major open question in circuit complexity.

(C) HM Frontier for MCSP$[2^{n^{1/2}}/10n, 2^{n^{1/2}}]$ **and** Almost-Formulas**:**

C1. MCSP$[\frac{2^{n^{1/2}}}{10n}, 2^{n^{1/2}}] \notin N^{0.01}$-Almost-Formula$[N^{1.01}]$ implies NP $\not\subseteq$ NC^1 (Section 3.3).
C2. MCSP$[\frac{2^{n^{1/2}}}{10n}, 2^{n^{1/2}}] \notin$ P/poly under standard cryptographic assumptions [51].
C3. PARITY $\notin N^{0.01}$-Almost-Formula$[N^{1.01}]$ (Section 3.3).
C4. MCSP$[2^{n^{1/2}}/10n, 2^{n^{1/2}}] \notin$ Formula$[N^{1.99}]$ ([26]; see also Reference [43]).

C. An almost-formula is a circuit with a bounded number of gates of fan-out larger than 1. More precisely, a γ-Almost-Formula$[s]$ is a circuit containing at most s AND, OR, NOT gates of fan-in at most 2, and among such gates, at most γ of them have fan-out larger than 1. Consequently, this class naturally interpolates between formulas and circuits. This magnification frontier can be seen as progress toward establishing magnification theorems for worst-case variants of MCSP in the regime of sub-quadratic formulas (see the discussion in Reference [43]).

(D) HM Frontier for MCSP$[2^{\sqrt{n}}]$ **and one-sided error randomized formulas:**

D1. MCSP$[2^{\sqrt{n}}] \notin$ GapAND$_{O(N)}$-Formula$[N^{2.01}] \Rightarrow$ NQP $\notin$ NC^1 (Section 3.2).
D2. MCSP$[2^{\sqrt{n}}] \notin$ P/poly under standard cryptographic assumptions [51].
D3.1. Andreev$_N \notin$ GapAND$_{O(N)}$-Formula$[N^{2.99}]$ (implicit in Reference [23]).
D3.2. MCSP$[2^n/n^4] \notin$ GapAND$_{O(N)}$-Formula$[N^{2.99}]$ (implicit in Reference [18]).
D4. MCSP$[2^{\sqrt{n}}] \notin$ GapAND$_{O(N)}$-Formula$[N^{1.99}]$ ([12], building on References [26, 43]).

D. GapAND$_N$ is the promise function on N bits such that it outputs 1 when all input bits are 1, and outputs 0 when at most 1/10 of the input bits are 1. GapAND$_{O(N)}$-Formula$[s]$ denotes circuits with GapAND$_{O(N)}$ gate at the top with formulas of size s being inputs of the top gate. Therefore, GapAND$_{O(N)}$-Formula can be seen as *randomized formulas with one-sided error*.[2] The most interesting aspect of this magnification frontier is that the gap between the known hardness result and the magnification threshold is *nearly-tight* ($N^{2-\varepsilon}$ versus $N^{2+\varepsilon}$).[3]

(E) HM Frontier for $(n-k)$-Clique **and** AC^0**:**

E1. If $(n-k)$-Clique $\notin$ AC$^0[m^{1.01}]$ for some $k = (\log n)^C$, then NP $\not\subseteq$ NC^1 (Section 3.4).
E2. (Non-uniform) ETH $\Rightarrow$ $(n-k)$-Clique $\notin$ P/poly for some $k = (\log n)^C$ (Section 3.4).
E3. Parity $\notin$ AC^0 [1, 21].
E4. $(n-k)$-Clique $\notin$ mP/poly for some $k = (\log n)^C$ ([4]; see Section 3.4).

E. The ℓ-Clique problem is defined on graphs on n vertices in the adjacency matrix representation of size $m = \Theta(n^2)$. (The statements above refers to the regime of very large clique detection.) The

[2]Suppose there is a GapAND$_{O(N)}$-Formula circuit computing a function $f: \{0,1\}^N \to \{0,1\}$. Consider a uniform distribution of all subformulas below the top GapAND$_{O(N)}$ gate. Then for any input x, if $f(x) = 1$, then a sample formula from that distribution always outputs 1 on x; otherwise, it outputs 0 with probability at least 0.9 on x. However, it is possible to derandomize a distribution of formulas computing f with one-sided error using a top GapAND$_{O(N)}$ gate.
[3]This tight threshold is first observed in Reference [12], we include it here to show that the barrier discussed in this article also applies to this particular setting.

class mP/poly refers to monotone circuits of polynomial size. In this frontier, we are modifying Item 4 from HM frontier so that instead of slightly weaker C^- we consider an incomparable C^-. This frontier is, however, particularly interesting, as items E1 and E4 connect hardness magnification to a basic question about the power of *non-monotone* circuits when computing *monotone* functions (see References [15, 22] and references therein): Is every monotone function in AC^0 computable by a monotone (unbounded depth) Boolean circuit of polynomial size? If this is the case, then $\mathsf{NP} \not\subseteq \mathsf{NC}^1$ would follow.

Note that these hardness magnification frontiers offer different approaches to proving lower bounds against NC^1.

Essential Questions. Do magnification theorems bring us closer to strong circuit lower bounds? To understand the limits and prospects of hardness magnification, the following questions are relevant.

- **Q1.** *Naturalization.* Is hardness magnification a *non-naturalizing* approach to circuit lower bounds? If we accept standard cryptographic assumptions, then non-naturalizability is a necessary property of any successful approach to strong circuit lower bounds.[4]
- **Q2.** *Extending known lower bounds.* Can we adapt an existing lower bound proof from Items 3 and 4 in some HM frontier to show the lower bound required from Item 1 in that HM frontier? Is it possible to establish the required lower bounds via a reduction from L to Q?
- **Q3.** *Improving existing magnification theorems.* Can we close the gap between Items 1 and 4 in HM frontier by establishing a magnification theorem that meets *known* lower bounds, such as the ones appearing in Item 4?

In the next sections, we present results that shed light into all these questions.

1.2 Hardness Magnification and Natural Proofs

The very existence of the natural proofs barrier provides a direction for proving strong circuit lower bounds: one can proceed by *refuting the existence of natural properties*.[5] In other words, a way to avoid natural proofs is to prove that there are no natural proofs. It is also easy to see that P/poly-natural properties useful against P/poly can be turned into natural properties with much higher constructivity, e.g., into linear-size natural properties useful against circuits of polynomial-size.[6] If read contrapositively, then this gives a form of hardness magnification.

The initial hardness magnification theorem of Oliveira and Santhanam [46] proceeds in a similar fashion. It proposes to approach $\mathsf{NP} \not\subseteq \mathsf{P/poly}$ by deriving slightly superlinear circuit lower bounds for specific problems such as an *approximate version* of MCSP, which asks to distinguish truth tables of Boolean functions computable by small circuits from truth tables of Boolean functions that are hard to approximate by small circuits. Interestingly, this approach does not seem to naturalize,

[4] We assume familiarity of the reader with the natural proofs framework of Reference [51]. Intuitively, the natural proofs barrier says that the existing circuit lower bounds are too strong in the sense that they give us not only the lower bound we want but also the so called natural property: an efficient circuit accepting many Boolean functions (represented by their truth tables) that are hard for the respective circuit class, and rejecting all easy functions. However, if such efficient circuits existed for strong circuit classes such as P/poly, then they would break the existence of strong pseudorandom generators. See Preliminaries (Section 2) for definitions.

[5] A similar perspective has been employed in proof complexity in attempts to approach strong proof complexity lower bounds by extending the natural proofs barrier (see References [38, 50]).

[6] A more constructive natural property is obtained from a less constructive natural property by applying the less constructive natural property on a suitably long prefix of the input.

as it appears to yield strong lower bounds only for certain problems, and not for most of them. (The same heuristic argument appears in Reference [3].) However, this is only an informal argument, and we would like to get stronger evidence that the natural proofs barrier does not apply here.

We show that hardness magnification for approximate MCSP can be used to conclude the *non-existence* of natural proofs against polynomial-size circuits. More precisely, we prove that if approximate MCSP requires slightly superlinear-size circuits, then there are no P/poly-natural properties against P/poly. This strongly suggests that the natural proofs barrier is not relevant to the magnification approach. Indeed, there remains the possibility that the weak circuit lower bound for MCSP in the hypothesis of the result can be shown using naturalizing techniques (as there are not any strong enough plausible cryptographic conjectures known that rule this out), and yet by using magnification to "break" naturalness, we could get strong circuit lower bounds and even conclude the non-existence of natural proofs![7]

The core of our proof is the following new hardness magnification theorem: If approximate MCSP requires slightly superlinear-size circuits, then not only $\mathsf{NP} \not\subseteq \mathsf{P/poly}$ but *it is impossible even to learn efficiently*. We can then refute the existence of natural proofs by applying the known translation of natural properties to learning algorithms [8]. Similar implications hold with a *worst-case gap version* of MCSP (in the sense of HM Frontiers B and C but with different parameters) instead of approximate MCSP, following an idea from Reference [24].

Interestingly, all the implications from the previous paragraph are actually *equivalences*. In particular, the existence of natural properties is equivalent to the existence of highly efficient circuits for computing approximate MCSP and worst-case gap MCSP with certain parameters (cf. Theorem 1). This extends a known characterization of natural properties: Carmosino et al. [8] showed that P/poly natural proofs against P/poly are equivalent to learning P/poly by subexponential-size circuits, which was in turn shown to be equivalent by Oliveira and Santhanam [45] to the non-existence of non-uniform pseudorandom function families of sub-exponential security. The connection of hardness magnification to learning and pseudorandom function generators might be of independent interest, since it extends the consequences of magnification into two central areas in Complexity Theory.

Theorem 1 (Equivalences for Hardness Magnification). *The following statements are equivalent*[8]*:*

(*a*) ***Hardness of approximate* MCSP *against almost-linear size circuits.***
There exist $c \geq 1$, $0 < \gamma < 1$, *and* $\varepsilon > 0$ *such that* $\mathsf{MCSP}[(n^c, 0), (2^{n^\gamma}, n^{-c})] \notin \mathsf{Circuit}[N^{1+\varepsilon}]$.

(*b*) ***Hardness of worst-case* MCSP *against almost-linear size circuits.***
There exists $c \geq 1$ *and* $\varepsilon > 0$ *such that* $\mathsf{MCSP}[n^c, 2^n/n^c] \notin \mathsf{Circuit}[N^{1+\varepsilon}]$.

(*c*) ***Hardness of sub-exponential size learning using non-adaptive queries.***
There exist $\ell \geq 1$ *and* $0 < \gamma < 1$ *such that* $\mathsf{Circuit}[n^\ell]$ *cannot be learned up to error* $O(1/n^\ell)$ *under the uniform distribution by circuits of size* $2^{O(n^\gamma)}$ *using non-adaptive membership queries.*

(*d*) ***Non-existence of natural properties against polynomial size circuits.***
For some $d \geq 1$ *there is no* $\mathsf{Circuit}[\mathrm{poly}(N)]$*-natural property useful against* $\mathsf{Circuit}[n^d]$.

[7] We remark that lower bounds from HM frontiers A3–E3 and A4–D4 do naturalize. For example, the lower bound from B3 naturalizes because of the existence of learning algorithms constructed in Reference [34]. The naturalization of C3 is obtained by using Lemma 52 in Reference [34], which shows that each function approximable by a formula of size $n^{1.99}$ (unlike a random function) has a non-trivial correlation with some parity function. This allows us to recognize many functions hard to approximate by $n^{1.99}$-size formulas (and thus hard for small almost-formulas) by checking their correlations with all parity functions.

[8] See Preliminaries (Section 2) for definitions.

(*e*) ***Existence of non-uniform* PRFs *secure against sub-exponential size circuits.*** *For every constant $a \geq 0$, there exists $d \geq 1$, a sequence $\mathfrak{F} = \{\mathcal{F}_n\}_{n\geq 1}$ of families $\mathcal{F}_n$ of n-bit Boolean functions $f_n \in \mathrm{Circuit}[n^d]$, and a sequence of probability distributions $\mathfrak{D} = \{\mathcal{D}_n\}_{n\geq 1}$ supported over $\mathcal{F}_n$ such that, for infinitely many values of n, $(\mathcal{F}_n, \mathcal{D}_n)$ is pseudo-random function family that $(1/N^{\omega(1)})$-fools (oracle) circuits of size $2^{a\cdot n}$.*

The proof of this result appears in Section 4.1. We highlight below the most interesting implications of Theorem 1. (Note that some of them have appeared in other works in similar or related forms.)

- $(a) \rightarrow (d)$: The initial hardness magnification result from Reference [46, Theorem 1] (stated for circuits) implies the *non-existence* of natural proofs useful against polynomial-size circuits, indicating that the natural proofs barrier might not be relevant to the magnification approach.
- $(a), (b) \leftrightarrow (d)$: *Any* P/poly natural property useful against P/poly can be transformed into an almost-linear size natural property that is simply the approximate $\mathrm{MCSP}[(n^c, 0), (2^{n^\gamma}, n^{-c})]$ or worst-case gap $\mathrm{MCSP}[n^c, 2^n/n^c]$. (Note the different regime of circuit-size parameters for these problems.)
- $(a), (b) \leftrightarrow (c)$: A weak-seeming hardness assumption for worst-case gap and approximate versions of MCSP implies a strong non-learnability result: polynomial-size circuits cannot be learned over the uniform distribution even non-uniformly in sub-exponential time.
- $(a), (b) \leftrightarrow (e)$: Hardness magnification for MCSP also yields cryptographic hardness in a certain regime.

We note that the use of non-adaptive membership queries in Theorem 4.1 Item (c) is not essential. It follows from Reference [8] that, in the context of learnability of polynomial size circuits under the uniform distribution in sub-exponential time, adaptive queries are not significantly more powerful than non-adaptive queries.[9]

Toward a more robust theory. While Theorem 1 formally connects hardness magnification and natural properties, it would be very interesting to understand to which extent different hardness magnification theorems are provably non-naturalizable. This would provide a more complete answer to Question Q1 asked above. For instance, Theorem 1 leaves open whether hardness magnification for worst-case versions of MCSP such as $\mathrm{MCSP}[n^c, 2^{n^\varepsilon}]$ refutes natural proofs as well. Note that one way of approaching this question would be to study reductions from $\mathrm{MCSP}[n^c, 2^{n^\gamma}]$ to its approximate version $\mathrm{MCSP}[(n^{c'}, 0), (2^{n^{\gamma'}}, n^{-c'})]$.[10] In Section 4.2, we observe that this question is related to the problem of basing *hardness of learning* on *worst-case assumptions* such as $\mathsf{P} \neq \mathsf{NP}$ (cf. Reference [5]). We refer to the discussion in Section 4.2 for more details.

1.3 The Locality Barrier

The results from the preceding section show that hardness magnification can go beyond natural proofs. Is there another barrier that makes it difficult to establish lower bounds via magnification? In this section, we present a general argument to explain why the lower-bound techniques behind A3–E3, A4–D4 in the magnification frontiers from Section 1.1 cannot be adapted (without

[9] In a bit more detail, one can easily extract a natural property from a learner that uses adaptive queries. In turn, closer inspection of the technique of Reference [8] shows that a non-adaptive learner can be obtained from a natural property.

[10] More precisely, the existence of a reduction from $\mathrm{MCSP}[n^c, 2^{n^\gamma}]$ to $\mathrm{MCSP}[(n^{c'}, 0), (2^{n^{\gamma'}}, n^{-c'})]$ shows that lower bounds for the former problem yield lower bounds for the latter. Since any such lower bound must be non-naturalizable by Theorem 1, we obtain the same consequence for $\mathrm{MCSP}[n^c, 2^{n^\gamma}]$. (Note that in the context of hardness magnification it is also important to have highly efficient reductions.)

significantly new ideas) to establish the required lower bounds in Items A1–E1, respectively. We refer to it as the *locality barrier*. While we will focus on these particular examples to make the discussion concrete, we believe that this barrier applies more broadly (and can be seen as a circuit-complexity analog of Relativization).

To explain the locality barrier, let us consider the argument behind the proof of B1 presented in Section 3.2. Recall that this result shows that if $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \notin \text{Formula-XOR}[N^{1.01}]$ then $\mathsf{NQP} \not\subseteq \mathsf{NC}^1$. This and other known hardness magnification theorems are established in the contrapositive. The core of the argument is to prove that there are highly efficient Formula-XOR circuits that reduce an input to $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ of length $N = 2^n$ to deciding whether certain strings of length N' (much smaller than N) belong to a certain language L'. Then, under the assumption that $\mathsf{NQP} \subseteq \mathsf{NC}^1$, one argues that L' has polynomial size formulas. Finally, since $N' \ll N$, we can employ such formulas and still conclude that $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ is in $\text{Formula-XOR}[N^{1.01}]$, which completes the proof.

Note that the argument above provides a *conditional* construction of highly efficient formulas for the original problem. Crucially, however, we can derive an *unconditional* circuit upper bound from this argument: If we stop right before we replace the calls to L' by an algorithm for L' (this is what makes the reduction conditional), then it unconditionally follows that $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ can be computed by highly efficient Formula-XOR circuits containing oracle gates of small fan-in, for some oracle. Similarly, one can argue that the problems in Items A1–E1 can be computed in the respective models by highly efficient Boolean devices containing oracles of small fan-in.

We stress that, as opposed to a magnification theorem, where one cares about the complexity of the oracle gates, in our discussion of the locality barrier, we only need the fact that there is *some way* of setting these oracles gates so that the resulting circuit or formula solves the original problem. (A definition of this model appears in Section 2.5.) A more exhaustive interpretation of magnification theorems as construction of circuits with small fan-in oracles can be found in Appendix A.2.

However, we argue that the lower-bound arguments from Items A3–E3 of the hardness magnification frontiers quite easily handle (in the respective models) the presence of oracles of small fan-in, *regardless of the function* computed by these oracles. Using a more involved argument, we can also localize lower bounds from items A4–D4. Consequently, these methods do not seem to be refined enough to prove the lower bounds required by A1–D1 without excluding oracle circuits that are unconditionally known to exist for the corresponding problems.

Following the example above, we state our results for the Magnification Frontier B.

Theorem 2 (Locality Barrier for HM Frontier B). *The following results hold.*

- (B1^{O}) (Oracle Circuits from Magnification): *For any* $\varepsilon > 0$, $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \in \text{Formula-}O\text{-XOR}[N^{1.01}]$ *for some oracle* O*, where every oracle gate has fan-in at most* N^{ε} *and appears in the layer right above the* XOR *leaves.*
- (B3^{O}) (Extension of Lower-bound Techniques above Magnification Threshold): *For any* $\delta > 0$, InnerProduct *over* N *input bits cannot be computed by* $N^{2-3\delta}$*-size* Formula-O-XOR *circuits with at most* $N^{2-3\delta}$ *oracle gates of fan-in* N^{δ} *in the layer right above the* XOR *leaves, for any oracle* O*.*
- (B4^{O}) (Extension of Lower-bound Techniques below Magnification Threshold): *There is a universal constant* c *such that for all constants* $\varepsilon > 0$ *and* $\alpha > 2$, $\mathsf{MCSP}[n^c, 2^{\varepsilon/\alpha \cdot n}]$ *cannot be computed by oracle formulas* F *with* $\mathsf{SIZE}_3(F) \leq N^{2-\varepsilon}$ *and adaptivity* $o(\log N/\log\log N)$.[11]

[11] That is, on any path from root to a leaf, there are at most $o(\log N/\log\log N)$ oracles.

Here, $\mathrm{Size}_t(F)$ denotes the size of the formula, if we replace every oracle O with fan-in β in F by a formula of size β^t, which reads all its inputs exactly β^{t-1} times (see Section 5.2.2 for the motivation of this definition).

The first two items of Theorem 2 are proved in Section 5.1.2. The third item is proved in Section 5.2.2. While Theorem 2 does not specify that, we actually localize all proofs of the lower bounds from B3 and B4 we are aware of. Interestingly, the localization of B4 allows us to refute the Antichecker Hypothesis from Reference [43] (and a family of potential hardness magnification theorems), cf. Section 5.2.2. We refer to Section 5 for analogous statements describing the locality barrier in frontiers A, C, D, and E.

The localizations of lower bounds A3–E3 and A4–D4 often go through for very clean reasons. For example, localizations of lower bounds A3–C3 based on algebraic methods boil down to the fact that any local oracle (computing any function over a small number of variables) can be simulated by low-degree polynomials over finite fields or the reals. Lower bounds from A4–D4 and D3–E3 based on random restrictions localize, typically, because random restrictions simplify local oracles—e.g., they can reduce the number of inputs of the oracle so that the restricted oracle can be represented by a small DNF or a shallow decision tree. However, the actual proofs of localized lower bounds can become more involved, as we need to rule out the existence of small circuits with many oracles that might depend on each other. In fact, the localization of B4 is technically the most involved contribution of the article. We note also that it seems rather surprising that we are always able to localize lower bounds with the right position of oracles—with exactly the same position as the one from the corresponding magnification theorem or (in the case of lower bounds below the magnification threshold) a potential magnification theorem.

The recent HM Frontier introduced by Reference [12] for $\mathrm{MCSP}[2^{\sqrt{n}}]$ and two-sided error randomised formulas is also subject to the locality barrier based on ideas analogous to the barrier for HM Frontier D (cf. Section 5.1.4). In another recent work, Reference [17] (and Reference [40]) prove magnification for $\mathrm{MCSP}[2^{o(n)}]$ against one-tape Turing machines by constructing highly efficient *uniform* oracle algorithms that make short oracle queries. They also prove a sub-quadratic time lower bound for $\mathrm{MCSP}[2^{(1-o(1))n}]$ against one-tape Turing machines and show that this technique can be localised; it is impossible to extend this technique to $\mathrm{MCSP}[2^{o(n)}]$, with the intention of obtaining magnification (or obtaining magnification from $\mathrm{MCSP}[2^{(1-o(1))n}]$ lower bounds against one-tape Turing machines using similar ideas).

Locality of Computations and Lower-bound Techniques. The fact that many lower-bound techniques extend to computational devices with oracles of small fan-in was observed already by Yao in 1989 on a paper on local computations [61]. According to Yao, a local function is one that can be efficiently computed using only localized processing elements. In our terminology, this corresponds to circuits with oracles of small fan-in. Among other results, Reference [61] argues that Razborov's monotone circuit-size lower bound for k-Clique [48] and Karchmer and Wigderson's monotone formula size lower bound for ST-CONN [35] extend to Boolean devices with monotone oracles of bounded fan-in. Compared to Yao's work, our motivation and perspective are different. While Yao is particularly interested in lower bounds that can be extended in this sense (see, e.g., Sections 2 and 6 in Reference [61]), here we view such extensions as a *limitation* of the corresponding arguments, meaning that they are not refined enough to address the locality barrier.[12]

[12]On a more technical level, we are interested in the regime of barely super-linear size circuits and formulas, and our results do not impose a monotonicity constraint on the oracle.

We note, however, that not every lower-bound technique extends to circuits with small fan-in oracles.[13] For instance, by the work of Allender and Koucký [3] (also a more recent work by Chen and Tell [14]), the parity function Parity_n over n input bits can be computed by a TC^0 circuit of size $O(n)$ (number of wires) containing $\leq n^{1-\varepsilon}$ oracle gates of fan-in $\leq n^{\varepsilon}$, provided that its depth $d = O(1/\varepsilon)$. However, it is known that $\text{Parity}_n \notin \text{TC}^0_d[n^{1+c^{-d}}]$ for a constant $c > 0$ [28] (again, the complexity measure is the number of wires). Since the latter lower bound is super-linear for every choice of d, it follows by the result of References [3, 14] that it cannot be extended to circuits containing a certain number of oracles of fan-in n^{ε}, for a large enough depth d that depends on ε. Incidentally, the hardness magnification theorems of References [3, 14] do not achieve a magnification frontier.

In Section 3.2, we identify one specific lower bound related to HM frontier D, which is both above the magnification threshold and provably non-localizable, cf. Theorem 50. In principle, there might be ways to overcome the locality barrier and match the lower bound with the magnification threshold. We refer to Section 1.4 below for additional discussion.

On Lower Bounds through Reductions. The discussion above has focused on the possibility of *directly* adapting existing lower bounds from Item 3 in HM frontier to establish the desired lower bound in Item 1. There is, however, an *indirect* approach that one might hope to use: *reductions*. For instance, in the context of the HM Frontier B discussed above, can we have a reduction from InnerProduct to $\text{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ that would allow us to show that $\text{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \notin \text{Formula-XOR}[N^{1.01}]$? The first thing to notice is that, for this approach to make sense, the reduction needs to have a specific form so that composing the reduction with a candidate Formula-XOR circuit for $\text{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ violates the hardness of InnerProduct. Is there any hope to design a reduction of this form?

The locality barrier presents a *definitive answer* in this case. Indeed, it is immediate from the first two items of Theorem 2 that such a reduction *does not exist*. For the same reason, it is not possible to use reductions to establish the required lower bounds in some other magnification frontiers, cf. Section 5.1.6. Essentially, under the constraints needed for a reduction to be meaningful, we end up with a class of reductions that produce circuits that are ruled out by the locality barrier.

Locality versus Natural Proofs and Relativization. How does the formal "strength" of the natural proofs and relativization barriers compare to that of the locality barrier? Locality can be seen as a circuit-complexity analog of the relativization barrier. While natural proofs are based on a global condition (cryptographic conjectures), relativization and locality show that some proof techniques "sweep in" additional results that are not true relative to appropriate oracles.

1.4 Concluding Remarks and Open Problems

Hardness magnification shows that obtaining a refined understanding of weak computational models is an approach to major complexity lower bounds, such as separating EXP from NC^1. As discussed in Sections 1.1 and 1.2 above, its different instantiations are connected to a few basic questions in Complexity Theory, including the power of non-monotone operations, learnability of circuit classes, and pseudorandomness.

One of our main conceptual contributions in this work is to identify a challenge when implementing this strategy for lower bounds. Quoting the influential article [51] that introduced the natural proofs barrier,

[13] Of course, any such discussion depends on parameters such as number of oracles and their fan-ins, so whether a technique avoids or not the locality barrier is relative to a particular magnification theorem.

"We do not conclude that researchers should give up on proving serious lower bounds. Quite the contrary, by classifying a large number of techniques that are unable to do the job we hope to focus research in a more fruitful direction."

Razborov and Rudich [51, Section 6]

We share a similar opinion with respect to hardness magnification and the obstruction identified in Section 1.3. While locality provides a unified explanation for the difficulty of adapting combinatorial lower-bound techniques to exploit most (if not all) known magnification frontiers, it might be possible to discover new HM frontiers whose associated lower-bound techniques in Item 3 are sensitive to the presence of small fan-in oracles. For instance, in the case of *uniform* complexity lower bounds, this has been achieved in Reference [42] via an indirect diagonalization that explores the theory of pseudorandomness.[14] Alternatively, it might be possible to establish magnification theorems using a technique that does not produce circuits with small fan-in oracles. Furthermore, recent works suggest approaches such as the *Explicit Obstructions* framework by Reference [12], or the *meta-computational view of PRG constructions* by Reference [25], as potential ways of bypassing the locality barrier. Even if one is pessimistic about these possibilities, we believe that an important contribution of the theory of hardness magnification is to break the divide between "weak" and "strong" circuit classes advocated by the natural proofs barrier, and that it deserves further investigation.

We finish with a couple of technical questions related to our contributions. First, we would like to understand if it is possible to strengthen items (*a*) and (*b*) in Theorem 1 to a wider range of parameters. For example, is hardness magnification for worst-case MCSP$[n^c, 2^{n^\gamma}]$ with $\gamma < 1$ non-naturalizable? The core of this question seems to be the problem of reducing worst-case MCSP from item (*a*) to approximate MCSP from item (*b*).

A related point is that Theorem 1 does not achieve an HM frontier. In fact, super-linear lower bounds for general circuits seem far out of reach at present. It would be thus desirable to obtain an HM frontier that is provably non-naturalizable, or at least a non-naturalizable hardness magnification theorem that would not require a non-naturalizable lower bound.

Another important direction is to show that hardness magnification avoids natural proofs also in the context of *non*-meta-computational problems. Interestingly, many magnification theorems from Reference [43] established for MCSP and variants were subsequently shown to hold for any sparse language in NP [11]. Could it be the case that hardness magnification overcomes natural proofs in a much broader sense?

Finally, it would be useful to investigate the locality of additional lower-bound techniques. Can we, for example, come up with non-localizable lower bounds similar to Theorem 50 that would be above the magnification threshold and work for a problem more closely related to the one from the corresponding HM frontier?

2 PRELIMINARIES

2.1 Notation

Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, $\mathrm{tt}(f)$ denotes the 2^n-bit string representing the truth table of f. However, for any string $y \in \{0,1\}^{2^n}$, define f_y as the function on n inputs such that $\mathrm{tt}(f_y) = y$.

[14] In other words, the magnification theorem discussed in Reference [42] admits a formulation for uniform randomized algorithms, and its proof provides an algorithm with oracle gates of small fan-in in the spirit of the oracle circuits discussed here. Nevertheless, the unconditional lower bound established in the same paper does not extend to algorithms with such oracle gates.

Circuit$[s]$ denotes fan-in two Boolean circuits (over ANDs, ORs, and NOTs) of size at most s, where we count the number of gates. Formula$[s]$ denotes formulas over the basis U_2 (fan-in two ANDs and ORs) of size at most s (counting the number of leaves) with input leaves labelled by literals or constants.

For a circuit class C, $C[s]$ denotes circuits from C of size at most s.

A function $f : \{0,1\}^n \to \{0,1\}$ is γ-approximated by a circuit C, if $\Pr_x[C(x) = f(x)] \geq \gamma$.

2.2 Complexity of Learning

Definition 3 (Learning). A circuit class C is learnable over the uniform distribution by circuits in $\mathcal{D}$ up to error ε with confidence δ if there are randomized oracle $\mathcal{D}$-circuits L^f such that for every Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ computable by a circuit from C, when given oracle access to f, input 1^n and the internal randomness $w \in \{0,1\}^*$, L^f outputs the description of a circuit satisfying

$$\Pr_w\{L^f(1^n, w) \ (1-\varepsilon)\text{-approximates } f\} \geq \delta.$$

L^f uses non-adaptive membership queries if the set of queries that L^f makes to the oracle does not depend on the answers to previous queries. If $\delta = 1$, then we omit mentioning the confidence parameter.

2.3 Natural Properties, MCSP, and Its Variants

Let $\mathcal{F}_n$ be the set of all functions on n variables. $\mathfrak{R} = \{\mathcal{R}_n \subseteq \mathcal{F}_n\}_{n\in\mathbb{N}}$ is a combinatorial property of Boolean functions.

Definition 4 (Natural Property [51]). Let $\mathfrak{R} = \{\mathcal{R}_n\}$ be a combinatorial property, C be a circuit class and Γ be a complexity class. Then, $\mathfrak{R}$ is a Γ-natural property useful against $C[s(n)]$, if there exists an $n_0 \in \mathbb{N}$ such that the following hold:

- **Constructivity**: For any function $f_n \in \mathcal{F}_n$, the predicate $f_n \stackrel{?}{\in} \mathcal{R}_n$ is computable in Γ in the size of the truth table of f_n.
- **Largeness**: For every $n \geq n_0$, $\Pr_{f_n \sim \mathcal{F}_n}\{f_n \in \mathcal{R}_n\} \geq \frac{1}{2^{O(n)}}$.
- **Usefulness**: For every $n \geq n_0$, $\mathcal{R}_n \cap C[s(n)] = \emptyset$.

The following result, which follows from Reference [8], connects the existence of natural properties useful against a class C to designing learning algorithms for C.

Theorem 5 (From Theorem 5.1 of Reference [8] and Lemma 14 of Reference [30]). *Let R be a* P/poly*-natural property useful against* Circuit$[n^d]$ *for some $d \geq 1$. Then, for each $\gamma \in (0,1)$, there are randomized, oracle circuits $\{D_n\}_{n\geq 1} \in$ Circuit$[2^{O(n^\gamma)}]$ that learn* Circuit$[n^k]$ *up to error $\frac{1}{n^k}$ using non-adaptive oracle queries to f_n, where $k = \frac{d\gamma}{a}$ and a is a universal constant that does not depend on d and γ.*

Definition 6 (Gap MCSP). Let $s, t : \mathbb{N} \to \mathbb{N}$, where $s(n) \leq t(n)$ and $0 \leq \varepsilon, \sigma < 1/2$. Define MCSP$[(s,\sigma),(t,\varepsilon)]$ on inputs of length $N = 2^n$, as the following promise problem:

- YES instances: $y \in \{0,1\}^N$ such that there exists a circuit of size $s(n)$ that $(1-\sigma)$-approximates f_y.
- NO instances: $y \in \{0,1\}^N$ such that no circuit of size $t(n)$ $(1-\varepsilon)$-approximates f_y.

We refer to MCSP$[(s,0),(t,0)]$ as MCSP$[s,t]$. Informally speaking, if $\varepsilon > 0$, we say that MCSP$[(s,0),(t,\varepsilon)]$ is an *approximate* version of MCSP. Otherwise, it is a *worst-case* version of MCSP.

Remark 7. In Definition 6, if $s(n) = t(n)$, we also require that $\sigma < \varepsilon$ for the yes and no instances to be disjoint.

Definition 8 (Succinct MCSP*).* For functions $s, t : \mathbb{N} \mapsto \mathbb{N}$, Succinct-MCSP$[s(n), t(n)]$ is the following problem. Given an input $\langle 1^n, 1^s, (x_1, b_1), \ldots, (x_t, b_t)\rangle$ where $x_i \in \{0,1\}^n, b_i \in \{0,1\}$, decide if there is a circuit C of size s such that $C(x_i) = b_i$ for all $i = 1, \ldots, t$.

2.4 Pseudorandom Generators

Definition 9 (Pseudorandom Function Families). For any circuit class $\mathcal{C}$, size functions $s(n), t(n) \geq n$, family $\mathcal{G}_n$ of n-bit Boolean functions and distribution $\mathcal{D}_n$ over $\mathcal{G}_n$, we say that a pair $(\mathcal{G}_n, \mathcal{D}_n)$ is a $(t(n), \varepsilon(n))$-pseudorandom function family (PRF) in $\mathcal{C}[s(n)]$, if each function in $\mathcal{G}_n$ is in $\mathcal{C}[s(n)]$ and for every randomized circuit $A^O \in \text{Circuit}^O[t(n)]$, where O denotes oracle access to a fixed Boolean function over n inputs, we have

$$\left| \Pr_{g \sim \mathcal{D}_n, w} \{A^g(w) = 1\} - \Pr_{f \sim \mathcal{F}_n, w} \{A^f(w) = 1\} \right| \leq \varepsilon(n),$$

where w represents the internal randomness of A^O.

Reference [45] states an equivalence between the non-existence of PRFs in a circuit class $\mathcal{C}$ and learning algorithms for $\mathcal{C}$. In particular, we care about the following direction, which they prove using a small-support version of Von-Neumann's Min-max Theorem.

Theorem 10 (No PRFs in $\mathcal{C}$ Implies Learning Algorithm for $\mathcal{C}$ [45]). *Let $t(n) \leq 2^{O(n)}$. Suppose that for every $k \geq 1$ and large enough n, there exists no* $(\text{poly}(t(n)), 1/10)$*-pseudorandom function families in $\mathcal{C}[n^k]$. Then, for every $\varepsilon > 0, k \geq 1$ and large enough n, there is a randomized oracle circuit in* $\text{Circuit}^O[2^{n^\varepsilon}]$ *that learns every function $f_n \in \mathcal{C}[n^k]$ up to error $1/n^k$ with confidence $1 - 1/n$, where O denotes membership query access to f_n.*

2.5 Local Circuit Classes

Our definition of local computation is somewhat similar to some definitions appearing in Reference [61].

Definition 11 (Local Circuit Classes). Let $\mathcal{C}$ be a circuit class (such as $\text{AC}^0[s]$, $\text{TC}^0_d[s]$, $\text{Circuit}[s]$, etc). For functions $q, \ell, a \colon \mathbb{N} \to \mathbb{N}$, we say that a language L is in $[q, \ell, a]\text{-}\mathcal{C}$ if there exists a sequence $\{E_n\}$ of oracle circuits for which the following holds:

(i) Each oracle circuit E_n is a circuit from $\mathcal{C}$.
(ii) There are at most $q(n)$ oracle gates in E_n, each of fan-in at most $\ell(n)$, and any path from an input gate to an output gate encounters at most $a(n)$ oracle gates.
(iii) There exists a language $O \subseteq \{0,1\}^*$ such that the sequence $\{E_n^O\}$ (E_n with its oracle gates set to O) computes L.

In the definition above, q stands for *quantity*, ℓ for *locality*, and a for *adaptivity* of the corresponding oracle gates.

2.6 Random Restrictions

Let $\rho : [N] \to \{0, 1, *\}$ be a *restriction*, and $\boldsymbol{\rho}$ be a *random restriction*, i.e., a distribution of restrictions. We say that $\boldsymbol{\rho}$ is p-regular if $\Pr[\boldsymbol{\rho}(i) = *] = p$ and $\Pr[\boldsymbol{\rho}(i) = 0] = \Pr[\boldsymbol{\rho}(i) = 1] = (1-p)/2$ for every $i \in [N]$. We also say $\boldsymbol{\rho}$ is k-wise independent if any k coordinates of $\boldsymbol{\rho}$ are independent. For a function $f : \{0,1\}^N \to \{0,1\}$, we use $f\restriction_\rho$ to denote the function $\{0,1\}^{|\rho^{-1}(*)|} \to \{0,1\}$ obtained by restricting f according to ρ in the natural way.

We need the following lemma stating that one can sample from a k-wise independent random restriction with a short seed, and moreover all restrictions have a small circuit description.

Lemma 12 ([27, 60]). *There exists a q-regular k-wise independent random restriction $\boldsymbol{\rho}$ distributed over $\rho : [N] \to \{0, 1, *\}$ samplable with $O(k \cdot \log(N) \log(1/q))$ bits. Furthermore, each output coordinate of the random restriction can be computed in time polynomial in the number of random bits.*

2.7 Technical Results

Lemma 13 (Hoeffding's Inequality). *Let $X_1, \ldots, X_n$ be independent random variables such that $0 \le X_i \le 1$ for every $i \in [n]$. Let $X = \sum_{i=1}^{n} X_i$. Then, for any $\varepsilon > 0$, we have*

$$\Pr\{|X - \boldsymbol{E}X| \ge \varepsilon n\} \le 2\exp(-2\varepsilon^2 n).$$

3 MAGNIFICATION FRONTIERS

3.1 EXP $\not\subseteq$ NC^1 and AC^0-XOR Lower Bounds for MKtP

In this section, we present the proofs of the new results stated in HM Frontier A. Recall that $\mathrm{Kt}(x)$ is defined as the minimum over $|M| + \log t$ such that a program M (simulated by a universal Turing machine) outputs x in t steps. For thresholds $\theta, \theta' : \mathbb{N} \to \mathbb{N}$, we denote by $\mathrm{MKtP}[\theta(N), \theta'(N)]$ the promise problem whose YES instances consist of the strings $x \in \{0,1\}^N$ such that $\mathrm{Kt}(x) \le \theta(N)$ and NO instances consist of the strings such that $\mathrm{Kt}(x) > \theta'(N)$.

We start with the hardness magnification theorem of HM Frontier A1.

Theorem 14. *There exists a constant c such that, for every large enough constant $d > 1$,*

$$\mathrm{MKtP}[(\log N)^d, (\log N)^d + c \log N] \notin \mathrm{AC}^0\text{-XOR}[N^{1.01}] \quad \textit{implies} \quad \mathrm{EXP} \not\subseteq \mathrm{NC}^1.$$

Proof. We prove the contrapositive. Assume that EXP $\subseteq$ NC^1. First, recall that any N-bit-input polynomial-size NC^1 circuit can be converted into a depth-d' AC^0 circuit of size $2^{N^{O(1/d')}}$ for every positive integer constant d' (see, e.g., Reference [2, Lemma 8.1]).

Oliveira, Pich, and Santhanam [43] showed that there exists a problem $L \in$ EXP such that $\mathrm{MKtP}[\theta(N), \theta(N) + c \log N] \in \mathrm{AND}_{O(N)}\text{-}L_{O(\theta(N))}\text{-XOR}$ for $\theta(N) \ge \log N$. (Here the subscript denotes the fan-in of a gate.) That is, the promise problem $\mathrm{MKtP}[\theta(N), \theta(N) + c \log N]$ can be computed by the following form of an L-oracle circuit: The output gate is an AND gate of fan-in $O(N)$, at the middle layer are L-oracle gates of fan-in $O(\theta(N))$, and at the bottom layer are XOR gates. Under the assumption that EXP $\subseteq$ NC^1, we can replace L-oracle circuits with depth-d' AC^0 circuits of size $2^{(\log N)^{O(d/d')}}$, which is smaller than $N^{0.01}$ by choosing a constant d' large enough. In particular, we obtain a depth-$(d' + O(1))$ almost linear size AC^0 circuit with bottom XOR gates that computes $\mathrm{MKtP}[\theta(N), \theta(N) + c \log N]$. □

The rest of this section is devoted to proving the following AC^0 lower bound for MKtP, which establishes HM Frontier A4.

Theorem 15. *For any $d = d(N)$, for some $\theta(N) = d \cdot \widetilde{O}(\log N)^3$ and any $\theta'(N) = N/\omega(\log N)^d$, it holds that $\mathrm{MKtP}[\theta(N), \theta'(N)] \notin \mathrm{AC}^0_d$.*

Note that Theorem 15 is only meaningful if $d = o(\log N / \log\log N)$, because otherwise the promise problem is not well-defined.

The idea of the proof is as follows: Trevisan and Xue [59] showed that there exists a pseudorandom restriction ρ of seed length polylog(N) that shrinks every polynomial-size depth-2 circuit into shallow decision trees. Moreover, the expected fraction of unrestricted variables $\rho^{-1}(*)$ is at least $p = \Omega(1/\log N)$. In particular, by composing d independent pseudorandom restrictions

$\rho_1, \ldots, \rho_d$, every depth-d circuit can be turned into a constant function, while still leaving at least p^d-fraction of inputs unrestricted. The seed length required to sample d independent pseudorandom restrictions is at most $d \times \mathrm{polylog}(N)$, and thus $\mathsf{Kt}(0^N \circ \rho) \leq \mathrm{polylog}(N)$. Here, $\rho_1 \circ \rho_2$ denotes ρ_2 extended by ρ_1. We stress that the exponent of the seed length does not depend on d. Since the circuit hit with the pseudorandom restriction becomes a constant function, it cannot distinguish $0^N \circ \rho$ with $U_N \circ \rho$, i.e., the distribution where the unrestricted variables of ρ are replaced with the uniform distribution U_N. Assuming that there remain sufficiently many unrestricted inputs (e.g., $N/O(\log N)^d \gg \mathrm{polylog}(N)$), the latter distribution has a large Kt complexity, which is a contradiction to the fact that an AC^0_d-circuit computes a gap version of MKtP.

We note that Cheraghchi, Kabanets, Lu, and Myrisiotis [18] used the pseudorandom restriction method to obtain an exponential-size AC^0 lower bound. A crucial difference in this work is that instead of optimizing the size of AC^0 circuits, we aim at minimizing the threshold θ of $\mathsf{MKtP}[\theta]$.

Following Reference [59], to generate a random restriction $\rho \in \{0,1,*\}^N$ that leaves a variable unrestricted with probability 2^{-q}, we regard a binary string $w \in \{0,1\}^{(q+1)N}$ as a random restriction ρ_w. Specifically:

Definition 16. For a string $w \in \{0,1\}^{(q+1)N}$, we define a restriction $\rho_w \in \{0,1,*\}^N$ as follows: Write w as $(w_1, b_1) \cdots (w_N, b_N)$, where $w_i \in \{0,1\}^q$ and $b_i \in \{0,1\}$. For each $i \in [N]$, if $w_i = 1^q$ then set $\rho_w(i) := *$; otherwise, set $\rho_w(i) := b_i$.

Note that this is defined so that $\Pr_w[\rho_w(i) = *] = 2^{-q}$ for every $i \in [N]$, when w is distributed uniformly at random.

Trevisan and Xue [59] showed that Håstad's switching lemma can be derandomized by using a distribution that fools CNFs. To state this formally, we need the following definitions. Define a t-width CNF as one that has at most t literals in each clause. We say that a distribution $\mathcal{D}$ over $\{0,1\}^n$ ε-fools a set of functions $\mathcal{S}_n$ over n variables if for every $f \in \mathcal{S}_n$, $|\Pr_{x\sim D}\{f(x) = 1\} - \Pr_{x\sim U_n}\{f(x) = 1\}| \leq \varepsilon$. Finally, define $\mathsf{DT}(f)$ as the depth of the smallest decision tree computing f.

Lemma 17 (Derandomized Switching Lemma [59, Lemma 7]). *Let φ be a t-width M-clause CNF formula over N inputs. Let $p = 2^{-q}$ for some $q \in \mathbb{N}$. Assume that a distribution $\mathcal{D}$ over $\{0,1\}^{(q+1)N}$ ε_0-fools $M \cdot 2^{t(q+1)}$-clause CNFs. Then,*

$$\Pr_{w\sim\mathcal{D}}[\mathsf{DT}(\varphi\restriction_{\rho_w}) > s] \leq 2^{s+t+1}(5pt)^s + \varepsilon_0 \cdot 2^{(s+1)(2t+\log M)}.$$

Theorem 18 (Based on References [59] and [58, Theorem 56]). *Let $s, M, d, N \in \mathbb{N}$ be positive integers. Let $p = 2^{-q}$ for some $q \in \mathbb{N}$ so that $1/128s \leq p < 1/64s$. Assume that there is a pseudorandom generator $G\colon \{0,1\}^r \to \{0,1\}^{(q+1)N}$ that ε_0-fools CNFs of size $M \cdot 2^s \cdot 2^{s(q+1)}$. Then, there exists a distribution $\mathcal{R}$ of random restrictions that satisfies the following:*

(1) *For every circuit C of size M and depth d over N inputs,*

$$\Pr_{\rho\sim\mathcal{R}}[\mathsf{DT}(C\restriction_\rho) > s] \leq M \cdot \left(2^{-s+1} + \varepsilon_0 \cdot 2^{(s+1)(3s+\log M)}\right).$$

(2) *For any parameter $\delta < 1$, with probability at least $1 - N(\delta + d\varepsilon_0)$, the number of unrestricted variables in $[N]$ is at least $\lfloor N \cdot p^{d-1}/64 \log(1/\delta) \rfloor$.*
(3) *$\mathcal{R}$ can be generated by a seed of length dr in polynomial time.*

Proof. We apply the derandomized switching lemma (Lemma 17) d times. In the first iteration, we set $p := 1/64$ (and $q := 6$) and generate $\rho_{G(z)[1,\ldots,(6+1)N]}$. (Here, we use the first $(6+1)N$ bits of $G(z)$ to generate $\rho_{G(z)}$.) This turns a circuit C of size M into a circuit whose bottom fan-in is at most s. For every other iteration i (where $i = 2, \ldots, d$), we set $p := 2^{-q}$ and turn a circuit C of depth

$d-i+2$ into a circuit of depth $d-i+1$. Our final pseudorandom restriction $\rho \sim \mathcal{R}$ is defined by the composition of the d independent pseudorandom restrictions $\rho_{G(z_1)[1,\ldots,(6+1)N]}, \rho_{G(z_2)}, \ldots, \rho_{G(z_d)}$.

Our proof is essentially the same with [58], except that (1) we apply the switching lemma d times (instead of $d-1$) to turn depth-d circuits into shallow decision trees, and (2) in References [58, 59], for the application of constructing a pseudorandom generator for AC^0, fixed bits of pseudorandom restrictions must be generated by using truly random bits, whereas in our case, we generate all the bits by using G.

In more detail, for each $i \in [d]$, let M_i be the number of the gates at level i in C (i.e., the gates whose distance from the input gates is i). At the first iteration, we set $p := 1/64 = 2^{-6}$ and $q := 6$. We then generate $\rho^1 := \rho_{G(z_1)[1,\ldots,(6+1)N]}$ by choosing a seed $z_1 \sim \{0,1\}^r$ uniformly at random. We regard C as a depth-$(d+1)$ circuit of bottom fan-in 1, and apply Lemma 17 to each gate at level 1 (in the original circuit C). The probability that there exists a gate at level 1 in $C\restriction_{\rho^1}$ that cannot be computed by a decision tree of depth s is bounded above by

$$M_1 \cdot \left(2^{s+1+1}(5/64)^s + \varepsilon_0 \cdot 2^{(s+1)(2+\log M)}\right).$$

In the complement event (i.e., if each gate at level 1 in $C\restriction_{\rho^1}$ can be computed by a decision tree of depth s), each gate at level 1 can be written as DNFs and CNFs of width s and, hence, can be merged into some gate at level 2. Thus, a circuit $C\restriction_{\rho^1}$ can be turned into a circuit of depth d and bottom fan-in s. Moreover, the number of gates at level 1 is bounded by $M \cdot 2^s$, which is an invariant preserved during the iterations.

For every other iteration i ($i = 2, \ldots, d$), we generate $\rho^i := \rho_{G(z_i)}$ by choosing a seed $z_i \sim \{0,1\}^r$ uniformly at random. Using the invariant that the number of gates at level $i-1$ is at most $M \cdot 2^s$, the probability that some gate at level i in $C\restriction_{\rho^1 \cdots \rho^i}$ cannot be computed by a decision tree of depth s is bounded above by

$$M_i \cdot \left(2^{s+s+1}(5ps)^s + \varepsilon_0 \cdot 2^{(s+1)(2s+\log(M2^s))}\right).$$

In the complement event, every gate at level i can be written as width-s CNFs or DNFs of size 2^s and, hence, can be merged into some gate at level $i+1$ (for $i < d$). At the last iteration (i.e., $i = d$), the circuit $C\restriction_{\rho^1 \cdots \rho^d}$ can be written as a decision tree of depth s. We define the pseudorandom restriction ρ as $\rho^d \circ \cdots \circ \rho^1$. Item 3 is obvious from this construction.

Overall, the probability that $\mathsf{DT}(C\restriction_\rho) > s$ is at most $M \cdot (2^{-s+1} + \varepsilon_0 \cdot 2^{(s+1)(3s+\log M)})$. This completes the proof of Item 1.

To see Item 2, we divide N input bits into k disjoint blocks $T_1, \ldots, T_k$ of size at least t (and, hence, $k = \lfloor N/t \rfloor$), where t is a parameter chosen later. We claim that each block must contain at least one unrestricted variable in $\rho \sim \mathcal{R}$ with high probability (and, hence, $|\rho^{-1}(*)| \geq \lfloor N/t \rfloor$). Fix any block $T = T_i$ for some $i \in [k]$. As in Reference [58], one can easily observe that the condition that every variable in T is restricted can be checked by a CNF of size at most $|T|$ ($\leq N$). By a simple hybrid argument, the concatenation of d independent pseudorandom distributions $G(z_1), \ldots, G(z_d)$ $d\varepsilon_0$-fools CNFs (cf. Reference [58, Corollary 55]). Therefore, the probability that every variable in T is restricted by $\rho \sim \mathcal{R}$ is bounded by $(1 - p^{d-1}/64)^t + d\varepsilon_0$, where the first term is an upper bound for the probability that every variable in T is restricted by a truly random restriction. Choosing $t = 64 \log(1/\delta)/p^{d-1}$ and using a union bound, the probability that some block T_i is completely fixed can be bounded above by $\lfloor N/t \rfloor \cdot (\delta + d\varepsilon_0)$, which completes the proof of Item 2. □

Corollary 19. *For every circuit C of size M ($\geq N$) and depth d over N inputs, there exists a restriction ρ such that*

(1) *$C\restriction_\rho$ is a decision tree of depth at most $s := 2 \log 8M$,*
(2) *$|\rho^{-1}(*)| \geq N/O(\log M)^d$, and*
(3) *$\mathsf{Kt}(\rho) \leq d \cdot \widetilde{O}((\log M)^3)$.*

Proof. Tal [58, Theorem 52] showed that there exists a polynomial-time pseudorandom generator G of seed length $r := \widetilde{O}(\log M_0 \cdot \log(M_0/\varepsilon_0))$ that ε_0-fools CNFs of size M_0. We set $M_0 := M \cdot 2^s \cdot 2^{s(q+1)}$, $s := 2\log 8M$, and $\varepsilon_0 := 2^{-9s^2}$. Then the seed length r of G is at most $r = \widetilde{O}(\log M_0 \cdot \log(M_0/\varepsilon_0)) = \widetilde{O}(\log M \cdot (\log M)^2)$. Applying Theorem 18, the probability that $\mathsf{DT}(C\restriction_\rho) > s$ is bounded by $\frac{1}{2}$. Choosing $\delta = 1/8N$, we also have that the probability that $|\rho^{-1}(*)| < \lfloor N \cdot p^{d-1}/64\log(1/\delta)\rfloor$ is at most $\frac{1}{4}$. Thus, there exists some restriction ρ in the support of $\mathcal{R}$ such that $\mathsf{DT}(C\restriction_\rho) \le s$ and $|\rho^{-1}(*)| \ge \Omega(N \cdot p^{d-1}/\log N) \ge N/O(\log M)^d$. □

Using the assumption that a circuit computes MKtP, we show that shallow decision trees must be a constant function.

Lemma 20. *Let C be a circuit and ρ be a restriction such that $C\restriction_\rho$ is a decision tree of depth s. If $\mathsf{MKtP}[O(s\log N) + \mathsf{Kt}(\rho)] \subseteq C^{-1}(1)$, then $C\restriction_\rho \equiv 1$.*

Proof. We prove the contrapositive. Assume that $C\restriction_\rho \not\equiv 1$, which means that there is a path $\pi\colon [N] \to \{0,1,*\}$ of a decision tree $C\restriction_\rho$ that assigns at most s variables so that $C\restriction_{\rho\pi} \equiv 0$. Note that $\mathsf{Kt}(\pi) \le O(s\log N)$, because one can specify each restricted variable of π by using $O(\log N)$ bits. Thus, we have $\mathsf{Kt}(0^N \circ \pi \circ \rho) \le O(s\log N) + \mathsf{Kt}(\rho)$. However, $C(0^N \circ \pi \circ \rho) = C\restriction_{\rho\pi}(0^N) = 0$. Therefore, we obtain $\mathsf{MKtP}[O(s\log N) + \mathsf{Kt}(\rho)] \not\subseteq C^{-1}(1)$. □

Now, we are ready to prove the main result of this section.

Proof of Theorem 15. Assume, by way of contradiction, that there is a circuit C of size $M := N^{O(1)}$ and depth d that computes $\mathsf{MKtP}[d\cdot\widetilde{O}(\log N)^3, N/\omega(\log N)^d]$. Using Corollary 19, we take a restriction ρ such that $C\restriction_\rho$ is a decision tree of depth $s = O(\log N)$. By Lemma 20, we have $C\restriction_\rho \equiv 1$, under the assumption that $O(s\log N) + \mathsf{Kt}(\rho) \le \theta(N)$, which is satisfied by choosing $\theta(N)$ large enough. Now, by counting the number of inputs accepted by $C\restriction_\rho$, we obtain

$$2^{N/O(\log N)^d} \le 2^{|\rho^{-1}(*)|} = |(C\restriction_\rho)^{-1}(1)| \le 2^{\theta'(N)+1},$$

where, in the last inequality, we used the fact that the number of strings whose Kt complexity is at most $\theta'(N)$ is at most $2^{\theta'(N)+1}$. However, the inequality contradicts the choice of $\theta'(N)$. □

3.2 NQP $\not\subseteq$ NC^1 and Formula-XOR or GapAND-Formula for MCSP

This section is devoted to proving HM Frontier B1 and HM Frontier D1. In fact, we provide two different proofs of HM Frontier B1, one based on Reference [46], another one based on Reference [11].

In both proofs, the hardness magnification is achieved by constructing an oracle circuit for MCSP. The most interesting part of the first proof is that it gives a *conditional* construction assuming QP $\subseteq$ P/poly. While the oracle circuit construction can be made *unconditional* (as in the second proof), it illustrates a potentially more applicable approach: proving the hardness magnification theorem while assuming the target circuit lower bound is false (i.e., NQP $\subseteq$ NC^1).

3.2.1 Reduction-based Approach from Reference [46]. In the initial magnification theorem [46, Theorem 1], approximate MCSP was shown to admit hardness magnification phenomena. Here, we present a similar hardness magnification theorem for a worst-case version of MCSP.

A natural way of reducing worst-case MCSP to approximate MCSP is to apply error-correcting codes. Error-correcting codes map a hard Boolean function to a Boolean function that is hard on average. A problem with this approach is that error-correcting codes do not guarantee that an easy Boolean function will be mapped to an easy Boolean function. Our main idea is to enforce

the latter property with an extra assumption QP $\subseteq$ P/poly. Here, QP denotes $\mathsf{TIME}[n^{\log^{O(1)} n}]$. Similarly, NQP will stand for $\mathsf{NTIME}[n^{\log^{O(1)} n}]$.

We will use the following explicit error-correcting code.

Theorem 21 (Explicit Linear Error-correcting Codes [33, 53]). *There exists a sequence* $\{E_N\}_{N\in\mathbb{N}}$ *of error-correcting codes* $E_N : \{0,1\}^N \to \{0,1\}^{M(N)}$ *with the following properties:*

- $E_N(x)$ *can be computed by a uniform deterministic algorithm running in time* $\mathrm{poly}(N)$.
- $M(N) = b \cdot N$ *for a fixed* $b \geq 1$.
- *There exists a constant* $\delta > 0$ *such that any codeword* $E_N(x) \in \{0,1\}^{M(N)}$ *that is corrupted on at most a* δ*-fraction of coordinates can be uniquely decoded to* x *by a uniform deterministic algorithm* D *running in time* $\mathrm{poly}(M(N))$.
- *Each output bit is computed by a parity function: for each input length* $N \geq 1$ *and for each coordinate* $i \in [M(N)]$*, there exists a set* $S_{N,i} \subseteq [N]$ *such that for every* $x \in \{0,1\}^N$*,*

$$E_N(x)_i = \bigoplus_{j\in S_{N,i}} x_j.$$

Under the assumption that QP $\subseteq$ P/poly, we present an efficient reduction from worst-case MCSP to approximate MCSP: Given the truth table of a function f, we simply map it to $E_N(\mathsf{tt}(f))$. The following lemma establishes the correctness of this reduction.

Lemma 22 (Reducing Worst-case MCSP to Approximate MCSP). *Assume* QP $\subseteq$ P/poly. *Then the error-correcting code* E_N *from Theorem 21 satisfies the following:*

(1) $f_n \in \mathsf{Circuit}[2^{n^{1/3}}] \Rightarrow E_N(\mathsf{tt}(f_n)) \in \mathsf{Circuit}[2^{\sqrt{m}}]$,[15]
(2) $f_n \notin \mathsf{Circuit}[2^{n^{2/3}}] \Rightarrow E_N(\mathsf{tt}(f_n))$ *is hard to* $(1-\delta)$*-approximate by* $2^{\sqrt{m}}$*-size circuits,*

where $m = \Theta(n)$.

Proof. For the first implication, we consider the map

$$C, i \mapsto E_N(\mathsf{tt}(C))_i,$$

where C is a circuit with n inputs and size $2^{n^{1/3}}$, $i \in \{0,1\}^m$, and $m = \log |E_N|$. The map takes an input of length $2^{O(n^{1/3})}$, and is computable in time $2^{O(n)}$; hence, the map is in QP $\subseteq$ P/poly. Thus, there exists a circuit F of size $2^{O(n^{1/3})}$ that, taking the description of a circuit C of size $2^{n^{1/3}}$ and $i \in \{0,1\}^m$ as input, outputs the ith bit of $E_N(\mathsf{tt}(C))$. Therefore, if f_n is computed by a circuit C of size $2^{n^{1/3}}$, then the function $i \mapsto E_N(\mathsf{tt}(f_n))_i$ is computable by a circuit $F(C, \text{-})$ of size $2^{O(n^{1/3})} < 2^{\sqrt{m}}$.

The second implication is obtained in a similar way by considering the map

$$C, i \mapsto D_N(\mathsf{tt}(C))_i,$$

where C is a circuit with $m = \log |E_N|$ inputs and size $2^{\sqrt{m}}$, $i \in \{0,1\}^n$ and D_N is an efficient decoder of E_N. The new map is computable in time $2^{O(m)}$ and again is in QP $\subseteq$ P/poly. Therefore, if $E_N(\mathsf{tt}(f_n))$ is $(1-\delta)$-approximated by a circuit C of size $2^{\sqrt{m}}$, f_n is computable by a circuit of size $2^{O(\sqrt{m})} < 2^{n^{2/3}}$. □

Since the error-correcting code of Theorem 21 can be computed by using one layer of XOR gates, we obtain the following corollary.

Corollary 23. *If* QP $\subseteq$ P/poly, *then* $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ *is reducible to* $\mathsf{MCSP}[(2^{\sqrt{n}}, 0), (2^{\sqrt{n}}, \delta)]$ *by using a many-one reduction computed by a linear-size circuit of* XOR *gates.*

[15] Here, we identify $E_N(\mathsf{tt}(f_n))$ with the function whose truth table is $E_N(\mathsf{tt}(f_n))$.

We are ready to prove the main result of this section:

Theorem 24 (Magnification for Worst-case MCSP via Error-correcting Codes). *Assume that* $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \notin \text{Formula-XOR}[N^{1+\varepsilon}]$ *for some constant* $\varepsilon > 0$. *Then either* $\mathsf{QP} \not\subseteq \mathsf{P/poly}$ *or* $\mathsf{NP} \not\subseteq \mathsf{NC}^1$.

Proof. We prove the contrapositive. Assume that $\mathsf{QP} \subseteq \mathsf{P/poly}$ and $\mathsf{NP} \subseteq \mathsf{NC}^1$. Reference [46, Lemma 16] shows that $\mathsf{NP} \subseteq \mathsf{NC}^1$ implies $\mathsf{MCSP}[(2^{\sqrt{n}}, 0), (2^{\sqrt{n}}, \delta)] \in \text{Formula}[N^{1+\varepsilon}]$ for any constant $\varepsilon > 0$. By combining this with Corollary 23, we obtain that $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \in \text{Formula-XOR}[O(N^{1+\varepsilon})]$. □

3.2.2 Kernelization-based Approach from Reference [11]. Now, we give another proof of HM Frontier B1 by adapting techniques from Reference [11]. In fact, the following proof implies (under a straightforward adjustment of parameters) both HM Frontier B1 and HM Frontier D1.

Theorem 25 (Magnification for Worst-case MCSP via Kernelization for GapAND-Formula-XOR). *Assume that* $\mathsf{MCSP}[2^{n^{1/3}}] \notin \text{GapAND}_{O(N)}\text{-Formula-XOR}[N^{\varepsilon}]$ *for some constant* $\varepsilon > 0$. *Then* $\mathsf{NQP} \not\subseteq \mathsf{NC}^1$.

Proof Sketch. The following proof is just an adaption of Theorem 3.4 of Reference [11].

Let $N = 2^n$ and $s = 2^{n^{1/3}} = 2^{(\log N)^{1/3}}$. Let $S = \mathsf{MCSP}[2^{n^{1/3}}]^{-1}(1)$ (that is, all yes instances of $\mathsf{MCSP}[2^{n^{1/3}}]$ on inputs of length N), and $m = |S|$. We have that $m \leq s^{O(s)}$. Let E_N be the error correcting code from Theorem 21. Recall that E_N maps from $\{0,1\}^N$ to $\{0,1\}^{b \cdot N}$ for a constant b.

Let $T = c_1 \cdot \log m$ for a large enough constant c_1. Suppose we pick T random indexes $I = (i_1, i_2, \ldots, i_T)$ from $[b \cdot N]$ independently and uniformly at random. Given $x \in \{0,1\}^N$, let $H_I(x) := (E_N(x)_{i_1}, E_N(x)_{i_2}, \ldots, E_N(x)_{i_T})$.

By a Chernoff bound and a union bound, we can see that with high probability over random choices of I, all inputs from S are mapped into distinct strings in $\{0,1\}^T$ by H_I. We fix such a good collection of indexes I_{good}.

Now, consider the following language:

$$L_{\mathsf{check}} : [b \cdot N]^T \times \{0,1\}^T \times [b \cdot N] \times \{0,1\} \to \{0,1\},$$

which takes as inputs I (hash function coordinates), w (hash value), i (index), and z (check-bit). $L_{\mathsf{check}}(I, w, i, z)$ guesses an input $y \in \{0,1\}^N$, and accepts if $H_I(y) = w$, $\mathsf{MCSP}[2^{n^{1/3}}](y) = 1$, and $E_N(y)_i = z$. It is easy to see that L_{check} is in NQP.

Given $x \in \{0,1\}^N$, we claim that $\mathsf{MCSP}[2^{n^{1/3}}](x) = 1$ iff $L_{\mathsf{check}}(I_{\text{good}}, H_{I_{\text{good}}}(x), i, E_N(x)_i) = 1$ for all $i \in [b \cdot N]$.

(1) When $\mathsf{MCSP}[2^{n^{1/3}}](x) = 1$, on the particular guess $y = x$, $L_{\mathsf{check}}(I_{\text{good}}, H_{I_{\text{good}}}(x), i, E_N(x)_i)$ accepts for all $i \in [b \cdot N]$.

(2) When $\mathsf{MCSP}[2^{n^{1/3}}](x) = 0$, we set $z = H_{I_{\text{good}}}(x)$. By our choice of I_{good}, there is at most one x' satisfying both $\mathsf{MCSP}[2^{n^{1/3}}](x') = 1$ and $H_{I_{\text{good}}}(x') = z$. If there is no such x', then all $L_{\mathsf{check}}(I_{\text{good}}, H_{I_{\text{good}}}(x), i, x_i)$ reject. Otherwise, we have $x \neq x'$. Let i be an index such that $E_N(x)_i \neq E_N(x')_i$. Then $L_{\mathsf{check}}(I_{\text{good}}, H_{I_{\text{good}}}(x), i, E_N(x)_i)$ rejects.

Moreover, in the second case, $L_{\mathsf{check}}(I_{\text{good}}, H_{I_{\text{good}}}(x), i, E_N(x)_i)$ indeed rejects at least for a constant fraction of $i \in [b \cdot N]$, since $E_N(x)$ is an error correcting code,

Now suppose $\mathsf{NQP} \subseteq \mathsf{NC}^1$ for the sake of contradiction. Since $H_{I_{\text{good}}}(x)$ can be computed by $T = N^{o(1)}$ many XOR gates (I_{good} is hardwired into the circuit), we can construct $b \cdot N$

Formula-XOR$[N^{o(1)}]$ circuits $C_1, C_2, \ldots, C_{b\cdot N}$, such that if MCSP$[2^{n^{1/3}}](x) = 1$ then $C_i(x) = 1$ for all x, and otherwise $C_i(x) = 0$ for a constant fraction of i's.

By a simple error reduction via random sampling, we construct $m = O(N)$ Formula-XOR$[N^{o(1)}]$ circuits $D_1, D_2, \ldots, D_m$, such that if MCSP$[2^{n^{1/3}}](x) = 1$ then $D_i(x) = 1$ for all x, and otherwise $D_i(x) = 0$ for at least a 0.9 fraction of inputs. Hence, we have MCSP$[2^{n^{1/3}}] \in$ GapAND$_{O(N)}$-Formula-XOR$[N^{o(1)}]$, a contradiction to the assumption. □

Remark 26. We note that GapAND$_{O(N)}$-Formula-XOR$[N^{\varepsilon}]$ circuits are a special case of both Formula-XOR$[N^{1+\varepsilon}]$ circuits and GapAND$_{O(N)}$-Formula$[N^{2+\varepsilon}]$ circuits. Therefore, the above proof implies both HM Frontier B1 and HM Frontier D1.

3.3 NP $\not\subseteq$ NC^1 and Almost-formula Lower Bounds for MCSP

Recall that near-quadratic *formula* lower bounds are known for MCSP$[2^{n^{o(1)}}, 2^{n^{o(1)}}]$. However, a hardness magnification obtained by a super efficient construction of anticheckers established in Reference [43] states that NP $\subseteq$ P/poly implies almost linear-size circuits for a worst-case version of parameterized MCSP$[2^{n^{o(1)}}, 2^{n^{o(1)}}]$. Consequently, if we could make the hardness magnification work for formulas, then NP $\not\subseteq$ NC^1 would follow. We make a step in this direction by showing that NP $\subseteq$ NC^1 implies the existence of almost-formulas of almost linear size solving the worst-case MCSP$[2^{n^{o(1)}}, 2^{n^{o(1)}}]$, cf. Theorem 29. This is established by a more detailed analysis of the proof from Reference [43] extended with an application of the Valiant-Vazirani Isolation Lemma (cf. Reference [6, Lemma 17.19]) in the process of selecting anticheckers. We also observe that almost-formulas of sub-quadratic size cannot solve PARITY, cf. Theorem 30. These results yield HM Frontier C1 and HM Frontier C3.

We start the presentation with a lemma needed to derive HM Frontier C1.

Lemma 27 (Anticheckers). *Assume* NP $\subseteq$ NC^1. *Then for any* $\lambda \in (0,1)$ *there are circuits* $\{C_{2^n}\}_{n=1}^{\infty}$ *of size* $2^{n+O(n^\lambda)}$, *which given* $\mathrm{tt}(f) \in \{0,1\}^N$, *output* $2^{O(n^\lambda)}$ n*-bit strings* $y_1, \ldots, y_{2^{O(n^\lambda)}}$ *together with bits* $f(y_1), \ldots, f(y_{2^{O(n^\lambda)}})$ *forming a set of anticheckers for* f, *i.e., if* f *is hard for circuits of size* 2^{n^λ} *then every circuit of size* $2^{n^\lambda}/2n$ *fails to compute* f *on one of the inputs* $y_1, \ldots, y_{2^{O(n^\lambda)}}$. *Moreover, each pair* $y_i, f(y_i)$ *is generated by a subcircuit of* C_{2^n} *with inputs* $y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1}), \mathrm{tt}(f)$ *whose only gates with fanout* > 1 *are* $y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1})$.

Proof. This proof follows [43]. Our contribution here is the "moreover" part, but we also give a more succinct self-contained proof. For each Boolean function f the desired set of anticheckers is known to exist, the only problem is to find it with a circuit of the desired size and formula-like form. To do so, we will simulate the proof of the existence of anticheckers but make the involved counting constructive by using linear hash functions and the assumption NP $\subseteq$ NC^1. Additionally, for the "moreover" part of the lemma, we will employ the Valiant-Vazirani Isolation Lemma (cf. Reference [6, Lemma 17.19]) in the process of selecting good anticheckers.

Let $\lambda \in (0,1)$ and f be a Boolean function with n inputs hard for circuits of size 2^{n^λ}. For j n-bit strings $y_1, \ldots, y_j$ and $s \in [0,1]$, define a predicate

$$P_f(y_1, \ldots, y_j)[s] \text{ iff } \leq s \text{ fraction of all circuits of size } 2^{n^\lambda}/2n \text{ compute } f \text{ on } y_1, \ldots, y_j.$$

Further, let $R_f(y_1, \ldots, y_j)$ be the number of circuits of size $2^{n^\lambda}/2n$ that do not make any error on $y_1, \ldots, y_j$ when computing f. Note that P_f and R_f depend on j values of f, not on the whole $\mathrm{tt}(f)$, but for simplicity we do not display $f(y_1), \ldots, f(y_j)$ among the parameters of P_f and R_f.

Suppose that given $\mathrm{tt}(f)$, we already generated $y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1})$ such that $P_f(y_1, \ldots, y_{i-1})[(1 - 1/4n)^{i-1}]$ holds. For $i = 1$ the generated set is empty. We want to find $y_i, f(y_i)$ such that $P_f(y_1, \ldots, y_i)[(1 - 1/4n)^i]$. To do so, we will construct a formula $F(y_1, \ldots, y_i, f(y_1), \ldots, f(y_i))$ of size $2^{O(n^\lambda)}$ (if $i \leq 2^{O(n^\lambda)}$) such that under the assumption $R_f(y_1, \ldots, y_{i-1}) \geq 2n^2$, both of the following hold:

$$F(y_1, \ldots, y_i, f(y_1), \ldots, f(y_i)) = 1 \quad \Rightarrow \quad P_f(y_1, \ldots, y_i)[(1 - 1/4n)^i],$$

$$P_f(y_1, \ldots, y_{i-1})[(1 - 1/4n)^{i-1}] \quad \Rightarrow \quad \exists y_i, F(y_1, \ldots, y_i, f(y_1), \ldots, f(y_i)) = 1.$$

Assume for now that we already have such a formula F. We first show how to find $y_i, f(y_i)$ given F by an exhaustive search through all n-bit strings in combination with Valiant-Vazirani Lemma.

Consider a $2^{O(n^\lambda)}$-size formula $F^{r,h}(y_1, \ldots, y_{i-1}, z, f(y_1), \ldots, f(y_{i-1}), f(z))$ computing the following predicate

$$F(y_1, \ldots, y_{i-1}, z, f(y_1), \ldots, f(y_{i-1}), f(z)) \wedge \text{“}h(z) = 0^r\text{”}, \tag{1}$$

where $z \in \{0,1\}^n$, $r \leq n + 2$ and $h \in \mathcal{H}_{n,r}$ for a pairwise independent efficiently computable hash function collection $\mathcal{H}_{n,r}$ from $\{0,1\}^n$ to $\{0,1\}^r$. Formula $F^{r,h}$ exists since $\mathsf{NP} \subseteq \mathsf{NC}^1$. By Valiant-Vazirani Lemma, for fixed $y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1})$, if h is chosen randomly from $\mathcal{H}_{n,r}$ and r randomly from $\{2, \ldots, n+1\}$, then with probability $\geq 1/8n$, there is a unique z satisfying (1). Therefore, the probability that none of $2^{O(n^\lambda)}$ many randomly chosen tuples r, h guarantees a unique solution is $< (1 - 1/8n)^{2^{O(n^\lambda)}} \leq 1/2^{2^{O(n^\lambda)}/8n}$. That is, there exist a set $\mathcal{R}$ of $2^{O(n^\lambda)}$ tuples r, h such that for each $y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1})$, at least one tuple r, h from $\mathcal{R}$ will guarantee a unique solution. Consequently, for each $y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1})$ for at least one $r, h \in \mathcal{R}$ the following $2^{n+O(n^\lambda)}$-size formula:

$$D_j^{r,h}(y_1, \ldots, y_{i-1}, f(y_1), \ldots, f(y_{i-1})) = \bigvee_{k=1,\ldots,2^n} (b_j^k \wedge F^{r,h}(y_1, \ldots, y_{i-1}, b^k, f(y_1), \ldots, f(y_{i-1}), f(b^k)),$$

where b_j^k is the jth bit of the kth n-bit string b^k (in the lexicographic order), outputs the jth bit of a good antichecker y_i. Since $\mathsf{NP} \subseteq \mathsf{NC}^1$, we can select the right y_i from the $2^{O(n^\lambda)}$ candidate strings corresponding to tuples r, h from $\mathcal{R}$ by applying a formula of size $2^{O(n^\lambda)}$ on top of them. Having y_i, a formula of size $poly(n)2^n$ with access to $\mathrm{tt}(f)$ can generate $f(y_i)$. See Figures 1 and 2 for the above oracle circuit construction.

Iteratively, a circuit of size $2^{n+O(n^\lambda)}$ will generate $y_1, \ldots, y_{2^{O(n^\lambda)}}, f(y_1), \ldots, f(y_{2^{O(n^\lambda)}})$ such that $P_f(y_1, \ldots, y_{2^{O(n^\lambda)}})[(1 - 1/4n)^{2^{O(n^\lambda)}}]$ as long as $R_f(y_1, \ldots, y_{2^{O(n^\lambda)}}) \geq 2n^2$. Deciding whether $R_f(y_1, \ldots, y_i) \geq 2n^2$ is in $\mathsf{NP} \subseteq \mathsf{NC}^1$ (on input $y_1, \ldots, y_i, f(y_1), \ldots, f(y_i), 1^{2^{n^\lambda}}$), so there are formulas of size $2^{O(n^\lambda)}$ for it. Since $(1-1/4n)^{2^{O(n^\lambda)}} \leq 1/2^{2^{O(n^\lambda)}/4n}$, we reach $R_f(y_1, \ldots, y_i) < 2n^2$ with $i \leq 2^{O(n^\lambda)}$. When this happens, the remaining $< 2n^2$ circuits of size $2^{n^\lambda}/2n$ can be generated by an $\mathsf{NP}^{\mathsf{coNP}}$ algorithm, and since $\mathsf{NP} \subseteq \mathsf{NC}^1$, by a formula of size $2^{O(n^\lambda)}$. Finally, for each of the remaining circuits, we can find an n bit string witnessing its error exhaustively by a formula of size $2^{n+O(n^\lambda)}$. Altogether, the desired antickeckers $y_1, \ldots, y_{2^{O(n^\lambda)}}$ with bits $f(y_1), \ldots, f(y_{2^{O(n^\lambda)}})$ will be generated by a circuit of size $2^{n+O(n^\lambda)}$. Note that this circuit will have the desired formula-like structure, because its only gates with fanout bigger than 1 are those computing tuples $y_i, f(y_i)$.

Claim 28. *If $P_f(y_1, \ldots, y_{i-1})[(1 - 1/4n)^{i-1}]$ and $R_f(y_1, \ldots, y_{i-1}) \geq 2n^2$, then for some y_i, $P_f(y_1, \ldots, y_i)[(1 - 1/4n)^{i-1}(1 - 1/2n)]$.*

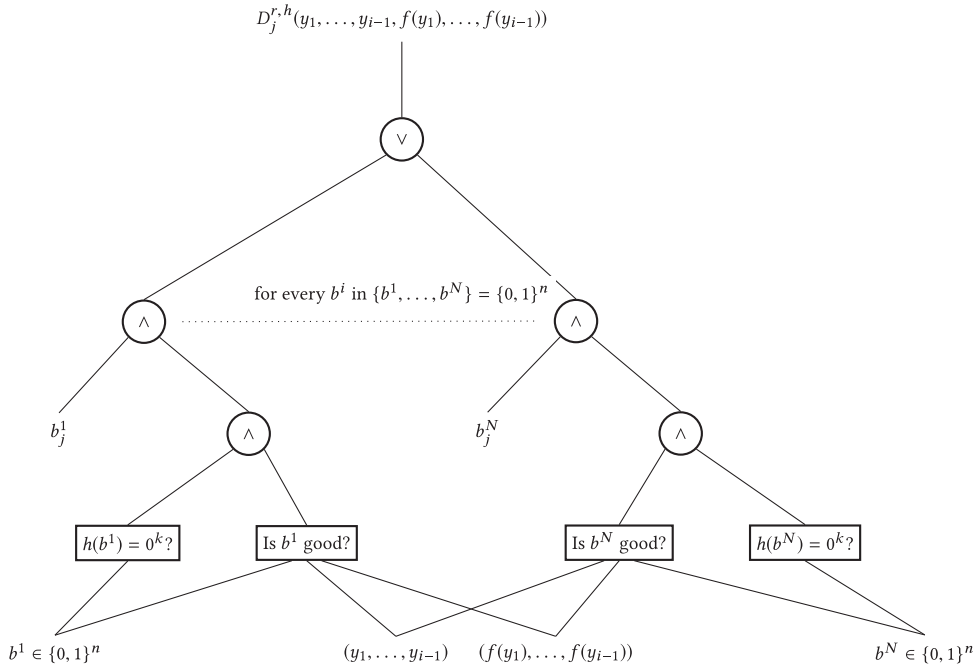

Fig. 1. The oracle circuit $D_j^{r,h}$ for any $(r,h)\in\mathcal{R}$ and $j\in\{0,1\}^{\log n}$. The NP-oracles, represented as boxes, are simulated in the proof by polynomial-sized formulas assuming $\mathsf{NP}\not\subseteq\mathsf{NC}^1$.

Claim 28 is proved by a standard counting argument, cf. Reference [43, Claim 22]. Observe that with Claim 28, we can construct the desired formula F. Here, we employ approximate counting with linear hash functions: If $X\subseteq\{0,1\}^m$ is a set of size s, then there are matrices $A_1,\ldots,A_{\log(4s^c)}$ such that each A_j defines a linear function mapping a Cartesian power X^c to $(s(1+\varepsilon))^c/\log(4s^c)$, for $c=2(\varepsilon^{-1}(\log\log s+\log\varepsilon^{-1}))$. Moreover, for each A_j there is $X_j^c\subseteq X^c$ satisfying $\forall x\in X_j^c\forall x'\in X^c$ $(x\neq x'\to A_j(x)\neq A_j(x'))$, and $\bigcup_j X_j^c=X^c$. Mapping $x\in X^c$ to $A_j(x)$ in the jth block of size $(s(1+\varepsilon))^c/\log(4s^c)$, for the first A_j with $x\in X_j^c$, thus defines an injection from X^c to $(s(1+\varepsilon))^c$, which witnesses that the size of X is $\leq s(1+\varepsilon)$. See, e.g., Reference [31, Section 3, 2nd paragraph] for details.

Therefore, once we have $P_f(y_1,\ldots,y_i)[(1-1/4n)^{i-1}(1-1/2n)]$, we can conclude that there are matrices $A_1,\ldots,A_{2^{O(n^\lambda)}}$ defining an injective mapping of a Cartesian power (with exponent of rate $poly(n)$) of the set of all circuits of size $2^{n^\lambda}/2n$ that compute f on $y_1,\ldots,y_i$ to the same Cartesian power of $(1-1/4n)^{i-1}(1-1/2n)(1+1/4n)\leq(1-1/4n)^i$ fraction of the set of all circuits of size $2^{n^\lambda}/2n$. The existence of such matrices, not only witnesses $P_f(y_1,\ldots,y_i)[(1-1/4n)^i]$ but is also an $\mathsf{NP}^{\mathsf{coNP}}$ property, and since $\mathsf{NP}\subseteq\mathsf{NC}^1$, decidable by a formula F of size $2^{O(n^\lambda)}$. □

Theorem 29 (Improved Magnification via Antickeckers). *Assume that* $\mathsf{MCSP}[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$ *is hard for circuits C (with 2^n inputs) of size $2^{n+O(n^{1/2})}$ with the following form. Given $\mathsf{tt}(f)$, subcircuits of C generate $y_1,\ldots,y_{2^{O(n^{1/2})}},f(y_1),\ldots,f(y_{2^{O(n^{1/2})}})$ so that each $y_i,f(y_i)$ is generated by a subcircuit of C with inputs $y_1,\ldots,y_{i-1},f(y_1),\ldots,f(y_{i-1}),\mathsf{tt}(f)$ whose only gates with fanout >1 are $y_1,\ldots,y_{i-1},f(y_1),\ldots,f(y_{i-1})$. Having $y_1,\ldots,y_{2^{O(n^{1/2})}},f(y_1),\ldots,f(y_{2^{O(n^{1/2})}})$, C applies a formula of size $2^{O(n^{1/2})}$ on top of these gates.*

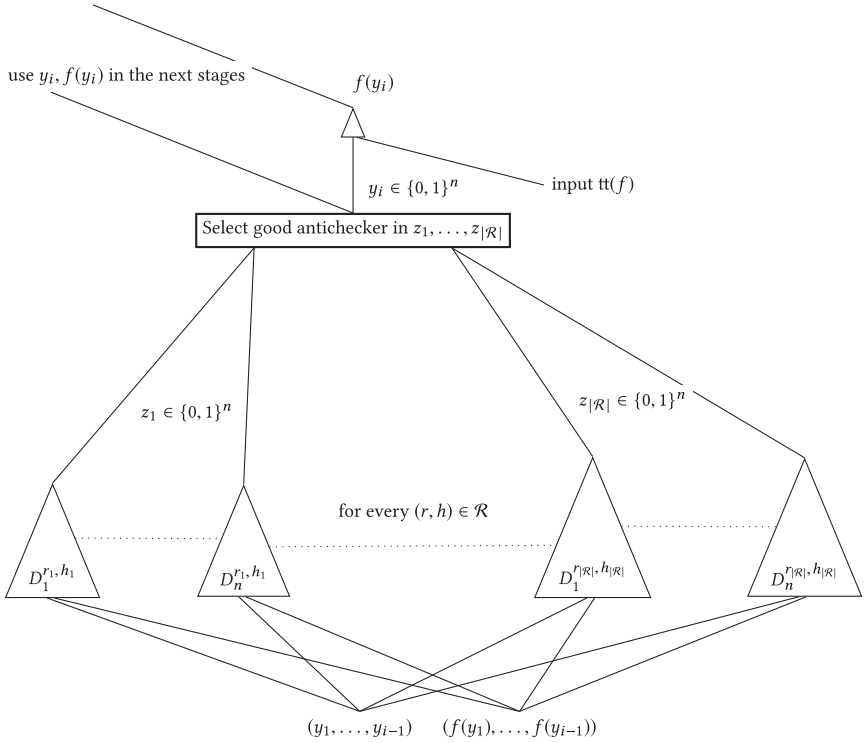

Fig. 2. The overall oracle circuit for any iteration $i \leq 2^{O(n^\lambda)}$, where $R_f(y_1, \ldots, y_i) \geq 2n^2$.

Then NP $\not\subseteq$ NC^1.

Proof. If NP $\subseteq$ NC^1, then MCSP$[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$ can be solved by circuits of size $2^{n+O(n^{1/2})}$ of the required form: given a Boolean function f, apply Lemma 27 to generate a set of its antichekers $y_1, \ldots, y_{2^{O(n^{1/2})}}$ together with bits $f(y_1), \ldots, f(y_{2^{O(n^{1/2})}})$ and using NP $\subseteq$ NC^1 decide whether f is hard for circuits of size $2^{n^{1/2}}/2n$ on $y_1, \ldots, y_{2^{O(n^{1/2})}}$. □

Note that circuits from the assumption of hardness magnification via antichekers, Theorem 29, are $2^{O(n^{1/2})}$-almost-formulas of almost linear size, which gives us HM Frontier C1. We can now complement it with HM Frontier C3.

Consider an s-almost-formula. Each gate G of F with fanout larger than 1 is computed by a formula with inputs being either the original inputs of F or gates of F with fanout larger than 1. We call any maximal formula of this form a *principal* formula of G.

Theorem 30. PARITY $\notin n^\varepsilon$-almost-Formula$[n^{2-9\varepsilon}]$, *if* $\varepsilon < 1$.

Proof Sketch. For the sake of contradiction, assume PARITY has n^ε-almost-formulas of size $n^{2-9\varepsilon}$. Since there are only n^ε gates of fanout > 1, we can replace these gates by appropriate constants and obtain formulas F_n of size $n^{2-8\varepsilon}$ computing PARITY with probability $\geq 1/2 + 1/2^{n^\varepsilon}$. In more detail, each formula F_n checks if the principal formulas compute the fixed constants. If this is the case, then F_n outputs the output of the original almost-formula (since gates with fan-out larger than 1 are fixed, the output can be computed by a formula). Otherwise, F_n outputs a fixed constant, whichever is better on the majority of the remaining inputs. This does not increase the

size of the resulting formula F_n by more than a constant factor. As pointed out by Komargodski-Raz [36], each Boolean function f on n input bits can be approximated by a real polynomial of degree $O(t\sqrt{L(f)}\frac{\log n}{\log\log n})$ up to a point-wise additive error of 2^{-t}, and this can be shown to imply that each formula of size $o((n/t)^2(\log\log n/\log n)^2)$ computes PARITY over n input bits with probability at most $1/2+1/2^{t+O(1)}$ (for large enough t). Taking $t=n^{2\varepsilon}$, we get a contradiction. □

3.4 NP $\not\subseteq$ NC^1 and AC^0 Lower Bounds for $(n-k)$-Clique

In this section, we discuss the proofs of some statements claimed in HM Frontier E from Section 1.1. Recall that we consider graphs on n vertices that are described in the adjacency matrix representation. The input graph is therefore represented using $m=\Theta(n^2)$ bits. We begin with the proof of the magnification result in HM Frontier E1.

Proposition 31. *Let $k(n)=(\log n)^C$ for some constant C. If there exists $\varepsilon>0$ such that for every depth $d\geq 1$, $(n-k)$-Clique $\notin \mathsf{AC}^0_d[m^{1+\varepsilon}]$, then* NP $\not\subseteq$ NC^1.

Proof. We use a straightforward reduction to the magnification theorem for k-Vertex-Cover established in Reference [46, Theorem 7]. (We state Proposition 31 in a slightly weaker form just for simplicity.) Indeed, a graph G on n vertices has a vertex cover of size $\leq k$ if and only if G has an independent set of size $\geq n-k$. In turn, the latter is true if and only if the complement graph $\overline{G}$ has a clique of size $\geq n-k$. Therefore, by negating input literals, the complexities of $(n-k)$-Clique and k-Vertex-Cover are equivalent with respect to AC^0 circuits. For this reason, the hardness magnification theorem of Reference [46] immediately implies Proposition 31. □

We state below conditional and unconditional lower bounds on the complexity of detecting very large cliques. The next proposition implies the lower bound claimed in HM Frontier E4.

Proposition 32 ([4]; see also Reference [32, Section 9.2]). *For $k(n)\leq n/2$, every monotone circuit for $(n-k)$-Clique requires $2^{\Omega(k^{1/3})}$ gates.*

Interestingly, the problem can be solved by (bounded depth) polynomial size monotone circuits if $k\leq\sqrt{\log n}$ [4].

Finally, by the observation employed in the proof of Proposition 31, for non-monotone computations the complexities of detecting large cliques and small vertex covers are equivalent. A consequence of this is that one can show the following result, which implies the statement in HM Frontier E2.

Proposition 33. *If* ETH *for non-uniform circuits holds, then $(n-k)$-Clique $\notin$ P/poly as long as $\omega(\log n)\leq k\leq n/2$.*

Indeed, under ETH the k-Vertex-Cover problem cannot be solved in time $2^{o(k)}\cdot\mathrm{poly}(m)$ (see References [29] and [19, Theorem 29.5.9]). Further discussion on the conditional hardness of k-Vertex-Cover that also applies to $(n-k)$-Clique appears in Reference [46].

4 HARDNESS MAGNIFICATION AND NATURAL PROOFS

4.1 Equivalences

The main contribution of this section is new hardness magnification results showing non-learnability of circuit classes from slightly super-linear lower bounds for the approximate version of MCSP and the gap version of MCSP. We use these magnification results to prove Theorem 1.

Lemma 34 (Hardness Magnification for Learnability from Lower Bounds for Approximate MCSP). *Let $s,t:\mathbb{N}\to\mathbb{N}$ be size functions such that $n\leq s(n)\leq t(n)$ and ε,δ be parameters such that $\varepsilon<1/2$, $0\leq\delta\leq 1/9$. If for infinitely many input lengths $N=2^n$,*

$\mathrm{MCSP}[(s,0),(t,\varepsilon)] \notin \mathrm{Circuit}[N\cdot\mathrm{poly}(t(n)/\varepsilon)]$, *then for infinitely many input lengths* n, $\mathrm{Circuit}[s(n)]$ *cannot be learnt up to error* $\varepsilon/2$ *with confidence* $1-\delta$ *by* $t(n)$*-size circuits using non-adaptive membership queries over the uniform distribution.*

We also show a related result that gives lower bounds for learnability of a circuit class C using C-circuits by starting with a lower bound against worst-case MCSP instead of the average-case.

Lemma 35 (Hardness Magnification for Learnability from Lower Bounds for Gap MCSP). *Let* $c \geq 1$ *be an arbitrary constant. If there is* $\varepsilon < 1/2$*, such that infinitely many input lengths* $N = 2^n$*,* $\mathrm{MCSP}[n^c, 2^n/n^c] \notin \mathrm{Circuit}[N^{1+\varepsilon}]$*, then for every* $\gamma \in (0,1)$*, for infinitely many input lengths* n*,* $\mathrm{Circuit}[n^c]$ *cannot be learnt up to error* $1/O(n^{2c})$ *with confidence* $1-1/n$ *by* $\mathrm{Circuit}[2^{O(n^\gamma)}]$*-circuits using non-adaptive membership queries over the uniform distribution.*

Lemmas 35 and 34 can be used to derive Theorem 1, which we recall below.

Theorem 36 (Theorem 1 Recalled). *The following statements are equivalent:*

(*a*) ***Hardness of approximate* MCSP *against almost-linear size circuits.***
There exist $c \geq 1$*,* $0 < \gamma < 1$*, and* $\varepsilon > 0$ *such that* $\mathrm{MCSP}[(n^c,0),(2^{n^\gamma},n^{-c})] \notin \mathrm{Circuit}[N^{1+\varepsilon}]$.

(*b*) ***Hardness of worst-case* MCSP *against almost-linear size circuits.***
There exists $c \geq 1$ *and* $\varepsilon > 0$ *such that* $\mathrm{MCSP}[n^c, 2^n/n^c] \notin \mathrm{Circuit}[N^{1+\varepsilon}]$.

(*c*) ***Hardness of sub-exponential size learning using non-adaptive queries.***
There exist $\ell \geq 1$ *and* $0 < \gamma < 1$ *such that* $\mathrm{Circuit}[n^\ell]$ *cannot be learned up to error* $O(1/n^\ell)$ *under the uniform distribution by circuits of size* $2^{O(n^\gamma)}$ *using non-adaptive membership queries.*

(*d*) ***Non-existence of natural properties against polynomial size circuits.***
For some $d \geq 1$ *there is no* $\mathrm{Circuit}[\mathrm{poly}(N)]$*-natural property useful against* $\mathrm{Circuit}[n^d]$.

(*e*) ***Existence of non-uniform* PRFs *secure against sub-exponential size circuits.***
For every constant $a \geq 0$*, there exists* $d \geq 1$*, a sequence* $\mathfrak{F} = \{\mathcal{F}_n\}_{n\geq 1}$ *of families* $\mathcal{F}_n$ *of* n*-bit Boolean functions* $f_n \in \mathrm{Circuit}[n^d]$*, and a sequence of probability distributions* $\mathfrak{D} = \{\mathcal{D}_n\}_{n\geq 1}$ *supported over* $\mathcal{F}_n$ *such that, for infinitely many values of* n*,* $(\mathcal{F}_n, \mathcal{D}_n)$ *is pseudo-random function family that* $(1/N^{\omega(1)})$*-fools (oracle) circuits of size* $2^{a\cdot n}$.

Proof of Theorem 1. The following implications establish the desired equivalences.

$(a) \implies (c)$: For the parameters c, γ, ε given by (a), we apply Lemma 34 for $s(n) = n^c$ and $t(n) = 2^{n^\gamma}$, to see that for some $\gamma' > 0$, $\mathrm{Circuit}[n^c]$ cannot be learned by circuits of size $2^{O(n^{\gamma'})}$ via non-adaptive queries up to an error $O(1/n^c)$.

$(c) \implies (d)$: We show the contrapositive of this implication. Suppose that for every $d \geq 1$, there exists a $\mathrm{Circuit}[\mathrm{poly}(n)]$-natural property that is useful against $\mathrm{Circuit}[n^d]$ for all large enough n. By Theorem 5, for every $c \geq 1$, we can learn $\mathrm{Circuit}[n^c]$ by a sequence of oracle $\mathrm{Circuit}[2^{O(n^{1/2})}]$-circuits up to an error of n^{-c}, by choosing $d = 2ac$ for the constant a from Theorem 5.

$(d) \implies (a)$, $(d) \implies (b)$: Trivial, using the fact that random functions are hard.

$(c) \implies (e)$: Follows from the contrapositive of Theorem 10.

$(e) \implies (c)$: Follows from the non-uniform version of Proposition 29 in Reference [45], using essentially the same proof.

$(b) \implies (c)$: For the parameter c given by (b), we apply Lemma 35 to see that $\mathrm{Circuit}[n^c]$ cannot be learned by circuits of size $2^{O(n^\gamma)}$ via non-adaptive queries up to an error $O(1/n^c)$, for any $\gamma \in (0,1)$. □

We now complete the proof of Theorem 1 by proving Lemmas 34 and 35. In both lemmas, we proceed by applying the learning algorithm as a "distinguisher," which helps to solve MCSP. This idea appeared already in Reference [45]. In more detail, we run the learning algorithm on an input of MCSP and check if the learning algorithm successfully learnt the function represented by the input. If the input was an easy instance of MCSP, then our algorithm will accept with high probability. Otherwise, it will reject. By standard amplification and derandomization procedures, we then obtain the desired circuit for MCSP. The crucial point is that the circuit will be, in fact, very efficient.

Proof of Lemma 34. For the promise problem MCSP$[(s,0),(t,\varepsilon)]$ over N inputs, define

$$\Pi_{yes} = \{y \in \{0,1\}^N \mid \exists \text{ circuit of size} \leq s(n) \text{ that computes } f_y\},$$
$$\Pi_{no} = \{y \in \{0,1\}^N \mid \text{no circuit of size} \leq t(n) \ (1-\varepsilon)\text{-approximates } f_y\}.$$

We prove the contrapositive of the statement, by showing a reduction from MCSP$[(s,0),(t,\varepsilon)]$ to a learning algorithm for Circuit$[s(n)]$ using non-adaptive membership queries over the uniform distribution. For a fixed $\varepsilon < 1/2$ and $0 \leq \delta \leq 1/9$, let $\{D_n\}_{n\geq 1} \in \text{Circuit}[t(n)]$ be the corresponding sequence of oracle circuits, which learns Circuit$[s(n)]$ up to error $\varepsilon/2$, where D_n makes non-adaptive queries to some function $f \in \text{Circuit}[s(n)]$ over n inputs.

Let $q = q(n) = \frac{200}{\varepsilon^2}$. Define $F_N : \{0,1\}^N \times \{0,1\}^{nq(n)} \times \{0,1\}^{t(n)} \to \{0,1\}$ as the sequence of randomized circuits such that:

$$z \in \Pi_{yes} \implies \Pr_{y_1,w}\{F_N(z,y_1,w)=1\} > 2/3,$$
$$z \in \Pi_{no} \implies \Pr_{y_1,w}\{F_N(z,y_1,w)=1\} < 1/3.$$

The reduction F_N does the following. Let $Y = (x_1,\ldots,x_{t(n)})$ be the set of queries made by D_n. F_N runs the learner D_n with input w as its source of internal randomness and answers its oracle queries to f_z by using the other input $z \in \{0,1\}^N$. If the output string of the learner cannot be interpreted as a $t(n)$-sized circuit, then F_N outputs 0. Otherwise, let h be the $t(n)$-sized circuit on n inputs, which can interpret the hypothesis output by the learner as a $t(n)$-sized circuit. F_N then interprets the random input y_1 as a sequence of q random examples $v_1,\ldots,v_q \in \{0,1\}^n$ and computes h on each of these. It then forms a string $u \in \{0,1\}^q$, where for every $i \in [q], u_i = 1$ if and only if $h(v_i) = f_z(v_i)$. Finally, it uses a threshold gate on T on $q(n)$ inputs to check if the Hamming weight of u is at least $((1-3\varepsilon/4)q)$.

We now show the correctness of the reduction. If $z \in \Pi_{yes}$, then f_z is computed by some circuit of size at most $s(n)$. Thus, for every random choice of y_1 and w, D_n can learn the function f_z and with probability at least $(1-\delta)$, output a hypothesis h that has an error of at most $\varepsilon/2$ with respect to f_z. Now, for the q samples given by y_1, by an application of Hoeffding's inequality (Lemma 13), the probability that the Hamming weight of $u \in \{0,1\}^q$ is lesser than $(1-0.6\varepsilon)\,q$ is at most $2\exp(-2q\varepsilon^2/100)$, which is at most $1/4$ for our choice of q. When $\delta \leq 1/9$, we see that $T(u) = 1$ with probability at least $(1-\delta)3/4 \geq 2/3$.

However, if $z \in \Pi_{no}$, then no circuit of size at most $t(n)$ can even $(1-\varepsilon)$-approximate f_z. Thus, for any random choice of y_1 and w, any hypothesis h which D_n outputs is a circuit of size at most $t(n)$ and thus is at least ε-far from f_z. By a similar application of Hoeffding's inequality, we see that the probability that the Hamming weight of $u \in \{0,1\}^q$ is greater than $(1-0.9\varepsilon)\,q$ is at most $2\exp(-2q\varepsilon^2/100) \leq 1/4$. Therefore, $T(u) = 0$ with probability $2/3$.

For the next step, we need to derandomize the circuits F_N. Define E_N as

$$E_N : \{0,1\}^N \times \left(\{0,1\}^{n\cdot q+t(n)}\right)^R \to \{0,1\},$$
$$E_N(z,y^{(1)},\ldots,y^{(R)}) = \mathsf{MAJ}_R(F_N(z,y^{(1)}),\ldots,F_N(z,y^{(R)})),$$

where $R = CN$ and each $y^{(j)} \in \{0,1\}^{n\cdot q+t(n)}$, for each $j \in [R]$.

When $z \in \Pi_{yes}$, then using Hoeffding's inequality, we see that with probability at most 2^{-2N} (for suitably chosen C), the string $(F_N(z, y^{(1)}), \ldots, F_N(z, y^{(R)}))$ has Hamming weight $\leq 3R/5$. Similarly, when $z \in \Pi_{no}$, with probability at most 2^{-2N}, the string $(F_N(z, y^{(1)}), \ldots, F_N(z, y^{(R)}))$ has Hamming weight $\geq 2R/5$. Thus, the majority gate differentiates between the two cases except with probability at most 2^{-2N}. We use Adleman's trick [6] to fix a string $\alpha \in \{0,1\}^{R\cdot(n\cdot q+t(n))}$ which correctly derandomizes F_N on all inputs in Π_{yes} and Π_{no} and call the resulting circuit as E_N^* which computes the function $E_N^* : \{0,1\}^N \to \{0,1\}$.

We next compute the size of E_N^*. Each $F_N(z, y^{(i)})$ is fixed to $F_N(z, \alpha^{(i)})$, where $\alpha^{(i)} \in \{0,1\}^{(n\cdot q+t(n))}$ is the ith section of the hardwired random string α. Observe that for the set of oracle queries Y made by D_n, it is enough to use appropriate literals from the input z whenever we need to access the truth table of f_z. Indeed, whenever D_n uses a random example, the randomness comes from $\alpha^{(i)}$ which is fixed non-uniformly and whenever it makes a membership query, the set of queries Y is fixed for D_n because of its non-adaptivity. Recall that the size of the circuit D_n is $t(n)$ and the hypothesis h output by the learner can be interpreted as a circuit and efficiently computed by another circuit of size $\text{poly}(t(n))$. Thus, the circuit size to compute $F_N(z, \alpha)$ is at most $\text{poly}(t(n)\cdot q)$ and the total circuit size to construct E_N^* is $O(N \cdot \text{poly}(t(n)/\varepsilon))$. □

Proof Sketch of Lemma 35. We show a two-sided error randomized reduction from MCSP $[n^c, 2^n/n^c]$ to $\{D_n\}_{n\geq 1}$. Let $q = q(n) = O(n^{3c})$. The reduction is almost the same as that of Lemma 34. Here, we use a threshold gate on $q(n)$ inputs which answers 1 whenever the Hamming weight of its input is greater than $(1 - 1/n^{1.5c})q(n)$.

When the input to MCSP$[n^c, 2^n/n^c]$ is a yes instance, with probability at least $(1 - 1/n)$, D_n outputs a hypothesis $h_n \in \text{Circuit}[2^{n^\gamma}]$ which has error at most $1/O(n^{2c})$. Now for the $q(n)$ samples drawn uniformly at random, the probability that h agrees with the input instance on at least a $(1 - 1/n^{1.5c})q(n)$ samples is at least $(1 - 1/n)2/3$.

When the input to MCSP$[n^c, 2^n/n^c]$ is a no instance, any hypothesis h which D_n outputs must have error greater than $1/O(n^{c+2})$. Indeed, if the error is less than $O(1/n^{c+2})$, then by hardwiring all the error inputs by using circuits of size at most $O(\frac{2^n}{n^{c+2}} \cdot n)$, we get a circuit of size at most $2^n/n^c$, which is a contradiction to the promise of the no instance. By Hoeffding's inequality, the probability that h agrees with the input instance on at most a $(1-1/n^{1.5c})q(n)$ samples is at least $2/3$.

The derandomization is the same as that of Lemma 34, obtained by repeating the above reduction $R = O(N)$ times and computing the majority over the R outputs of the reduction. The circuit size to compute MCSP$[n^c, 2^n/n^c]$ is thus $O(N \cdot 2^{O(n^\gamma)}n^{3c}) = O(N^{1+\varepsilon})$, for $\varepsilon = o(1)$. □

4.2 Toward a More Robust Theory

The question of non-naturalizability of hardness magnification for worst-case versions of MCSP is connected to the question of basing hardness of learning on the assumption NP $\not\subseteq$ Circuit$[2^{O(n^\gamma)}]$. For simplicity, we illustrate the connection on the case of hardness magnification for MCSP$[n^c/2n, n^c]$, but with a different choice of parameters, we could similarly consider versions of worst-case MCSP closer to those appearing in the existing HM frontiers such as HM frontier C.

Proposition 37. *Assume that there is $d \geq 2$ such that* NP $\not\subseteq$ Circuit$[2^{O(n^{1/d})}]$ *implies hardness of learning* Circuit$[n^d]$ *by $2^{n^{1/d}}$-size circuits with error $1/n^d$. Then, there is a constant e such that for every $\gamma \in (0,1)$ and $c \geq 1$,* MCSP$[n^c/2n, n^c] \notin$ Circuit$[N^{1+e\gamma c}]$ *implies that there is no* P/poly*-natural property against* P/poly.

Proof. By Theorem 5, P/poly-natural property against P/poly implies that for every d there are $2^{n^{1/d}}$-size circuits learning Circuit$[n^d]$ with error $1/n^d$. By our assumption, this implies

$\mathsf{NP} \subseteq \mathsf{Circuit}[2^{O(n^{1/d})}]$, for some $d \geq 2$. We can now use $\mathsf{NP} \subseteq \mathsf{Circuit}[2^{O(n^{1/d})}]$ as the assumption in the proof of Theorem 29 to conclude that there is a constant e independent of γ such that for $c \geq 1$, $\mathsf{MCSP}[n^c/2n, n^c] \in \mathsf{Circuit}[N^{1+e\gamma c}]$. □

A form of the opposite implication (i.e., non-naturalizable hardness magnification for a worst-case version of MCSP implying that we can base hardness of learning on an assumption such as $\mathsf{NP} \not\subseteq \mathsf{P/poly}$) holds as well. However, we need to assume NP-completeness of MCSP. Moreover, instead of the non-naturalizability of hardness magnification, we need to assume a reduction from worst-case MCSP to approximate MCSP. Note that such a reduction, if implemented very efficiently, could be used to obtain a non-naturalizable hardness magnification for worst-case MCSP from a non-naturalizable hardness magnification for approximate MCSP.

Definition 38. A p-time algorithm A k-reduces $\mathsf{MCSP}[s,t]$ to $\mathsf{MCSP}[(s,0),(t,\varepsilon)]$ if it maps an instance of $\mathsf{MCSP}[s,t]$ to an instance of $\mathsf{MCSP}[(s,0),(t,\varepsilon)]$ and

(1) For $f \in \mathsf{Circuit}[s]$, $A(\mathsf{tt}(f))$ is the truth table of a Boolean function in $\mathsf{Circuit}[s^k]$.
(2) For $f \notin \mathsf{Circuit}[t]$, $A(\mathsf{tt}(f))$ is not $(1-\varepsilon)$-approximable by circuits of size $t^{1/k}$.

Proposition 39. *Assume there is a p-time algorithm k-reducing $\mathsf{MCSP}[s,t]$ to $\mathsf{MCSP}[(s,0),(t,\varepsilon)]$ and that for all $0 < \alpha < \beta < 1$, $\mathsf{MCSP}[2^{\alpha n}, 2^{\beta n}]$ is NP-complete. If for every sufficiently small $\alpha > 0$ there is $\beta < 1/k$ and $2^{\beta n}$-size circuits learning $\mathsf{Circuit}[2^{\alpha n}]$ with error ε, then $\mathsf{NP} \subseteq \mathsf{P/poly}$.*

Proof. Let A be the p-time k-reduction from the statement and $\alpha > 0$ be sufficiently small. Assume we can learn in $2^{\beta n}$-size $\mathsf{Circuit}[2^{k\alpha n}]$ with error ε and $k\alpha < \beta < 1/k$. This implies that $\mathsf{MCSP}[(2^{k\alpha n},0),(2^{\beta n},\varepsilon)]$ is in $\mathsf{P/poly}$. Since A reduces an NP-complete problem $\mathsf{MCSP}[2^{\alpha n}, 2^{k\beta n}]$ to $\mathsf{MCSP}[(2^{k\alpha n},0),(2^{\beta n},\varepsilon)]$, this shows that $\mathsf{NP} \subseteq \mathsf{P/poly}$. □

5 THE LOCALITY BARRIER

5.1 Lower Bounds above Magnification Threshold

5.1.1 The Razborov-Smolensky Polynomial Approximation Method. In this section, we observe that the lower-bound techniques of Razborov and Smolensky [49, 54] can be "localized." The following proposition instantiates the locality barrier for HM Frontier A.

Proposition 40 (Locality Barrier for HM Frontier A). *The following results hold.*

- ($\mathrm{A1}^O$) (Oracle Circuits from Magnification): $\mathsf{MKtP}[n^c, 2n^c] \in \mathsf{AND}\text{-}\mathcal{O}\text{-}\mathsf{XOR}[N^{1.01}]$. *More precisely, $\mathsf{MKtP}[n^c, 2n^c]$ is computed by circuits with $N^{1.01}$ gates and of the following form: the output gate is an AND gate of fan-in $O(N)$, at the middle layer are oracle gates of fan-in $\mathrm{poly}(n)$, and at the bottom layer are XOR gates.*
- ($\mathrm{A3}^O$) (Extension of Lower-bound Techniques): *For a constant d, assume that $O_1, \ldots, O_d \in \mathbb{N}$ satisfy $\prod_{i=1}^{d} O_i \leq \sqrt{N}/\omega(\log N)^d$. Then* Majority *cannot be computed by a depth-d polynomial-size oracle $(\mathsf{AC}^0[\oplus])^O$ circuit whose oracle gates on the ith level have fan-in at most O_i.*

The first item is immediate from the proof of Theorem 14 in Section 3.1. In what follows, we prove the second item of Proposition 40.

Recall that the proof techniques of Razborov and Smolensky [49, 54] consist of two parts: The first lemma shows that any low degree polynomial cannot approximate Majority. (A simple proof sketch can be found in, e.g., Reference [37].)

Lemma 41. *For any polynomial $p \in \mathbb{F}_2[x_1, \ldots, x_N]$ of degree $\leq \sqrt{N}/4$,*

$$\Pr_{x \sim \{0,1\}^N}[p(x) \neq \mathsf{Majority}(x)] \geq \frac{1}{4}.$$

The second lemma shows that $\mathsf{AC}^0[\oplus]$ circuits can be approximated by low degree polynomials. We show that this argument can be localized.

Lemma 42. *Let C be a depth-d polynomial-size oracle $\mathsf{AC}^0[\oplus]$ circuit whose oracle gates on the ith level have fan-in at most O_i. Then there exists a polynomial $p \in \mathbb{F}_2[x_1, \ldots, x_N]$ of degree $\leq O(\log N)^d \cdot \prod_{i=1}^{d} O_i$ such that $\Pr_{x \sim \{0,1\}^N}[p(x) \neq C(x)] < \frac{1}{4}$.*

Proof Sketch. We convert each layer of the circuit C into a low degree probabilistic polynomial p that approximates C.

Consider the ith level of a circuit C. NOT, OR, AND, and XOR gates can be converted into a probabilistic polynomial of degree $O(\log N)$ and error $1/\mathrm{poly}(N)$ in the standard way [49]. To represent an oracle gate O as a low-degree polynomial, we simply take the multilinear extension of the oracle gate O. Note that, at the ith level, the fan-in of the oracle gate O is bounded by O_i; thus, the oracle gate at the ith level can be represented as a polynomial of degree $\leq O_i$. Thus, in either cases, any gate at ith level can be represented as a probabilistic polynomial of degree $\max\{O(\log N), O_i\}$. Continuing this for $i = 1, \ldots, d$ and composing resulting polynomials, we obtain a probabilistic polynomial of degree $\prod_{i=1}^{d} \max\{O(\log N), O_i\}$ that approximates C. This implies via standard techniques the existence of a (deterministic) polynomial of the same degree that correctly computes the circuit on most inputs. □

These two lemmas immediately imply the Majority lower bound for $(\mathsf{AC}^0[\oplus])^O$:

Proof of $(A3^O)$ of Proposition 40. Suppose that there exists a depth-d polynomial-size oracle $\mathsf{AC}^0[\oplus]$ circuit that computes Majority and satisfies the condition of Proposition 40. By Lemma 42, there exists a polynomial p of degree at most $O(\log N)^d \cdot \prod_{i=1}^{d} O_i \leq o(\sqrt{N})$ that approximates Majority. However, this contradicts Lemma 41. □

Finally, we mention that an incomparable bound can be obtained by using a lower bound for $\mathsf{AC}^0[\oplus]$ interactive compression games.

Proposition 43 ([44, Corollary 5.3]). $(\mathrm{A3}^O)$ $\mathrm{Majority} \notin (\mathsf{AC}^0[\oplus])^O[\mathrm{poly}(n)]$ *if the total number of input wires in the circuit feeding the O-gates is $N/(\log N)^{\omega(1)}$.*

5.1.2 The Formula-XOR *Lower Bound of Reference [57].* This section captures an instantiation of the locality barrier for HM Frontier B. Throughout this section, we use the $\{-1, 1\}$ realization of the Boolean domain (that is, -1 represents True and 1 represents False). Let Formula-XOR on variables $x_1, \ldots, x_n$ be the class of formulas where the input leaves are labeled by parity functions of arbitrary arity over $x_1, \ldots, x_n$.

Proposition 44 (Locality Barrier for HM Frontier B). *The following results hold.*

- $(\mathrm{B1}^O)$ (Oracle Circuits from Magnification): *For any $\varepsilon > 0$, $\mathrm{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}] \in$ Formula-O-XOR$[N^{1.01}]$, where every oracle O has fan-in at most N^{ε} and appears in the layer right above the* XOR *leaves.*
- $(\mathrm{B3}^O)$ (Extension of Lower-bound Techniques): *For any $\delta > 0$,* InnerProduct *over N input bits cannot be computed by $N^{2-3\delta}$-size* Formula-O-XOR *circuits with at most $N^{2-3\delta}$ oracle gates of fan-in N^{δ} in the layer right above the* XOR *leaves, for any oracle O.*

To prove item 2 of Proposition 44, we adapt Tal's [57] lower bound for bipartite formulas,[16] for which we need the following results.

[16] A bipartite formula on variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ is a formula such that each leaf computes an arbitrary function in either $(x_1, \ldots, x_n)$ or $(y_1, \ldots, y_n)$. Formula-XOR circuits are a subset of bipartite formulas as one can always write $\oplus(x_1, \ldots, x_{2n})$ as the parity of $\oplus(x_1, \ldots, x_n)$ and $\oplus(x_{n+1}, \ldots, x_{2n})$.

Lemma 45 ([52, 57]). *Let F be a De Morgan formula of size s that computes $f : \{-1,1\}^n \to \{1,1\}$. Then, there exists a multilinear polynomial p over $\mathbb{R}$ of degree $O(\sqrt{s})$, such that for every $x \in \{-1,1\}^n$, $p(x) \in [F(x) - 1/3, F(x) + 1/3]$.*

For any function $f : \{-1,1\}^n \to \{-1,1\}$, f is ε-correlated with a parity $p_S(x) = \prod_{i \in S} x_i$, if $|\mathbb{E}_{x \in \{-1,1\}^n}[f(x) \cdot p_S(x)]| \geq \varepsilon$.

Lemma 46. *For any $\delta > 0$, let $F(x_1, \ldots, x_n)$ be a* Formula-O-XOR *formula of size s, where every oracle O has fan-in at most n^δ and appears in the layer right above the* XOR *leaves. Then the following hold true:*

(1) *There exists a multi-linear polynomial $p(x_1, \ldots, x_n)$ over $\mathbb{R}$ with at most $s^{O(\sqrt{s})} \cdot 2^{n^\delta \cdot O(\sqrt{s})}$ monomials such that for every $x \in \{-1,1\}^n$, $sign(p(x)) = F(x)$.*
(2) *There exists a parity function $f_T(x_1, \ldots, x_n)$ that is at least $(\frac{1}{s^{O(\sqrt{s})} \cdot 2^{n^\delta O(\sqrt{s})}})$-correlated with F.*

Proof. We assume that the oracle function is a Boolean function on n^δ inputs. Let $t \leq s/n^\delta$ be the number of oracle gates in F. Let $p_1, \ldots, p_s$ be the leaves of F, where each p_i is an XOR gate over $x_1, \ldots, x_n$ and every oracle gate $g_1, \ldots, g_t$ is such that $g_i(x) = O(p_{i1}(x), \ldots, p_{i\ell}(x))$, where $\ell = n^\delta$ and $p_{ij} \in \{p_1, \ldots, p_t\}$ for every $i \in [t], j \in [\ell]$.

Let F' be a De morgan formula obtained by replacing oracle gates in F with new variables z_i (for notational simplicity, we assume that every leaf is an input to some oracle gate), for $i \in [t]$. We now use Lemma 45 on F' to get a degree $d = O(\sqrt{t})$ polynomial $q(z)$ such that for every $z \in \{-1,1\}^t$, $\text{sign}(q(z)) = F'(z)$. Expanding $q(z)$ as a multilinear polynomial:

$$q(z) = \sum_{S \subseteq [t], |S| \leq d} \hat{q}(S) \prod_{i \in S} z_i.$$

To prove the first item, we replace each z_i by the original leaf, and we get that for every $x \in \{-1,1\}^n$,

$$\begin{aligned} F(x) &= \text{sign}\left(\sum_{S \subseteq [t], |S| \leq d} \hat{q}(S) \prod_{i \in S} g_i(x) \right) \\ &= \text{sign}\left(\sum_{S \subseteq [t], |S| \leq d} \hat{q}(S) \prod_{i \in S} \left(\sum_{U \subseteq [\ell]} \hat{O}(U) \prod_{j \in U} p_{ij}(x) \right) \right) \\ &= \text{sign}\left(\sum_{\substack{S \subseteq [t] \\ S = \{i_1, \ldots, i_{|S|}\} \\ |S| \leq d}} \sum_{U_{i_1}, \ldots, U_{i_{|S|}} \subseteq [\ell]} \hat{q}(S) \cdot \left(\prod_{1 \leq k \leq |S|} \hat{O}(U_{i_k}) \prod_{j \in U_{i_k}} p_{i_k j}(x) \right) \right), \end{aligned}$$

where the second equality uses the fact that any Boolean function on ℓ inputs can be represented by a multilinear polynomial of degree at most ℓ where each coefficient is between $[-1,1]$. Clearly, the number of monomials is at most $s^{O(\sqrt{s})} \cdot 2^{n^\delta \cdot O(\sqrt{s})}$.

To prove the second item, first observe that for every $z \in \{-1,1\}^t$, $q(z) \cdot F'(z) \in [2/3, 4/3]$, because $|q(z) - F'(z)| \leq 1/3$ for every z. This also means that for the polynomial $r(x) = q(g_1(x), \ldots, g_t(x))$, $\mathbb{E}_{x \in \{-1,1\}^n}[r(x) \cdot F(x)] \geq 2/3$.

Given that $\hat{q}(S) = \mathbb{E}_{z\in\{-1,1\}^s}\left[q(z)\prod_{i\in S} z_i\right]$, we see that $|\hat{q}(S)| \le 4/3$. We have

$$\begin{aligned}
2/3 &\le \mathop{\mathbb{E}}_{x\in\{-1,1\}^n}[F(x)\cdot r(x)] \\
&= \mathop{\mathbb{E}}_{x\in\{-1,1\}^n}\left[F(x)\cdot\sum_{S\subseteq[t],|S|\le d}\hat{q}(S)\prod_{i\in S}g_i(x)\right] \\
&\le \sum_{\substack{S\subseteq[t]\\ S=\{i_1,\ldots,i_{|S|}\}\\ |S|\le d}}\ \sum_{U_{i_1},\ldots,U_{i_{|S|}}\subseteq[\ell]}\hat{q}(S)\prod_{1\le k\le|S|}\hat{O}(U_{i_k})\mathop{\mathbb{E}}_{x\in\{-1,1\}^n}\left[F(x)\prod_{1\le k\le|S|}\prod_{j\in U_{i_k}}p_{i_kj}(x)\right].
\end{aligned}$$

Since $|\hat{q}(S)| \le 4/3$ for every $S \subseteq [t]$ and $|\hat{O}(U)| \le 1$ for every $U \subseteq [\ell]$, we see that there exists a set S of size at most d and sets $U_{i_1},\ldots,U_{i_{|S|}}$ such that $|\mathbb{E}_{x\in\{-1,1\}^n}[F(x)\cdot\prod_{1\le k\le|S|}\prod_{j\in U_{i_k}}p_{i_kj}(x)]| \ge \frac{1}{t^{O(\sqrt{t})}\cdot 2^{n^\delta O(\sqrt{t})}} \ge \frac{1}{s^{O(\sqrt{s})}\cdot 2^{n^\delta O(\sqrt{s})}}$. Taking p_T be the parity of the parities given by $p_T = \prod_{1\le k\le|S|}\prod_{j\in U_{i_k}}p_{i_kj}(x)$, we see that p_T is $\frac{1}{s^{O(\sqrt{s})}\cdot 2^{n^\delta O(\sqrt{s})}}$-correlated with F. □

Define the Inner Product modulo 2 function, $\mathsf{InnerProduct}_n : \{-1,1\}^n\times\{-1,1\}^n \to \{-1,1\}$ as $IP_n(x,y) = (-1)^{\sum_{i=1}^n (1-x_i)(1-y_i)/4}$.

Proof Sketch of Proposition 44. The first item follows from an inspection of the proof of Theorem 25 in Section 3.2. Theorem 24 gives the same oracle circuit construction (with different oracles) under the assumption $\mathsf{QP}\subseteq\mathsf{P/poly}$.

The second item follows from Lemma 46. We observe that three different techniques used to show Formula-XOR lower bounds localize. First, Tal's lower bound based on sign rank shows that the sign rank of any Formula-XOR circuit F is at most the number of monomials in the polynomial p given by the first item of lemma 46. Since this is at most $s^{O(\sqrt{s})}\cdot 2^{n^\delta\cdot O(\sqrt{s})}$ and InnerProduct has a sign rank that is at least $2^{n/2}$ [20], the lower bound follows. Second, Tal's lower bound based on the discrepancy of a function also localizes, as he shows that the discrepancy of F is at least a constant times the correlation of F with the parity f_T given by item 2 of Lemma 46, which is at least $\Omega(\frac{1}{s^{O(\sqrt{s})}\cdot 2^{n^\delta O(\sqrt{s})}})$, whereas the discrepancy of the inner product is at most $1/2^{n/2}$ (cf. Reference [32, Lemma 14.5]), thus proving the given lower bound for inner product. Finally, we also observe that the lower-bound technique of showing high correlation of F with some parity f_T and the fact that inner product has exactly $2^{-n/2}$-correlation with any parity also localizes to give the same lower bound. □

5.1.3 Almost-formula Lower Bounds. This section captures an instantiation of the locality barrier for HM Frontier C. We recall the following definition. Consider an s-almost-formula. Each gate G of F with fanout larger than 1 is computed by a formula with inputs being either the original inputs of F or gates of F with fanout larger than 1. We call any maximal formula of this form a *principal* formula of G.

Theorem 47 (Locality Barrier for HM Frontier C). *The following results hold.*

- ($\mathrm{C1}^O$) (Oracle Circuits from Magnification): $\mathsf{MCSP}[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$ *is computable by* $2^{O(n^{1/2})}$*-almost-formulas of size* $2^{n+O(n^{1/2})}$ *with oracles of fanin* $2^{O(n^{1/2})}$ *at the bottom layer of principal formulas computing gates with fanout larger than 1.*
- ($\mathrm{C3}^O$) (Extension of Lower-bound Techniques): *For every* $\varepsilon < 1$, PARITY *is not in* n^ε-almost-$\mathsf{Formula}[n^{2-9\varepsilon}]$ *even if the almost-formulas are allowed to use arbitrary oracles of fanin* $< n^\varepsilon$ *at the bottom layer of principal formulas computing gates with fanout larger than* 1.

Proof. The first item follows by inspecting the proof of Theorem 29. It is not hard to see that $\mathrm{MCSP}[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$ is computable by $2^{O(n^{1/2})}$-almost-formulas F_N of size $2^{n+O(n^{1/2})}$ with local oracles of fanin $2^{O(n^{1/2})}$. Moreover, the only gates of fanout larger than 1 are the gates computing antichecker $y_1, \ldots, y_{2^{O(n^{1/2})}}$ with bits $f(y_1), \ldots, f(y_{2^{O(n^{1/2})}})$. We want to show that the local oracles are at the bottom of principal formulas generating gates with fanout larger than 1. To achieve this, we need to modify formulas F_N a bit.

First, note that F_N contains an oracle that is applied on top of antichecker $y_1, \ldots, y_{2^{O(n^{1/2})}}$ with bits $f(y_1), \ldots, f(y_{2^{O(n^{1/2})}})$. To ensure that this oracle is at the bottom of a principal formula computing a gate with fanout bigger than 1, we simply add dummy negation gates to the output gate and the gates computing antichecker $y_1, \ldots, y_{2^{O(n^{1/2})}}$ with bits $f(y_1), \ldots, f(y_{2^{O(n^{1/2})}})$, if necessary.

Second, note that each y_{i+1}, $f(y_{i+1})$ is generated as follows: (1) if $R_f(y_1, \ldots, y_i) \geq 2n^2$ then a subformula F' generates antichecker y_{i+1}, $f(y_{i+1})$, and (2) if $R_f(y_1, \ldots, y_i) < 2n^2$ then a subformula F'' generates antichecker y_{i+1}, $f(y_{i+1})$. In both cases, we replace predicates $R_f(y_1, \ldots, y_i) < 2n^2$ by oracles. In case 1, subformulas of F' with oracles at the bottom compute predicates $F^{r,h}$ from the proof of Lemma 27. This process generates a set of $2^{O(n^{1/2})}$ potential antichecker. F' chooses the right antichecker by applying another oracle. To ensure that this top oracle is at the bottom of a principal formula, we add dummy negation gates to the gates generating the potential antichecker. This increases the number of gates with fanout larger than 1 only by $2^{O(n^{1/2})}$. In case 2, y_{i+1}, $f(y_{i+1})$ is generated by oracles outputting circuits that have not been killed yet and evaluating them on all possible inputs. Here, we ensure that the oracles are at the bottom by asking them to perform both tasks: choose the next alive circuit and evaluate it on a given input. The oracle selecting the right antichecker from the set of potential antichecker is treated in the same way as in case 1. All in all, we obtain the desired oracle almost-formulas.

The second item is proved analogously to Theorem 30. For the sake of contradiction assume PARITY has n^{ε}-almost-formulas of size $n^{2-9\varepsilon}$ with local oracles at the bottom of principal formulas. Since there are only n^{ε} gates of fanout > 1, we can replace these gates by constants and obtain formulas F_n of size $n^{2-8\varepsilon}$ with local oracles at the bottom computing PARITY with probability $\geq 1/2 + 1/2^{n^{\varepsilon}}$. Let $L'(f)$ be the size (i.e., the number of leafs) of the smallest formula with local oracles at the bottom computing f. Since oracles have fanin $< n^{\varepsilon}$ and are located at the bottom, each function $f : \{-1,1\}^n \mapsto \{-1,1\}$ can be approximated by a polynomial of degree $O(t\sqrt{L'(f)}\frac{\log n}{\log\log n}n^{\varepsilon})$ up to point-wise error of 2^{-t}. This implies that each formula of size $o((n/t)^2(\log\log n/\log n)^2(1/n^{\varepsilon})^2)$ with local oracles at the bottom computes PARITY with probability at most $1/2 + 1/2^{t+O(1)}$ (for large enough t). Taking $t = n^{2\varepsilon}$, we get a contradiction. □

5.1.4 GapAND-Formula *Lower Bounds.* This section captures an instantiation of the locality barrier for HM Frontier D.

Theorem 48 (Locality Barrier for HM Frontier D). *The following results hold.*

(1) ($\mathrm{D1}^O$) (Oracle Circuits from Magnification): $\mathrm{MCSP}[2^{\sqrt{n}}] \in \mathrm{GapAND}_{O(N)}$-$O_{N^{o(1)}}$-Formula$[N^2]$.
(2) ($\mathrm{D3}^O$) (Extension of Lower-bound Techniques): *For* $0 < \beta < \varepsilon < 1$, $\mathrm{Andreev}_N \notin \mathrm{GapAND}_{O(N)}$-$O_{N^{\beta}}$-Formula$[N^{3-\varepsilon}]$.
(3) ($\mathrm{D3}^O$): *Furthermore,* $\mathrm{MCSP}[2^n/n^4] \notin \mathrm{GapAND}_{O(N)}$-$O_{N^{\beta}}$-Formula$[N^{3-\varepsilon}]$, *for* $0 < \beta < \varepsilon < 1$.

Item 1 of the theorem above follows directly from Theorem 25.

Next, we show that the classical $N^{3-o(1)}$-formula size lower bound for the Andreev's function [23, 56] localizes, even in the presence of a GapAND gate of bounded fan-in at the top of the formula.

Proof of Item 2. Let $m = N/2$, recall that Andreev_N is defined on a $2m$-bit string $z = x \circ y$, where $x, y \in \{0,1\}^m$. For simplicity, we assume m is a power of 2 in the following.

$\text{Andreev}_N(x,y)$ first partitions x into $\log m$ blocks $x_1, x_2, \ldots, x_{\log m}$, each of length $m/\log m$. After that, it computes $i \in \{0,1\}^{\log m}$ as $i = \text{PARITY}(x_1) \circ \text{PARITY}(x_2) \circ \ldots \text{PARITY}(x_{\log m})$. It then treats i as an integer from $[m]$, and outputs y_i.

Now, suppose there is a $\text{GapAND}_{O(N)}$-O_{N^β}-Formula$[N^{3-\varepsilon}]$ formula for Andreev_N. Suppose we fix the y variables to a string $w \in \{0,1\}^m$, and apply a random restriction keeping exactly one variable from each block alive to x variables, then w.p. 0.9, we obtain a $\text{GapAND}_{O(N)}$-O_{N^β}-Formula$[N^{1-\varepsilon} \cdot \text{polylog}(N)]$ formula computing $f_w : \{0,1\}^{\log m} \to \{0,1\}$ [56].

That is, for all $w \in \{0,1\}^m$, there exists an O_{N^β}-Formula$[N^{1-\varepsilon} \cdot \text{polylog}(N)]$ formula 0.8-approximating f_w. Note that there are at most $2^{N^{1-\varepsilon+\beta} \cdot \text{polylog}(N)}$ such O_{N^β}-Formula$[N^{1-\varepsilon} \cdot \text{polylog}(N)]$ formulas, and there are 2^N possible w's (Note that O is a fixed oracle that does not depend on w). Since each O_{N^β}-Formula$[N^{1-\varepsilon} \cdot \text{polylog}(N)]$ formula can only 0.8-approximate $2^{\alpha \cdot N}$ many functions from $\{0,1\}^{\log m} \to \{0,1\}$ for a constant $\alpha < 1$, there must exist a w such that f_w cannot be 0.8-approximated by such formulas, contradiction. □

Next, we observe that the $N^{3-o(1)}$-formula lower bound for MCSP [18] also localizes.

Proof of Item 3. We first observe that the PRG construction of Reference [18] also works for oracle formulas. (We omit the details of this proof.)

Claim 49 ([18]). *For $0 < \beta < \varepsilon < 1$, there is $M = N^{1-\Omega_{\beta,\varepsilon}(1)}$ and a PRG $G : \{0,1\}^M \to \{0,1\}^N$ such that the following hold.*

(1) *For each fixed $z \in \{0,1\}^M$, $G(z)$, when interpreted as a function from $\{0,1\}^{\log N} \to \{0,1\}$, can be computed by a circuit of size $N^{1-\Omega(1)}$.*
(2) *For all O_{N^β}-Formula$[N^{3-\varepsilon}]$ formulas C, we have*

$$\left| \Pr_{z \in \{0,1\}^N}[C(z) = 1] - \Pr_{z \in \{0,1\}^M}[C(G(z)) = 1] \right| \leq 0.01.$$

Now, suppose $\text{MCSP}[2^n/n^4]$ on $N = 2^n$ bits can be computed by a $\text{GapAND}_{O(N)}$-O_{N^β}-Formula$[N^{3-\varepsilon}]$ formula C. Let $C_1, C_2, \ldots, C_{b \cdot N}$ be the O_{N^β}-Formula$[N^{3-\varepsilon}]$ subformulas of C under the top GapAND gate, where b is a constant.

We know that

$$\Pr_{z \in \{0,1\}^N}[\text{MCSP}[2^n/n^4](z) = 1] = o(1).$$

Since C computes $\text{MCSP}[2^n/n^4]$, and $C(x) = 0$ implies $C_i(x) = 0$ for at least a 0.9 fraction of $i \in [b \cdot N]$. We have that

$$\Pr_{i \in [b \cdot N],\, z \in \{0,1\}^N}[C_i(z) = 1] \leq 0.2.$$

On the other side, by the definition of $\text{MCSP}[2^n/n^4]$, and the Item (1) of Claim 49, it follows that

$$\Pr_{z \in \{0,1\}^M}[\text{MCSP}[2^n/n^4](G(z)) = 1] = 1.$$

Again, since C computes $\text{MCSP}[2^n/n^4]$, and $C(x) = 1$ implies $C_i(x) = 1$ for all $i \in [b \cdot N]$. We have that

$$\Pr_{i \in [b \cdot N],\, z \in \{0,1\}^M}[C_i(G(z)) = 1] = 1.$$

Therefore, there must exist an i such that

$$\left| \Pr_{z \in \{0,1\}^N} [C_i(z) = 1] - \Pr_{z \in \{0,1\}^M} [C_i(G(z)) = 1] \right| \geq 0.5,$$

which is a contradiction to Item (2) of Claim 49. □

Finally, we show that there is a language in E that cannot be computed by $\mathsf{GapAND}_{O(N)}$-Formula$[N^{3-\varepsilon}]$ formulas, but it *can* be computed by an $O_{N^{o(1)}}$-Formula$[N^2]$ formula. Therefore, this lower bound does not localize in the sense of Theorem 48.

Theorem 50. *There is a language $L \in \mathsf{E}$, such that $L \notin \mathsf{GapAND}_{O(N)}$-Formula$[N^{3-\varepsilon}]$ for all constants $\varepsilon > 0$, but $L \in O_{N^{o(1)}}$-Formula$[N^2]$.*

Proof. The function L is very similar to the Andreev's function. On an input x of length N, let $m = \log N$ (we assume N is a power of 2 for simplicity). To avoid the second input to $\mathsf{Andreev}_N$, we want to find a function $f_{\mathsf{hard}} : \{0,1\}^m \to \{0,1\}$ that cannot be 0.8-computed by $N^{1-\varepsilon/2}$ formulas in $2^{O(N)}$ time (such a function exists by a simple counting argument). To find f_{hard}, we simply enumerate all possible functions $f : \{0,1\}^m \to \{0,1\}$, and check whether it can be 0.8-approximated by an $N^{1-\varepsilon/2}$-size formula.

There are $2^{2^m} = 2^N$ possible functions on m bits, and $(N^{1-\varepsilon/2})^{O(N^{1-\varepsilon/2})} = 2^{N^{1-\varepsilon/2} \cdot \mathrm{polylog}(N)}$ many formulas of $N^{1-\varepsilon/2}$ size. Hence, a straightforward implementation of the algorithm runs in $2^{O(N)}$ time.

Next, L partitions x into m blocks $x_1, x_2, \ldots, x_m$, each of length N/m. After that, it computes $i \in \{0,1\}^m$ as $i = \mathsf{PARITY}(x_1) \circ \mathsf{PARITY}(x_2) \circ \ldots \mathsf{PARITY}(x_m)$. It then outputs $f_{\mathsf{hard}}(i)$.

Now, suppose there is a $\mathsf{GapAND}_{O(N)}$-Formula$[N^{3-\varepsilon}]$ for L. We apply a random restriction keeping exactly one variable from each block alive, then w.p. 0.9, we obtain a $\mathsf{GapAND}_{O(N)}$-Formula$[N^{1-\varepsilon} \cdot \mathrm{polylog}(N)]$ formula for f_{hard} [56], which implies that there is an $N^{1-\varepsilon} \cdot \mathrm{polylog}(N)$-size formula 0.8-approximating f_{hard}, contradiction.

Finally, it is easy to verify that $L \in \mathsf{E}$ and $L \in O_{N^{o(1)}}$-Formula$[N^2]$. □

5.1.5 AC^0 Lower Bounds via Random Restrictions. This section states and proves a result capturing an instantiation of the locality barrier for HM Frontier E.

Proposition 51 (Locality Barrier for HM Frontier E). *The following results hold.*

- ($\mathrm{E1}^O$) (Oracle Circuits from Magnification): *For each $k = (\log n)^C$ and every large enough depth d, $(n-k)$-Clique $\in (\mathsf{AC}^0_d)^O[m^{1+\varepsilon_d}]$, where $\varepsilon_d \to 0$ as $d \to \infty$, and the corresponding circuit employs a single oracle gate O of fan-in at most $O((\log n)^{4C})$.*
- ($\mathrm{E3}^O$) (Extension of Lower-bound Techniques): *Parity $\notin (\mathsf{AC}^0)^O[\mathrm{poly}(n)]$ if the total number of input wires in the circuit feeding the O-gates is $n/(\log n)^{\omega(1)}$.*

Proof. The first item is established by inspection of the proof of Proposition 31, which relies on the circuit construction from Reference [46] and a straightforward translation between vertex cover and clique detection. Recall that the circuit in Reference [46] simulates a well-known kernelization algorithm for k-Vertex-Cover. This algorithm produces a graph H containing $O(k^2)$ vertices and a new parameter $k_H \leq k$. This graph can be described by a string of length $O(k^4)$, and the pair (H, k_H) becomes the input string to the single oracle O that is necessary in the oracle circuit construction. (If O solves vertex cover, then the resulting oracle circuit correctly solves $(n-k)$-Clique.)

The second item easily follows by simulating oracle circuits via interactive compression games (see, e.g., Reference [44, Section 5]). In other words, one can view a circuit with oracles as an interactive protocol between two parties, where one of them has unbounded computational

power, and the other is restricted to computations in a fixed circuit class. The total number of wires feeding the oracle gates corresponds to the number of bits sent to the unbounded party. The desired lower bound for oracle circuits then follows immediately from the main result from Reference [10], which shows that the random restriction method can be extended to establish limitations on circuits with oracle gates of large fan-in. □

Informally, the main difficulty with the use of random restrictions in connection to HM Frontier E is that as soon as one simplifies a Boolean circuit so that the oracle gate O is directly fed by input literals, one can fix just $(\log n)^{O(C)}$ input variables and eliminate this gate. Sacrificing such a small number of coordinates will not affect a typical worst-case lower bound based on the random restriction method.

5.1.6 Lower Bounds through Reductions. Consider a reduction of PARITY to MCSP$[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$ by subquadratic-size n^{ε}-almost-formulas with $n^{\varepsilon'}$ MCSP (possibly non-local) oracles at the bottom of each principal formula computing a gate with fanout > 1. By Theorem 47, such a reduction would imply MCSP$[2^{n^{1/2}}/2n, 2^{n^{1/2}}] \notin n^{\varepsilon}$-almost-Formula$[N^{1.1}]$ assuming that after replacing all oracles by n^{ε}-almost-formulas of size $N^{1.1}$ the total size of the resulting circuit remains $< N^{2-9(\varepsilon+\varepsilon')}$. In combination with hardness magnification, this would give us NP $\not\subseteq$ NC^1. Unfortunately, Theorem 47 rules this possibility out.

Corollary 52. PARITY *is not computable by subquadratic-size* n^{ε}*-almost-formulas with* $n^{\varepsilon'}$ *oracle gates computing* MCSP$[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$*, assuming that after replacing all oracles by* n^{ε}*-almost-formulas of size* $N^{1.1}$ *the total size of the resulting circuit remains* $< N^{2-9(\varepsilon+\varepsilon')}$ *for* $\varepsilon + \varepsilon' < 1$.

Proof. Assume the reduction in question exists. By Theorem 29, for every $\varepsilon > 0$ and all sufficiently big n, MCSP$[2^{n^{1/2}}/2n, 2^{n^{1/2}}]$ is computable by $N^{1.1}$-size n^{ε}-almost-formulas with local oracles at the bottom of principal formulas computing gates with fanout > 1. By the assumption, if we replace the MCSP oracles in the reduction by almost-formulas with local oracles, then the resulting circuit is an $n^{\varepsilon+\varepsilon'}$-almost-formula of size $N^{2-9(\varepsilon+\varepsilon')}$ with oracles of bounded fan-in. This contradicts the second item of Theorem 47. □

Analogous arguments rule out the possibility of establishing strong lower bounds via reductions also in other HM frontiers.

5.2 Lower Bounds below Magnification Threshold

The localizations presented in this section show that one cannot obtain strong circuit lower bounds by "lowering the threshold" in certain hardness magnification proofs (because such hardness magnification theorems are false). In other words, the localisations of the lower-bound techniques in this section rule out a family of potential approaches for establishing strong lower bounds by magnifying an *already known* (weak) lower bound for a variant of MCSP or MKtP. As a consequence of one of our results (Theorem 60 in Section 5.2.2), we also refute the Antichecker Hypothesis from Reference [43].

5.2.1 AC^0 *Lower Bounds via Pseudorandom Restrictions.* In this section, we show that the AC^0 lower bounds proved for MCSP (MKtP) via pseudorandom restrictions [18] (see also Section 3.1) localize in a very strong sense. Consequently, this excludes magnification theorems that can be proved by approaches that unconditionally give AC^0 circuit constructions with local oracles for MCSP or MKtP.

We use AC$^0_d[O_1, O_2, \ldots, O_d]$ to denote AC^0_d circuits extended with arbitrary oracles, such that oracle gates on the ith level (the gates whose distance from the inputs is i) have fan-in at most O_i.

Theorem 53. *There is a constant c such that for all $\varepsilon > 0$, constants d, and $O_1, O_2, \ldots, O_d$ such that $\prod_{i=1}^{d} O_i \leq N/(\log N)^{\omega(1)}$, $\mathsf{MCSP}[n^c, n^{2c}] \notin \mathsf{AC}^0_d[O_1, O_2, \ldots, O_d][\mathrm{poly}(N)]$.*

Remark 54. We remark that the constraint on oracles in the above theorem is incomparable to the second item of Proposition 51. Here, we focus on the maximum oracle fan-in at each level, while there the focus is on the total fan-in of all oracles. A lower-bound result for an explicit problem with parameters similar to Theorem 53 is not known for AC^0 oracle circuits extended with parity gates (see *Reference [44]* for results in this direction).

We are going to apply Lemma 17, together with the following well-known results on k-wise independence fooling CNFs.

Lemma 55 ([7, 58]). *$k = O(\log(M/\varepsilon) \cdot \log(M))$-wise independent distribution ε-fools M-clauses CNFs.*

Combining Lemmas 17 and 55, we have the following lemma.

Lemma 56. *Let φ be a t-width M-clause CNF formula over N inputs, $p = 2^{-q}$ for some $q \in \mathbb{N}$, and $\varepsilon_0 > 0$ be a real. There is a p-regular*

$$k = \Theta(\log(M \cdot 2^{t(q+1)}/\varepsilon_0) \cdot \log(M \cdot 2^{t(q+1)}) \cdot q^{-1})\text{-wise}$$

independent random restriction $\boldsymbol{\rho}$ such that

$$\Pr_{\rho \sim \boldsymbol{\rho}}[\mathrm{DT}(\varphi|_\rho) > s] \leq 2^{s+t+1}(5pt)^s + \varepsilon_0 \cdot 2^{(s+1)(2t+\log M)}.$$

Moreover, $\boldsymbol{\rho}$ is samplable with $O(t \cdot q \cdot \mathrm{polylog}(M, N) \cdot \log(1/\varepsilon_0))$ bits, and each output coordinate of the random restriction can be computed in time polynomial in the number of random bits.

The moreover part follows from standard construction of k-wise independent distributions (see, e.g., Reference [60]).

We also need the following lemma, which states that an arbitrary oracle with inputs being small-size decision trees shrinks to a small-size decision tree with high probability, under suitable pseudorandom restrictions.

Lemma 57. *Let $O : \{0,1\}^T \to \{0,1\}$ be an arbitrary function, and $D_1, D_2, \ldots, D_T$ be T k-query decision trees on variables $x_1, x_2, \ldots, x_N$. Let $F := O \circ (D_1, D_2, \ldots, D_T)$ be their compositions. For $s \in \mathbb{N}$, and all $k(s+1)$-wise independent $1/(T \cdot k^2)$-regular random restriction $\boldsymbol{\rho}$, we have*

$$\Pr_{\rho \sim \boldsymbol{\rho}}[\mathrm{DT}(F|_\rho) > s] \leq \left(\frac{k(s+1)}{2e^2}\right)^{-(s+1)}.$$

Proof. We focus on the following particular decision tree for evaluating $\{D_1, D_2, \ldots, D_T\}$ with respect to a restriction $\rho : [N] \to \{0, 1, *\}$:

Algorithm $\mathrm{Eval}(\rho, D_1, D_2, \ldots, D_T)$.

- For i from 1 to T:
  - Simulate decision tree D_i with restriction ρ. That is, when D_i queries an index j, we feed ρ_j to D_i if $\rho_j \in \{0,1\}$, and query the jth bit otherwise.
- Let α_i be the output of the ith decision tree, we output $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_T)$.

To obtain a decision tree for $F|_\rho$, we can run $\mathrm{Eval}(\rho, D_1, D_2, \ldots, D_T)$ to obtain α first and output $F(\alpha)$ at the end.

Let $\widetilde{\mathsf{DT}}(F\restriction_\rho)$ be the query complexity of the above decision tree. Since $\mathsf{DT}(F\restriction_\rho) \le \widetilde{\mathsf{DT}}(F\restriction_\rho)$, where $\mathsf{DT}(F\restriction_\rho)$ is the minimum complexity among all decision trees computing $F\restriction_\rho$, it suffices to bound

$$\Pr_{\rho\sim\boldsymbol{\rho}}[\widetilde{\mathsf{DT}}(F\restriction_\rho) > s].$$

Consider the event that $\widetilde{\mathsf{DT}}(F\restriction_\rho) > s$, it is equivalent to that there exists a string $w \in \{0,1\}^s$, such that if we fix the first s queried unrestricted bits in ρ according to w, then Eval ends up querying $> s$ bits. (Note that since we only care about whether $\widetilde{\mathsf{DT}}(F\restriction_\rho) > s$, we can force the algorithm to abort if it tries to make the $(s+1)$th query.)

Now, suppose we fix the string w, then the number of queries made by Eval only depends on ρ. Suppose the algorithm has queried at least $s+1$ bits, we let $D'_1, D'_2, \ldots, D'_t$ ($t \le s+1$) be the decision trees in which the algorithm made queries during the first $s+1$ queries. This implies that if we run $\mathsf{Eval}(\rho, D'_1, D'_2, \ldots, D'_t)$ with respect to the same string w, the algorithm also makes at least $s+1$ queries.

Now, since $\boldsymbol{\rho}$ is $k(s+1)$-wise independent. The probability that $\mathsf{Eval}(\rho, D'_1, D'_2, \ldots, D'_t)$ makes at least $s+1$ queries with respect to the fixed string w is bounded by

$$\begin{aligned}
&(T\cdot k^2)^{-(s+1)} \cdot \binom{t\cdot k}{s+1} \\
\le &(T\cdot k^2)^{-(s+1)} \cdot \left(\frac{t\cdot k\cdot e}{s+1}\right)^{s+1} \\
\le &\left(\frac{T\cdot k\cdot (s+1)}{t\cdot e}\right)^{-(s+1)} \le \left(\frac{T\cdot k}{e}\right)^{-(s+1)}.
\end{aligned}$$

Putting everything together, we have

$$\begin{aligned}
&\Pr_{\rho\sim\boldsymbol{\rho}}[\widetilde{\mathsf{DT}}(F\restriction_\rho) > s] \\
\le &2^s \cdot \left(\frac{T\cdot k}{e}\right)^{-(s+1)} \cdot \sum_{t=0}^{s+1}\binom{T}{t} \\
\le &2^s \cdot \left(\frac{T\cdot k}{e}\right)^{-(s+1)} \cdot \left(\frac{T\cdot e}{s+1}\right)^{s+1} \\
\le &\left(\frac{k\cdot (s+1)}{2e^2}\right)^{-(s+1)}.
\end{aligned}$$

□

Remark 58. Clearly, Lemma 57 also holds when $\boldsymbol{\rho}$ is $k(s+1)$-wise independent and p-regular, for $p \le \frac{1}{T\cdot k^2}$.

Now, we are ready to prove Theorem 53.

Proof of Theorem 53. We assume N and $\log N$ are both powers of 2 for simplicity. Let $p = 1/\log^5 N$, $\varepsilon_0 = 2^{-\log^6 N}$, $s = t = 10\log^2 N$, $M = 2^s \cdot N^{\log N}$, and $\boldsymbol{\rho}$ be the k-wise independent p-regular random restriction guaranteed by Lemma 56. Note that we have $k = \omega(\log^6 N)$ and $k = \log^{O(1)} N$.

Let $C \in \mathsf{AC}^0_d[O_1, O_2, \ldots, O_d]$ be a circuit with S gates computing $\mathsf{MCSP}[n^c, n^{2c}]$. For each $i \in [d]$, let S_i be the number of gates at level i (i.e., the gates whose distance from the input gates is i). Recall that O_i is the maximum oracle fan-in at level i. We are going to prove the stronger claim that $S = \Omega(N^{\log N})$. Now, suppose for the sake of contradiction that $S \le N^{\log N}/8$.

Now, we proceed in d iterations. We will ensure that at the end of the ith iteration, all gates at level i become s-query decision trees with high probability. At the ith iteration, we apply ρ

$$\tau_i = \lceil \log_{1/p} O_i \rceil + 1$$

times. It is straightforward to see that the composition of τ_i independent restrictions from $\boldsymbol{\rho}$ is a k-wise independent p_i-regular random restriction for $p_i = p^{\tau_i} \leq \frac{1}{O_i \cdot \log^5 N}$.

Note that each oracle gate at original level i has inputs computed by s-query decision trees (at the first step, one can treat the input variables as 1-query decision trees). By Lemma 57 and noting that $k \geq s(s+1)$ and $O_i \cdot \log^5 N \geq O_i \cdot s^2$, with probability at least

$$1 - S_i \cdot \left(\frac{s(s+1)}{2e^2}\right)^{-(s+1)} \geq 1 - S_i \cdot N^{-\log N},$$

all oracle gates at level i become s-query decision trees after these τ_i restrictions.

Similarly, note that each AND/OR gate at level i are equivalent to a CNF or DNF with width-s and size at most $2^s \cdot S$. By Lemma 56, again with probability at least

$$\begin{aligned}
&1 - S_i \cdot \left(2^{s+t+1}(5pt)^s + \varepsilon_0 \cdot 2^{(s+1)(2t+\log M)}\right) \\
\geq\ &1 - S_i \cdot \Big(2^{20\log^2 N+1}(5 \cdot (1/\log^5 N) \cdot 10\log^2 N)^{10\log^2 N} \\
&\qquad + 2^{-\log^6 N} \cdot 2^{(10\log^2 N+1)(20\log^2 N+\log(N^{\log N} \cdot 2^{10\log^2 N}))}\Big) \\
\geq\ &1 - S_i \cdot N^{-\log N},
\end{aligned}$$

all AND/OR gates at level i become s-query decision tree after these τ_i restrictions.

Finally, note that in total, we have applied $\boldsymbol{\rho}$ at most

$$\tau_{\text{total}} = 2d + \log_{1/p}\left(\prod_{i=1}^{d} O_i\right) = \log_{1/p} N - \omega(1)$$

times, and the final output gate shrinks to an s-query decision tree with probability at least

$$1 - 2 \cdot S \cdot N^{-\log N}.$$

Since $S \leq N^{\log N}/8$, with probability at least 3/4, after all these restrictions, C is equivalent to an s-query decision tree.

Now let $p_{\text{end}} = p^{\tau_{\text{total}}} = N^{-1} \cdot p^{-\omega(1)}$. By Chebyshev's inequality, the number of unrestricted variables at the end of the restriction is at least $N_{\text{remain}} = \frac{1}{2} \cdot p_{\text{end}} \cdot N = (\log N)^{\omega(1)}$ with probability at least 1/2. Therefore, with probability at least 1/4, at the end of the restrictions, it holds that the remaining circuit C is equivalent to an s-query decision tree D, and the number of unrestricted variables is at least N_{remain}.

Suppose we fix all these remaining unrestricted variables to be 0 to get an input x^*, since each restriction from $\boldsymbol{\rho}$ can be computed by a $\text{poly}(n)$-size circuit, x^* has a circuit of $\text{poly}(n) \cdot \log N = \text{poly}(n) \leq n^c$ size (now, we set c). Let S be the set of input variables that D queries on the input x^*. Note that there are at least $2^{N_{\text{remain}}-|S|}$ ways of assigning values to unrestricted variables while keeping variables in S all 0. And, we can see that F's output on x^* is the same as its output on all of these assignments. But there must exist at least one assignment such the MCSP value is at least $(\log N)^{2c} = n^{2c}$ ($2^{N_{\text{remain}}-|S|} = 2^{n^{\omega(1)}}$), contradiction to the assumption that C computes $\text{MCSP}[n^c, n^{2c}]$. □

5.2.2 The Nearly Quadratic-formula Lower Bound of Reference [26]. In this section, we prove that the nearly quadratic-formula lower bound of Reference [26] localizes, and thereby proving the third item of Theorem 2. This localization indeed refutes a family of possible approaches to establish circuit lower bounds through hardness magnification via "lowering the threshold."

More concretely, consider the following hypothesized approach. Suppose we can compute $\mathsf{MCSP}[2^{\sqrt{n}}]$ by a formula F with NP oracles, such that when we replace every oracle O with fan-in β in F by a formula of size β^k that reads all its inputs exactly β^{k-1} times, the size of the new formula is less than $N^{1.99}$. Then, we know that NP cannot be computed by formulas of size n^k that reads all its inputs exactly n^{k-1} times, as otherwise, we get an $N^{1.99}$-size formula for $\mathsf{MCSP}[2^{\sqrt{n}}]$, which is a contradiction to the lower bound in Reference [26]. If this holds for all $k > 0$, then we would have $\mathsf{NP} \not\subset \mathsf{Formula}[n^k]$ for all k.

In the following, by localizing Reference [26], we show that there is no such oracle formula construction for MCSP even if the oracles can be arbitrary. This excludes magnification theorems obtained by approaches that unconditionally produce circuits with oracles, and essentially addresses a question from Reference [43]. It also suggests that the consideration of almost-formulas in HM Frontier C is unavoidable.

A Size Measure on Oracle Formulas and a Potential Approach to Formula-size Lower Bound

We first introduce a size measure Size_t on oracle formulas to formalize the previous discussion.

For a parameter t and an oracle formula F, we define $\mathsf{Size}_t(F)$ as the size of the formula, if we replace every oracle O with fan-in β in F by a formula of size β^t that reads all its inputs exactly β^{t-1} times.

More formally,

$$\mathsf{SIZE}_t(F) := \begin{cases} \mathsf{SIZE}_t(F_1) + \mathsf{SIZE}_t(F_2) & F = F_1 \wedge F_2 \text{ or } F = F_1 \vee F_2, \\ \beta^{t-1} \cdot \left(\sum_{i=1}^{\beta} \mathsf{SIZE}_t(F_i)\right) & F = O(F_1, F_2, \ldots, F_\beta). \end{cases}$$

Proposition 59. *For a constant $k > 0$, if there is an* NP *oracle formula F (all oracles are languages in* NP*) for* $\mathsf{MCSP}[2^{\sqrt{n}}]$ *such that* $\mathsf{SIZE}_{k+1}(F) \leq N^{2-\varepsilon}$ *for a constant $\varepsilon > 0$, then* $\mathsf{NP} \not\subseteq \mathsf{Formula}[n^k]$.

Proof. Suppose $\mathsf{NP} \subseteq \mathsf{Formula}[n^k]$ for the sake of contradiction. Then, in particular, each NP language can be computed by a size-n^{k+1} formula that reads all its inputs exactly n^k times by adding some dummy nodes in the formula. Therefore, by replacing all NP oracles in F by such formulas, we have an $N^{2-\varepsilon}$-size formula for $\mathsf{MCSP}[2^{\sqrt{n}}]$, in contradiction to the lower bound in Reference [26]. □

Localization of Reference [26]

Our following theorem shows that the above approach is not viable even with $k = 3$ by localizing Reference [26], with a moderate constraint on the adaptivity of the oracle circuits.

Theorem 60. *There is a universal constant c such that for all constants $\varepsilon > 0$ and $\alpha > 2$,* $\mathsf{MCSP}[n^c, 2^{(\varepsilon/\alpha)\cdot n}]$ *cannot be computed by oracle formulas F with* $\mathsf{SIZE}_3(F) \leq N^{2-\varepsilon}$ *and adaptivity* $o(\log N / \log\log N)$ *(that is, on any path from root to a leaf, there are at most* $o(\log N / \log\log N)$ *oracles).*

Remark 61. It is not hard to see that the adaptivity can be at most $O(\log N)$ given the condition $\mathsf{SIZE}_3(F) \leq N^{2-\varepsilon}$.

Before proving Theorem 60, we first show it refutes the Antichecker Hypothesis (restated below) from Reference [43].

The Antichecker Hypothesis. *For every $\lambda \in (0,1)$, there are $\varepsilon > 0$ and a collection $\mathcal{Y} = \{Y_1, \ldots, Y_\ell\}$ of sets $Y_i \subseteq \{0,1\}^n$, where $\ell = 2^{(2-\varepsilon)n}$ and each $|Y_i| = 2^{n^{1-\varepsilon}}$, for which the following holds.*

If $f : \{0,1\}^n \mapsto \{0,1\}$ and $f \notin \mathsf{Circuit}[2^{n^\lambda}]$, then some set $Y \in \mathcal{Y}$ forms an antichecker for f: For each circuit C of size $2^{n^\lambda}/10n$, there is an input $y \in Y$ such that $C(y) \neq f(y)$.

Corollary 62. *The Antichecker Hypothesis is false.*

Proof. It is easy to see that, assuming the Antichecker Hypothesis, we can solve $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ with a formula F of $N^{2-\varepsilon}$ size that uses $N^{2-\varepsilon}$ oracles of fan-in $\mathrm{poly}(n)2^{n^{1-\varepsilon}} = \mathrm{polylog}(N) \cdot 2^{(\log N)^{1-\varepsilon}} = N^{o(1)}$ only at the layer above the inputs, for some $\varepsilon > 0$. However, since $\mathsf{SIZE}_3(F) \leq N^{2-\varepsilon+o(1)}$, F cannot compute $\mathsf{MCSP}[2^{n^{1/3}}, 2^{n^{2/3}}]$ by Theorem 60, contradiction. □

Now, we are ready to prove Theorem 60.

Proof of Theorem 60. Let $k = \log^3 N$, and $\boldsymbol{\rho}$ be the k-wise independent $(1/\sqrt{k})$-regular random restriction guaranteed by Lemma 12.

For an oracle formula F and a subformula G of it, we say G is a maximal subformula if G is an entire subtree rooted at either the root, an oracle gate, or a gate whose father is an oracle.

We are going to apply $t = \Theta(\log_k N)$ independent pseudorandom restrictions $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_t$, each distributed identically to $\boldsymbol{\rho}$, where t will be set precisely later.

The Overall Proof Structure

To analyze the size of the oracle formula under the random restriction sequence $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_t$, we define a potential function Φ inductively for all maximal subformulas of the given formula F. As it will be clear from the definition, Φ is not only a function of the structure of the oracle formula but also depends on the history of the pseudorandom restrictions.

Formally, for each maximal subformula G of the given formula F, and for each integer $0 \leq i \leq t$, we define a random variable $\Phi_{G,i}$, which denotes the potential function of G after the first i pseudorandom restrictions and only depends on $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_i$.

Definition of Tiny Formulas and Blow up. For an oracle formula, if the top gate is an oracle, then we say it is tiny if it depends on at most $\log N$ variables. Otherwise, we say it is tiny if it depends on at most $c_{\mathsf{tiny}} \cdot k$ variables, for a constant c_{tiny} to be specified later.

After each pseudorandom restriction, for a formula with an oracle gate at the top, when it depends on at most $b = 20$ variables, we blow it up to a formula of size $B = 2^b$ (note that if there are two oracle gates u, v such that u and v both depend on at most b variables and u is an ancestor of v, then it suffices to only blow up u).

The above two definitions (tiny formulas and the process of blowing up) may not seem easy to understand at first. Let us explain the motivation behind them. The key difficulty of the proof is to handle the oracle gates properly. The process of blowing up ensures that whenever an oracle becomes too small, we just replace it with a constant size normal formula, so it becomes easier to deal with.

The definition of tiny formulas is more subtle. As it will be clear in Case II and Case III of the inductive definition of Φ, setting the threshold of being tiny to $\log N$ for oracle formulas with top oracle gates ensures that the corresponding event of becoming tiny happens with high probability, which is indeed crucial in our proof.

The properties of Φ. We require the following properties on Φ.

(1) For an oracle formula F, Φ is multiplied by a factor of $\frac{c_F}{k}$ under $\boldsymbol{\rho}$ in expectation, where c_F depends on F but it is upper bounded by a universal constant.
(2) With probability $1 - p_F$, for all stages, and all maximal subformulas G of F, $\Phi = 0$ for G implies that G is tiny, where p_F depends on F but it is upper bounded by N^{-2}.
(3) It holds that either $\Phi = 0$ or $\Phi \geq 1$. Together with the second item, it implies that if the oracle formula is not tiny then $\Phi \geq 1$.

With these carefully designed properties of Φ, the overall proof is straightforward. We first show that Φ of F is closely related to $\mathsf{SIZE}_3(F)$, and our conditions on the oracle formula imply that Φ of the whole oracle formula is bounded by $N^{2-\varepsilon+o(1)}$ at the beginning. Then after roughly $t \approx \log_k(N^{2-\varepsilon+o(1)})$ rounds of restrictions from $\boldsymbol{\rho}$, Φ becomes 0 with a good probability, which also implies the whole oracle formula becomes tiny (only depend on polylog(N) bits).

But then, we argue that after t rounds of restrictions from $\boldsymbol{\rho}$, with high probability the number of unrestricted variables is still at least $N^{\Omega(1)}$. Using a similar argument as from References [26, 43, 46], we show that the tiny oracle formula left behind cannot compute $\mathsf{MCSP}[n^c, 2^{\varepsilon/\alpha \cdot n}]$ on the remaining variables, which concludes the proof.

The Inductive Definition of the Potential Function Φ

In the following, we gradually develop the definition of the potential function Φ. We remark that Cases I and II below are actually special cases of Cases III and IV, respectively. We discuss them first in the hope that they provide some intuitions and make it easier to understand the more complicated Cases III and IV.

Case I: Φ for a Pure Formula. We begin with the simplest case of pure formulas F (formulas with no oracles) of size S. We define

$$\Phi = \begin{cases} S & S \geq 100 \cdot k, \\ 0 & \text{otherwise.} \end{cases}$$

It follows from the shrinkage lemma [23], formula decomposition [56, Claim 6.2], and the k-wise independence of $\boldsymbol{\rho}$ that, when $S \geq 100 \cdot k$, the *expected size* of S drops by a factor of at least k/c_{Tal}, for a universal constant c_{Tal} (we can set $c_F = c_{\mathsf{Tal}}$). Otherwise, the formula is tiny. It is straightforward to verify that all three properties of Φ are satisfied (we can set $p_F = 0$ in this case).

Case II: Φ for a Pure Oracle. Next, we consider the case that F is a pure oracle O with fan-in T (pure oracle means each input to O is just a variable). We set

$$\Phi = T^2 \cdot k^3$$

at the beginning. And set $\Phi \leftarrow \Phi/k$ after each $\boldsymbol{\rho}$. Whenever it happens $\Phi < 1$, we set $\Phi = 0$ afterwards. Here, we can simply set $c_F = 1$.

Now, we argue that with probability at least $1 - N^{-5}$ (that is, we set $p_F = N^{-5}$), when $\Phi = 0$, O only depends on at most $\log N$ variables and therefore becomes tiny.

Note that $\Phi = 0$ means at least $\log_k T^2$ rounds of random restrictions have been applied.[17] Their composition is a k-wise independent restriction that keeps a variable unrestricted with probability at most T^{-1}. Therefore, the probability that the number of alive variable is larger than

[17] Note that for this argument, a potential function of $T \cdot k$ already suffices. We use $T^2 \cdot k^3$ here to make it consistent with Case III.

$\log N$ is smaller than

$$\binom{T}{\log N} \cdot T^{-\log N} \leq \left(\frac{e \cdot T}{\log N}\right)^{\log N} \cdot T^{-\log N} \leq \left(\frac{e}{\log N}\right)^{\log N} \leq N^{-5}.$$

Note that in the above inequalities, we can safely assume $T > \log N$.

Case III: Φ for an Oracle Formula with an Oracle Top Gate. Then, we move to the case of a maximal subformula F with an oracle top gate O with fan-in T. Let Φ_i be the corresponding potential function of the maximal subformula with root being the ith input to O. We set

$$\Phi = \max\left(\sum_{i=1}^{T} \Phi_i, 1/k\right) \cdot T^2 \cdot k^4,$$

at the beginning.

When $\sum_{i=1}^{T} \Phi_i > 0$, we still let $\Phi = (\sum_{i=1}^{T} \Phi_i) \cdot T^2 \cdot k^4$. When $\sum_{i=1}^{T} \Phi_i$ first becomes 0 (this could happen before the first restriction, if $\sum_{i=1}^{T} \Phi_i = 0$ at the beginning), we set $\Phi = T^2 \cdot k^3$ and decrease it by a factor of k during each later restriction, and set it to 0 if it becomes < 1.

Here, we set c_F to be the maximum of $c_{F'}$ for all maximal subformulas F' whose root is an input to the top oracle gate O in F.

First let us argue that Φ is multiplied by a factor of $\frac{c_F}{k}$ after each $\boldsymbol{\rho}$ in expectation. When $\sum_{i=1}^{T} \Phi_i = 0$, it is evident from the way we set Φ (note that $c_F \geq 1$). When $\sum_{i=1}^{T} \Phi_i > 0$, it follows from the induction as each Φ_i is multiplied by a factor of $\frac{c_F}{k}$ after each $\boldsymbol{\rho}$ in expectation. In the borderline case when $\sum_{i=1}^{T} \Phi_i > 0$ before $\boldsymbol{\rho}$ and becomes 0 afterwards. One can see Φ drops from at least $T^2 \cdot k^4$ to at most $T^2 \cdot k^3$.

Moreover, when $\sum_{i=1}^{T} \Phi_i = 0$, with probability at least $1 - \sum_{i=1}^{T} p_{F_i}$ (F_i is the ith subformula whose root is an input to the top oracle gate O in F) all the subformulas are tiny, so at this time the oracle depends on at most $O(T \cdot k)$ variables.

Therefore, when Φ drops to 0, with probability at least $1 - \sum_{i=1}^{T} p_{F_i} - N^{-5}$ the whole oracle formula becomes tiny, by a calculation similar to the pure oracle case. Therefore, we can set $p_F = \sum_{i=1}^{T} p_{F_i} + N^{-5}$.

Case IV: Φ for a Formula with Oracle Leaves. Finally, we deal with the most complicated case when the maximal subformula F is a formula with oracle leaves. Suppose F is a formula of size S with m oracle leaves. Let Φ_i be the potential function of the subformula corresponding to the ith oracle leaf. Also, let c_{drop} be the maximum of the c_F's of all the subformulas corresponding to the oracle leaves.

The difficulty in analyzing this case is that there could be many oracles that are tiny but have not blown up yet, and we have to keep track of the number of such oracles. Let N_{active} be the number of remaining active tiny oracles (oracles that are tiny but have not blown up). Clearly, $N_{\text{active}} \leq S$ at the beginning.

We set

$$\Phi = S + N_{\text{active}} \cdot k^2 + \sum_{i=1}^{m} \Phi_i \cdot k^4$$

at the beginning. When $S \leq 100 \cdot k$ happens, we change Φ to be

$$N_{\text{active}} \cdot k^2 + \sum_{i=1}^{m} \Phi_i \cdot k^4$$

afterwards.

After each $\boldsymbol{\rho}$, if $S \geq 100 \cdot k$, the expected size of S becomes at most

$$c_1 \cdot S/k + c_2 \cdot k \cdot \left(\sum_{i=1}^{m} \Phi_i + N_{\text{active}} \right),$$

for two universal constants c_1 and c_2. This bound holds because, by Claim 4.4 of Reference [27], a formula of size S can be decomposed into $6S/k$ subformulas, each of size at most k, and each formula has at most two subformula children.

The number of active oracle leaves (that are not blown up) is at most $\sum_{i=1}^{m} \Phi_i + N_{\text{active}}$. Hence, at least $6 \cdot S/k - \sum_{i=1}^{m} \Phi_i - N_{\text{active}}$ subformulas do not contain an active oracle leaf, and their total expected size is $O(S/k)$ after $\boldsymbol{\rho}$ (by Lemmas 4.1 and 4.3 of Reference [27], and Reference [56]). For those subformulas containing active oracle leaves, their total size is at most $(\sum_{i=1}^{m} \Phi_i + N_{\text{active}}) \cdot O(k)$ after $\boldsymbol{\rho}$ (this takes account of the worst case situation that all these active oracle leaves blow up).

Also, we can see that after $\boldsymbol{\rho}$, N_{active} becomes at most

$$N_{\text{active}}/k^2 + \sum_{i=1}^{m} \Phi_i$$

in expectation. This is because for a tiny active oracle depending on at most $\log N$ variables, the probability that it does not blow up after $\boldsymbol{\rho}$ is at most

$$\binom{\log N}{b} \cdot k^{-b/2} \leq (\log N)^{b-1.5 \cdot b} = (\log N)^{-10} \leq 1/k^2.$$

By induction, we also have that $\sum_{i=1}^{m} \Phi_i$ is multiplied by a factor of $\frac{c_{\text{drop}}}{k}$ in expectation as well after each $\boldsymbol{\rho}$. Therefore, after $\boldsymbol{\rho}$, the expectation of Φ can be bounded by

$$\begin{aligned} & c_1 \cdot S/k + c_2 k \left(\sum_{i=1}^{m} \Phi_i + N_{\text{active}} \right) + \left(N_{\text{active}}/k^2 + \sum_{i=1}^{m} \Phi_i \right) \cdot k^2 + \left(\sum_{i=1}^{m} \Phi_i \right) \cdot \frac{c_{\text{drop}}}{k} \cdot k^4 \\ \leq \; & S \cdot \frac{c_1}{k} + N_{\text{active}} \cdot k^2 \cdot \frac{c_2 + 1/k}{k} + \sum_{i=1}^{m} \Phi_i \cdot k^4 \cdot \left(\frac{c_{\text{drop}} + c_2/k^2 + 1/k}{k} \right). \end{aligned}$$

We can set

$$c_F = \max(c_1, c_2 + 1/k, c_{\text{drop}} + c_2/k^2 + 1/k).$$

Recall that when $S \leq 100 \cdot k$ happens, we change Φ to be

$$N_{\text{active}} \cdot k^2 + \sum_{i=1}^{m} \Phi_i \cdot k^4$$

afterwards.

By the previous discussion, after this Φ still drops by a factor of k/c_F in expectation after each $\boldsymbol{\rho}$. Note that when $\Phi = 0$, we can see the size of the whole formula is smaller than $B \cdot 100 \cdot k = O(k)$, therefore it is tiny (here, we set $c_{\text{tiny}} = B \cdot 100$). This is because $\Phi = 0$ implies $S \leq 100 \cdot k$ happened at some point, and also $N_{\text{active}} = \sum_{i=1}^{m} \Phi_i = 0$. They together imply that all oracles have blown up, and the size bound follows, since each oracle adds at most B leaves.

Let F_i be the subformula with root being the ith oracle leaf. In this case, we can set $p_F = \sum_{i=1}^{m} p_{F_i}$.

The MCSP Lower Bound

Let F be an oracle formula with $\mathsf{SIZE}_3(F) \le N^{2-\varepsilon}$ and adaptivity $\tau = o(\log N/\log\log N)$. We first need to verify that c_F is upper bounded by a universal constant. One can upper bound

$$c_F \le \max(c_1, c_2 + 1/k, c_{\mathsf{Tal}}) + \tau \cdot (c_2/k^2 + 1/k) \le \max(c_1, c_2 + 1/k, c_{\mathsf{Tal}}) + o(1) = O(1).$$

We can also upper bound p_F by $p_F \le N^{-5} \cdot N^2 = N^{-3}$.

By the inductive definition of the potential function Ψ on maximal subformulas, it is not hard to show that

$$\Phi \le \mathsf{SIZE}_3(F) \cdot k^{O(\tau)} \le N^{2-\varepsilon+o(1)}.$$

Note that this inequality crucially employs the definition of $\mathsf{SIZE}_3(\cdot)$.

After each $\boldsymbol{\rho}$, Φ is reduced by a factor of k/c_F. After

$$t = \lceil \log_{k/c_F} \Phi \rceil + 2$$

rounds of $\boldsymbol{\rho}$, the expected Φ of the overall formula becomes $< 1/10$, which means with probability $0.9 - p_F \ge 0.8$ it is tiny and only depends on at most $O(k) = O(\log^3 N)$ variables.

Note that by definition

$$(k/c_F)^t \le \Phi \cdot k^3,$$

and therefore

$$k^t \le \Phi \cdot k^3 \cdot (c_F)^t \le N^{2-\varepsilon+o(1)},$$

as $(c_F)^t = (c_F)^{O(\log N/\log\log N)} = N^{o(1)}$.

The composition of t independent $\boldsymbol{\rho}$ keeps a variable unrestricted with probability $k^{-t/2} \ge N^{-1+\varepsilon/2-o(1)}$, and is clearly pairwise independent. By Chebyshev's inequality, after t restrictions from $\boldsymbol{\rho}$, with probability 0.5, at least

$$1/2 \cdot N \cdot N^{-1+\varepsilon/2-o(1)} \ge N^{\varepsilon/2-o(1)}$$

variables remain active. So with probability at least 0.3, after t restrictions from $\boldsymbol{\rho}$, the remaining formula F only depends on $O(\log^3 N)$ variables, and the number of remaining unrestricted variables is at least $N^{\varepsilon/2-o(1)}$.

Suppose we fix all these remaining unrestricted variables to be 0 to get an input x^*. Since each restriction from $\boldsymbol{\rho}$ can be computed by a $\mathrm{poly}(n)$-size circuit, x^* has a circuit of $\mathrm{poly}(n) \cdot t = \mathrm{poly}(n) \le n^c$ size (here, we set c). Let S be the set of input variables that F depends on. Note that there are at least $2^{N^{\varepsilon/2-o(1)}-|S|}$ ways of assigning values to unrestricted variables while keeping variables in S all 0. Since F only depends on S, F's output on x^* is the same as its output on all of these assignments. But there must exist at least one assignment such the MCSP value is at least $N^{\varepsilon/\alpha} = 2^{(\varepsilon/\alpha)\cdot n}$ as $\alpha > 2$. Therefore, F cannot compute $\mathrm{MCSP}[n^c, 2^{(\varepsilon/\alpha)\cdot n}]$. □

APPENDIX

A REVIEW OF HARDNESS MAGNIFICATION IN CIRCUIT COMPLEXITY

A.1 Previous Work

We focus on some representative examples. For definitions and more details, check Section 2 or consult the original papers.

Srinivasan [55] (Informal). If there exists $\varepsilon > 0$ such that $n^{1-o(1)}$-approximating MAX-CLIQUE requires Boolean circuits of size at least $m^{1+\varepsilon}$ (where $m = \Theta(n^2)$), then $\mathrm{NP} \not\subseteq \mathrm{Circuit}[\mathrm{poly}]$.

Allender-Koucký [3] and Chen-Tell [14]. The following results hold.

- Let $\Pi \in \{\text{BFE}, \text{W}_{\text{S}_5}, \text{W5-STCONN}\}$. Suppose that for each $c > 1$ there exist infinitely many $d \in \mathbb{N}$ such that TC^0 circuits of depth d require more than $n^{1+c^{-d}}$ wires to solve Π. Then, $\text{NC}^1 \not\subseteq \text{TC}^0$.
- Suppose that for each $c > 1$ there exist infinitely many $d \in \mathbb{N}$ such that MAJ cannot be computed by ACC^0 circuits of depth d with $n^{1+c^{-d}}$ wires. Then $\text{MAJ} \notin \text{ACC}^0$, and consequently $\text{TC}^0 \not\subseteq \text{ACC}^0$.

Lipton-Williams [39]. If there is $\varepsilon > 0$ such that for every $\delta > 0$ we have $\text{CircEval} \notin \text{Size-Depth}[n^{1+\varepsilon}, n^{1-\delta}]$, then for every $k \geq 1$ and $\gamma > 0$ we have $\text{CircEval} \notin \text{Size-Depth}[n^k, n^{1-\gamma}]$ (in particular $\text{P} \not\subseteq \text{NC}$).

Oliveira-Santhanam [46]. The following results hold.

- Let $s(n) = n^k$ and $\delta(n) = n^{-k}$, where $k \in \mathbb{N}$. If $\text{MCSP}[(s,0),(s,\delta)] \notin \text{Formula}[N^{1+\varepsilon}]$ for some $\varepsilon > 0$, then there is $L \in \text{NP}$ over m-bit inputs and $\delta > 0$ such that $L \notin \text{Formula}[2^{m^\delta}]$.
- Suppose there exists $k \geq 1$ such that for every $d \geq 1$ there is $\varepsilon_d > 0$ such that $\text{MCSP}[(s,0),(s,\delta)] \notin \text{AC}^0_d[N^{1+\varepsilon_d}]$, where $s(n) = n^k$ and $\delta(n) = n^{-k}$. Then $\text{NP} \not\subseteq \text{NC}^1$.
- Let $k(n) = n^{o(1)}$. If there exists $\varepsilon > 0$ such that k-Vertex-Cover $\notin \text{DTISP}[m^{1+\varepsilon}, m^{o(1)}]$, where the input is an n-vertex graph represented by an adjacency matrix of bit length $m = \Theta(n^2)$, then $\text{P} \neq \text{NP}$.
- Let $k(n) = (\log n)^C$, where $C \in \mathbb{N}$ is arbitrary. If for every $d \geq 1$ there exists $\varepsilon > 0$ such that k-Vertex-Cover $\notin \text{AC}^0_d[m^{1+\varepsilon}]$, then $\text{NP} \not\subseteq \text{NC}^1$.

Oliveira-Pich-Santhanam [43] and McKay-Murray-Williams [40] (Informal). If there exists $\varepsilon > 0$ such that for every small enough $\beta > 0$,

- $\text{MCSP}[2^{\beta n}] \notin \text{Circuit}[N^{1+\varepsilon}]$, then $\text{NP} \not\subseteq \text{Circuit}[\text{poly}]$.
- $\text{MKtP}[2^{\beta n}] \notin \text{TC}^0[N^{1+\varepsilon}]$, then $\text{EXP} \not\subseteq \text{TC}^0[\text{poly}]$.
- $\text{MKtP}[2^{\beta n}] \notin U_2\text{-Formula}[N^{3+\varepsilon}]$, then $\text{EXP} \not\subseteq \text{Formula}[\text{poly}]$.
- $\text{MKtP}[2^{\beta n}] \notin B_2\text{-Formula}[N^{2+\varepsilon}]$, then $\text{EXP} \not\subseteq \text{Formula}[\text{poly}]$.
- $\text{MKtP}[2^{\beta n}] \notin \text{Formula-XOR}[N^{1+\varepsilon}]$, then $\text{EXP} \not\subseteq \text{Formula}[\text{poly}]$.
- $\text{MKtP}[2^{\beta n}] \notin \text{BP}[N^{2+\varepsilon}]$, then $\text{EXP} \not\subseteq \text{BP}[\text{poly}]$.
- $\text{MKtP}[2^{\beta n}] \notin (\text{AC}^0[6])[N^{1+\varepsilon}]$, then $\text{EXP} \not\subseteq \text{AC}^0[6]$.

Many results for MKtP admit analogues for MrKtP, which considers a randomized version of Kt complexity introduced by Reference [42]. An advantage of MrKtP is that strong unconditional lower bounds against uniform computations are known, while the hardness of problems such as MCSP and MKtP currently relies on cryptographic assumptions.

Reference [40] also show magnification from weak lower bounds against one-pass streaming algorithms to separating P and NP, which was later observed to extend to weak lower bounds against one-tape Turing machines by Reference [17].

Chen-McKay-Murray-Williams [13] and Chen-Jin-Williams [11] (Informal). The following results hold.

- If there is $\varepsilon > 0$, $c \geq 1$, and an n^c-sparse language $L \in \text{NP}$ such that $L \notin \text{Circuit}[n^{1+\varepsilon}]$, then $\text{NE} \not\subseteq \text{Circuit}[2^{\delta \cdot n}]$ for some $\delta > 0$.
- If there is $\varepsilon > 0$ such that for every $\beta > 0$ there is a 2^{n^β}-sparse language $L \in \text{NTIME}[2^{n^\beta}]$ such that $L \notin \text{Circuit}[n^{1+\varepsilon}]$, then $\text{NEXP} \not\subseteq \text{Circuit}[\text{poly}]$.

More recently, Reference [11] established that many hardness magnification theorems for problems such as MCSP and MKtP hold in fact under the assumption that a *sufficiently sparse and explicit family of languages in* NP admits weak lower bounds in many models previously considered. We refer to their work for more details.

Chen-Jin-Williams [12] If there exists $\varepsilon > 0$ such that for every small enough $\beta > 0$,

- $\mathsf{MCSP}[2^{\beta n}]$ does not have $N^{2+\varepsilon}$-size probabilistic formulas, then $\oplus\mathsf{P} \not\subseteq \mathsf{NC}^1$.
- $\mathsf{MKtP}[2^{\beta n}]$ does not have $N^{2+\varepsilon}$-size probabilistic formulas, then $\mathsf{EXP} \not\subseteq \mathsf{NC}^1$.

Interestingly, Reference [12] also shows an *unconditional lower bound* for MCSP and MKtP against $N^{2-\varepsilon}$-sized probabilistic formulas (with the same size parameter). In other words, they show an arbitrarily small gap between the magnification threshold and known lower bounds for probabilistic formulas, which is known for very few other magnification results.

Hardness Magnification results have also been established in related areas, such as Frege lower bounds in proof complexity [41] and non-commutative arithmetic circuit lower bounds [9].

A.2 Hardness Magnification through the Lens of Oracle Circuits

We can view the results from Appendix A.1 as *unconditional* upper bounds on the size of small fan-in oracle circuits solving the corresponding problems, for a certain choice of oracle gates. In a magnification theorem, it is important to upper bound the uniform complexity of the oracle gates. For our discussion, this is not going to be relevant.

We repeat here a definition from Section 2, for convenience of the reader.

Definition 63 (Local Circuit Classes). Let C be a circuit class (such as $\mathsf{AC}^0[s]$, $\mathsf{TC}^0_d[s]$, $\mathsf{Circuit}[s]$, etc). For functions $q, \ell, a \colon \mathbb{N} \to \mathbb{N}$, we say that a language L is in $[q, \ell, a]\text{–}C$ if there exists a sequence $\{E_n\}$ of oracle circuits for which the following holds:

(i) Each oracle circuit E_n is a circuit from C.
(ii) There are at most $q(n)$ oracle gates in E_n, each of fan-in at most $\ell(n)$, and any path from an input gate to an output gate encounters at most $a(n)$ oracle gates.
(iii) There exists a language $O \subseteq \{0,1\}^*$ such that the sequence $\{E_n^O\}$ (E_n with its oracle gates set to O) computes L.

In the definition above, q stands for *quantity*, ℓ for *locality*, and a for *adaptivity* of the corresponding oracle gates.

The fact that existing magnification theorems produce such circuits is a consequence of the algorithmic nature of the underlying proofs, which show how to reduce an instance of a problem to shorter instances of another related problem. By inspection of each proof, it is possible to establish a variety of upper bounds. We explicitly state some of them below.

Proposition 64. *The following results hold.*

- [3] *For every* $\Pi \in \{\mathsf{BFE}, \mathsf{W}_{\mathsf{S}_5}, \text{W5-STCONN}\}$ *and every* $\beta > 0$, $\Pi_n \in [O(n^{1-\beta}), n^\beta, O(\frac{1}{\beta})]\text{–}\mathsf{TC}^0[O(n)]$.
- [39] *For every* $\delta > 0$, $\mathsf{CircEval}_n \in [n \cdot \mathrm{poly}(\log n), n^\delta, n^{1-\delta}]\text{–}\mathsf{Circuit}[n \cdot \mathrm{poly}(\log n)]$.
- [46] *For every constructive function* $n \le s(n) \le 2^n/\mathrm{poly}(n)$ *and parameter* $0 < \delta(n) < 1/2$, $\mathsf{MCSP}[(s,0),(s,\delta)] \in [N, \mathrm{poly}(s/\delta), 1]\text{–}\mathsf{Formula}[N \cdot \mathrm{poly}(s/\delta)]$.
- [46] *Let* $k = (\log n)^C$, *where* $C \in \mathbb{N}$. *Then* k-Vertex-Cover $\in [1, (\log n)^{4C}, 1]\text{–}\mathsf{AC}^0_d[m^{1+\varepsilon}]$, *where* $\varepsilon_d \to 0$ *as* $d \to \infty$.
- [43] *For every* $\beta > 0$ *and for every constructive function* $s(n) \le 2^{\beta n}$, Gap-$\mathsf{MKtP} \in [N, \mathrm{poly}(s), 1]\text{–}\mathsf{Formula\text{-}XOR}[N \cdot \mathrm{poly}(s)]$.

- [43] *For every constructive function* $s(n) \leq 2^n/\mathrm{poly}(n)$*, it follows that* Gap-MCSP $\in [N \cdot \mathrm{poly}(s), \mathrm{poly}(s), \mathrm{poly}(s)]-\mathrm{Circuit}[N \cdot \mathrm{poly}(s)]$.
- [40] *For every constructive function* $s(n) \leq 2^n/\mathrm{poly}(n)$*, we have* $\mathrm{MCSP}[s(n)] \in [O(N/\mathrm{poly}(s)), \mathrm{poly}(s), O(n/\log(s))]-\mathrm{Circuit}[N/\mathrm{poly}(s)]$.

We stress, however, that not every hardness magnification theorem needs to lead to an unconditional construction of efficient oracle circuits. (All the proofs that we know of produce such circuits though.)

ACKNOWLEDGMENTS

Part of this work was completed while some of the authors were visiting the Simons Institute for the Theory of Computing. We are grateful to the Simons Institute for their support. We would like to thank anonymous reviewers for their helpful comments.

REFERENCES

[1] Miklós Ajtai. 1983. Σ_1^1-formulae on finite structures. *Ann. Pure Appl. Logic* 24, 1 (1983), 1–48.
[2] Eric Allender, Lisa Hellerstein, Paul McCabe, Toniann Pitassi, and Michael E. Saks. 2008. Minimizing disjunctive normal form formulas and AC^0 circuits given a truth table. *SIAM J. Comput.* 38, 1 (2008), 63–84. https://doi.org/10.1137/060664537
[3] Eric Allender and Michal Koucký. 2010. Amplifying lower bounds by means of self-reducibility. *J. ACM* 57, 3 (2010), 14:1–14:36.
[4] Alexander E. Andreev and Stasys Jukna. 2008. Very large cliques are easy to detect. *Discrete Math.* 308, 16 (2008), 3717–3721. https://doi.org/10.1016/j.disc.2007.07.036
[5] Benny Applebaum, Boaz Barak, and David Xiao. 2008. On basing lower-bounds for learning on worst-case assumptions. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS'08)*. 211–220.
[6] Sanjeev Arora and Boaz Barak. 2009. *Computational Complexity: A Modern Approach*. Cambridge University Press.
[7] Louay M. J. Bazzi. 2009. Polylogarithmic independence can fool DNF formulas. *SIAM J. Comput.* 38, 6 (2009), 2220–2272. https://doi.org/10.1137/070691954
[8] Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonina Kolokolova. 2016. Learning algorithms from natural proofs. In *Proceedings of the Conference on Computational Complexity (CCC'16)*. 10:1–10:24. https://doi.org/10.4230/LIPIcs.CCC.2016.10
[9] Marco L. Carmosino, Russell Impagliazzo, Shachar Lovett, and Ivan Mihajlin. 2018. Hardness amplification for non-commutative arithmetic circuits. In *Proceedings of the 33rd Computational Complexity Conference (CCC'18) (LIPIcs)*, Rocco A. Servedio (Ed.), Vol. 102. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 12:1–12:16. https://doi.org/10.4230/LIPIcs.CCC.2018.12
[10] Arkadev Chattopadhyay and Rahul Santhanam. 2012. Lower bounds on interactive compressibility by constant-depth circuits. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS'12)*. 619–628.
[11] Lijie Chen, Ce Jin, and Ryan Williams. 2019. Hardness magnification for all sparse NP Languages. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS'19)*.
[12] Lijie Chen, Ce Jin, and R. Ryan Williams. 2020. Sharp threshold results for computational complexity. In *Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC'20)*, Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy (Eds.). ACM, 1335–1348. https://doi.org/10.1145/3357713.3384283
[13] Lijie Chen, Dylan M. McKay, Cody D. Murray, and R. Ryan Williams. 2019. Relations and equivalences between circuit lower bounds and Karp-Lipton theorems. In *Proceedings of the Computational Complexity Conference (CCC'19)*.
[14] Lijie Chen and Roei Tell. 2019. Bootstrapping results for threshold circuits "just beyond" known lower bounds. In *Proceedings of the Symposium on Theory of Computing (STOC'19)*.
[15] Xi Chen, Igor Carboni Oliveira, and Rocco A. Servedio. 2017. Addition is exponentially harder than counting for shallow monotone circuits. In *Proceedings of the Symposium on Theory of Computing (STOC'17)*. 1232–1245.
[16] Mahdi Cheraghchi, Elena Grigorescu, Brendan Juba, Karl Wimmer, and Ning Xie. 2018. $AC^0 \circ MOD_2$ lower bounds for the Boolean Inner Product. *J. Comput. Syst. Sci.* 97 (2018), 45–59.
[17] Mahdi Cheraghchi, Shuichi Hirahara, Dimitrios Myrisiotis, and Yuichi Yoshida. 2021. One-tape turing machine and branching program lower bounds for MCSP. In *Proceedings of the 38th International Symposium on Theoretical Aspects of Computer Science (STACS'21) (LIPIcs)*, Markus Bläser and Benjamin Monmege (Eds.), Vol. 187. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 23:1–23:19. https://doi.org/10.4230/LIPIcs.STACS.2021.23

[18] Mahdi Cheraghchi, Valentine Kabanets, Zhenjian Lu, and Dimitrios Myrisiotis. 2019. Circuit lower bounds for MCSP from local pseudorandom generators. In *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP'19)*. 39:1–39:14. https://doi.org/10.4230/LIPIcs.ICALP.2019.39

[19] Rodney G. Downey and Michael R. Fellows. 2013. *Fundamentals of Parameterized Complexity*. Springer. https://doi.org/10.1007/978-1-4471-5559-1

[20] Jürgen Forster. 2002. A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. Syst. Sci.* 65, 4 (2002), 612–625. https://doi.org/10.1016/S0022-0000(02)00019-3

[21] Merrick L. Furst, James B. Saxe, and Michael Sipser. 1984. Parity, circuits, and the polynomial-time hierarchy. *Math. Syst. Theory* 17, 1 (1984), 13–27. https://doi.org/10.1007/BF01744431

[22] Mika Göös, Pritish Kamath, Robert Robere, and Dmitry Sokolov. 2019. Adventures in monotone complexity and TFNP. In *Proceedings of the Innovations in Theoretical Computer Science Conference (ITCS'19)*. 38:1–38:19.

[23] Johan Håstad. 1998. The shrinkage exponent of de Morgan formulas is 2. *SIAM J. Comput.* 27, 1 (1998), 48–64. https://doi.org/10.1137/S0097539794261556

[24] Shuichi Hirahara. 2018. Non-black-box worst-case to average-case reductions within NP. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS'18)*. 247–258.

[25] Shuichi Hirahara. 2020. Non-disjoint promise problems from meta-computational view of pseudorandom generator constructions. In *Proceedings of the 35th Computational Complexity Conference (CCC'20) (LIPIcs)*, Shubhangi Saraf (Ed.), Vol. 169. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 20:1–20:47. https://doi.org/10.4230/LIPIcs.CCC.2020.20

[26] Shuichi Hirahara and Rahul Santhanam. 2017. On the average-case complexity of MCSP and its variants. In *Proceedings of the Computational Complexity Conference (CCC'17)*. 7:1–7:20. https://doi.org/10.4230/LIPIcs.CCC.2017.7

[27] Russell Impagliazzo, Raghu Meka, and David Zuckerman. 2019. Pseudorandomness from shrinkage. *J. ACM* 66, 2 (2019), 11:1–11:16. https://doi.org/10.1145/3230630

[28] Russell Impagliazzo, Ramamohan Paturi, and Michael E. Saks. 1997. Size-depth tradeoffs for threshold circuits. *SIAM J. Comput.* 26, 3 (1997), 693–707.

[29] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. 2001. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.* 63, 4 (2001), 512–530. https://doi.org/10.1006/jcss.2001.1774

[30] Russell Impagliazzo and Avi Wigderson. 2001. Randomness vs time: Derandomization under a uniform assumption. *J. Comput. Syst. Sci.* 63, 4 (2001), 672–688. https://doi.org/10.1006/jcss.2001.1780

[31] Emil Jerábek. 2009. Approximate counting by hashing in bounded arithmetic. *J. Symb. Log.* 74, 3 (2009), 829–860. https://doi.org/10.2178/jsl/1245158087

[32] Stasys Jukna. 2012. *Boolean Function Complexity—Advances and Frontiers*. Springer.

[33] Jørn Justesen. 1972. Class of constructive asymptotically good algebraic codes. *IEEE Trans. Inf. Theory* 18, 5 (1972), 652–656. https://doi.org/10.1109/TIT.1972.1054893

[34] Valentine Kabanets, Sajin Koroth, Zhenjian Lu, Dimitrios Myrisiotis, and Igor Carboni Oliveira. 2021. Algorithms and lower bounds for de morgan formulas of low-communication leaf gates. *ACM Trans. Comput. Theory* 13, 4 (2021), 23:1–23:37. https://doi.org/10.1145/3470861

[35] Mauricio Karchmer and Avi Wigderson. 1990. Monotone circuits for connectivity require super-logarithmic depth. *SIAM J. Discrete Math.* 3, 2 (1990), 255–265. https://doi.org/10.1137/0403021

[36] Ilan Komargodski and Raz Ran. 2013. Average-case lower bounds for formula size. In *Proceedings of the Symposium on Theory of Computing (STOC'13)*.

[37] Swastik Kopparty. 2011. On the complexity of powering in finite fields. In *Proceedings of the Symposium on Theory of Computing (STOC'11)*. 489–498. https://doi.org/10.1145/1993636.1993702

[38] Jan Krajíček. 2011. *Forcing with Random Variables and Proof Complexity*. Cambridge University Press.

[39] Richard J. Lipton and Ryan Williams. 2013. Amplifying circuit lower bounds against polynomial time, with applications. *Comput. Complex.* 22, 2 (2013), 311–343.

[40] Dylan M. McKay, Cody D. Murray, and R. Ryan Williams. 2019. Weak lower bounds on resource-bounded compression imply strong separations of complexity classes. In *Proceedings of the Symposium on Theory of Computing (STOC'19)*.

[41] Moritz Müller and Ján Pich. 2020. Feasibly constructive proofs of succinct weak circuit lower bounds. *Ann. Pure Appl. Log.* 171, 2 (2020). https://doi.org/10.1016/j.apal.2019.102735

[42] Igor Carboni Oliveira. 2019. Randomness and intractability in Kolmogorov complexity. In *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP'19)*. 32:1–32:14.

[43] Igor Carboni Oliveira, Ján Pich, and Rahul Santhanam. 2019. Hardness magnification near state-of-the-art lower bounds. In *Proceedings of the Computational Complexity Conference (CCC'19)*.

[44] Igor Carboni Oliveira and Rahul Santhanam. 2015. Majority is incompressible by $AC^0[p]$ circuits. In *Proceedings of the Conference on Computational Complexity (CCC'15)*. 124–157.

[45] Igor Carboni Oliveira and Rahul Santhanam. 2017. Conspiracies between learning algorithms, circuit lower bounds, and pseudorandomness. In *Proceedings of the Computational Complexity Conference (CCC'17)*. 18:1–18:49. https://doi.org/10.4230/LIPIcs.CCC.2017.18

[46] Igor Carboni Oliveira and Rahul Santhanam. 2018. Hardness magnification for natural problems. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS'18)*. 65–76. https://doi.org/10.1109/FOCS.2018.00016
[47] Ninad Rajgopal. 2021. *The Complexity of Meta-Computational Problems*. University of Oxford, UK.
[48] Alexander A. Razborov. 1985. Lower bounds on the monotone complexity of some Boolean functions. *Doklady Akademii Nauk SSSR* 281 (1985), 798–801. English translation in: Soviet Mathematics Doklady 31:354–357, 1985.
[49] Alexander A. Razborov. 1987. Lower bounds on the size of constant-depth networks over a complete basis with logical addition. *Mathematicheskie Zametki* 41, 4 (1987), 598–607.
[50] Alexander A. Razborov. 2015. Pseudorandom generators hard for k-DNF resolution and polynomial calculus resolution. *Ann. Math.* 181, 2 (2015), 415–472.
[51] Alexander A. Razborov and Steven Rudich. 1997. Natural proofs. *J. Comput. Syst. Sci.* 55, 1 (1997), 24–35. https://doi.org/10.1006/jcss.1997.1494
[52] Ben W. Reichardt. 2011. Reflections for quantum query algorithms. In *Proceedings of the Symposium on Discrete Algorithms (SODA'11)*. 560–569.
[53] Michael Sipser and Daniel A. Spielman. 1996. Expander codes. *IEEE Trans. Information Theory* 42, 6 (1996), 1710–1722. https://doi.org/10.1109/18.556667
[54] Roman Smolensky. 1987. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the Symposium on Theory of Computing (STOC'87)*. 77–82.
[55] Aravind Srinivasan. 2003. On the approximability of clique and related maximization problems. *J. Comput. Syst. Sci.* 67, 3 (2003), 633–651.
[56] Avishay Tal. 2014. Shrinkage of De Morgan formulae by spectral techniques. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS'14)*. 551–560. https://doi.org/10.1109/FOCS.2014.65
[57] Avishay Tal. 2017. Formula lower bounds via the quantum method. In *Proceedings of the Symposium on Theory of Computing (STOC'17)*. 1256–1268. https://doi.org/10.1145/3055399.3055472
[58] Avishay Tal. 2017. Tight bounds on the fourier spectrum of AC^0. In *Proceedings of the Computational Complexity Conference (CCC'17)*. 15:1–15:31. https://doi.org/10.4230/LIPIcs.CCC.2017.15
[59] Luca Trevisan and Tongke Xue. 2013. A derandomized switching lemma and an improved derandomization of AC^0. In *Proceedings of the Conference on Computational Complexity (CCC'13)*. 242–247. https://doi.org/10.1109/CCC.2013.32
[60] Salil P. Vadhan. 2012. Pseudorandomness. *Found. Trends Theor. Comput. Sci.* 7, 1–3 (2012), 1–336. https://doi.org/10.1561/0400000010
[61] Andrew Chi-Chih Yao. 1989. Circuits and local computation. In *Proceedings of the Symposium on Theory of Computing (STOC'89)*. 186–196. https://doi.org/10.1145/73007.73025

Received July 2021; revised January 2022; accepted May 2022